UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Head Office: Università degli Studi di Padova

Department of Industrial Engineering

Ph.D. COURSE IN INDUSTRIAL ENGINEERING

CURRICULUM: CHEMICAL AND ENVIRONMENTAL ENGINEERING

34th SERIES

DIGITALIZING PHARMACEUTICAL DEVELOPMENT AND MANUFACTURING: ADVANCED MATHEMATICAL MODELING FOR OPERATION DESIGN, PROCESS MONITORING AND PROCESS CONTROL

Coordinator: Ch.mo Prof. Giulio Rosati

Supervisor: Ch.mo Prof. Massimiliano Barolo

Ph.D. student: Francesco Destro

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Industrial Engineering)
Curriculum: Chemical and Environmental Engineering

at the
University of Padova
2021

A mia sorella Caterina

Foreword

The research project presented in this Dissertation has involved the intellectual and financial support of many people and institutions, to whom the author is enormously grateful. Most of the research activity reported and discussed in the Dissertation has been conducted at CAPE-Lab, Computer-Aided Process Engineering Laboratory (Department of Industrial Engineering, University of Padova, Italy), under the supervision of Prof. Massimiliano Barolo. Part of the work has been carried out at Process Systems Enterprise Ltd. (London, UK) during a 3-month stay under the supervision of Prof. Costas Pantelides, and part has been carried out at Purdue University (West Lafayette, IN; USA) during a 9-month stay under the supervision of Prof. Zoltan Nagy.

The research discussed in this Dissertation has been supported financially by Fondazione CARIPARO. Additional financial support has been provided by Fondazione Aldo Gini, by the European Commission through an Erasmus+ grant, by Process Systems Enterprise Ltd., and by the United States Food & Drug Administration through grant U01FD006738. All sponsors are gratefully acknowledged by the author.

All the material presented in this Thesis is original, except for explicit references provided by the author.

A benchmark simulator of a continuous carousel for intensified filtration-drying of crystallization slurries for Quality-by-Design and Quality-by-Control studies stem out from the research work, and will be made publicly available at the following link:

www.github.com/francescodestro/ContCarSim

The full list of publications originated from this Dissertation is reported below.

CONTRIBUTIONS IN PEER-REVIEWED JOURNALS

Destro, F., I. Hur, V. Wang, M. Abdi, X. Feng, E. Wood, S. Coleman, P. Firth, A. Barton, M. Barolo, and Z. K. Nagy (2021). Mathematical modeling and digital design of an intensified filtration-washing-drying unit for pharmaceutical continuous manufacturing. *Chem. Eng. Sci.* **244** (23), 116803.

Destro, F., S. García Muñoz, F. Bezzo, and M. Barolo (2021). Improving powder feeding monitoring in continuous solid dosage forms manufacturing through state estimation. *Int. J. Pharm.*, **605**, 120808.

Destro, F., P. Facco, S. García-Muñoz, F. Bezzo, and M. Barolo (2020). A hybrid framework for process monitoring: Enhancing data-driven methodologies with state and parameter estimation. *J. Process Control*, **92**, 333-351.

CONTRIBUTIONS IN PEER-REVIEWED JOURNALS (in preparation)

Destro, F., Z. K. Nagy, and M. Barolo. A benchmark simulator for quality-by-design and quality-by-control studies in continuous pharmaceutical manufacturing – Intensified filtration-drying of paracetamol/ethanol slurries. *In preparation*

Destro, F. and M. Barolo. Trends and perspectives on the modernization of pharmaceutical development and manufacturing: the role of mathematical modeling. *In preparation*

Destro, F., Z. K. Nagy, and M. Barolo. Quality-by-Control of intensified continuous filtration-drying of active pharmaceutical ingredients. *In preparation*

CONTRIBUTIONS IN PEER-REVIEWED CONFERENCE PROCEEDINGS

Destro, F., V. Wang, M. Abdi, X. Feng, E. Wood, S. Coleman, P. Firth, A. Barton, M. Barolo, and Z. K. Nagy. An intensified unit for continuous integrated filtration, washing and drying of drug substances: mathematical modelling and design space identification. *IFAC PapersOnLine*, **54**, 85-90.

Destro, F., V. Wang, M. Abdi, X. Feng, E. Wood, S. Coleman, P. Firth, A. Barton, M. Barolo, and Z. K. Nagy (2021), Quality-by-Control of continuous drug substance isolation: study on a novel unit for integrated filtration-drying, In: *Computer-Aided Chemical Engineering* **50**, *Proc. of the 31st European Symposium on Computer Aided Process Engineering* (M. Türkay, R. Gani, Eds.), Elsevier, Amsterdam (The Netherlands), 1363-1369.

Destro, F., A. J. Salmon, P. Facco, C. C. Pantelides, F. Bezzo, M. Barolo (2020). Monitoring a segmented fluid bed dryer by hybrid data-driven/knowledge-driven modeling. *IFAC-PapersOnLine*, **53**, 11638-11643.

CONTRIBUTIONS IN BOOK CHAPTERS

Laky, D. J., D. M. Casas-Orozco, F. Destro, M. Barolo, Z. K. Nagy (In press), Integrated synthesis, crystallization and drying of active pharmaceutical ingredients: a model-based framework for process optimization and control. Z. K. Nagy, R. Ramandrachan, P. Pardalos (Ed.), In: *Optimization of Pharmaceutical processes, models and methods*, Springer Nature, Basingstoke (UK).

CONFERENCE PRESENTATIONS

Destro, F., V. Wang, M. Abdi, X. Feng, E. Wood, S. Coleman, P. Firth, A. Barton, M. Barolo, and Z. K. Nagy (2021), An intensified unit for continuous integrated filtration, washing and drying of drug substances: mathematical modelling and design space identification. Oral presentation at: Advanced Control of Chemical Processes 2021 – IFAC ADCHEM 2021), June 13-16 (Virtual).

Destro, F., V. Wang, M. Abdi, X. Feng, E. Wood, S. Coleman, P. Firth, A. Barton, M. Barolo, and Z. K. Nagy (2021), Quality-by-Control of continuous drug substance isolation: study on a novel unit for integrated filtration-drying. Oral presentation at: 31st European Symposium on Computer Aided Process Engineering, June 6-9 (Virtual).

Destro, F., S. García-Muñoz, F. Bezzo, and M. Barolo (2020). Monitoring loss-in-weight feeders through state estimation. Presented at: the 2020 Virtual AIChE Annual Meeting, November 16-20 (Virtual).

Destro, F., I. Hur, D. M. Casas-Orozco, S. L. Lee, M. Abdi, X. Feng, E. Wood, M. Barolo, and Z. K. Nagy (2020). Digital design of an intensified filtration-drying unit for pharmaceutical upstream manufacturing. Oral presentation at: the 2020 Virtual AIChE Annual Meeting, November 16-20 (Virtual).

Destro, F., A. J. Salmon, P. Facco, C. C. Pantelides, F. Bezzo, and M. Barolo (2020). Monitoring a segmented fluid bed dryer by hybrid data-driven/knowledge-driven modeling. Oral presentation at: 21st IFAC World Congress, July 12-17 (Virtual).

Destro, F., P. Facco, S. García-Muñoz, F. Bezzo, and M. Barolo (2019). A hybrid approach to process monitoring. Poster presentation at: Convegno GRICU 2019. June 30-July 3, Mondello (Palermo, Italy).

AWARDS

Young Author Award Finalist**,** 11th IFAC Symposium on Advanced Control of Chemical Processes 2021 – IFAC ADCHEM 2021, June 13-16 (2021), Virtual

Best Poster Award, Convegno GRICU 2019. June 30-July 3, Mondello (Palermo, Italy).

Abstract

For pharmaceutical companies, the economic return on investments on research and development has been decreasing in recent years, mainly due to the large cost (~$2 billion) and time (~10 years) for bringing a new product to the market in the latest years. At the same time, an alarming number of drug shortages and recalls for quality concerns has been registered by regulators. These events affect companies, from the financial side, but also patients, who might experience increasingly large costs for drugs, or unavailability of essential medicines. The lack of adoption of modern technology and approaches for pharmaceutical development and manufacturing is acknowledged as a main actor for the occurrence of drug shortages and recalls, and for the overwhelming increase of the time- and cost-to-market for new products. A recent example is the sluggish rollout of COVID-19 vaccines, which has been significantly affected by technological limitations in pharmaceutical development and manufacturing, especially regarding process scale-up.

In the early 2000s, a modernization momentum of the sector was initiated. Regulators such as the United States Food & Drug Administration and the European Medicines Agency created a series of initiatives, the most impactful one being the Quality-by-Design (QbD) framework. Within the QbD initiative, regulators defined a novel pharmaceutical development and manufacturing paradigm, rooted in product and process understanding and based on sound science and quality risk management. Under the previous so-called Quality-by-Testing approach, the end-product was directly tested for verifying its quality. In a QbD approach, instead, the quality is proactively designed into the product. This transition represents a main step forward, as it comes with the realization that increased testing does not inherently bring improvement to the product quality.

Since the release of the QbD initiative, pharmaceutical development and manufacturing have undergone a strong modernization. However, much effort is still needed by the pharmaceutical sector to catch up with other industries on the adoption of modern development and manufacturing technologies. Recently, QbD is evolving towards a new phase, that features the adoption of novel emerging technology, the most important ones being continuous processing, active (i.e., closed-loop) quality control and increased use of mathematics.

Mathematical modeling can be used for developing digital tools pivotal to the efficient and rapid implementation of QbD, and its adoption has also been recommended by regulators with dedicated guidelines. Mathematical methodologies can support all stages of the pharmaceutical life cycle, and enable the implementation of continuous processing and active quality control.

Within this context, the role and expertise of chemical engineers, especially of the process systems engineering field, are of utmost importance.
The objective of this Dissertation is to promote the use of advanced mathematical modeling techniques within pharmaceutical development and manufacturing environments to:

- reduce pharmaceutical development time and cost;
- increase the efficiency and the robustness of pharmaceutical manufacturing.

These objectives are achieved by developing and/or implementing mathematical methodologies in key areas of pharmaceutical development and manufacturing: operation design, process monitoring and process control. The case studies space across the whole pharmaceutical flowsheet, but are particularly focused on continuous manufacturing processes. Applications of mathematical modeling are specifically addressed to tackle current bottlenecks towards the transition to end-to-end continuous pharmaceutical processing.

With respect to **innovative applications of mathematical modeling for pharmaceutical process monitoring**, a proof of concept of a novel approach for monitoring the powder mixture composition in continuous direct compression lines is presented. The composition of the powder mixture fed to direct compression lines is typically not directly measured, but obtained as calculation from the mass flow delivered by each of the feeders supplying the blend ingredients. The powder mass flow is, in turn, estimated numerically from the time series of (noisy) measurements of powder net weight in the hopper of the feeders. Proprietary algorithms embedded in the feeder software filter such mass flow estimations through statistical filters, to counteract noise propagation from the net weight measurements. In the Dissertation, an alternative approach to mass flow (and, in turn, to powder mixture composition) estimation is proposed, based on state estimation. It is demonstrated through data from a pilot plant that the state estimator can effectively reconcile noisy measurements, and that it provides more accurate powder composition estimations, when compared to traditional approaches relying on statistical filtering of noise.

A **novel methodology for process monitoring**, based on *hybrid* data-driven/knowledge-driven modelling, is introduced in this Dissertation. The proposed approach merges traditional standalone data-driven and knowledge-driven process monitoring approaches, taking the advantages of both. In practice, the hybrid monitoring approach consists in data-driven monitoring of a data matrix that includes i) "actual" data coming from sensor measurements, and ii) "digital" data, coming from a state estimator (based on a first-principles model of the system under investigation). When benchmarked against conventional monitoring techniques, the conceived hybrid monitoring approach shows superior fault detection and diagnosis performances, even when the first-principles model is affected by process/model mismatch. Hybrid monitoring is particularly interesting for pharmaceutical applications, as it perfectly fits

to the QbD paradigm of factoring the available knowledge on the physics of the process into the control strategy.

On the **digital operation design** side, a mechanistic model for a novel continuous integrated filter-dryer carousel system, scarcely studied before, is developed, and used for determining the probabilistic design space of the unit (i.e., the region of operating conditions and raw material properties that, under a certain probability, allow attaining the product quality). The carousel can continuously filter, wash and dry crystallization slurry streams into dry crystals cakes, tackling a current technological gap for the implementation of **end-to-end continuous processing**. For a set of feed conditions and control inputs, the developed model calculate the solvents and impurities content in the cake (product critical quality attributes) across carousel operation. Before of using the model for the digital design of the unit operation, filtration and drying experiments are carried out, respectively, on a Nutsche filter and on a thermogravimetric analyser, for model calibration and validation.

Following a **life cycle approach to process modeling**, the model used for designing the carousel operation is then further developed and tailored for **process control** applications on the unit. An advanced real time simulator is developed, with computational routines simulating sensors and actuators present in physical carousels. The simulator supports the implementation of control loops for control strategy testing. Filtration and drying experiments on a paracetamol/ethanol slurry system are carried out on a pilot scale carousel, to calibrate and validate the simulator. For the first time, a closed-loop control strategy is proposed for this type of units, based on a novel Quality-by-Control framework. The proposed control strategy also includes advanced model-based routines, such as state estimation and real time optimization, which are based on the developed mechanistic model of the carousel. The conceived control system is tested on the simulator, under a set of disturbances known to affect the unit operation (e.g., filter mesh fouling), and demonstrated superior control performance compared to traditional QbD control strategies, based on open-loop quality control within the design space.

The results presented and discussed in this Dissertation make several steps forward in the journey to adopt model-based methodologies for modernizing pharmaceutical development and manufacturing. Enabling technologies for the novel Quality-by-Control paradigm and for the transition to end-to-end continuous manufacturing have been developed. In particular, the presented results are expected to promote the adoption of advanced fault detection and diagnosis, digital operation design and closed-loop quality control routines in the pharmaceutical industry.

Riassunto

L'industria farmaceutica ha dato un enorme contributo alla società negli ultimi decenni, con un aumento dell'aspettativa di vita globale di 20 anni negli ultimi 50 anni ed una quota significativa del prodotto interno lordo mondiale generata dalle aziende farmaceutiche (per esempio, circa il 2,2% nel 2017).

Recentemente, il settore farmaceutico è stato sottoposto ad un numero crescente di sfide. Per le aziende farmaceutiche, il ritorno economico sugli investimenti in ricerca e sviluppo è diminuito negli ultimi anni, soprattutto a causa dei crescenti costi (~2 miliardi di dollari) e tempi (~10 anni) per portare un nuovo prodotto sul mercato negli ultimi anni. Allo stesso tempo, le autorità regolatorie hanno riportato un numero allarmante di carenze e richiami di farmaci per problemi di qualità. Questi eventi colpiscono le aziende farmaceutiche, dal lato finanziario, ma anche i pazienti, che rischiano di essere sottoposti a costi sempre più elevati per i farmaci, o possono sperimentare indisponibilità di farmaci essenziali. La mancanza di adozione di tecnologie e approcci moderni per lo sviluppo e la produzione di farmaci è riconosciuta come una causa principale per il verificarsi di carenze e richiami di farmaci e per l'aumento considerevole del tempo e dei costi di commercializzazione di nuovi prodotti. Consideriamo, per esempio, il recente caso del lancio dei vaccini COVID-19. Nonostante l'entusiasmo iniziale dopo la scoperta dei vaccini a fine 2020/inizio 2021, ci si è resi conto rapidamente che l'immunità di gregge su scala globale sarebbe stata raggiunta solo dopo anni. Un ruolo importante nel ritardo del lancio dei vaccini è stato giocato dalle attuali limitazioni tecnologiche nello sviluppo e nella produzione farmaceutica, specialmente per quanto riguarda lo scale-up del processo.

Gli attuali limiti tecnologici dello sviluppo e della produzione farmaceutica sono noti da tempo. Nei primi anni 2000, è stato avviata un'iniziativa di modernizzazione del settore farmaceutico, chiamata *Quality-by-Design* (QbD), da parte di enti regolatori come la *United States Food & Drug Administration* e la *European Medicines Agency.* Nell'ambito dell'iniziativa QbD, i regolatori hanno definito un nuovo paradigma di sviluppo e produzione farmaceutica, radicato nella comprensione del prodotto e del processo e basato su una solida scienza e sulla gestione del rischio di qualità. Secondo il precedente approccio chiamato *Quality-by-Testing*, la qualità del prodotto veniva testata direttamente sul prodotto finale. In un approccio QbD, invece, la qualità è progettata proattivamente nel prodotto. Questa transizione rappresenta un importante passo avanti.

Dal lancio dell'iniziativa QbD, lo sviluppo e la produzione farmaceutica hanno subito una forte modernizzazione. Tuttavia, sono ancora necessari molti sforzi da parte del settore farmaceutico per mettersi al passo con le altre industrie sull'adozione delle moderne tecnologie di sviluppo e

produzione. Recentemente, l'iniziativa QbD sta evolvendo verso una nuova fase, che prevede l'adozione di nuove tecnologie emergenti: la produzione in continuo, il controllo ad anello chiuso della qualità del prodotto, e un maggiore uso della modellazione matematica per lo sviluppo e la produzione farmaceutici.
La modellazione matematica può essere utilizzata per sviluppare strumenti digitali fondamentali per l'implementazione efficiente e rapida dell'iniziativa QbD, e la sua adozione è stata raccomandata anche dagli enti regolatori con linee guida dedicate. Le metodologie matematiche possono supportare tutte le fasi del ciclo di vita farmaceutico, e consentire l'implementazione della produzione in continuo e del controllo ad anello chiuso della qualità del prodotto. In questo contesto, il ruolo e le competenze degli ingegneri chimici sono di massima importanza.
L'obiettivo di questa Dissertazione è sviluppare e applicare metodi matematici per:
- ridurre i tempi e i costi dello sviluppo farmaceutico;
- aumentare l'efficienza e la robustezza della produzione farmaceutica.
Questi obiettivi sono raggiunti sviluppando e/o implementando metodologie matematiche in aree chiave dello sviluppo e della produzione farmaceutica: progettazione delle operazioni produttive, monitoraggio di processo e controllo di processo. I casi di studio di questa Dissertazione sono particolarmente focalizzati su processi produttivi in modalità continua.

Per quanto riguarda **applicazioni innovative della modellazione matematica al monitoraggio di processi farmaceutici**, viene presentato un nuovo approccio per il monitoraggio della composizione della miscela di polveri in linee di pastigliatura continua. In impianti industriali, ogni ingrediente della miscela di polveri da inviare alla pastigliatrice viene fornito al processo attraverso un alimentatore separato. La composizione della miscela di polvere alimentata a valle dagli alimentatori non viene misurata direttamente, ma ottenuta dal rapporto delle portate massive erogate da ciascun alimentatore. La portata massiva della polvere è, a sua volta, stimata numericamente da misurazioni (rumorose) del peso netto della polvere nella tramoggia degli alimentatori. Algoritmi proprietari incorporati nel software degli alimentatori filtrano e uniformano tali stime di portata massiva attraverso filtri statistici, nel tentativo di contrastare la propagazione del rumore dalle misurazioni del peso netto della polvere. Nella Dissertazione, viene proposto un approccio alternativo alla stima della portata di polvere (e, a di conseguenza, della composizione della miscela di polveri), basato sulla stima di stato. La stima dello stato è una famiglia di tecniche matematiche che, dato un modello matematico di un processo e misure in tempo reale da un impianto, stimano gli stati non misurati del Sistema, e forniscono stime delle misure di impianto attenuate dal rumore. In questa Dissertazione, per la prima volta viene testato uno stimatore di stato su dati di una linea pilota di pastigliatura, per monitorare la composizione della miscela di polveri nell'impianto. Si dimostra che lo stimatore di stato può riconciliare efficacemente le misurazioni rumorose, e che

fornisce stime più accurate della composizione della polvere, rispetto agli approcci tradizionali basati sull'utilizzo di filtra statistici per il rumore.

In questa Dissertazione, viene anche introdotta una **nuova metodologia per il monitoraggio del processo, basata su modellazione ibrida**. L'approccio proposto fonde i tradizionali approcci di monitoraggio di processo basati unicamente su modelli a base di dati o su modelli a principi primi, prendendo i vantaggi di entrambi. L'approccio di monitoraggio ibrido consiste nel convenzionale monitoraggio basato su dati, ma di una matrice che include i) dati "reali" provenienti di sensori di processo, e ii) dati "digitali", provenienti da uno stimatore di stato (basato su un modello di principi primi del sistema in esame). Confrontato con tecniche di monitoraggio convenzionali, l'approccio di monitoraggio ibrido proposto dimostra migliori prestazioni rispetto al rilevamento e alla diagnosi dei guasti, anche quando il modello a principi primi presenta errori di modellazione significativi. Il monitoraggio ibrido è particolarmente interessante per le applicazioni farmaceutiche, in quanto si adatta perfettamente al paradigma QbD di incorporare la conoscenza disponibile sulla fisica del processo nel sistema di controllo.

Per quanto riguarda la **progettazione digitale delle operazioni produttive**, è stato sviluppato un modello a principi primi per un nuovo sistema continuo integrato di filtrazione-essiccamento a carosello. Il modello è stato utilizzato per determinare il *design space* dell'unità (cioè, la regione delle condizioni operative e delle proprietà delle materie prime che, con una determinata probabilità, permettono di raggiungere la qualità del prodotto desiderata). Il carosello può filtrare, lavare ed essiccare in maniera continua *slurries* prodotti da cristallizzatori, trasformandoli in torte di cristalli secchi. Il carosello è una nuova tecnologia che mira a superare l'assenza di unità di processo funzionanti in continuo per la filtrazione ed essiccamento di *slurries*, operazioni necessarie per sviluppare processi farmaceutici completamente continui, dall'inizio alla fine. Fissate le condizioni operative e di processo, il modello sviluppato calcola il contenuto di solventi e impurità nella torta (specifiche di qualità del prodotto) durante il funzionamento del carosello.

Il modello usato per progettare le operazioni produttive del carosello è stato poi ulteriormente sviluppato ed utilizzato per applicazioni di **controllo di processo** sull'unità. È stato sviluppato un simulatore avanzato, con funzioni digitali che simulano i sensori e gli attuatori presenti nei caroselli fisici. Il simulatore può essere utilizzato per testare differenti strategie di controllo. Esperimenti di filtrazione ed essiccamento su uno *slurry* di paracetamolo/etanolo sono stati eseguiti su un carosello in scala pilota, per calibrare e convalidare il simulatore. Per la prima volta, è stata proposta una strategia di controllo ad anello chiuso per questo tipo di unità, basata sul recente paradigma *Quality-by-Control*. La strategia di controllo proposta include anche metodologie avanzate basate su modello, come la stima dello stato e l'ottimizzazione in tempo reale. Il sistema di controllo concepito è stato testato sul simulatore, in presenza di una serie di

disturbi noti per influenzare il funzionamento dell'unità (ad esempio, lo sporcamento delle maglie dei filtri del carosello), ed ha dimostrato prestazioni di controllo migliori rispetto a strategie di controllo QbD tradizionali, basate sul controllo di qualità ad anello aperto.

I risultati presentati e discussi in questa Dissertazione rappresentano un significativo passo avanti per l'adozione di metodologie basate su modelli per modernizzare lo sviluppo e la produzione farmaceutica. Sono state sviluppate tecnologie per il nuovo paradigma *Quality-by-Control* e per la transizione alla produzione farmaceutica in continuo. In particolare, i risultati presentati promuovono l'adozione di metodologie avanzate per il rilevamento e la diagnosi avanzata di guasti, per la progettazione digitale delle operazioni produttive e per il controllo di processo ad anello chiuso nell'industria farmaceutica.

Table of contents

List of symbols

Acronyms

API	=	active pharmaceutical ingredient
BLA	=	biologic license application
CAGR	=	compound annual growth rate
cGMPs	=	current good manufacturing practices
CMA	=	critical material attribute
CQA	=	critical quality attribute
CPP	=	critical process parameter
CTA	=	clinical trial application
CSD	=	crystals size distribution
CV	=	control variable
DD	=	data-driven
DS	=	design space
EFPIA	=	European Federation of Pharmaceutical Industries and Associations
EIOT	=	extended iterative optimization technology
EKF	=	Extended Kalman Filter
EMA	=	European Medicines Agency
EU	=	European Union
FDA	=	United States Food & Drug Administration
FMEA	=	failure mode effects analysis
FMECA	=	failure mode effects and criticality analysis
FOPDT	=	first-order-plus-dead-time
FPM	=	first-principles model
GSA	=	global sensitivity analysis
HPLC	=	high performance liquid chromatography
ICH	=	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
IFPMA	=	International Federation of Pharmaceutical Manufacturers & Associations
IPA	=	isopropyl alcohol
IVIVC	=	in vivo vs in vitro correlation

IND	=	investigational new drug application
KASA	=	knowledge-aided assessment & structured application
KD	=	knowledge-driven
LVM	=	latent-variables model
MAA	=	marketing authorisation application
MBDoE	=	model-based design of experiments
MHE	=	moving-horizon estimator
MLE	=	maximum likelihood estimation
MPC	=	model predictive control
MSPC	=	multivariate statistical process control
NDA	=	new drug application
NIR	=	near-infrared
NOC	=	normal operating conditions
PAT	=	process analytical technology
PCM	=	paracetamol
PDE	=	partial differential equation
QbC	=	Quality-by-Control
QbD	=	Quality-by-Design
QTTP	=	quality target product profile
PC	=	principal component
PCA	=	principal component analysis
PET	=	poly-ethylene terephthalate
PLS	=	projection to latent structures
PRBS	=	pseudo-random binary sequences
R&D	=	research & development
ROI	=	return on investment
RTO	=	real time optimization
RTRT	=	real time release testing
SPC	=	statistical process control
SPE	=	squared prediction error
SSE	=	sum of squared errors
TGA	=	thermogravimetric analyzer
UK	=	United Kingdom of Great Britain and Northern Ireland
USA	=	United States of America

Chapter 1

Motivation and state of the art*

This Chapter provides an overview of the current state of pharmaceutical development and manufacturing, and produces the motivation of this Dissertation. Current state and future trends of the pharmaceutical industry from a socio-economic perspective are first described. After briefly introducing the pre-launch and post-launch life stages of a new pharmaceutical product, the pharmaceutical regulatory framework is outlined and its historical evolution is addressed. Pharmaceutical emerging technology and the enabling role of process systems engineering for its implementation are then evaluated. In particular, the role of mathematical modeling for the modernization of pharmaceutical development and manufacturing is critically discussed. An outlook on the state-of-the-art mathematical modeling for supporting operation design, process monitoring and process control is provided. Finally, the innovative objectives of this research are described, and a roadmap for the Dissertation is drawn.

1.1 The pharmaceutical industry: a socio-economic outlook

The pharmaceutical industry gave a tremendous contribution to the well-being of people in the last decades, with the global life expectancy increasing as much as 20 years over the last 50 years (IFPMA, 2021), and the life expectancy of European citizens rising of even 30 years across the last century (EFPIA, 2021). High quality medicine and vaccines have a paramount role in the sustainability of well-functioning healthcare systems, as they significantly contribute to cost reduction in other healthcare spending sectors, such as mid- and long term hospitalizations costs (IFPMA, 2021). In addition to the immediate benefits to the patients' lives, the pharmaceutical industry gives a substantial contribution to the world's economy. In 2017, the pharmaceutical sector generated, through direct and indirect effects, about 2.2% of the world's gross domestic product, corresponding to approximately $1,838 billion (WifOR, 2020). The workforce directly employed by the pharmaceutical industry includes about 5.5 million people on a global scale, while figures rise to 74.3 million when the employment

* Destro, F. and M. Barolo. Trends and perspectives on the modernization of pharmaceutical development and manufacturing: the role of mathematical modeling. *In preparation*

induced in the supply chain and in other sectors is considered.

Despite the COVID-19 pandemic significantly impacted the healthcare sector and the economic forecasts for 2020 and 2021, the pharmaceutical industry economy is expected to follow a solid growth trend (EvaluatePharma, 2020a; IFPMA, 2021; IQVIA Institute, 2020). Due to the initial rush into the development of vaccines and treatments for COVID-19, almost 160 clinical trials for new drugs have been put on hold in 2020, but most of the research that had been suspended resumed since then (EvaluatePharma, 2020b). Also the 2020 drop in pharmaceutical product sales, forecasted by analysts to be of about $4.9 billion for the top 15 companies (EvaluatePharma, 2020a), is estimated to be temporary. Overall, the pre-pandemic factors influencing pharmaceutical products use and spending in the near- and middle term future remain substantially unchanged (IQVIA Institute, 2020).

The world drug market is expected to reach a size of $1.6 trillion in 2025 (Table 1.1), growing at a compound annual growth rate (CAGR) of 3-6% through 2025. These estimations exclude spending on COVID-19 vaccines, which are projected to be of about $157 billion, and to mainly involve the vaccination campaign to finish by 2022 (IQVIA Institute, 2020). The pharmaceutical market will grow at 7-10% CAGR in the so-called "pharmerging countries" (Table 1), 21 countries, such as Mexico, Poland and Argentina, with per capita income below $30,000 and an aggregate pharmaceutical growth in five years over $1 billion (IFPMA, 2021).

Table 1.1. *Worldwide pharmaceutical market: spending and growth rate in 2016-2020 and forecasts for 2021-2025 for selected countries and for the total market. Adapted from IQVIA Institute (2020). CAGR = compound annual growth rate.*

Countries	2020 SPENDING [$bn]	2016-2020 CAGR	2025 SPENDING [$bn]	2021-2025 CAGR
Global	1,265.2	4.6%	1580–1610	3–6%
Developed	959.5	3.8%	1130–1160	1.5–4.5%
10 Developed	847.2	3.8%	990–1020	1.5–4.5%
United States	527.8	4.2%	605–635	2–5%
Japan	88.2	-0.2%	75–95	-2–1%
Germany	54.9	5.3%	65–85	3.5–6.5%
France	36.3	2.4%	43–47	1–4%
Italy	33.3	4.2%	38–42	2–5%
United Kingdom	30.2	5.3%	38–42	2.5–5.5%
Spain	25.7	4.6%	28–32	1.5–4.5%
Canada	22.8	4.8%	28–32	2–5%
South Korea	16.2	6.8%	18–22	4.5–7.5%
Australia	11.8	3.3%	13–17	1–4%
Other developed	112.3	4.2%	125–155	2.5–5.5%
Pharmerging	290.8	7.4%	415–445	7–10%
China	134.4	4.9%	170–200	4.5–7.5%
Brazil	28.7	10.7%	43–47	7.5–10.5%
Russia	17.5	10.8%	33–37	11–14%
India	21.1	9.5%	28–32	7.5–10.5%
Other Pharmerging	89.1	9.6%	120–150	8.5–11.5%
Lower income Countries	15.0	3.9%	18–22	3–6%

Developed countries will grow at a lower 1.5-4.5% CAGR pace (Table 1.1), as gains from newly approved products will not outweigh losses due to patents expirations (IQVIA Institute, 2020). Despite the lower CAGR, developed countries will be responsible for the largest part of the market size increase, even though comparisons with the CAGRs in 2016-2020 (Table 1.1) show the growing importance of pharmerging countries for the pharmaceutical market. Looking at the prescription drug sales and market share for the top 10 pharma players in 2018 and to the forecasts for 2026 (Table 1.2), it emerges that, despite all companies present a positive CAGR, the total market share of the top 10 companies will shrink. Patent expiration is the main reason for this phenomenon. Pharmaceutical companies will have to diversify their portfolios by bringing new products to the market, to compensate for the loss of brand exclusivity. However, the return on investment (ROI) on research and development (R&D) for new medicines has been constantly decaying over the last years. Deloitte (2021) reports that, for a selected cohort of 15 biopharmaceutical companies, the ROI on R&D decreased from 7.2% in 2014 to just 1.6% in 2019, with a constant decay. A modest recovery to 2.5% was measured in 2020, but this result does not imply an inversion of the trend. As a matter of fact, the cost to bring a new product to the market is steadily growing (+67% from 2010 to 2019; Deloitte, 2019), due to increasing complexity of product and process development, and to the expanding cycle time for launching a new product.

Table 1.2. *Worldwide pharmaceutical market: prescription drug sales and market share in 2018 and forecasts for 2026 for the top 10 companies and for the total market. Adapted from EvaluatePharma, 2020b.*

Rank	Company	Prescription sales [$bn]			Market share		Rank change
		2019	2026	CAGR	2019	2026	
1	Roche	48.2	61.0	+3.4%	5.5%	4.4%	+0
2	Johnson & Johnson	40.1	56.1	+4.9%	4.6%	4.0%	+3
3	Novartis	46.1	54.8	+2.5%	5.3%	3.9%	-1
4	Merck & Co	40.9	53.2	+3.8%	4.7%	3.8%	+0
5	AbbVie	32.4	52.7	+7.2%	3.7%	3.8%	+2
6	Pfizer	43.8	51.1	+2.2%	5.0%	3.7%	-3
7	Bristol-Myers Squibb	25.2	44.7	+8.6%	2.9%	3.2%	+3
8	Sanofi	34.9	41.7	+2.6%	4.0%	3.0%	-2
9	AstraZeneca	23.2	41.0	+8.5%	2.7%	2.9%	+2
10	GlaxoSmithKline	31.3	40.8	+3.9%	3.6%	2.9%	-2
	Total Top 10	366.1	497.1	+4.5%	42.0%	35.8%	
	Other	505.8	893.1	+8.5%	58.0%	64.2%	
	Total	871.8	1,390.3	+6.9%	100.0%	100.0%	

To contrast the decreasing ROI on R&D and maintain their competitiveness, pharmaceutical companies are required to innovate their approaches to all the stages of the pharmaceutical life cycle. An additional crucial driver for pharmaceutical industry modernization comes from the large number of medicine shortages and recalls that occurred in recent years with severe economic and health consequences (ISPE, 2017; Natof and Pellegrini, 2021). More modern approaches to pharmaceutical development and manufacturing would lead to significantly less

alarming tolls of shortages and recalls, as also warned by the regulatory agencies (Yu and Kopcha, 2017). The change of paradigm consists in adopting Industry 4.0 and process systems engineering enabling technologies that have been revolutionizing the manufacturing sector in recent years (Isaksson et al., 2018; Kagermann et al., 2011). Immediate benefits of pharmaceutical industry modernization are, but are not limited to:

- cost reduction for drug discovery, product and process development and manufacturing;
- decrease of the cycle time between patent filing and product launch;
- faster and more efficient scale-up from the development to the manufacturing scale;
- greater product quality assurance;
- decrease of medicine shortages and recalls.

The advantages of this innovation effort are both for companies, in terms of ROI and of economic gain, and for patients, who can expect to have broader access to innovative and high quality medicines and vaccines at a lower price, with less incidence of potentially life-threatening events such as shortages and recalls. As a proof, let us consider the recent case of the COVID-19 vaccines. Shortly after that regulatory agencies across the world emitted emergency use authorizations for different sera in late 2020 and early 2021, it became clear that pharmaceutical companies where not ready to ramp up the vaccines production to quickly meet the worldwide demand (ECDC, 2021). The task was indeed difficult, considering the physical lack of an adequate number of suitable facilities. However, the shortage was also caused by the poor diffusion in the pharmaceutical sector of efficient scale-up approaches from the development to the manufacturing scale, which are instead established in most of the other manufacturing industries (Liu et al., 2020; Sarkis et al., 2021). In addition, outdated manufacturing technology still in use in many pharmaceutical plants can play a role in multiple batch rejections (Johnson & Johnson, 2021; Torjesen, 2021), further delaying vaccines rollout with tremendous implications on the health of the population and on the global economy.

The recent urgency of innovation of the pharmaceutical industry meets a modernization momentum in the sector that has been growing since the early 2000s, under several initiatives promoted by regulatory agencies, such as the United States Food & Drug Administration (FDA) and the European Medicines Agency (EMA), to improve pharmaceutical quality assurance. Actually, as the pharmaceutical industry is a strongly regulated sector, every modernizing change to a stage of the product life cycle must be designed in accordance with, and cannot be understood if decoupled from, the national and international regulations. For this reason, after a brief outline of the pre-launch and post-launch life stages of a pharmaceutical product in §1.2, in §1.3 an overview is provided on the pharmaceutical regulatory framework for the countries adhering to the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). The ICH was founded in Brussel in 1990 by representatives of the regulatory agencies of the United States of America (USA), of the European Union (EU) and of Japan, and has since then acquired new members.

1.2 The journey of a new pharmaceutical product

The launch of new molecular entitities and new biologicals entities, usually simply referred to as *biologicals*, is a costly process (Table 1.3), of average duration of 12-13 years since the discovery of the new active substance (EFPIA, 2021). Since approved drugs must compensate for the R&D expense of candidate drugs that do not obtain the final approval for commercialization, the average R&D expense for new drugs in 2019 was larger than the impressive amount of $3 billion (Table 1.3). Overall, the pharmaceutical industry is estimated to have invested in R&D about $179 billion in 2018, which has been claimed to be 7.3 times the R&D expense in the same year of the aerospace and defense industries, 6.5 times the one of the chemical industry, and 1.5 times that of the computer services industry (IFPMA, 2021). Nevertheless, pharmaceutical companies continue to imposingly invest in R&D to widen their portfolios, to maintain and increase their competitiveness. In 2020, more than 8000 active substances were in the different stages of the development pipeline.

Table 1.3. *Drugs approved by the FDA in the years 2007-2019: number and R&D expense per drug, calculated dividing the total R&D expense per year by the number of approved drugs. The 3 year lag R&D expense considers the R&D expense for 3 years prior to the drug approval instead of the same year. Adapted from EvaluatePharma* (2020a).

Year	# drugs approved by FDA	R&D expense per drug [$ bn]	R&D expense per drug – 3-year lag [$ bn]]
2007	25	4.9	-
2008	31	4.2	-
2009	35	3.6	-
2010	26	4.9	4.7
2011	35	3.9	3.7
2012	44	3.1	2.9
2013	35	4.0	3.7
2014	51	2.8	2.7
2015	56	2.7	2.4
2016	27	5.9	5.1
2017	55	3.1	2.6
2018	62	2.9	2.4
2019	53	3.5	3.0

Figure 1.1 presents the pre-launch and post-launch life stages of a pharmaceutical product, together with the breakdown into the different stages of the budget to be invested for the new product. Research involves mainly the first phase of the process, and includes two steps: *i)* drug discovery and *ii)* pre-clinical testing. During drug discovery, researchers attempt to find promising agents for fighting a disease using basic and innovative scientific methods. Then, the identified compounds are patented, and pre-clinical (or pre-human) testing starts. The active substance is tested for safety and efficacy *in-silico* and *in-vivo*, to decide whether to continue

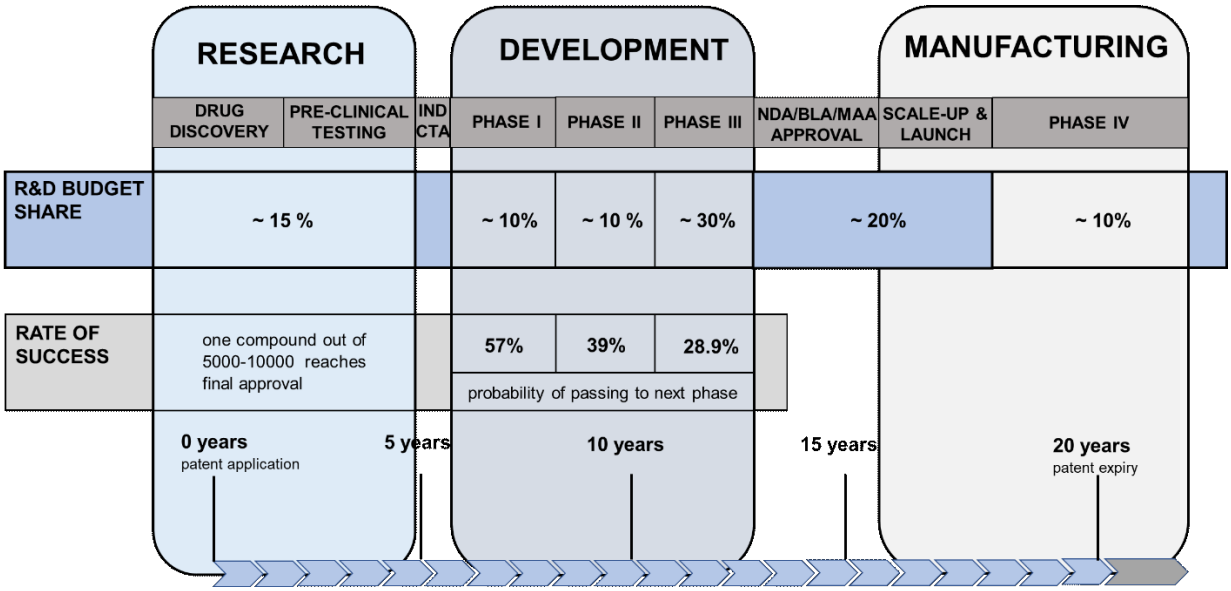


Figure 1.1. *Life stages of a pharmaceutical product. Data from EFPIA (2021) and IFPMA (2021).*

the approval process with clinical tests on humans. Examples of pre-clinical studies are tests on the acute and chronic toxicity properties and the pharmacokinetics of the compound. On average, only one out of 5,000-10,000 compounds that enter pre-clinical studies reaches the final approval (Lipsky and Sharp, 2001). Overall, the duration of the research phase is of 3-6 years, and it consumes, on average, 15.7% of the budget for bringing to the market a new product (EFPIA, 2021).

Compounds that seem promising after the pre-clinical tests enter the development stage, which consists of clinical trials on humans carried out in three phases. Before starting the actual trials, an application must be filed with regulators, such as the investigational new drug application (IND) with the FDA or the clinical trial application (CTA) with the EMA. Regulators review the application, containing information such as the pre-clinical tests results and the clinical trials design, and decide whether to approve the research on humans. During development, both the product and the process to manufacture it are designed. In phase I of the clinical trials, the candidate compound is tested on 20-100 volunteers, to understand the safety of the compound on humans and to design a suitable dose. Compounds passing phase I (average rate of success of 57%; IFPMA, 2021) go on to phase II, where the main focus is understanding if the candidate drug is effective. In phase II, the group of volunteers is typically of 100-500 patients, with an average success rate of 39% of passing to phase III (IFPMA, 2021). Successful results for phase II allow starting phase III clinical trials, during which the candidate drug is tested for safety and efficacy on a large number of volunteers (1,000-5,000). During phase III, candidate drugs are also benchmarked against the current standard treatment available in the market, if there is one. Chances for a new compound to pass phase III clinical trials are, on average, of 68% for phase III (IFPMA, 2021). Phase III clinical trials are the most expensive phase of the launch of a new

product (~30% of the budget), and require to manufacture a large amount of the candidate drug. The needed amount of active pharmaceutical ingredient (API) or drug product rapidly passes from a few grams for pre-clinical/phase I tests to hundreds of kilograms in phase III, due to the significant number of volunteers involved in this phase. The clinical material production ramp-up is typically conducted with an empiric approach through a series of scale-ups, aimed at delivering the supplies for the trials as fast as possible, rather than at developing robust manufacturing processes (Suresh and Basu, 2008). At the end of the trials, process development is frozen and the process is quickly scaled-up to the manufacturing scale for the final validation. Hence, the lack of robustness of the process in development scale propagates to the process in manufacturing scale, with severe consequences in manufacturing efficiency.

Overall, the duration of the development phase is of 6-7 years, and absorbs, on average, 47.4% of the budget for launching a new product (EFPIA, 2021). The cost for manufacturing the clinical trials supplies is a significant part of the overall cost of conducting the trial themselves, which also includes expenses such as reimbursements to volunteers and clinical staff. Suresh and Basu (2008) estimated that the cost for conducting clinical trials makes up for about 50-60% of the development budget alone, while the remaining 40-50% is the cost for product development, including for the manufacture of clinical trial supplies.

For compounds successfully passing the clinical trials, a submission for approval is submitted to regulators, such as the new drug application (NDA) or the biologic license application (BLA) to the FDA or the marketing authorisation application (MAA) to the EMA. This phase takes about 4.3% of the budget for the product launch. Regulators may take up to 2 years for reviewing the application, and deciding whether to give the launch authorization. In case of positive response, the approved drug enters the manufacturing stage. Periodic reviews (phase IV clinical trials) of safety and efficacy must be carried out after the product launch, and take up to 11.2% of the total budget (EFPIA, 2021). A patent for a drug expires 20 years after having been granted, although in some cases an extension of five years might be given. After patent expiration, profits decay because of competition with generics (IQVIA Institute, 2020).

Pharmaceutical companies, in order to increase their profit and the ROI on the R&D expense, should aim at *i)* reducing the R&D cost (up to 65% cumulative cost for product launch) , *ii)* reducing the product launch time and *iii)* modernizing scale-up and manufacturing approaches, also for clinical supplies manufacturing in the development stage. Product launch time reduction (aim *ii*) is particularly relevant: IBM (2005) estimated that, for a product with $1 billion annual peak sales in its life cycle, bringing the sales peak from the current time interval of 10 years after launch to 5 years after launch would increase the revenues from the product of $1.6 billion. Suresh and Basu (2008) estimated that, for the whole industry, development modernization can create between $12 and $18 billion a year of additional revenue, just by decreasing the time to market of new products. The introduction of science-based approaches for process development and scale-up (aim *iii*) is also pivotal to maximizing the manufacturing

efficiency, and minimizing the occurrence of shortages and recalls. Moving from empirical to science-based process development and scale-up since the development stage is fundamental to reduce the (significant) cost for making clinical trials supplies, and to deliver robust manufacturing processes. It has been estimated that cost savings in manufacturing for the whole industry in the range of $20 to $50 billion every year (Suresh and Basu, 2008) can be achieved by developing robust manufacturing processes.

There is consensus among analysts and experts of the field (Deloitte, 2020a, 2019a; IBM, 2005; McKinsey & Company, 2021) that the strategies and tools for achieving aims (*i-iii*) are:

- artificial intelligence, digitalization and mathematical modeling as accelerators for drug discovery (Deloitte, 2019b);
- cloud computing for speeding up information sharing among research collaborators, across companies and with health authorities worldwide (Deloitte, 2020a). Ten top biopharmaceutical companies sponsored the foundation of Accumulus Synergy, a cloud computing platform formed in 2020 to reduce regulatory review times and innovate data exchange between the biopharmaceutical industry and global health authorities (Accumulus Sinergy, 2021);
- adoption of innovative approaches for clinical trials design and execution (Deloitte, 2020b). These include the use of artificial intelligence, of telemedicine technology for trials decentralization and of real world evidence and data (data from patients databases or other private or public registries) for enhancing trials design and fastening the regulatory approval;
- implementation of the Quality-by-Design (QbD) paradigm (§1.3) in pharmaceutical development and manufacturing (IBM, 2005; Kourti and Davis, 2012; Yu and Kopcha, 2017);
- the use of emerging technologies (Fisher et al., 2016; O'Connor et al., 2016; Troup and Georgakis, 2013) in pharmaceutical development and manufacturing for advanced process design, monitoring and control, including continuous manufacturing, process modeling, data analytics, active process control, real-time release testing, PAT and advanced monitoring systems. These technologies will be detailed in §1.4.

1.3 The pharmaceutical regulatory framework

1.3.1 Historical evolution

The first regulations for the pharmaceutical industry came in response to fatal incidents. Safety studies for new drugs became mandatory in the United States only with the 1938 Food, Drug and Cosmetic Act (Ballentine, 1981), after Elixir Sulfanilamide, a drug for streptococcal

infections, caused more than 100 casualties in the United States in 1937. Efficacy studies for new medicines became a requirement in the United States only in 1962, as a consequence of the public interest in the topic drawn by the thalidomide-related birth defects of the early 1960's (Hamburg, 2012). Although regulations were put in place for the first time to assure safety and efficacy of drugs for the first time, the pharmaceutical industry started to become a highly regulated environment, reluctant to innovation and relying on brand exclusivity for making profits. The direct consequence became that in the early 2000s, as reported by the Wall Street Journal (Abboud and Hensley, 2003), in spite of cutting-edge new drug discoveries, pharmaceutical manufacturing technologies were still lagging far behind those of potato chips and soap industries. Regulators acknowledged this issue, and started pushing for a general modernization of the industry. The FDA laid its vision for pharmaceutical manufacturing innovation with the document "Current Good Manufacturing Practices (cGMPs) for the 21st century" (FDA, 2004a), aiming to promote more efficient and flexible manufacturing approaches that produce high quality drugs without the burden of extensive regulatory oversight. With the subsequent process analytical technology (PAT; FDA, 2004b) and QbD (ICH, 2009; Yu, 2008; Yu et al., 2014) initiatives, the FDA, together with the other regulators of the ICH, defined a novel pharmaceutical development and manufacturing paradigm, rooted in product and process understanding and based on sound science and quality risk management. The quality of a pharmaceutical product can be defined as the absence of contamination and the suitability for the therapeutic use promised in the label to the consumer (Woodcock, 2004). Under the previous so-called Quality-by-Testing approach, the end-product was directly tested for verifying its quality. In a QbD approach, instead, the quality is proactively designed into the product (ICH, 2009; Juran, 1992). This transition represents a main step forward, as it comes with the realization that increased testing does not inherently bring improvement to the product quality. Regulators understood that they can assess the pharmaceutical quality delivered by a running plant only during inspections (Yu and Woodcock, 2015), and that it is far more convenient to encourage manufacturers to develop processes capable of robustly building the quality into the product.

Since the launch of the QbD initiative, substantial progress has been made towards this direction. Still, much challenges remain ahead. During the global heparin crisis of 2007-2008, a contaminant included in the drug product caused several casualties (Liu et al., 2009). The contaminant had passed the quality control system undetected, proving that there is still much space of improvement in pharmaceutical quality control. Another critical issue involves drug shortages and recalls, which recently have been registered in unacceptably large numbers (FDA, 2013a; O'Connor et al., 2016). The unprecedented shortages and recalls volume indicates failures in routine manufacturing operations and scale-up procedures, often combined with the use of old-fashioned equipment and improper quality management systems (Yu and Woodcock, 2015). Actually, manufacturing and quality issues were found to cause more than 60% of all

biologicals shortages (FDA, 2013a). In response to these and other events, regulators recently made further moves to promote the modernization of pharmaceutical development and manufacturing, with a renewed focus on quality. The FDA created a new Office of Pharmaceutical Quality within the Center for Drug Evaluation and Research, with the purpose of integrating review, inspection and research activities for boosting quality enhancement across the pharmaceutical life cycle. The FDA has also established a new pharmaceutical quality assessment (Knowledge-aided Assessment & Structured Application, KASA; Yu et al., 2019) system, to modernize the procedure of quality assessment of regulatory drug applications. Under the KASA system, modern structured and quantitative approaches are adopted for evaluating drug applications, with the purpose of increasing the consistency among regulatory approvals for different products and the effectiveness of the regulatory oversight across the product life cycle.

The ultimate aim encouraged by regulators for tackling the current challenges in pharmaceutical quality control is the implementation of robust pharmaceutical processes, under the so-called "six sigma quality" (Yu and Kopcha, 2017). Under a six sigma quality system, established in most manufacturing industries, there are six process standard deviations between the process mean of each quality variable and the relevant nearest specification limit. Six sigma quality eventually results in no more than 3.4 quality failures occurring per million opportunities, compared to 308,537 defects per million opportunities under two sigma quality. Current pharmaceutical manufacturing is, in most of the cases, two to three sigma quality (Yu and Kopcha, 2017). To achieve the six sigma standard, it is necessary to introduce novel technology in pharmaceutical development and manufacturing. Such emerging technology, new for the pharmaceutical industry, but not necessarily new to other industries, has been encouraged through the recent Emerging Technology Program (FDA, 2017; O'Connor et al., 2016) by the FDA, and through the EU Innovation Network in Europe (EMA, 2018). There is agreement among regulators, practitioners and academics (Am Ende, 2019; Fisher et al., 2016; Su et al., 2017) that the emerging technologies to focus on to deliver next-generation pharmaceutical quality are:

a) transition to a more continuous processing mode (Fisher et al., 2019; Lee et al., 2015; Plumb, 2005);
b) process simulation and modeling (Chatterjee et al., 2017; Chen et al., 2020);
c) advanced data analytics (Severson et al., 2018; Steinwandter et al., 2019);
d) active process control on quality, including advanced techniques such as model predictive control (MPC; Rawlings et al., 2017);
e) enabling technology for real-time release testing, namely for online monitoring of the product quality (as opposed to end testing), including PAT (De Beer et al., 2011; Simon et al., 2015) and soft-sensors based on mathematical techniques (Pantelides and Renfro, 2013; Sinner et al., 2021);

f) routines for automatic detection of special cause variability that might compromise the product quality (ICH, 2011; Kourti, 2019).

These technologies have gradually been introduced in the pharmaceutical industry, due the initiatives of regulators for improving pharmaceutical quality assurance (Yu and Kopcha, 2017) and to the economic advantages that come to manufacturers (Isaksson et al., 2018). However, the implementation is far from being well established yet. For this purpose, the contribution from the fields of chemical and process systems engineering, possessing key expertise in digitalization strategies (Collins, 2018; Troup and Georgakis, 2013), will be crucial.

It is important to remark here that the original QbD guidelines (ICH, 2012a, 2009, 2008, 2005) already allowed (and promoted) the introduction of the novel technologies that have emerged lately. Hence, the recent pharmaceutical modernization trends should be seen as an evolution of QbD from its early 2000s implementation version (Yu, 2008) to a contemporary smart manufacturing framework (Davis et al., 2015), which someone has referred to as "Pharma 4.0" (Boni, 2016; Steinwandter et al., 2019), but that is still fully adherent to the initial QbD cornerstones.

1.3.2 The Quality-by-Design framework

An overview of the main ICH regulations defining the QbD initiative is provided in Table 1.4. The ICH Q7 guideline (ICH, 2000) introduced the cGMPs, lying the foundations for QbD. The core concepts of QbD are outlined by the ICH guidelines Q8(R2), Q9 and Q10 (ICH, 2009, 2008, 2005). The ICH guideline Q8(R2) presents the QbD approach to drug product development. The ICH guideline Q9 introduces the quality risk management system, and provides information on the risk-based approach to pharmaceutical quality attainment promoted by QbD. A pharmaceutical quality system framework for effective quality management across the product life cycle is presented in the ICH guideline Q10. ICH guideline Q8(R2), Q9 and Q10 have a special focus on pharmaceutical development and manufacturing of drug products. The ICH guideline Q11 (ICH, 2012a) clarifies how to apply the QbD paradigm to development and manufacturing of drug substances, and provides further descriptions of the QbD concepts outlined in the previous guidelines. ICH guideline Q12 (ICH, 2019) addresses the manufacturing phase of the product life cycle, and introduces a QbD framework to facilitate post-approval changes and continual improvement. Recently, ICH released for public consultation the draft version of guideline Q13 (ICH, 2021), which introduces definitions and tools to facilitate the transition of continuous pharmaceutical manufacturing.

The common ground across all guidelines is the ultimate aim of QbD, namely inherently building the quality into the product with a scientific and risk-based approach, to be adopted during all the product life stages. This is achieved through enhanced product and process understanding and through critical investigation of the sources of variability affecting the

Table 1.4. *Main ICH quality guidelines defining the QbD framework.*

Date	Guideline	Reference	Status
10/11/2000	Q7 – Good manufacturing practice guide for active pharmaceutical ingredients	ICH (2000)	Final guidance
09/11/2005	Q9 – Quality risk management system.	ICH (2005)	Final guidance
04/06/2008	Q10 – Pharmaceutical quality system	ICH (2008)	Final guidance
01/08/2009	Q8(R2) – Pharmaceutical development	ICH (2009)	Final guidance
11/11/2010	Q8, Q9, Q10 Questions and Answers	ICH (2010)	Final guidance
06/12/2011	Q8, Q9, Q10 Points to consider	ICH (2011)	Final guidance
01/05/2012	Q11 – Development and manufacture of drug substances (chemical entities and biotechnological/biological entities)	ICH (2012)	Final guidance
23/08/2017	Q11 Questions and answers	ICH (2017)	Final guidance
20/11/2019	Q12 – Technical and regulatory considerations for pharmaceutical product lifecycle management	ICH (2019)	Final guidance
27/07/2021	Q13 – Continuous manufacturing of drug substances and drug products	ICH (2021)	Draft guidance

product quality. QbD is opposed to the traditional Quality-by-Testing, consisting in quality tests on the end product and in empiric product and process development. Regulators summarize the main goals of QbD as (Yu et al., 2014):

- to reach the target product quality specifications, based on the evaluated clinical performance;
- to increase process capability and reduce product variability and defects through enhanced product and process design, understanding and control;
- to enhance development and manufacturing efficiencies;
- to improve root cause analysis for special variability occurrences and to enhance post-approval change management.

These goals can be achieved through the different elements of QbD (Yu et al., 2014):

1. determination of a quality target product profile (QTPP) that identifies the product critical quality attributes (CQAs). CQAs are the characteristics of the drug that are critical to quality from the patient's perspective;
2. product design and understanding, including the assessment of the raw material properties that have a critical impact on the product quality (referred to as critical material attributes, CMAs);
3. process design and understanding, including identification of critical process parameters (CPPs), and of functional relationships linking CMAs and CPPs to CQAs;

4. design of a control strategy that includes specifications for the drug substance(s), excipient(s), and drug product as well as controls for each step of the manufacturing process;
5. process capability and post-approval continual process improvement.

Before going on, it is necessary to provide clarification for the introduced QbD jargon (ICH, 2017, 2009):

- *QTPP*: a summary of the product quality characteristics that are deemed necessary to achieve the desired safety and efficacy. According to the ICH guideline Q8(R2) (ICH, 2009), a QTPP should include the following:
    - intended use in clinical setting, route of administration, dosage form, delivery systems;
    - dosage strength(s);
    - container closure system;
    - therapeutic moiety release or delivery and attributes affecting pharmacokinetic characteristics (e.g. dissolution, aerodynamic performance) appropriate to the drug product dosage form being developed;
    - drug product quality criteria (e.g. sterility, purity, stability and drug release) appropriate for the intended marketed product.
- *CQA*: "physical, chemical, biological, or microbiological property or characteristic that should be within an appropriate limit, range, or distribution to ensure the desired product quality" (ICH, 2009). CQAs are identified from the QTPP using as criterion the criticality to safety and efficacy attainment. CQAs are usually properties of the drug product, of the drug substance, and of intermediates. For the sake of simplicity, in this Dissertation, the CQAs of the intermediates will (improperly) be referred to as "product" CQAs, as often done in the literature;
- *CMA*: properties of the raw materials that affect the product CQAs;
- *CPP*: "a process parameter whose variability has an impact on one or more CQAs and therefore should be monitored or controlled to ensure the process produces the desired quality" (ICH, 2009). In most of the cases, CPPs are those variables that in conventional control science jargon are referred to as manipulated variables;
- *control strategy*: set of planned actions, routines and controls, derived from product and process understanding, to be followed during manufacturing to ensure process performance and product quality. A control strategy can include control and monitoring plans for the CPPs and CMAs related to drug substance and drug product and intermediate, the operating conditions of the facility and of the equipment, in-process controls, specifications of the finished product, and the methodologies and frequency for implementing monitoring and control activities. Note that the definition of control

strategy in the pharmaceutical context differs from the standard engineering meaning, according to which the term *control* is related to the concept of *process control*;

- *process capability*: ability of a process to realize a product that meets the target quality requirements. The concept of process capability can be defined in statistical terms, as the number of standard deviations between the process mean and the nearest specification limit (Yu and Kopcha, 2017);

Essentially, QbD consists in the determination of the characteristics of the drug that are critical for attaining the desired clinical performance, in their translation into CQAs, in the design of a product that can deliver them and in a robust process to manufacture it, after having obtained a critical understanding of the relation between product formulation/manufacturing variables and CQAs. Under a QbD approach, the process control strategy is designed with a risk-based approach and relies on solid process understanding. In addition, the process capability should be continuously improved, by reducing the process variability through increased product and process understanding.

The stages of the product life cycle (§1.2) relevant for QbD implementation are (ICH, 2008): *i)* pharmaceutical development, *ii)* technology transfer (i.e., the scale-up from the development to the manufacturing scale, or the transfer between manufacturing sites for marketed products*) iii)* commercial manufacturing, and *iv)* product discontinuation, for which a suitable approach based on process and product understanding should be followed for activities such as continued complaint handling and retention of documentation and samples. Current QbD guidelines provide recommendations and frameworks mostly for the pharmaceutical development and commercial manufacturing stages. Technology transfer and product discontinuation are addressed only in a general fashion, recommending to adopt science- and risk-based approaches, but lacking of technical information on how to do so to leave more flexibility to companies. Hence, in the remainder of this paragraph we summarize the key concepts of the QbD ICH guidelines that apply to the stages of pharmaceutical development and commercial manufacturing. Table 1.5 shows the main differences between a minimal and an enhanced (QbD) approach to pharmaceutical development and manufacturing. It is important to remember that minimal and enhanced approaches are not mutually exclusive, and that they can be used in combination. For instance, one CQA can be tested through traditional end testing, while enhanced approaches can be used for testing other one(s).

Pharmaceutical development aims at (ICH, 2009) "designing a quality product and its manufacturing process to consistently deliver the intended performance of the product". The minimal approach to pharmaceutical development to be followed when filing a marketing authorization consists of the following steps (ICH, 2009):

- step #1: *QTTP definition*, based on the desired clinical performance;

- step #2: *CQAs identification*. CQAs of the drug product, of the drug substance and of the components of the drug product formulation are identified from the previously defined QTTP;
- step #3: *manufacturing process design*, including process synthesis, operation design and equipment selection;
- step #4: *control strategy design*.

An QbD approach to pharmaceutical development (Figure 1.2) includes one or more of the additional following enhanced elements (ICH, 2009):

- step #3 enhanced:
  a) *risk assessment* for carrying out the systematic evaluation, understanding and design of product formulation and manufacturing process, including: *i)* identification of CMAs and CPPs that can have an effect on the product CQAs and *ii)* determination of functional relationships linking CMAs and CPPs to product CQAs;
  b) description of a *design space* (DS) from the determined relations between CMAs and CPPs and CQAs;
- step #4 enhanced:
  using the enhanced product and process understanding and quality risk management to establish an *enhanced control strategy*, which can include advanced features, such as a DS and/or real time release testing (RTRT).

Hence, the enhanced approach to pharmaceutical development can include the design of two key QbD enabling technologies:

- *DS*: "multidimensional combination and interaction of input variables (e.g., material attributes) and process parameters that have been demonstrated to provide assurance of quality" (ICH, 2009). In other words, the DS is the multivariate space of CPPs and CMAs that guarantees that the desired product CQAs are met. When a DS is approved by regulators, the manufacturer is allowed to move process operation within the DS during manufacturing. Operation outside of the DS requires to initiate a regulatory post approval change process. The introduction of the concept of DS represents a revolution in pharmaceutical development and manufacturing, as previously process operation was allowed only at the fixed combination of CPPs and CMAs contained in the application to regulators;
- *RTRT*: the ability to evaluate and ensure the quality of intermediates and/or final product during process operation based on process data, including a valid combination of CMAs, CPPs and CQAs (ICH, 2009). RTRT can replace end product testing, even though for batch release the review and quality control steps of cGMP are still necessary.

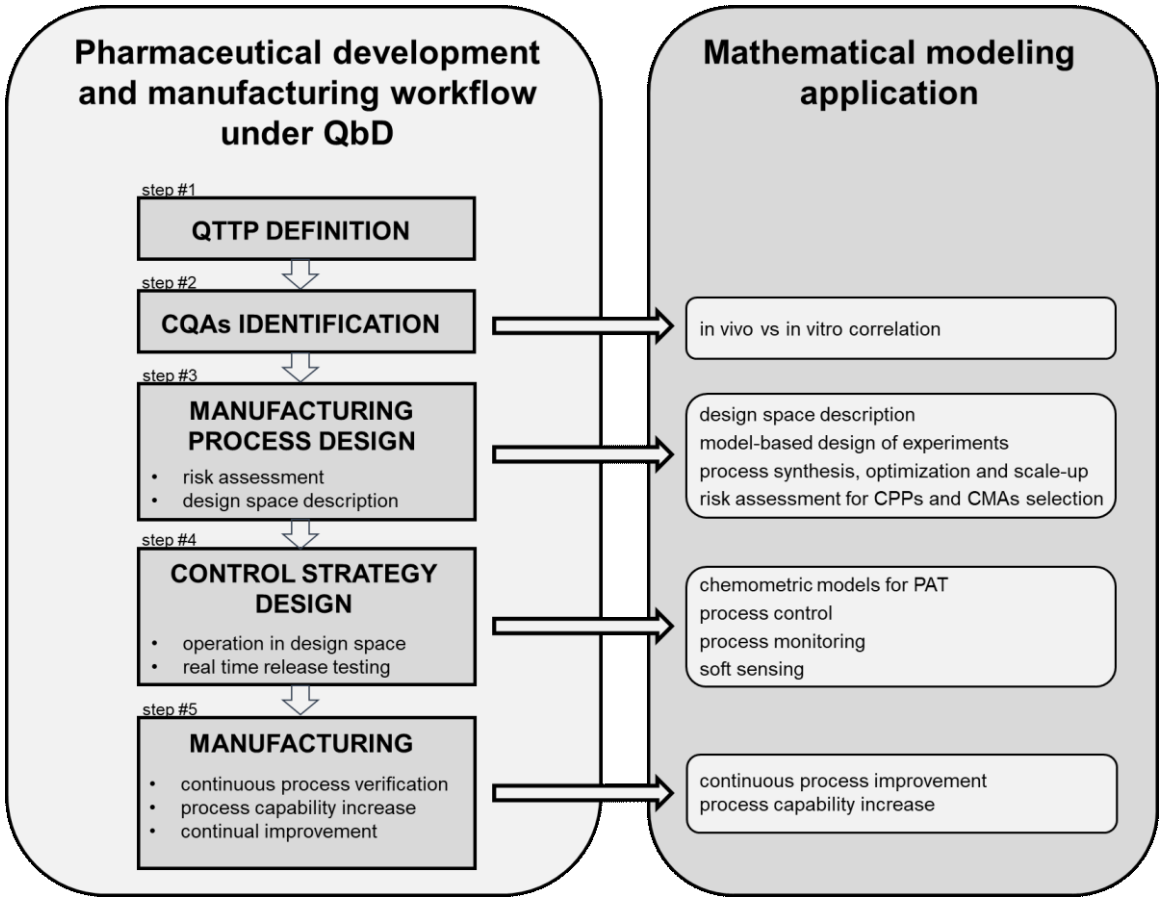


Figure 1.2. *Pharmaceutical development and manufacturing workflow under the QbD framework. Mathematical models can be used for supporting each stage, as sketched on the right and detailed in §1.4.*

Both DS identification and RTRT design require having gained systematic product and process understanding, which is the foundation of the QbD approach. In order to do so, multivariate experiments are typically carried out, as opposed to univariate experiments that are the common choice when the minimal approach is adopted. Note that, even though the DS is a revolutionary concept introduced by QbD, DS description is only a (non-mandatory) activity within control strategy design under QbD, which generally includes careful development of all routines and control actions for conducting the process. Moreover, DS description is not strictly necessary for developing an enhanced control strategy, although it is strongly encouraged.

Commercial manufacturing aims at "achieving product realization, establishing and maintaining a state of control and facilitating continual improvement (ICH, 2008)". In the manufacturing phase, in a minimal approach:

- step #5: manufacturing
  a) operating conditions are fixed;
  b) quality is mainly tested on intermediates and on the end-product (end-product testing);

Table 1.5. *Main differences between minimal and enhanced (QbD) approaches to pharmaceutical development and manufacturing. Adapted from ICH guideline Q8(R2) (ICH, 2009).*

Aspect	Minimal approaches	Enhanced approaches - QbD
Pharmaceutical development	• mainly empirical • univariate experiments for product and process understanding	• systematic, relating mechanistic understanding of CMAs and CPPs to product CQAs • multivariate experiments for product and process understanding
Manufacturing process	• fixed operating conditions • process performance validated on few initial batches • focus on reproducibility	• operating conditions can move within DS • process performance validated multiple times across life cycle, ideally continuously (continuous process verification) • focus on control strategy and robustness • process monitoring with statistical process control
Control strategy	• control strategy developed with empiric approaches to reproduce conditions found to deliver target quality during development • product quality verified through intermediates and end product testing	• control strategy developed with solid product and process understandings • product quality inherently built into the product • possibility or real time release testing, or of reduced end product testing
Lifecycle management	• reactive: problem solving and corrective action	• proactive: preventive action • continual improvement facilitated

c) the process performance is validated only for few initial batches.

Enhanced approaches, instead, feature an enhanced control strategy:

- step #5 enhanced
  a) the operating conditions can be moved within the DS, if it has been approved;
  b) RTRT often replaces end-product testing for one or more CQAs;
  c) *continuous process verification* is carried out, continuously evaluating the process performance to assess that it is realizing the desired quality.

Enhanced approaches can also include the following additional elements:

d) *continual process improvement*, achieved by using data collected on the manufacturing line for improving and optimizing the current process. Revision and maintenance activities are also required when operating under an enhanced control strategy, for example to validate the chemometric models of PAT used for RTRT or to review mathematical models used for design space description;

e) *process capability increase,* namely reduction of the process variability. This is achieved by identifying the sources of common cause variability, and implementing mitigating actions. Process capability increase is a main enabler of six sigma quality (Yu and Kopcha, 2017). Univariate statistical process control (SPC), or multivariate statistical process control (MSPC; Kresta et al., 1991), is encouraged by ICH

guidelines (ICH, 2011, 2009) to promptly detect the occurrence of special cause variability that might affect the product quality;

As a note to elements *a-b)* of enhanced approaches, as already stressed, for pharmaceutical products it is not enough to guarantee that the final product CQAs are within specification, but it must also be assessed that CPPs and CMAs were within the DS during the product manufacture. This requirement is based on the acknowledgement that only consistent paths will assure overall quality consistency, as different paths to a same end point for certain quality properties may result in different values for secondary properties, such as downstream processability (Kourti, 2019). Even though this reasoning is indeed correct, this conservative approach is based on the belief that comprehensive product and process understanding has not been reached. Theoretically, if all the product quality variables were known and it would be possible to verify that a given product meets the target specifications for all the quality attributes, it would not matter the operating conditions path that led to obtaining the product.

1.3.3 Current state of implementation of Quality-by-Design

Since the launch of the initiative, QbD elements have started to be introduced not only in applications for launching new products, but also in processes for products at a later development stage and for legacy products (Cook et al., 2014; Grangeia et al., 2020; ter Horst et al., 2021). Merck paved the way with the first application containing QbD elements approved by the FDA, the NDA for the drug Januvia (FDA, 2006). The first application approved by the FDA containing a formal DS description was approved in 2013 (FDA, 2013b): the BLA submitted by Genentech (Roche) for Gazyva, which is also the first approved BLA that follows QbD principles. In the EU, the EMA approved the first application with QbD principles and DS description in 2012: the MAA for Kalydeco, submitted by Vertex Pharmaceuticals. In 2012, a survey (Cook et al., 2014) evaluated the answers from 149 anonymous respondents, mainly from industry. Most of respondents (54% to 76%) reported widespread utilization of most QbD elements introduced with the ICH guideline Q8(R2) (ICH, 2009). The QbD tools most used by respondents were design of experiments (DoE; 76% of respondents), risk assessment (72% of respondents) and QTTP identification (66% of respondents). The survey (Cook et al., 2014) also reported frequent (55% of respondents) use of DS description, nowadays considered a cornerstone for QbD implementation. Grangeia et al. (2020) recently analyzed 60 research papers on QbD applications, most of which authored or co-authored by industrial practitioners. The authors reported widespread use of risk assessment and DoE tools, with DS description assuming always more relevance in latest years. If the industrial surveys and research contributions (Cook et al., 2014; Grangeia et al., 2020) demonstrate that the pharmaceutical industry is familiar with the QbD concepts, QbD application is still not fully established in the submissions for regulatory authorization filed in recent years. A recent analysis (ter Horst et al.,

2021) of the EMA reports for MAAs reviewed and assessed in the years 2014-2019 showed that only 38% of full applications for new drugs contained QbD elements. Just 35% of them also contained a DS description in 2018 (ter Horst et al., 2021). A similar trend was found in the assessment reports of the Japan regulatory agencies: despite the inclusion of QbD elements in new applications rose from 9% in 2009 to 71% in 2018 (Kajiwara et al., 2020), only 2% of applications contained a DS description, and just 3% of applications for drug products included RTRT procedures. Use of QbD for new molecular entities development is larger than for biologicals (ter Horst et al., 2021), mainly due to the higher complexity of biotechnological processes (Rathore et al., 2018). A recent survey (Reinhardt et al., 2020) confirmed that the implantation of industry 4.0 technology into the pharmaceutical industry is proceeding at a slow pace.

Different causes contribute to the sluggish implementation of QbD in applications submitted to regulatory agencies. In the industrial survey by Cook et al. (2014), more than 50% of responders expressed a neutral or negative feedback with respect to the impact of QbD on ROI. Actually, even though it has widely been demonstrated that QbD implementation has paramount middle- and long- term economic benefits for pharmaceutical companies (IBM, 2005; McKinsey & Company, 2021; Yu and Kopcha, 2017), it is also true that the initial costs for companies new to QbD are not negligible (Vishwasrao and Singh, 2016). QbD also usually requires large experimental campaigns for gaining product and process understanding. Such campaigns are sometimes more expensive than those carried out for product development under the traditional approach, even if they enable more flexibility in the manufacturing stage. Another finding is that companies are sometimes discouraged to introduce QbD elements in regulatory submissions even if QbD was applied upon development, due to the perception that receiving an approval for a QbD application is more difficult (Milmo, 2014). Concerns have been expressed that the larger amount of data coming from QbD experimental campaigns could highlight phenomena difficult to explain, which would rise questions from regulators. Further questions from regulators could arise from misalignment in the terminology or misinterpretation of the guidelines.

The response to the slow adoption of QbD and to skepticism lies in innovation. The pharmaceutical industry will benefit from the introduction of modern technologies under the QbD framework, as occurred to other industries previously (Isaksson et al., 2018). The pharmaceutical emerging technology (§1.3.1 and §1.4), especially the use of mathematical modeling, can play a crucial role in boosting QbD implementation, through cost reduction and enhanced robustness that inherently come with the digitalization of pharmaceutical development and manufacturing.

1.4 Pharmaceutical emerging technology: future trends of Quality-by-Design

Within the recent modernization trend, the pharmaceutical industry is attempting to catch up with state-of-the-art technologies that are established in other manufacturing industries. This transition requires adopting simultaneously *i)* technology from the novel Industry 4.0 paradigm (Kagermann et al., 2011) and *ii)* technology that has been established for a long time in other industries, but not in pharma, yet.

The initiative "Industrie 4.0" originated in Germany in 2011 (Kagermann et al., 2011) to promote the innovation of the manufacturing industry. Industry 4.0 aims at creating an efficient smart manufacturing environment based on cyber-physical systems, where the physical equipment is connected and exchanges information in real time with a digital counterpart that autonomously runs process operation through advanced monitoring and control algorithms. Industry 4.0 does not only involve process automation and digitalization, which actually already occurred in most of the manufacturing sectors, except for pharma. It is the autonomous decision-making feature of Industry 4.0 that makes it different from previous industrial revolution technology.

In the remainder of this paragraph, we briefly outline the most important emerging pharmaceutical technology, grouped in three sub-sections: continuous processing, active process control, and mathematical modeling. Please note that a detailed overview of modern sensors, knowledge management and data exchange systems, although of paramount importance for Industry 4.0, is out of the scope of this Dissertation. The interested reader may find these technologies reviewed elsewhere (De Beer et al., 2011; Gyürkés et al., 2020; Kim et al., 2021; Pantelides and Renfro, 2013; Reinhardt et al., 2020; Suresh et al., 2010).

1.4.1 Continuous manufacturing

Traditionally, pharmaceutical manufacturing has been carried out predominantly in batch mode. In recent years, many studies from regulators, academics and practitioners (Burcham et al., 2018; Collins, 2018; Fisher et al., 2019, 2016; Ierapetritou et al., 2016; Lee et al., 2015) demonstrated that pharmaceutical continuous manufacturing has many advantages compared to batch processing. Such benefits are both for manufacturers and for patients, and include reduced manufacturing time and cost, decreased facility footprint, greater product consistency and process flexibility, potential to reduce shortages, and simplified scale-up procedures. Continuous manufacturing is also an enabler of other pieces of pharmaceutical emerging technology, especially of active (i.e., engineering) process control. Actually, continuous processes are inherently more controllable than their batch counterparts, and they are more prone to the implementation of advanced monitoring and data analytics routines.

Regulators are promoting the transition to continuous manufacturing through several initiatives, aimed at clarifying how to develop a continuous process for a new medicine or vaccine under the QbD framework. In the novel ICH guideline Q13 (ICH, 2021) and in a recent document by the FDA (FDA, 2019), regulators outlined their quality considerations for continuous manufacturing, highlighting features that are peculiar to continuous processes, rather than to batch processes. One important recommendation given by the ICH and the FDA is the necessity of implementing material traceability approaches, namely to track the material flow through the integrated continuous system. Residence time distribution models (Kruisz et al., 2018; Sencar et al., 2020) are suggested for this purpose. Material traceability, combined with process monitoring routines, is needed for identifying and tracking nonconforming material generated in the plant, to be diverted in suitable diversion points that have to be designed into the plant. The implementation of active process control and RTRT in continuous processes is also strongly encouraged by the FDA (FDA, 2019). RTRT can be enabled by mathematical models used as surrogates of direct measurements.

The idea of transitioning to continuous processing has been in the pharmaceutical industry for long (Plumb, 2005). Many examples of single pharmaceutical unit operations exist in the academic and industrial literature. Recent developments in continuous flow chemistry (Bana et al., 2017; Baumann and Baxendale, 2015) allow continuous synthesis of small molecules, even through multi-step reactions. The first biopharmaceutical product manufactured through continuous perfusion was approved by FDA in 1993, and to date about 20 biologicals are synthetized through continuous processing (Fisher et al., 2019). Continuous implementations also exist for most of the other pharmaceutical unit operations, including crystallization (Acevedo et al., 2016; Wood et al., 2019a; Yang et al., 2015), powder feeding (Hanson, 2018), powder blending and tableting (Nagy et al., 2017; Su et al., 2019a). Still, the integration of all unit operations in an end-to-end continuous pharmaceutical process is little explored in the literature (Domokos et al., 2020; Mascia et al., 2013), and even less in the industrial practice. As of December 2020, only six drug products manufactured continuously have been approved by the FDA: Daurismo by Pfizer, Orkambi, Symdeko/Symkevi and Trikafta by Vertex Pharmaceuticals, Prezista by Johnson & Johnson and Verzenio by Eli Lilly (GlobalData Healthcare, 2020). In 2020, the first application for continuous manufacturing of an API and the first continuous biomanufacturing process were approved (FDA, 2020). Recently, Continuus Pharmaceuticals, a spin-out from an MIT-Novartis collaboration, obtained a $69.3 million government contract to build the first FDA approved end-to-end manufacturing facility using its proprietary technology for the production of small-molecule medicines (Continuus Pharmaceuticals, 2021).

Several challenges remain ahead for the widespread implementation of continuous manufacturing, despite the more flexible regulatory environment that is being drawn in recent years. Among them, the main ones are:

- *capital cost*. Even if it has been proven that continuous processing is economically convenient with respect to batch processing (Schaber et al., 2011), the large capital costs are sometimes a limitation for converting an already built batch process into a continuous one. High-volume of production is the main driver for converting a batch process to a continuous one, as happened for the Johnson & Johnson's plant producing Prezista (Centers, 2015);
- *end-to-end continuous integration.* Technology for connecting in a continuous fashion the single unit operations is not mature yet. In particular, a gap exists between upstream (drug substance) and downstream (drug product) manufacturing, for which continuous solutions already exist, and the intermediate purification section. Focusing on small molecules manufacturing, API purification is typically carried out through subsequent crystallization, filtration, washing and drying steps. However, continuous technology for filtration, washing and drying of crystallization slurries (§4.2), and for their continuous integration with crystallization, is still lagging behind (Burcham et al., 2018; McWilliams et al., 2018);
- *active process control.* Continuous processes are more easily controllable than their batch counterparts. As a result, product quality attainment is achieved more consistently and at a lower cost, when a control system is in place. However, the know-how on control of continuous pharmaceutical processes is still limited, especially when end-to-end integrated plants are considered. Even though there is growing interest in the pharmaceutical sector towards implementing active process control (Su et al., 2019b), pharmaceutical operation is typically still run at open-loop, as a result of the "golden batch" tradition. More advancement is needed both from the regulatory agencies, the industrial and the academic sides to enable the implementation of active process control, without which continuous processing is far less attractive.

1.4.2 Active process control: towards Quality-by-Control

A control strategy (§1.3.2) for a pharmaceutical process can include three levels of controls Yu et al., 2014), as depicted in Figure 1.3:

- *Level 1*. A Level 1 control strategy makes uses of process control techniques for actively controlling the product quality. Active process control is defined as "a system consisting of hardware and software architecture, mechanisms, and algorithms that automatically adjusts a process to maintain the process output within a desired range." (FDA, 2019), and in practice coincides with the engineering definition of process control (Seborg et al., 2017). Under a Level 1 control strategy, the CMAs are monitored, and the product CQAs are kept close to the target values. This is achieved by automatically adjusting the CPPs, in response to CMAs changes (feedforward control). In the case in which

CQAs are monitored in real time, CPPs can be adjusted also in response to CQAs changes (feedback control). In other words, the CPPs are manipulated in response to disturbances to track the desired CQAs set-points. Advanced techniques can be implemented at Level 1, such as MPC (Rawlings et al., 2017), real time optimization (RTO; Biegler et al., 2015), and state estimation (Ray, 1981). Level 1 control can also enable RTRT, which can be achieved not only through PAT, but also making use of soft sensors based on mathematical models (FDA, 2019; ICH, 2011; Yu et al., 2014);

- *Level 2*. Under a Level 2 control strategy, process operation is conducted at open-loop within the established DS. As when the process is operated in the DS the product quality is guaranteed, reliance on end-product testing is decreased with a Level 2 control strategy;
- *Level 3*. A Level 3 control strategy consists in operating the process under tightly constrained intervals of CMAs and CPPs, as there is limited understanding of the sources of variability and of the impact of CMAs and CPPs on the product CQAs. Hence, extensive end-product testing is also conducted.

A Level 3 control strategy corresponds to the traditional Quality-by-Testing approach, while a Level 2 control strategy, possibly mixed with some Level 3 control strategy elements, is aligned

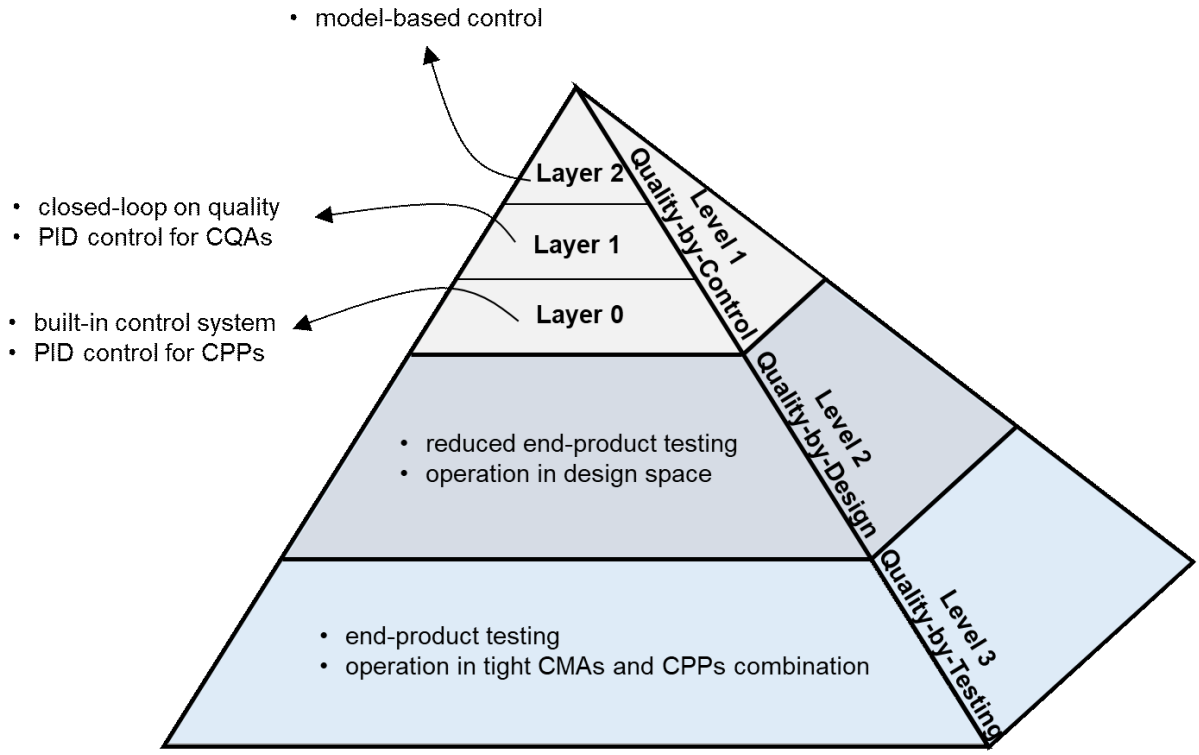


Figure 1.3. *Three levels of control strategies of a pharmaceutical process (*Yu et al., 2014). *A Level 1 control strategy can consist of layers: Layer 0, Layer 1 and Layer 2* (Su et al., 2019b).

to the QbD paradigm. Level 3 control strategies are still the most adopted in pharmaceutical manufacturing, even though Level 2 strategies are rapidly gaining pace (§1.3.3).

The ultimate aim of the QbD initiative is to achieve the implementation of Level 1 strategies, capable of guaranteeing quality assurance with the maximum efficiency. The benefits of implementing a modern Level 1 control strategy are both for companies and for patients, as stressed by regulators, practitioners and academics (Collins, 2018; Fisher et al., 2019; Troup and Georgakis, 2013; Yu et al., 2014). McKinsey (2021) estimates that the adoption of a smart quality control system for process development and manufacturing could have a tangible impact on profit, reduce the product launch time by more than 30%, increase manufacturing and supply chain capacity and responsiveness by 20 to 30% and prevent major compliance issues by reducing manual errors and variability.

Recently, Level 1 control strategies have started to appear more and more frequently in academic publications (Mesbah et al., 2017; Rehrl et al., 2016; Sen et al., 2014), originating the so-called novel Quality-by-Control (QbC) paradigm (Su et al., 2019a). QbC represents an advanced version of QbD, in which active process control is the core of the control strategy. A hierarchical process control approach (Figure 1.3), structured according to the ISA-95 Enterprise-Control System Integration Standard, has been proposed to guide the development of QbC control strategies (Su et al., 2019a, 2017). The proposed hierarchical structure presents three layers (Layers 0-2), not to be confused with the levels of control strategies outlined in Yu et al. (Yu et al., 2014). Layer 0 control consists of the built-in control systems of the unit operations, which are used for controlling the CPPs of the process. Layer 0 control makes use of programmable logic controllers (PLCs) or simple PID controllers (Seborg et al., 2017). Layer 1 control relies on PID loops to control the CQAs and reduce the impact of disturbances. Practically speaking, Layer 1 supervises Layer 0 through cascade control, typically with a single input single output (SISO) approach, and can span across single or multiple unit operations. PAT tools are usually resorted to at Layer 1 for measuring the CQAs. Layer 2 uses model-based techniques to achieve advanced process control and monitoring. Techniques commonly implemented at Layer 2 include MPC, RTO, state estimation and advanced fault detection and diagnosis routines (Venkatasubramanian et al., 2003a, 2003b). Note that a control system is considered to be of Level 1 if it includes at least Layers 0-1, since the implementation of Layer 0 alone does not close the loop on the CQAs.

All this considered, a tight interconnection arises between the main emerging technologies in QbD: continuous manufacturing, active process control and mathematical modeling. The transition to continuous manufacturing requires the QbC framework for systematic control system design, while QbC implementation is gaining momentum under the increased interest in continuous processing. Mathematical modeling is as an invaluable tool for boosting the implementation of both continuous processing and active process control (Collins, 2018; Fisher et al., 2016). *In-silico* simulators based on mathematical models can be used for testing different

control strategies and for optimizing the process operating conditions. In addition, mathematical models are necessary for implementing the advanced process control and monitoring routines that constitute Level 2 of the QbC framework.

1.4.3 The role of mathematical modeling in the QbD framework

A mathematical model is a description of a system in the form of equations, variables and parameters. Based on the amount of knowledge on the physics of the system embedded in the equations, mathematical models can be broadly divided into three different categories (Bonvin et al., 2016; Figure 1.4): knowledge-driven, data-driven or hybrid.

Knowledge-driven models (also referred to as mechanistic, first-principles or white-box) are the mathematical representation of the available physical understanding of a system. Knowledge-driven models are made up of conservation and/or constitutive equations. The variables of a knowledge-driven model include inputs, outputs and, most importantly, states, which fully characterize the internal state of the system and, for dynamic models, allow monitoring and forecasting its time evolution. The parameters of the model have a physical meaning. Examples of knowledge-driven models are, but are not limited to: mass and energy balances, thermodynamic models, population balance models, and mass/heat transfer models. Discrete element and finite element methods are typically used for solving the equations of knowledge-driven models numerically. Pharmaceutical applications of computational

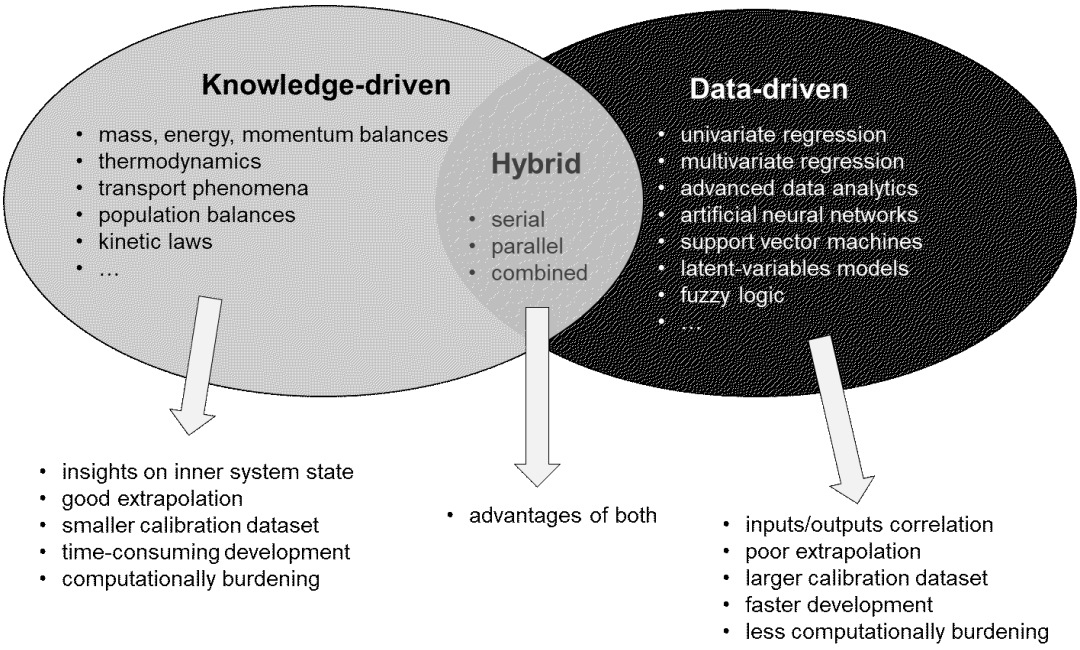


Figure 1.4. *Classification of mathematical models based on the amount of knowledge on the physics of the system under investigation.*

fluid dynamics, combining knowledge-driven modeling and numerical routines for simulating fluid and/or powders flow, are also gaining momentum (Babnik et al., 2020; Sarkar et al., 2019).
Data-driven models (also referred to as empirical) are a mathematical representation of the available data for a system. Data-driven models do not carry any information on the physical mechanisms occurring in the system under investigation. The equations simply relate the inputs and the outputs of the dataset used for model calibration. Hence, the variables of data-driven models are the inputs and the outputs of the dataset, and no additional information can be obtained on the internal system state. The parameters of the models do not have a physical meaning, and they express how the equations should be tuned for representing the system of interest. Examples of data-driven models are, but are not limited to: univariate and multivariate regression models, advanced data analytics, artificial neural networks, genetic algorithms, support vector machines and latent-variables models (LVMs).
Compared to data-driven models, knowledge-driven models are typically more time consuming to develop and computationally burdening, although they usually need less data for calibration. However, knowledge-driven models provide a comprehensive insight on the state of the system, and usually have good extrapolation properties, while data-driven models do not have these features. Hybrid models (also referred to as semi-empirical or gray box) combine knowledge-driven and data-driven components to take the advantages of both (Sansana et al., 2021; von Stosch et al., 2013). Hybrid models components can be arranged in a serial structure, in a parallel one or into a more complex serial/parallel architecture. An example of a hybrid model is a first-principles model of a bioreactor, in which a kinetic rate is estimated with a data-driven model (e.g., artificial neural networks), since the actual kinetics is not fully understood from a first-principles perspective.
The ICH guideline "Q8, Q9, Q10 Points to consider" (ICH, 2011) clarifies the role of mathematical models within the QbD framework. For the purposes of regulatory submissions, models are classified based on their contribution in assuring the quality of the product as:

a) low impact: models used to support product and/or process development (e.g., formulation optimization);
b) medium impact: models used for product quality assurance, but in combination with other indicators of quality (e.g., most design space models);
c) high impact models: models whose predictions are main indicators of the product quality (e.g., a chemometric model for analyzing the product quality).

Regulators provide a step-by-step guide for model development and implementation within the QbD framework (ICH, 2011):

- *Preliminary phase*
    1. definition of the purpose of the model
    2. decision of model type and of experiments to support model development;
    3. selection of the variables of the model;

4. evaluation of the limitations of the model assumptions;

- *Experimental phase*
    5. if historical data are not available, new experiments have to be carried out to support model development at either the laboratory, pilot or manufacturing scale;
- *Model development*
    6. development of the equations of the model and parameter estimation;
- *Model validation and pre-implementation activities*

  Validation procedures are tighter for higher impact models. The following elements are appropriate for high impact models, and are considered on a case-by-case basis for low and medium impact models:
    7. setting the criteria for a successful validation and/or verification;
    8. testing the model accuracy through cross-validation;
    9. testing the model accuracy using an external dataset;
    10. verifying the model accuracy by parallel testing with a reference method. Models used to support a DS or that are part of the control strategy have to be verified at the manufacturing scale, too;
    11. evaluating the impact of the uncertainty of the model prediction on the product quality. If appropriate, defining an approach to reduce the associated residual risk;
    12. documentation of the model development and validation activities. If suitable, preparation of a verification plan across life cycle;
- *Model implementation and maintenance during the life cycle*
    13. release for usage;
    14. after implementation, model verification continues throughout the product life cycle for high impact models, and possibly for medium impact models, too. Verification consists in the repetition of steps 8-10, or a subset of them deemed suitable for the considered model. Model maintenance and update might be needed as result of model verification.

The model development and implementation procedure of Steps 1-14 couples traditional scientific methodologies for mathematical modeling with the QbD focus on quality risk analysis. More specifically, modelers are invited to select the variables of the model based on risk considerations, to account for the risk associated with modeling assumptions (step 4), to estimate the uncertainty of the model, to evaluate its impact on the product quality (step 11), and to introduce mitigation strategies, if possible.

The encouragement of regulators (ICH, 2011, 2009, 2008) to adopt mathematical modeling for modernizing the pharmaceutical industry, together with the economic benefits, resulted in an increasing interest in the topic across the pharmaceutical community (Reklaitis et al., 2017;

Sarkis et al., 2021). A survey to 21 pharmaceutical industry professionals (Troup and Georgakis, 2013) reported use of mathematical modeling for process analytics, process monitoring, plant-wide information systems, unit operation modeling, quality control, and process optimization. All respondents declared to use SPC for process monitoring, while 67% use also MSPC for process monitoring, and LVMs to analyze historical datasets. On the modeling side, only 44% of respondents reported that more than 10% of unit operations had a knowledge-driven model developed. Instead, 33% of the companies routinely developed data-driven models for 80–100% of unit operations, and another third declared development of data-driven models for 40–60% of all unit operations. When asked for the rationale of the predominance of data-driven modeling, respondents mentioned the large time and resources needed for developing a knowledge-driven model, including the requirement for modeling expertise and the extent of resources necessary for model validation. Nonetheless, respondents also acknowledged the increased process understanding, robustness and extrapolation properties of knowledge-driven models. More recently, Rogers and Ierapetritou (2015a) revealed advances in the modeling of particulate processes, which boosted a renewed interest in knowledge-driven and hybrid modeling in the pharmaceutical sector. In 2017, a comprehensive review on QbD implementation (Reklaitis et al., 2017) reported that knowledge-driven and hybrid models were used always more often throughout the pharmaceutical life cycle, although data-driven models were still the most chosen approaches for process monitoring, control and data analytics.

Current research trends are focused on further enabling the life cycle approach to the use of mathematical modeling, with a special attention to promote the adoption of continuous processing and active process control. The target is having a mathematical model to support the initial process synthesis activities and the subsequent development and manufacturing phases, including DS description and real time operation. Transitioning across the different stages, the complexity of the mathematical model can be tuned to meet the level of accuracy and computational burden needed for a particular application. Model-based process synthesis is almost unexplored in the pharmaceutical sector (Casas-Orozco et al., 2021; Gernaey and Gani, 2010; Papadakis et al., 2018), despite it is an established methodology in the broader process industry (Grossmann and Daichendt, 1996). On the design space side, recent research and review articles (Bano et al., 2018b, 2018a; Boukouvala et al., 2010; Chatterjee et al., 2017; García-Muñoz et al., 2015; Reklaitis et al., 2017; Tomba et al., 2013) demonstrate that mathematical methodologies for DS description are now mature, including techniques for both knowledge-driven, hybrid and data-driven models. Future trends in operation design and DS description require addressing the bottlenecks of end-to-end continuous systems, and the presence of active process control. Regarding real time operation, very recent reviews (Chen et al., 2020; Steinwandter et al., 2019) report that online application of mathematical models in the pharmaceutical industry is gaining pace, for process monitoring, process control and

operation optimization.

Real time use of mathematical models is evolving towards the *digital twin* framework (Kritzinger et al., 2018). A digital twin (Grieves, 2014) of one or multiple unit operations is made up of: *i)* a physical component (i.e., the equipment), *ii)* a virtual component (i.e., a series of computer routines), and *iii)* an automated data exchange infrastructure between the physical and virtual components. During operation, the virtual component receives data from the process and, based on the collected data and model-based algorithms, autonomously communicates to the physical components decisions on how to conduct the process (e.g., through MPC). The establishment of the digital twins in the pharmaceutical sector is currently limited by real time model computational time, by model maintenance, by real time data exchange and by concerns in data security and confidentiality (Chen et al., 2020; Steinwandter et al., 2019). Actually, the most common scenario consists in the so-called *digital shadows* (Udugama et al., 2021), in which the data communication proceeds only from the physical component to the virtual one for the purpose of process monitoring.

An overview of mathematical modeling applications for supporting pharmaceutical development and manufacturing is outlined in the next paragraph, to give an idea of different uses of mathematical models across the product life cycle.

1.5 Mathematical modeling for supporting pharmaceutical development and manufacturing

1.5.1 Overview

Considering the workflow for implementing QbD in pharmaceutical development and manufacturing (§1.4 and Figure 1.2), examples of applications of mathematical models at different stages are (Chatterjee et al., 2017; ICH, 2011; Kourti et al., 2015):

- step #2: *CQAs identification*
  - *in vivo* vs *in vitro correlation* (IVIVC) models. Through IVIVC models, an i*n vitro* property of a drug (e.g., dissolution) can be used as surrogate for an *in vivo* performance (e.g., bioavailability), and it can be classified as a CQA (Rossi et al., 2007);
  - models for optimizing the product formulation;
- step #3: *manufacturing process design*
  - model-based design of experiments (MBDoE; Franceschini and Macchietto, 2008), used to maximize the process understanding gained from an experimental campaign;

  - use of models for supporting process synthesis, optimization and scale-up (e.g., use of process simulators; Papadakis et al., 2018);
  - support selection of CPPs and CMAs through quantitative or semi-quantitative approaches to risk assessment, such as failure mode effects analysis (FMEA) or failure mode effects and criticality analysis (FMECA);
  - global sensitivity analysis (Saltelli et al., 2008) for aiding selection of CPPs and CMAs;
  - models for operation design (§1.5.1), namely design space description and scale-up;
- step #4: *control strategy design*
  - chemometric models for PAT tools. In the most typical case, an LVM is used for correlating a spectrum obtained from a spectroscopy tool to a variable of interest, such as the API content in a tablet;
  - soft sensors, namely mathematical models or model-based techniques that estimate the value of (unmeasured) variables of interest, inferring them from process data (Pantelides and Renfro, 2013). Soft sensors can be used for end-point determination and they can enable RTRT;
  - models and model-based techniques for process monitoring, namely fault detection and diagnosis (Venkatasubramanian et al., 2003a, 2003b). SPC and MSPC are the most employed methodologies for this purpose. Process monitoring can ensure that the process in state of control, supporting RTRT (ICH, 2011). Soft sensors are often part of the process monitoring framework. Additional details on the use of models for process monitoring are presented in §1.5.2;
  - models for process control, which include, among others, models for offline optimization of manipulated variable trajectories, models for feedforward control and models for MPC. The state-of-the-art on the use of mathematical models for process control in pharmaceutical development and manufacturing is given in §1.5.3.
- step #5: *manufacturing*

  The manufacturing stage makes use of the models that are part of the designed control strategy (step #4). Other uses of mathematical modeling in the manufacturing stage are, but are not limited to:
  - continuous process verification: process monitoring methodologies (e.g., MSPC) can demonstrate that the process is in normal operating conditions and that the process variability is within the common cause limits;

- continual improvement: use of models, especially of data-driven type, for investigating historical data, with the purpose of troubleshooting and process improvement;
- process capability increase: MSPC can be used for understanding the main causes of the common cause variability, which can then be tackled with the purpose of achieving six sigma quality.

This spectrum of applications demonstrates the potential impact of mathematical modeling on the pharmaceutical industry modernization and digitalization, lying the basis for the Pharma 4.0 paradigm (Boni, 2016; Steinwandter et al., 2019).

In the next sub-paragraphs, state-of-the-art mathematical modeling for operation design, process monitoring and process control in the context of pharmaceutical development and manufacturing is provided. Model-based pharmaceutical process synthesis (Papadakis et al., 2018) and design of experiments (De-Luca et al., 2020; Shahmohammadi and McAuley, 2020) are of pivotal importance, too. However, they are still little addressed in the literature, and fall outside the scope of this Dissertation.

1.5.2 Operation design

Operation design consists in the determination of the operating conditions at which a process should be conducted to assure the product quality. In some cases, an additional interest is to optimize a given objective function, such as the economic profit (to be maximized), or the risk of obtaining non-compliant products (to be minimized).

In the context of QbD, operation design is strictly linked with the regulatory concept of DS. DS description consists in the determination of the multivariate space of CMAs and CPPs that allow attaining the desired product CQAs. As clarified by regulatory guidelines (ICH, 2009), proven acceptable ranges of individual CPPs and CMAs varied keeping the other parameters and attributes constant do not constitute a DS, as they do not account for the mutual interaction. In regulatory submissions, the DS is usually described in graphical terms (Figure 1.5) within the knowledge space, namely the region of CPPs and CMAs that has been explored in the DS description exercise. The normal ranges of operating conditions within the DS are defined as *control space*. The literature on DS description techniques and applications is a wide corpus, whose thorough review is out of the scope of this Dissertation. Design space scale-up (García-Muñoz, 2009; Liu et al., 2011) and maintenance (Bano et al., 2019), even though of paramount importance, have been little addressed in the literature so far, and are also out of the purpose of this Dissertation. In the remainder of this section we outline the main techniques for DS description, and the most recent trends and applications in relation to continuous processing and active process control.

1.5.2.1 Mathematical techniques for design space description

DS description is an activity inherently based on mathematical modeling. Even when the DS is submitted to regulatory agencies in graphical form as in Figure 1.5, it has previously been obtained through a mathematical model. The traditional approach to DS description is based on simple data-driven modeling, through the following procedure (Boukouvala et al., 2010; Chatterjee et al., 2017; Huang et al., 2009):

1. extensive experimental campaign in a pre-defined knowledge space of raw material attributes and process parameters;
2. risk assessment to determine process parameters that critically impact CQAs, namely the CPPs;
3. use of simple data-driven modeling and optimization for describing the DS boundaries;
4. confirmatory experiments to validate the DS before regulatory submission.

Multivariate data analysis on DoE data and response surface methodology (RSM; Box and Wilson, 1992) are the most common approaches for carrying out step 3. A multivariate model, such as multiple linear regression, is built upon the experimental data. The DS is then obtained in mathematical form through inversion of the mathematical model, or, more commonly, represented in graphical form through contour plots spanning the knowledge space (Goyal and Ierapetritou, 2002). RSM is still based on multivariate modeling of DoE data, but it also includes: *i)* an iterative modeling procedure, and *ii)* an optimization step. Following an RSM, experimental and modeling activities are iteratively executed until a satisfactory fitting is achieved. An optimization strategy is then followed to assess the values of CMAs and CPPs that produce the target product CQAs. This traditional approach requires long and expensive experimental campaigns and strives with describing high-dimensional DSs, especially if presenting collinearities among the inputs. Nonetheless, RSM has long represented the most

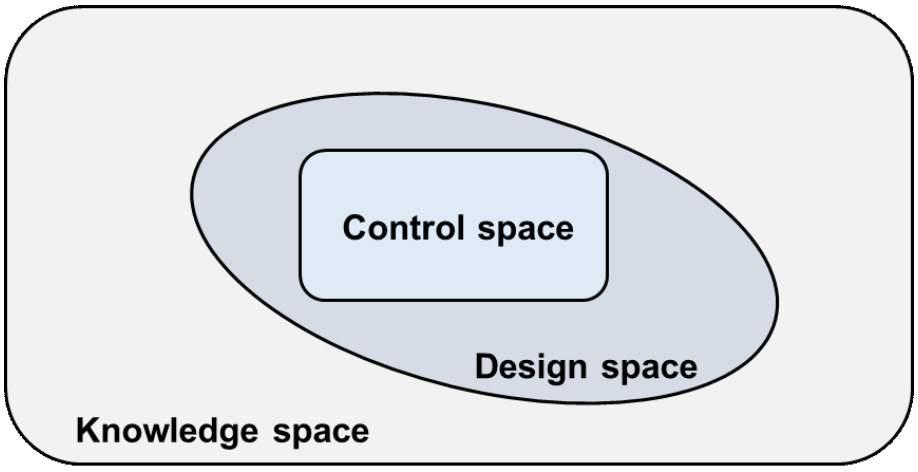


Figure 1.5. *Graphical representation of a design space within the knowledge space. An illustrative control space inside the design space is also reported. Adapted from MacGregor and Bruwer (2008).*

common approach to DS description, and still has widespread use in industry and academia (Belmir et al., 2021; Khafagy et al., 2020; Minatovicz et al., 2021), thanks to the easiness of implementation and to the availability of dedicated software.

In recent years, DS description evolved into more mature forms, including one or both of the following elements:

i) use of *knowledge-driven models*. Knowledge-driven modeling for supporting DS description is nowadays the preferred choice for well-known processes, as they can provide a comprehensive description of complex relations between CMAs, CPPs and product CQAs. Moreover, knowledge-driven models allow to increase process understanding, a pillar of QbD. A practical advantage is that knowledge-driven models can be used for *in-silico* experiments, which significantly reduce the (physical) experimental campaign burden. Despite the advantages of knowledge-driven models, traditional DS description based on RSM is still widespread, especially for new and not well understood processes, due to the aforementioned facility of implementation;

ii) accounting for *model uncertainty*. Traditional DS description procedures associate unique values of product CQAs to every combination of CMAs and CPPs, without accounting for model uncertainty. Traditional multivariate analysis and RSM on DoE data do not yield a probabilistic DS. Neither does the deterministic DS obtained from a knowledge-driven model account for model uncertainty. To fulfill the risk-based mandate of the QbD paradigm, a DS should be determined under a probabilistic approach, associating to every combination of CMAs and CPPs a probability distribution of product CQAs, or, at best, the probability that the desired product CQAs will be met. Next, some procedures for probabilistic DS description are outlined.

Formally, a probabilistic DS has been defined (Castagnoli et al., 2010; Lebrun et al., 2012) as the set of CMAs and CPPs, subset of the knowledge space, where the probability that the CQAs will meet the target specifications is above a certain probability threshold. A part from the experimental activities, probabilistic DS description requires two mathematical elements:

a) a model to relate the CMAs and the CPPs to the product CQAs;
b) a mathematical methodology that provides the probabilistic DS, given the desired CQAs specifications and the model of point *(a)*, also considering the model uncertainty.

Regarding the mathematical modeling activity *(a)*, any among data-driven, knowledge-driven and hybrid models can be resorted to. Use of knowledge-driven models is preferred, if first-principles knowledge is available for the system.

The many mathematical methodologies *(b)* for probabilistic DS description in the literature can be categorized in two approaches (Figure 1.6): *i) inverse* techniques based on optimization, such as flexibility analysis (Grossmann et al., 2014) and model inversion, such as LVMs

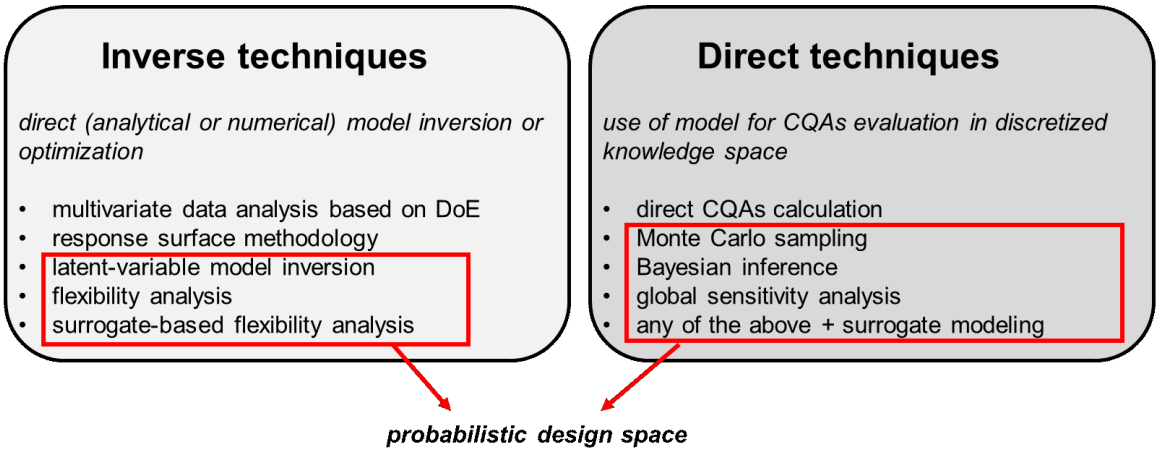


Figure 1.6. *Summary of inverse and direct techniques for design space description. The highlighted techniques can account for model uncertainty, leading to a probabilistic design space. The other techniques, instead, yield a deterministic design space, not considering the model uncertainty.*

approaches (Tomba et al., 2012), and *ii) direct* techniques with explicit sampling of the knowledge space (García-Muñoz et al., 2015).

With inverse techniques, the probabilistic DS is obtained through direct (analytical or numerical) inversion of the mathematical model, or by solving an optimization problem in which the model is expressed as a set of constraints. LVMs inversion is a popular example of inverse technique for DS description (García-Muñoz et al., 2010). In an LVM framework (Wise and Gallagher, 1996), few latent variables are extracted from the set of original variables, expressing the linear combinations of CMAs and CPPs that most affect the product CQAs. The envelope of CMAs and CPPs that allow satisfying the desired product CQAs with a given probability (i.e., the probabilistic DS) is achieved by inverting the LVM, and by back-propagating its uncertainty (Bano et al., 2017; Facco et al., 2015; Tomba et al., 2012). DS description through latent-variables modeling was first introduced as an improvement to traditional RSM, even without uncertainty back-propagation, to reduce the experimental burden and to deal with DSs of high dimensionality. The advantage is that the obtained DS can be represented in the latent space, which usually has a more intuitive graphical representation than the higher-dimensional space of the original variables. Applications include DS description of different unit operations, for instance tableting (Liu et al., 2011b) and granulation (Tomba et al., 2013) processes. However, the use of knowledge-driven models for DS description is preferred, as they can give a more accurate and comprehensive process representation, especially when nonlinearities are involved.

Flexibility analysis is a more advanced inverse technique for DS description (Floudas et al., 2001; Grossmann and Morari, 1983). It is one of the most popular techniques for operation design in process systems engineering, related to the analogous concepts of resilience, feasibility and operability (Grossmann et al., 2014). The flexibility test (Halemane and

Grossmann, 1983) and the flexibility index (Swaney and Grossmann, 1985) are the two classic approaches to flexibility analysis. For DS description, a set of constraints is first fixed, specifying the desired product CQAs explicitly through the model in terms of CMAs and CPPs. Then, multi-level (nonlinear) programs are solved to assess the portion of the knowledge space that satisfies such constraints under uncertainty in the model parameters. Solving the multi-level optimization problems can be challenging, especially when the simulations use a comprehensive knowledge-driven model, or when data-driven constraints are included. Hence, DS description is often carried out by developing a computationally efficient surrogate of the original model with data-driven techniques. Such hybrid model is then used for solving the flexibility problem instead of the original model (Bhosekar and Ierapetritou, 2018). A trade-off between surrogate accuracy and computational burden has to be achieved. Examples of DS description with surrogate-based flexibility analysis are a study on a roller compactor (Rogers and Ierapetritou, 2015b) and one on a whole continuous tableting line (Wang et al., 2017). When the knowledge space presents high dimensionality, even surrogate-based flexibility analysis might become computationally untreatable. Recently, Bano and coworkers (Bano et al., 2018b) coupled latent-variables modeling with surrogate-based flexibility analysis to tackle this issue. The methodology was tested on a simulated continuous direct compaction line. More recently, Laky et al. (2019) proposed two algorithms, extending the flexibility test and of the flexibility index to compute the probabilistic DS more efficiently. The proposed approaches were also tested on an industrial Michael addition reaction case study. A novel flexibility index formulation (Ochoa et al., 2021) has also recently been proposed to select a hyperrectangular operating conditions region inside the DS (i.e., the control space). Case studies on pharmaceutical reactive systems demonstrated very promising computational times.

Differently from inverse approaches, direct techniques for DS description are based on the intuitive idea of creating a fine grid in the knowledge space and of using the model through simulations for verifying, for every combination of CMAs and CPPs (model inputs), whether the product CQAs (model outputs) are acceptable or not. The DS obtained with this procedure is referred to as *deterministic* DS. Switching to a probabilistic DS requires evaluating how the variability on the product CQAs is influenced by the variability (i.e., uncertainty) of the model parameters. This can be done by running several (at least 200-500) Monte Carlo realizations at every discretization point (García-Muñoz et al., 2015). At each realization, the model parameters are sampled from their probability distribution, previously obtained during model calibration. The (frequentist) probability of meeting the desired product CQAs for a given combination of CMAs and CPPs is given by the fraction of the realizations that yielded compliant CQAs.

As an alternative to the frequentist approach, Bayesian inference has been used to propagate the uncertainty in the model parameters to the product CQAs for probabilistic DS description (Peterson, 2008, 2004). A multivariate regression model relating CMAs and CPPs to product

CQAs is first developed from DoE experimental data, or from data generated *in-silico* from a detailed knowledge-driven model. The knowledge space is then discretized, and for each discretization point the joint posterior distribution of the product CQAs is computed through Monte Carlo techniques. The probabilistic DS is finally defined as the region of the knowledge space where the posterior predictive probability of the CQAs to meet the target specifications is acceptable. DS description with a Bayesian approach finds applications in crystallization processes (Castagnoli et al., 2010) and solid oral dosage form manufacturing (Chatzizacharia and Hatziavramidis, 2014). The Bayesian framework is particularly interesting within the context of pharmaceutical development. During the initial development stages, the DS can be computed based on the few data available. Every time new data become available during the progress of the development phase, the DS can be updated in a very computationally efficient manner through Bayesian inference, which is based on the concept of updating a prior belief into a posterior distribution through the use of data. Tabora et al. (2019) recently provided a perspective on the role of Bayesian modeling in pharmaceutical development.

Global sensitivity analysis (GSA; Saltelli et al., 2008) is another important family of direct approaches to DS description. Traditionally, the main application of GSA in DS determination is reducing the DS dimensionality and the computational time for DS description by assessing which inputs (model parameters, process parameters and raw material attributes) do not significantly affect the product CQAs. For this purpose, recent applications (Öner et al., 2020; Wang et al., 2017) made use of Sobol's methods (Sobol, 1993), a type of GSA technique that decomposes the product CQAs variability into the contributions of the different inputs. After the critical inputs are selected with GSA, any direct or inverse technique can be used for DS description. However, recent works made use of GSA for directly describing the DS (Kotidis et al., 2019; Öner et al., 2020). The approach consists in spanning the model inputs space to determine the probability distributions of the product CQAs, and the region where the constraints on quality are satisfied. GSA for DS description is expected to gain interest in the near future, thanks to the increasing availability of automated software, such as the freely available easyGSA toolbox (Al et al., 2019) and the GSA routines implemented in the advanced modeling framework gPROMS®.

The straightforwardness of direct techniques for probabilistic DS description make them very appealing to the pharmaceutical community. Unfortunately, the computational burden grows exponentially with the number of CMAs and CPPs, due to the increasing dimensionality of the knowledge space. Especially when using detailed and accurate knowledge-driven models, this approach can quickly become computationally untreatable. GSA is a useful tool for reducing the dimensionality problem, but it is often not sufficient. Hybrid modeling has been proposed as a tool for tackling this issue, through the creation of (faster) data-driven surrogates of detailed knowledge-driven models. In the context of the frequentist approach to uncertainty propagation for DS description, Kucherenko et al. (2020) recently demonstrated a two-order of magnitude

reduction of computational time with respect to traditional Monte Carlo sampling with the use of surrogate modeling and adaptive sampling. Within the Bayesian approach, the computational problem related to high-dimensional knowledge spaces has been tackled by Bano et al. (2018a). The authors introduced latent-variables modeling within the Bayesian framework, to reduce the problem dimensionality. Recently, Kusumo et al. (2020) proposed a nested sampling approach to compute the Bayesian evidence, which led to improved computational times compared to the traditional Monte Carlo sampling.

Despite the large number of direct and inverse techniques, DS description is still not an established practice in regulatory submissions (§1.3.3). The main reason is the computational complexity of the techniques for DS description, whose implementation requires highly trained personnel. Automated software for DS description with user-friendly interfaces is expected to widely boost the adoption of DS description practices by industry.

1.5.2.2 Design space description and continuous pharmaceutical processing

The recent interest towards end-to-end continuous pharmaceutical manufacturing is making the description of plant-wide DSs an increasingly relevant problem. DS description for individual continuous units has been explored in many publications, including, for instance, flow synthesis (Armstrong et al., 2019), crystallization (Kishida and Braatz, 2012; Wang and Lakerveld, 2017) and twin-screw granulation (Liu et al., 2017). With the shift to continuous processing, developing DSs for whole manufacturing lines is extremely important, to assess how the variability in the process inputs propagates downstream, across the whole plant. Nonetheless, the extension to multiple units or to a whole manufacturing line is scarcely studied in the literature, due to the required computational burden and modeling expertise. Bano et al. (2018b) recently proposed a DS description for integrated milling and tableting operations, using a technique combining latent-variables modeling and surrogate-based flexibility analysis. Wang et al. (2017) presented a DS for a whole tableting line, comprising feeding, milling, blending and tableting. Also in this case, DS description was performed through surrogate-based flexibility analysis. A DS for a whole continuous tableting line was also proposed by García-Muñoz et al. (2018), who used a comprehensive knowledge-driven model of the plant to assess the propagation of disturbances from the feeding section into the downstream part of the line. The probabilistic DS of a pharmaceutical upstream process for ibuprofen synthesis has been presented by Montes et al. (2018), who made use of knowledge-driven modeling, Monte Carlo simulation and global sensitivity analysis for DS description. Recently, Lee et al. (2020) studied an active pharmaceutical ingredient (API) continuous manufacturing process, made up of continuous flow reactors in series and a downstream continuous thin film evaporator. A comprehensive knowledge-driven model for the process was first developed and calibrated with an experimental campaign. *In-silico* experiments were then carried out on a surrogate of the

model, to quantify the relations between CPPs and product CQAs for subsequent DS description.

Despite the reported examples of DS description for the standalone upstream and downstream sections, DS determination for an end-to-end continuous pharmaceutical process is still missing in the literature, to the best of the author's knowledge. An important step towards this objective is constituted by the development of dynamic models encompassing multiple unit operations, as the ones recently proposed for pharmaceutical and biopharmaceutical processes (Benyahia et al., 2012; Smiatek et al., 2020; Velayudhan, 2014), and by their use for propagating the CQAs distributions across the flowsheet (Zahel et al., 2017). Actually, in most of the aforementioned examples of DS description for processes encompassing multiple unit operations, the first step in the DS determination procedure is the development of a model for a process. *In-silico* experiments for DS description were then carried out, directly through the model or with a surrogate, instead of performing more complex physical experiments across multiple units. In this direction, an important bottleneck to be addressed is the research and technological gap in continuous integrated implementations of the steps that usually connect upstream and downstream manufacturing: crystallization, slurry filtration, cake washing and cake drying (Burcham et al., 2018; McWilliams et al., 2018). Few technologies in the market allow carrying out continuously the steps following crystallization, namely the API isolation steps (Gursch et al., 2016; Liu et al., 2019). Hence, physical implementation and DS description for the API isolation section of the process represents an enabling step towards the description of the DS for a whole manufacturing line.

1.5.2.3 Design space description and active process control

In all the examples of DS descriptions reported so far, the process was operating at open-loop with respect to quality, namely under a Level 3 control strategy (§1.4.2). Within the recent interest in moving towards Level 2 and Level 1 control strategies, the impact on the DS of the closure of the loop on quality becomes an important research topic. MacGregor and Bruwer (2008) first acknowledged that the control system should be designed before of DS description activities, as the DS is influenced by the control system. The authors analyzed the effect of closed-loop control on the DS of a pharmaceutical process, which resulted expanded upon introduction of active process control (Figure 1.7). Actually, a basic assumption of control engineering is that, to achieve tight quality control, a larger variability must be accepted in the manipulated variables. In practice, closed-loop control transfers the variability from the quality variables into the manipulated variables, i.e., from where it hurts to where it does not (Luyben and Luyben, 1997). Formally speaking, the control loop on quality can be closed through *i)* feed-forward control and/or *ii)* feedback control (Seborg et al., 2017). Under feed-forward control, the process operation is automatically adjusted in response to registered changes in the

raw materials properties or environment conditions. Under feedback control, instead, the process parameters are adjusted in response to measured (or inferred) changes in the CQAs.

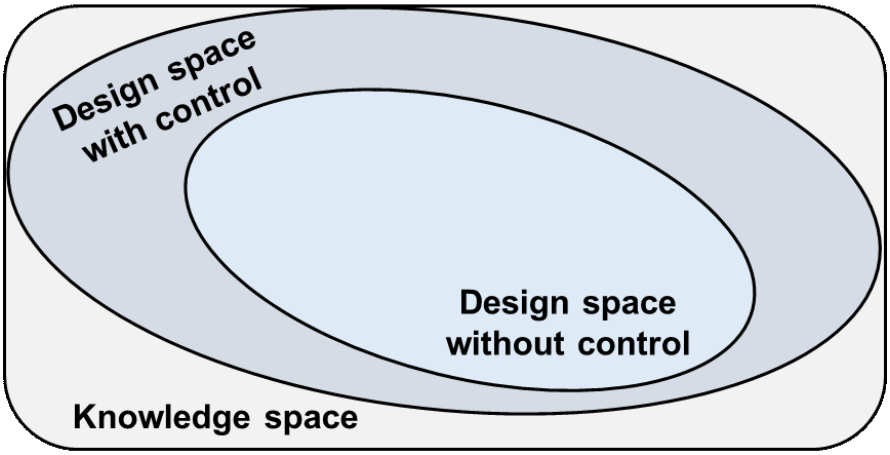


Figure 1.7. *Comparison of DS with and without closed-loop control on quality. Adapted from MacGregor and Bruwer* (MacGregor and Bruwer, 2008).

García-Muñoz et al. (2010) first presented the implications of introducing feed-forward control on quality in a pharmaceutical process. The authors demonstrated that the feed-forward controller led to an extremely widened acceptance region for the particle size distribution of the incoming material in a wet granulation process. More recently, Boukouvala et al. (2017) demonstrated the enlargement of the DS for a roller compaction process upon implementation of feedback control. Apart from these examples, few contributions highlight the impact of closed-loop control on the DS, despite the increasing applications of active process control in pharmaceutical development and manufacturing (Su et al., 2019b). This is, however, a study of paramount importance, since within the pharmaceutical regulatory framework (§1.3.2) it is important to validate the path of operating conditions that brought to attaining the final product quality, beside the actual final quality. Recently, Harinath et al. (2016) even suggested that control system design and DS description should be carried out simultaneously. The authors proposed an approach based on robust optimal control for simultaneously maximizing the DS and designing a linear feedback controller, for a continuous time system described by a state-space model with linear fractional uncertainties. Actually, the idea of joint DS determination and control system design is not entirely new, despite not being applied in pharma. The flexibility analysis theory already envisions the optimization of the control system within the multi-level optimization problem to be solved for operation design (Floudas et al., 2001; Grossmann et al., 2014). Research towards systematic application of these concepts to pharmaceutical development and manufacturing can have a tremendous impact on pharmaceutical quality and efficiency. A synergic effort between academia, industry and regulators is needed for these innovations to happen (Myerson et al., 2015).

1.5.3 Process monitoring

In the manufacturing industry, process monitoring consists in a series of activities for determining whether a process is operating under normal operating conditions, or if a fault occurred (*fault detection*). In the latter case, the cause for the occurrence of the fault is also explored (*fault diagnosis*), or identified among a set of known fault scenarios (*fault identification*). Techniques based on knowledge-driven (Gao et al., 2015; Venkatasubramanian et al., 2003a), data-driven (Jiang et al., 2019; Venkatasubramanian et al., 2003b) and hybrid models (Sansana et al., 2021) can be used for fault detection, diagnosis and identification.
Process monitoring is a key task in the pharmaceutical industry, where the state-of-control of the process must be assured for quality reasons. Although in the manufacturing industry process monitoring is primarily centered around the concept of fault, within the pharmaceutical sector it is traditionally intended with the slightly different meaning of measuring, either directly or through inference, the CMAs, CPPs and product CQAs. The aim is to verify that the process is being conducted within the operating conditions approved by regulators, and to support RTRT, as opposed to the traditional end-testing of the product CQAs. This is usually achieved through SPC, namely through univariate control charts reporting the control limits identified in normal operating conditions for the considered variable (Figure 1.8). Within the recent innovation trend

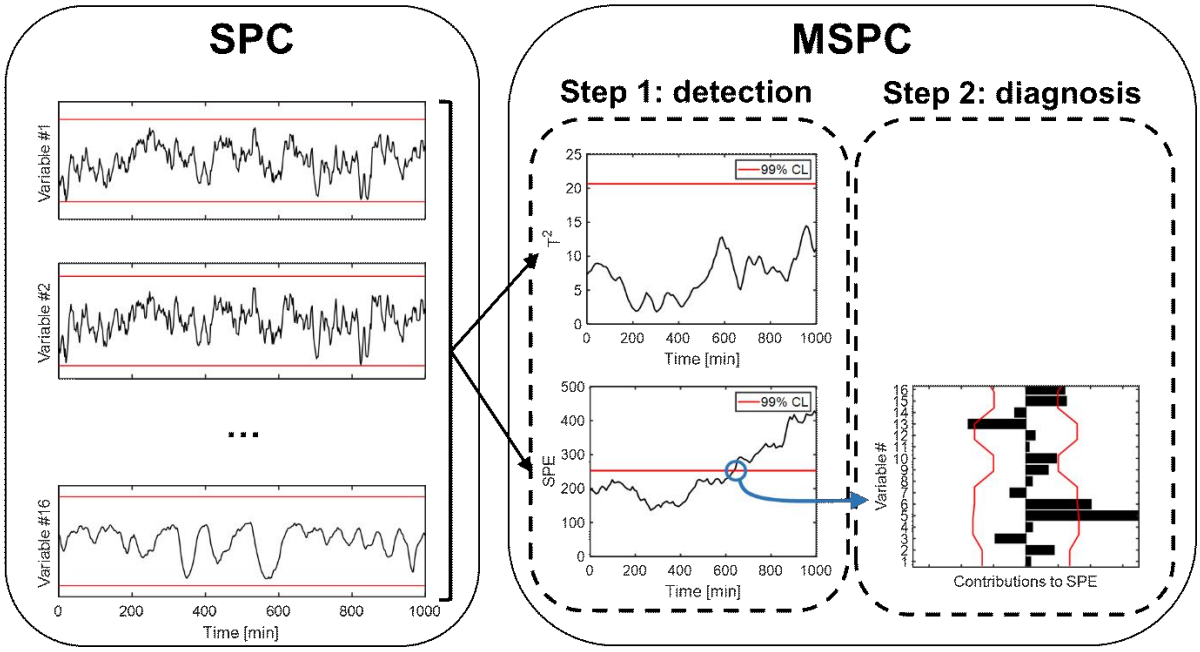


Figure 1.8. *Process monitoring: SPC vs MSPC. SPC: fault detection and diagnosis performed through univariate charts on process variables. MSPC: i) fault detection through two multivariate charts (T^2 and SPE), summarizing the variability of all the process variables, ii) fault diagnosis through contribution plots, calculated at time instant when a fault is detected on T^2 or/and SPE chart. CL = confidence limit.*

of the pharmaceutical industry, traditional process monitoring approaches have been modernized through the use of *i)* PAT tools, *ii)* soft sensors, and *iii)* advanced data-driven monitoring.

The introduction of PAT for monitoring CMAs, CPPs and CQAs (Claßen et al., 2017; De Beer et al., 2008; Sarraguça et al., 2014) represented the first stage of innovation of pharmaceutical manufacturing within the PAT and QbD initiatives. PAT requires the use of mathematical modeling, namely of chemometric models for relating the spectra collected by advanced sensors such as near-infrared (NIR) spectroscopy or Raman spectroscopy to the variable that is being measured. Projection to latent structures (PLS) regression (Geladi and Kowalski, 1986; Wise and Gallagher, 1996) is the chemometric method most commonly used for PAT. Machine vision, including multivariate image analysis (Geladi et al., 1989) and textural analysis (García-Muñoz and Carmody, 2010), has also been used as a PAT tool in pharmaceutical processes. Recent applications are for coating uniformity assessment (García-Muñoz and Gierer, 2010) and for measuring the API content in a powder mixture within a continuous blending process (Galata et al., 2021).

When physical sensors are not available for measuring a variable of interest, or the measurement is too expensive, sporadic or delayed, soft sensing can be resorted to. Soft sensors (Souza et al., 2016) infer the value of a secondary variable from a set of primary measurements making use of a mathematical model, which can be knowledge-driven, data-driven or hybrid. Literature contributions on soft-sensing for pharmaceutical development and manufacturing are abundant, and only selected applications are discussed here. A common use of soft sensors is for measuring the concentration of API in powder mixtures along continuous direct compression lines and in the final drug product. For this purpose, Rehrl et al. (2018) proposed soft sensors based on (data-driven) residence time distribution models. Recently, Kamyar et al. (2021) used a combination of knowledge-driven and data-driven soft sensors for measuring the blend composition along a direct compression line and the potency of the final tablets. Other applications of soft sensors involve measuring species concentrations in bioreactors (Sagmeister et al., 2013), moisture content in fluid bed dryers (Gagnon et al., 2017; Pla et al., 2018) and multiple variables in freeze drying processes (Drăgoi et al., 2013). The applications of soft-sensors in the pharmaceutical industry mentioned so far follow a so-called “open-loop” approach, in which the mathematical model receives input measurements from the process and, based on only that, predicts the variable(s) of interest. More advanced soft sensors are, instead, at “closed-loop”, namely based on state estimation (Ray, 1981). They estimate the variable(s) of interest making use not only of the process input measurements, but also of process output measurements. The output measurements are used as a feedback from the plant, to improve the soft sensor estimation and tackle the mismatch between the mathematical model and the process (plant/model mismatch; Figure 1.9). Soft sensors implementing state estimation routines have found multiple recent applications, including in bioreactors (Sinner et al., 2021),

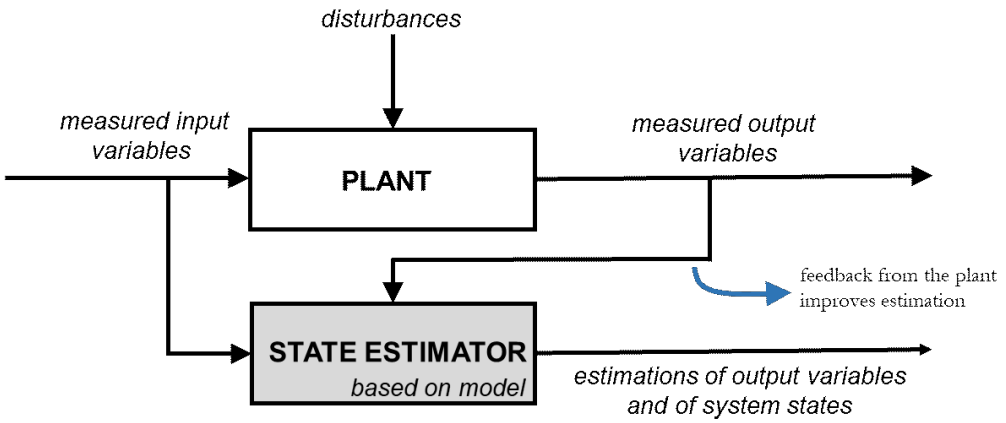


Figure 1.9. *Conceptual scheme of the rationale of a state estimator.*

direct compression lines (Liu et al., 2018) and freeze-drying (Bosca and Fissore, 2011). Next generation soft-sensors for pharmaceutical applications will have to make increasing use of knowledge-driven modeling and state estimation, following the QbD paradigm of developing robust processes relying on solid, and possibly physics-based, process understanding.

Even though process monitoring in the pharmaceutical industry has traditionally been carried out in a univariate fashion with more focus on quality assurance than on fault analysis, several recent contributions report the implementation of more advanced fault detection, diagnosis and identification routines in pharmaceutical processes. To this purpose, data-driven techniques have been resorted to, especially MSPC based on LVMs. An LVM can detect when new measurements coming from a plant are foreign to the stochastic structure of the normal operating conditions. Under MSPC, multivariate control charts (Figure 1.8) are used for detecting faults effectively, and contribution plots (Miller et al., 1998) are then used for fault diagnosis, to assess which of the measurements used for building the LVM are most responsible for the fault conditions (more details multivariate process monitoring by latent-variables modeling are provided in §3.2.2). The penicillin bioreactor simulator by Birol et al. (2002) has been widely employed for testing fault detection, diagnosis and identification routines in biomanufacturing (Monroy et al., 2012; Ündey et al., 2003). Kirdar et al. (2007) demonstrated that MSPC charts were useful for fault detection and diagnosis during routine manufacturing for a cell culture unit operation. García-Muñoz and Settell (2009) tested MSPC on a pilot-scale spray drying process. The monitoring system proved effective fault detection. Kona et al. (2013) proposed an MSPC system for fault detection in a batch fluid bed granulation process. Gupta et al. (2013) applied to a continuous granulation line, consisting of two feeders, a continuous blender and a roller compactor, an integrated framework for fault detection and diagnosis. The framework made use of multiple techniques, including wavelet analysis (Addison, 2017), MSPC, and qualitative trend analysis. More recently, MSPC has been applied to an API

synthesis process (Dumarey et al., 2019) and to a mixing process for manufacturing a pharmaceutical ointment (Bostijn et al., 2019). In the latter case, measurements from PAT were also included in the data matrix on which the LVM for MSPC was built. Measurements from PAT were added into the data matrix for MSPC also by Tahir et al. (2019), for monitoring a hot-melt extrusion process producing a paracetamol-affinisol extrudate. The authors also included within the MSPC framework the API concentration at the end of the extruder as predicted by a soft sensor. For MSPC of a freeze-drying process for pharmaceutical products, Colucci et al. (2021) even included in the data matrix information collected from an infrared camera and from the textural analysis carried out on RGB images of the final product. In all the applications mentioned so far, principal components analysis (PCA; Jackson, 1991) or PLS are the types of LVMs used for MSPC (Kresta et al., 1991). Instead, Quatrini et al. (2020) developed an MSPC monitoring system for a granulator using canonical variate analysis (Li et al., 2019) as LVM. The authors demonstrated effective fault detection and identification performances, and claim that canonical variate analysis allowed to better manage the time dynamics and nonlinearities of the process compared to the more commonly used PCA.

As discussed in §1.5.1, pharmaceutical manufacturing is always more often adopting MSPC monitoring systems such as in the examples described so far. The main limitation of data-driven approaches is that, even though they proved much effective for fault detection, fault diagnosis might, on the other hand, be challenging. This is especially true when the variables representing the root-cause of a faulty condition are not measured and, hence, not included in the data-driven monitoring system.

The use of knowledge-driven models for process monitoring has been proposed to overcome this limitation (Venkatasubramanian et al., 2003a). Knowledge-driven models have the advantage of embedding the fundamental knowledge on the mechanisms driving the process under investigation. This piece of information can enhance the fault detection and, even more, the fault diagnosis performance of the monitoring system. The main knowledge-driven monitoring approaches rely on parity relations (Gertler, 1998) or on state estimators (Caccavale et al., 2009; Deshpande et al., 2009). Both approaches are based on the underlying idea of detecting a fault when the process measurements depart from the predictions of the measurements given by a knowledge-driven model that is being run in parallel to the process. Recent studies (Liu et al., 2018; Moreno et al., 2018) presented a knowledge-driven monitoring framework for a direct continuous tableting line, to mitigate the effects of random measurement errors and for detection of non-random sensors malfunctions. Apart from this example, there are almost no pharmaceutical applications of process monitoring based on knowledge-driven modeling. Even though state estimators have found applications in pharmaceutical development and manufacturing, they are used as soft sensors, without fault detection and diagnosis routines implemented (Liu et al., 2018; Sinner et al., 2021). The complexity of knowledge-driven models development and the presence of process-model mismatch are the main reason for the low

popularity of knowledge-driven monitoring. Even the integration of knowledge-driven and data-driven approaches into a hybrid monitoring framework has found almost no applications in pharmaceutical manufacturing, except for soft-sensing. Considered the great interest within a QbD framework in developing physics-based process understanding, future research should be directed in the synthesis of monitoring systems using more knowledge-driven elements.

1.5.4 Process control

The traditional industrial approach to controlling process variables consists in the use of feedback control (Figure 1.10a), most frequently through PID controllers. At a given time, a PID controller keeps a controlled variable as close as possible to a reference set-point, by adjusting the value of a manipulated variable in response to a measured or inferred difference between the controlled variable and the set-point at the same time. The wide success of PID control in the process industry is due to its effectiveness in tackling most control problems, coupled to the easiness of implementation. PID controllers are commonly employed in the pharmaceutical industry for controlling the CPPs (Level 3 control strategy); however they have rarely been used for controlling the product CQAs (Level 1 or 2 control strategy). The main disadvantage of feedback control is that no corrective action is initiated until after a deviation of the controlled variable from the setpoint is registered, hence deviations from the setpoint will

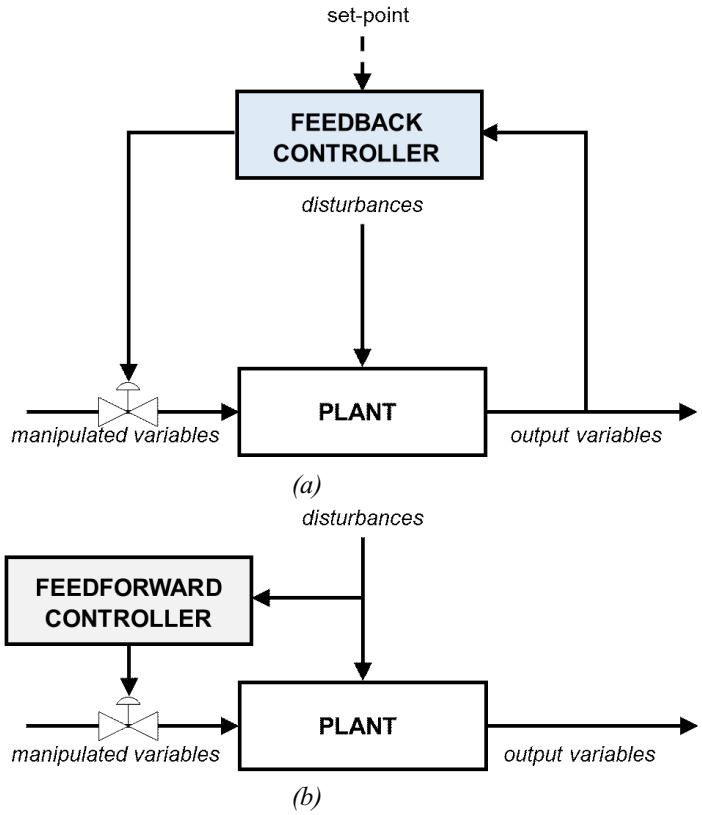


Figure 1.10. *Schematics of a system under (a) feedback control and (b) feedforward control.*

always occur with disturbances or setpoint change. Even in the case of known or measured disturbances, PID control does not take any preventive action. In contrast, when feedforward control (Figure 1.10b) routines are implemented, one or more disturbance variables are measured, and corrective actions can be immediately taken, ideally before they compromise the process. Note that a system might also be under simultaneous feedback/feedforward control. While feedback control is based on a PID control law, feedforward control requires a model of the process, to quantify how both the disturbance and the manipulated variable affect the controlled variable. Feedforward control has found different applications in pharmaceutical manufacturing, both alone and in combination with feedback control, including in continuous flow synthesis (A E Cervera-Padrell et al., 2012), high-shear wet granulation (García-Muñoz et al., 2010; Mathe et al., 2020), herbal drug manufacturing (Yan et al., 2014), fluid bed coating (Zhao et al., 2019) and continuous tableting (Singh et al., 2015). Combined feedback and feedforward control allow achieving good control performances in many situations. In addition to feedforward control, another model-based application for control strategy development consists in developing process simulators, which enable one to test and optimize the control loops. Plant-wide design of the control strategy of a continuous pharmaceutical plant has been demonstrated by Lakerveld et al. (2013).

MPC has been proposed as a more advanced model-based control methodology (Rawlings et al., 2017), which fulfills the QbD paradigm of controlling pharmaceutical processes based on solid scientific understanding. An MPC (Figure 1.11) uses a model to forecast the process states

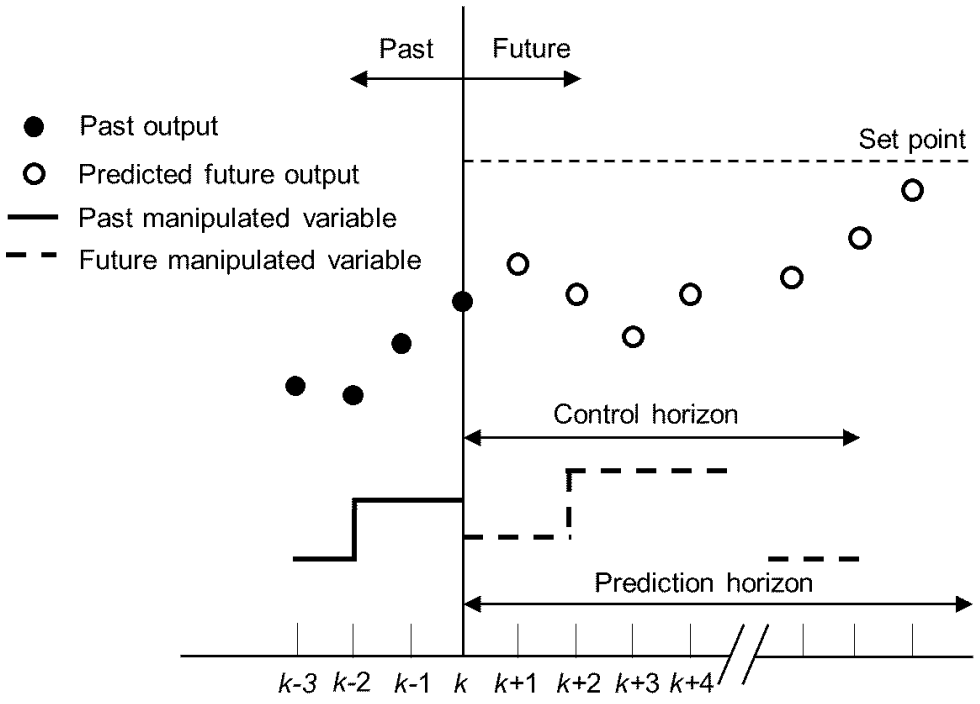


Figure 1.11. *Operational mode of an MPC. Adapted from Seborg et al. (2017).*

trajectory over a prediction horizon, comprising a fixed number of time steps. Based on the predictions, an optimization problem is solved to determine the optimal actions over a control horizon, contained within the prediction horizon. Only the control action for the next time step is applied, and the optimization problem is repeated at the following time step to obtain the subsequent control action to implement. The objective function of the optimization problem usually has two terms: *i)* the weighted sum of errors between the predicted trajectory of the controlled variable(s) and the reference setpoint trajectory and *ii)* the weighted sum of error between the control actions and the reference control actions. MPC leads to improved setpoint tracking compared to PID control, both for disturbance rejection and for setpoint change. Moreover, MPC is inherently more suitable for controlling systems with multiple inputs and multiple outputs. The main drawbacks are the implementation complexity and the computational burden, which often prevent MPC practical implementation. Nevertheless, MPC is always more applied in manufacturing systems, especially in its linear form.

In the pharmaceutical sector, MPC has rarely been applied to physical systems, although many recent simulation studies are demonstrating its effectiveness and advantages (Jelsch et al., 2021). MPC is the enabling technology for achieving robust Level 1 control strategies. Applications of MPC to pharmaceutical synthesis have recently been proposed. Nikolakopoulou et al. (2020) presented, through a simulation study, a fast MPC framework for the startup of a compact modular reconfigurable system for continuous-flow synthesis. The MPC was based on a linear model. Wong et al. (2018) studied, *in-silico*, the implementation of an MPC, based on artificial neural networks, to a complex reaction network in a CSTR. Regarding the purification section, MPC has extensively been applied to crystallizers (Nagy et al., 2013; Nagy and Braatz, 2012; Orehek et al., 2020). A very common implementation involves the optimization of the crystallizer cooling profile for controlling the crystal size distribution. To date, the only reported study on MPC for integrated purification operations, namely crystallization and the subsequent filtration and drying, is the *in-silico* work proposed by Sen et al. (2014). MPC for purification through a semi-continuous chromatographic process was presented by Papathanasiou et al. (2019, 2017). Many MPC applications to drug product manufacturing have recently been proposed. Rehrl et al. (2016) tested an MPC *in-silico*, on a feeding-blending system, demonstrating improved performances compared to PID control. The mathematical model used within the MPC framework was made up of simple knowledge-driven models (mass balances) and transfer functions. In subsequent work (Celikovic et al., 2020; Kirchengast et al., 2019), MPC was successfully tested on physical equipment, highlighting the benefits of model-based control with respect to standard PID control and feedforward control. Nonlinear MPC on a batch fluidized bed dryer for pharmaceutical wet granules has been demonstrated by Gagnon et al. (2017), both in a simulation study and in a pilot-scale system. A moving-horizon estimator was also implemented, to evaluate the system states and an unmeasured moisture disturbance. MPC applied to a tablet press was proposed by Haas et al.

(2017) in a simulation study, and by Su et al. (2019b) in physical equipment. Recently, Huang et al. (2021) presented on a simulation study a combined state estimation/nonlinear MPC approach to handle process-model mismatch in a rotary tablet press. A hybrid PID-MPC control strategy was successfully tested on a physical continuous tableting line (Singh et al., 2016, 2014a, 2014b), modeling the units through population balance models and discrete elements method. PAT tools were also used within the MPC framework. An additional process monitoring layer was also implemented on the line, through MSPC routines based on PCA.

Few studies address MPC applied to plant-wide pharmaceutical processes. Mesbah et al. (2017) used a nonlinear plant simulator for evaluating a plant-wide MPC for the end-to-end continuous pharmaceutical manufacturing plant. Paulson et al. (2018) recently showed in a simulation study that a fast stochastic MPC applied to an end-to-end continuous line could control the product CQAs effectively, with 25 times less variability than with standard MPC and with computational times feasible with real time implementation.

The next steps in development of advanced control strategies will have to focus on plantwide MPC for pharmaceutical processes, tackling current bottlenecks in connecting the different process sections and for operations lagging behind, such as purification. At the same time, an acceleration on the implementation on physical plants is needed, too. An increasing use of comprehensive nonlinear models within MPC frameworks is also envisioned positively within a QbD perspective. RTO (Biegler et al., 2015; Seborg et al., 2017) is another advanced model-based technique that can be included within a control strategy. Under RTO, the set points of the controlled variables are calculated by solving an optimization problem, either maximizing the economic profit or other objective functions of interest. RTO is often, but not always, coupled to a lower level MPC that attempts to minimize the error between the controlled variables and their optimized set points. RTO routines are called at a larger time interval compared to lower level control actions, and the optimization problem is usually based on a steady-state model of the process. Few RTO applications exist in pharmaceutical manufacturing (Giridhar and Reklaitis, 2020), and more work is needed in this direction.

Overall, plant simulators for control strategies testing, MPC, RTO, soft sensors, and fault analysis algorithms are the enabling technologies for reaching the ultimate objective of developing a complete digital twin, to be run in parallel with the process, continuously exchanging information for active process control and monitoring.

1.6 Objectives of the research

In recent years, the pharmaceutical industry has been experiencing a decreasing ROI on R&D, and increasing time and costs for bringing new products to the market (Deloitte, 2021; EFPIA, 2021). At the same time, an alarming number of drug shortages and recalls has been registered

(ISPE, 2017; Natof and Pellegrini, 2021). These phenomena are affecting both companies, from the financial side, and patients, who might experience high costs for drugs or unavailability of essential medicines. This is, for instance, the case for the COVID-19 vaccines rollout: after the initial enthusiasm for the vaccine discovery, the world quickly realized that it would have taken years to reach herd immunity on a global scale, due to the current technological limitations in pharmaceutical development and manufacturing, especially on the scale-up side. The lack of adoption of modern technology and approaches for pharmaceutical development and manufacturing has a critical role for the occurrence of shortages and recalls, and also in the increasing time- and cost-to-market for new products.

Pharmaceutical development and manufacturing have undergone a deep modernization in the last decade. However, there is still much work to be done by the pharmaceutical industry to catch up with the other manufacturing sectors on the adoption of modern Industry 4.0 technology.

The general objective of the PhD research is to aid the development of emerging pharmaceutical technology to:

- reduce pharmaceutical development time and cost;
- increase the efficiency and the robustness of pharmaceutical manufacturing.

To achieve these objectives, it is pivotal to transition to model-based pharmaceutical development and manufacturing. Even though the ICH guidelines promote the use of mathematical modeling, its use is still not widespread in the industry. This PhD research aims at tackling this delay by conceiving and applying mathematical models in key areas of pharmaceutical development and manufacturing: operation design, process monitoring and process control. The transitions to continuous pharmaceutical manufacturing and active process control are also addressed, in relation to the role of mathematical models for the implementation of these emerging technologies. The case studies span across the whole pharmaceutical train of operations, including drug substance manufacturing, purification and drug product manufacturing.

The specific objectives pursued in the PhD research are:

1. **advanced process monitoring**: development and implementation of modern pharmaceutical monitoring systems, for prompt fault detection and smooth fault diagnosis to inherently improve the efficiency and robustness of the manufacturing process (main objective *(b)*). This will be achieved through the implementation in the monitoring framework of knowledge-driven models and of state estimators, making a step forward from current data-driven monitoring approaches. The use of knowledge-driven models, even if they present process-model mismatch, enhances the physical understanding of the system embodied in the monitoring framework. This, in turn, provides a better insight on the inner state of the system for fault detection and diagnosis. In addition, the implementation of state estimators within the monitoring framework,

for soft-sensing of the system states and for measurement reconciliation, increases the robustness of the monitoring system to disturbances and to measurement error, and tackles the modeling error;

2. **digital operation design**: synthesis and use of dynamic simulators for pharmaceutical processes in the development stage. The simulator, upon validation with experimental data, can be used for performing quickly operation design and process scale-up, replacing extensive experimental campaigns and yielding robust processes. This is immediately beneficial to both main objectives *(a-b)*;
3. **control-relevant modeling**, to boost the adoption of active process control in the pharmaceutical industry. Mathematical simulators can be used for conceiving and testing *in-silico* the control system even from the pharmaceutical development phase, leading to more robust manufacturing processes (main objective *(b)*) and to potentially manufacturing faster and at a lower cost the drug for clinical trials (main objective *(a)*). Mathematical models are also enablers for advanced process control methodologies, such as MPC and RTO;

When developing and applying mathematical methodologies for specific objectives 1-3, the following additional specific objectives are pursued:

4. aiding the transition to continuous manufacturing: several case studies on mathematical modeling for design, monitoring and control of continuous pharmaceutical processes are developed in this Dissertation. In particular, special attention is dedicated to continuous unit operations for API crystals isolation from crystallization slurries, a bottleneck for end-to-end continuous pharmaceutical manufacturing;
5. knowledge-driven modeling: ICH guidelines encourage the use of physics-based models over data-driven models, where possible, as they provide a deeper degree of understanding of the process. These indications are followed in the Dissertation;
6. life cycle approach to mathematical modeling: (even simple) process models developed in early stages of pharmaceutical development can be improved and used for advanced purposes across the whole life cycle. For instance, a model used for preliminary operation design during development can be the basis for a more advanced model to be used for (nonlinear) MPC, or it can even be factored in a digital twin of the process.

1.7 Dissertation roadmap

The roadmap of this Dissertation is sketched in Figure 1.12, and additional details on the content of each chapter are provided in Table 1.6. Chapters 2 and 3 focus on process monitoring. Chapter 4 deals with operation design, while Chapter 5 focuses on process control.

Chapter 2 presents the proof of concept for the implementation of a state estimator (moving-

horizon estimator) to the feeding section of a direct compression line. The feeding section is composed by five loss-in-weight powder feeders operating in parallel, each one feeding one component of the tablet formulation to the downstream blending and compression units. The state estimator uses powder net weight measurements from load cells placed under each feeder for estimating the powder flowrate fed downstream, and, in turn, the composition of the powder mixture fed to downstream units. The net weight measurements are very noisy, due to the nature of the load cell setup. It is shown that the state estimator can effectively reconcile the noisy measurements, and provides more accurate powder composition estimations, compared to traditional approaches relying on statistical filtering.

Chapter 3 introduces a novel hybrid framework that bridges traditional standalone data-driven and knowledge-driven process monitoring approaches, to exploit the advantages of both. The process monitoring system features a data-driven model that includes two different data types: i) "actual" data coming from sensor measurements, and ii) "virtual" data coming from a state estimator, based on a first-principles model of the system under investigation. The hybrid monitoring framework is benchmarked on three simulated case studies, namely a fed-batch fermentation process for the manufacturing of penicillin, a segmented fluid bed dryer for wet pharmaceutical granules, and a continuous polycondensation process. The proposed hybrid monitoring approach shows superior fault detection and diagnosis performances when compared against conventional monitoring techniques, even when the first-principles model is affected by process/model mismatch.

Chapter 4 outlines a mathematical model for a novel integrated filter-dryer carousel system, designed for continuously filtering, washing and drying a crystallization slurry stream into a dry crystals cake. For a set of feed conditions and control inputs, the model tracks the solvents and impurities content in the cake (product CQAs) as the process evolves with time. For the isolation of paracetamol from a crystallization slurry, the filtration and drying model parameters are identified through experiments, respectively, on a Nutsche filter and on a thermogravimetric analyzer. The calibrated model is then used for determining the probabilistic design space and the maximum throughput for the isolation of paracetamol from a multi-component slurry, containing a non-volatile impurity, too.

The mathematical modeling framework of Chapter 4 is further developed into a real time simulator, which includes computational routines simulating sensors and actuators present in physical carousels, and supports the implementation of control loops. To calibrate the simulator, filtration and drying experiments are carried out on a pilot scale carousel for a paracetamol/ethanol slurry system. A closed-loop control strategy for the unit, based on the QbC paradigm, is then conceived and tested on the simulator, under a set of disturbances known to affect the unit operation. The proposed control strategy features model-based routines, such as real time optimization and state estimation, both based on the developed mathematical model

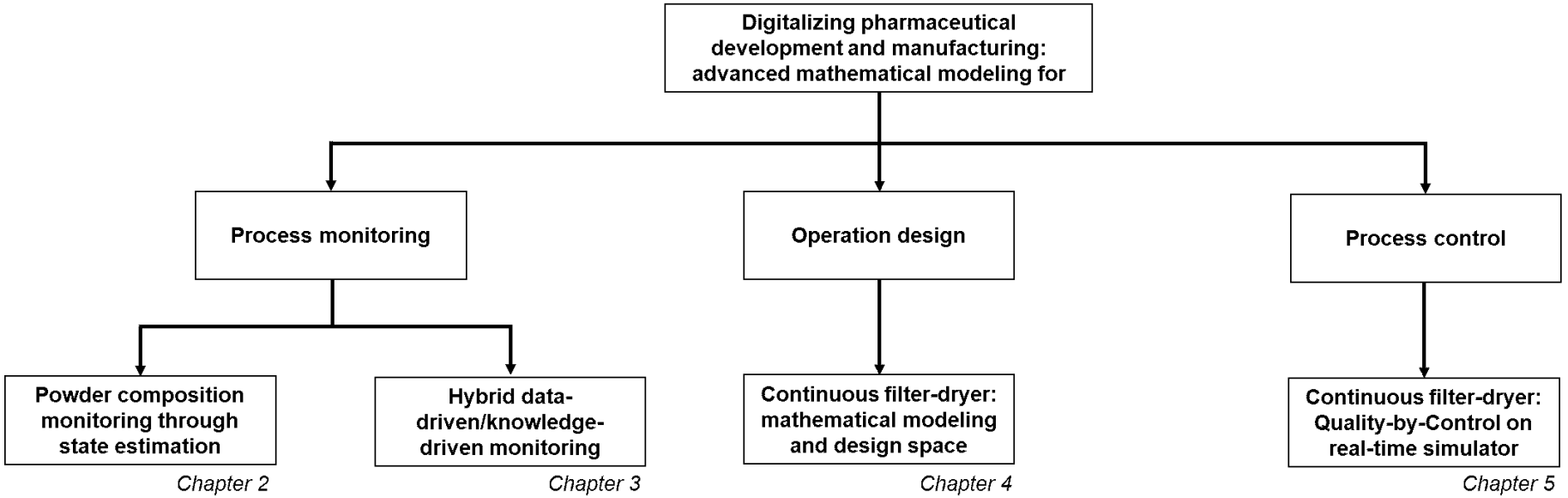


Figure 1.12. *Dissertation roadmap: scheme.*

Table 1.6. *Dissertation roadmap: aims and details of the research presented in each chapter.*

Chapter	Specific aim	Main application	Model	Case study	Data source	Novelty
2	1, 4, 5	Process monitoring	From literature	Continuous tableting	Industrial collaborator	Application
3	1, 4, 5	Process monitoring	From literature	Batch penicillin manufacturing Continuous fluid bed drying Continuous polycondensation	Simulated	Methodology
4	2, 4, 5, 6	Operation design	Developed	Continuous filter-dryer for API isolation	Experimental campaign	Model+Application
5	3, 4, 5, 6	Process control	Developed	Continuous filter-dryer for API isolation	Experimental campaign	Model+Control system

of the carousel. The designed QbC framework is the first proposed closed-loop control strategy for this type of unit, and it shows superior control performance when benchmarked against traditional QbD control strategies, which are based on open-loop quality control.

Table 1.6 provides a synoptic view of the content of each chapter. In Chapters 2 and 3, the employed mathematical models are taken from the literature. Model development activities, instead, have been carried out in Chapters 4 and 5. Knowledge-driven models are used in all chapters, alone, or in a hybrid fashion, combined with data-driven components.

All chapters include elements of continuous processing. Chapters 2, 4 and 5 are entirely focused on continuous processes, while Chapter 3 contains both batch and continuous case studies. Experimental data for the study of Chapter 2 have been kindly provided by an industrial collaborator. Chapter 3 contains only *in-silico* studies, while experiments for data collection have been carried out during the research activities presented in Chapters 4 and 5.

The novelty claimed for each chapter is of different type. While Chapter 2 presents the application of an established methodology (state estimation) to a new field (powder composition monitoring), Chapter 3 proposed a novel methodology for process monitoring (the hybrid monitoring framework). Chapters 4 and 5 present studies on the continuous carousel for intensified filtration, washing and drying of crystallization slurries, which is, by itself, a novel concept, scarcely studied before, and an enabler of end-to-end continuous processing. While Chapter 4 presents the first design space description for this type of unit, in Chapter 5, a closed-loop control system for the unit is proposed and tested for the first time.

Note that, since many different processes and mathematical methodologies are discussed across the Dissertation, the materials and methods information and the mathematical background are directly provided in the chapters where they are relevant, instead than in a separate chapter. Moreover, due to the many mathematical expressions contained in the Dissertation, the symbols for the variables, parameters and arrays encountered in the equations are chapter-dependent, and are thoroughly defined upon the first appearance in each chapter. Acronyms, despite being consistent across the Dissertation, are re-introduced again in each chapter, for sake of clarity.

Chapter 2

Powder composition monitoring in pharmaceutical continuous manufacturing through state estimation*

In this Chapter, we propose the use of a soft sensor, rooted on a physics-based model, for monitoring the composition of the powder fed to a continuous solid-dosage line. The soft sensor is a moving-horizon state estimator, which carries out model-based reconciliation of the feeder mass measurements, thus enabling accurate composition estimation of the powder mixture. Experimental datasets from a direct compression line are used to validate the methodology. Results demonstrate improvement in monitoring the powder composition in loss-in-weight feeding with respect to the traditional industrial approach, relying on statistical filters.

2.1 Introduction

Continuous manufacturing is gaining interest in the pharmaceutical community, due to the many advantages with respect to the so-far-dominant batch production mode (Lee et al., 2015; Plumb, 2005). Although many pharmaceutical operations are already carried out continuously, their integration in an end-to-end continuous processing framework still remains challenging. Focusing on solid-dosage manufacturing, an important bottleneck for the transition to continuous processing lays on powder feeding and blending. In the direct compression and direct encapsulation of powders, the powder composition cannot be changed after the feeders. Appropriate estimation and monitoring of the fed composition are therefore fundamental to develop robust control strategies in a continuous line.

Besides the active pharmaceutical ingredient (API), pharmaceutical tablets and capsules include excipients and lubricants. The materials must be fed according to formulations specifically designed for meeting either target physicochemical properties or proper manufacturability (e.g.,

*Destro, F., S. García-Muñoz, F. Bezzo, M. Barolo (2021). Powder composition monitoring in continuous pharmaceutical solid-dosage form manufacturing using state estimation – Proof of concept. *Int. J. Pharm.*, **695**, 120808.

API dissolution profile and powder flowability). In industrial operations, each material is supplied to the process through a separate loss-in-weight feeder. The powder, stored in a hopper, is gently pushed by a horizontal agitator into a rotating screws system, located below the hopper, from where the material is pushed out of the feeder.

Monitoring feeding operations is a challenging task. Despite recent progress in image analysis for powder composition monitoring (Facco et al., 2017; Galata et al., 2021) and the increasing availability of PAT tools for powder composition assessment along the manufacturing line, these advanced technologies are not established or widespread yet. As a result, the dispensed powder mixture composition is not directly measured in real time in pharmaceutical manufacturing. Instead, the mass composition of a given ingredient (e.g., the API) in the mixture is obtained by calculation, namely by dividing the ingredient mass flow by the sum of the mass flows of all ingredients. The mass flow of each ingredient is itself not measured, but is estimated numerically by the feeder as the ratio between a powder mass weight difference (loss in weight) in the hopper and the time interval across which this difference is measured. This estimated mass flow can exhibit a low signal-to-noise ratio, due to the propagation of noise from the powder net weight measurements (Gyürkés et al., 2020), and is further filtered and smoothed by proprietary algorithms embedded in the equipment. The effects of variations in the API fed concentration can potentially be seen (after axial mixing) through downstream concentration assessment (García-Muñoz et al., 2018) obtained by high-performance liquid chromatography (HPLC) or spectroscopic measurements (De Beer et al., 2011). However, as will be discussed in detail in Section 2.2, the dynamics of the variation of the API fed concentration, as calculated from the mass flow estimations, is not always in good agreement with the one assessed downstream. This issue is even more relevant with cohesive powders that are difficult to feed, and for which the net weight measurements are very noisy.

Better instantaneous estimation of feeder powder mass flows is desirable for improving the control strategy of a solid-dosage manufacturing line. Within the Quality-by-Design (Food & Drug Administration, 2004; Fisher et al., 2016) framework, practitioners are encouraged to modernize pharmaceutical manufacturing through the adoption of control strategies relying on risk analysis and solid process understanding. Fully understanding how the variability of process inputs (raw material properties and process operating conditions) impacts the process outputs (product critical quality attributes) is of paramount importance. In this perspective, improving the estimation of feeder powder mass flow is critical to implementing Quality-by-Design in solid-dosage manufacturing. However, this monitoring problem is largely unexplored in the literature, and industrial practice leaves this aspect up to the equipment vendors.

Recent studies on pharmaceutical powder feeding and blending design, monitoring and control for solid-dosage manufacturing are comprehensively reviewed elsewhere (Su et al., 2020). Several contributions address the selection of suitable feeder configurations (e.g., hopper capacity, screws characteristics) for given powder systems and process requirements

(Blackshields and Crean, 2018; Cartwright et al., 2013; Engisch and Muzzio, 2012). From the design space description side, García-Muñoz et al. (2017) assessed the impact of the feeders mass flow variability onto the drug product composition in a continuous direct compression process using a mathematical model for the process. Other studies focus on the improvement of feedback and feedforward control. Hanson (2018) demonstrated the advantages of introducing a ratio controller in the feeding system, linking the set-points of the mass flow controllers to the current API powder mass flow. Model predictive control of a feeding-blending system in a continuous pharmaceutical tableting plant was reported by Singh et al. (2014). The authors showed improved performance in the control of the API content in the tablets for servo problems, with respect to traditional proportional-integral-derivative control. Nonetheless, fluctuations in the API concentration in the products were still observed, mostly because of the propagation of measurement inaccuracies into the control loop. In Rehrl et al. (2016), the benefits of model predictive control implementation in a blending system are further outlined through a simulation study. The authors point out that, for applications in physical units, it is necessary to put in place a state estimator, to provide accurate estimations of unmeasured state variables (states), such as mass flows and composition.

State estimation (Ray, 1981) is a family of mathematical techniques that, given a mathematical model of a process and real time measurements from a plant, estimate the unmeasured system states and provide noise-attenuated estimations of the actual measurements. Moving-horizon estimation (Rao, 2000; Rao and Rawlings, 2002) estimates the current system states with an optimization-based strategy. For nonlinear systems, moving-horizon estimation is known to outperform alternative state estimation techniques in many applications (Haseltine and Rawlings, 2005).

Very few studies address state estimation in solid-dosage form manufacturing, despite the great advantages it can bring. Recently, Liu et al. (2018) tested a robust moving-horizon estimator (MHE) on a pharmaceutical feeding-blending system. They conducted an *in-silico* analysis under the ideal assumption of no plant/model mismatch. The state variables relevant to the process (powder hold-ups in feeders and blenders, powder mass flows and powder mixture composition across the system) were successfully reconstructed by the estimator, also in the presence of gross measurement errors, such as outliers and drifts.

In this study, we apply state estimation techniques to obtain a much improved instantaneous estimate of the mass flow of material being dispensed by a screw feeder in a continuous line. We develop a novel MHE-based powder composition monitoring scheme for the feeding and blending section of a real continuous tableting line, following a Quality-by-Design approach rooted in enhanced process understanding. The powder composition monitoring scheme developed encompasses the system from feeder to powder in the die, and does not include the compaction system in the tablet press. The process, detailed elsewhere (García-Muñoz et al., 2018), is made up of three sections: feeding, blending, and compression. An MHE is

implemented in the feeding section, based on a first-principles mathematical model describing loss-in-weight feeder operation (Bascone et al., 2020). The MHE estimates the outlet powder mass flow from each feeder using a model of the feeders and measurements of the motor speeds and load cell masses. This in turn enables estimating the API concentration fed downstream, so that a comparison with actual measurements is possible. The intermediate blending effect is accounted for through an additional mathematical model calibrated with tracing experiments or step change experiments. We benchmark the proposed monitoring approach against a typical industrial one that uses API concentration values calculated from the mass flow estimations as obtained from the software embedded in the feeders.

The remainder of this chapter is organized as follows. The problem is stated formally in Section 2.2. Section 2.3 describes the experimental setup and provides mathematical background to MHE and feeders modeling. The workflow for implementing the proposed monitoring system on the direct compression line is discussed in Section 2.4. Realtime monitoring of the line is demonstrated in Section 2.5, before of the concluding remarks.

2.2 The need for a better flowrate estimation

Let us consider a direct compression line made by F feeders, each one providing a different ingredient to the powder mixture. The mass flow estimations are provided to the user by the software interface of feeder i (i = 1, …, F). Namely, the following estimation signals are available for our particular case study:

- $\dot{m}_i^{inst}$, the raw instantaneous mass flow calculated as ratio between the difference of two consecutive measurements of powder mass weight in the hopper (loss in weight) and the difference between their respective sampling times;
- $\dot{m}_i^{avg}$, a smoothed version of $\dot{m}_i^{inst}$, calculated using a thirty-second moving average;
- $\dot{m}_i^{smooth}$, an additional smoothed version of $\dot{m}_i^{inst}$, obtained by a strong smoothing of $\dot{m}_i^{inst}$ through proprietary algorithms embedded in the equipment.

The API theoretical fed concentration $c_{API}^{theoretical}$ is calculated for this case study using the moving-average mass flow signals, according to:

$$c_{API}^{theoretical} = \frac{\dot{m}_i^{avg}}{\sum_{i=1}^{F} \dot{m}_i^{avg}} \quad , \tag{2.1}$$

where the API is fed from the first feeder (i = 1). For comparison, we define c_{API}^{inst} as the API concentration calculated from the raw instantaneous signals:

$$c_{API}^{inst} = \frac{\dot{m}_i^{inst}}{\sum_{i=1}^{F} \dot{m}_i^{inst}} \quad . \tag{2.2}$$

Additionally, spectroscopic and HPLC measurements are taken downstream, on both the powder mixture entering the tablet press feed frame and the final tablets. It was observed that certain downstream powder composition variations measured with spectroscopic tools and HPLC are not registered by c_{API}^{inst}, which is a quite noisy signal (Figure 2.1a). On the other hand, the API concentration calculated from $\dot{m}_i^{smooth}$ (not shown in Figure 2.1) cannot fully capture the composition variation dynamics due to oversmoothing. For this reason, a moving average mass flow $\dot{m}_i^{avg}$ can be carefully configured to make these dynamics evident upstream. Nonetheless, the API fed concentration inferred from $\dot{m}_i^{avg}$ (i.e., $c_{API}^{theoretical}$) still shows a low signal-to-noise ratio compared to the direct composition assessment, and lags behind fast-changing variations, as a consequence of the moving-average smoothing (Figure 2.1b).

Better instantaneous estimates of mass flow for improved quality assurance are therefore desirable, as the feeders are critical elements of the control strategy.

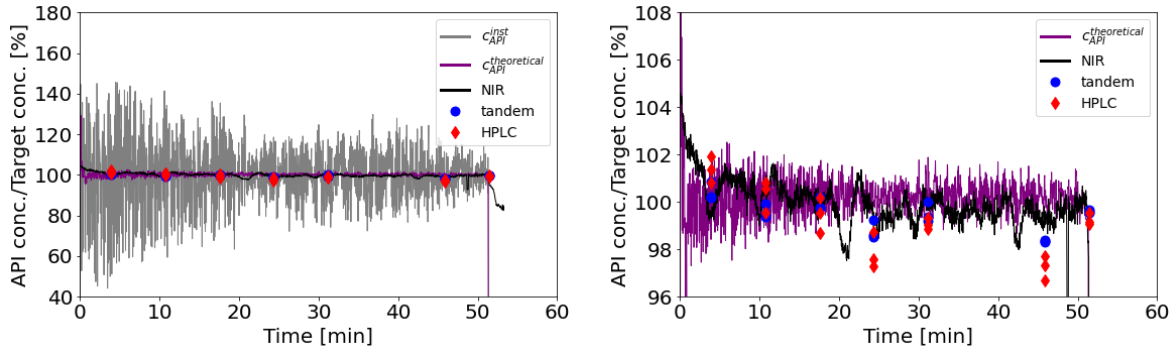


Figure 2.1. *Time profiles of the ratio between actual API concentrations and target API concentrations from a direct compression line (data from Dataset A discussed in Section 2.3). (a) Full view; (b) zoom-in. The concentration values are determined according to different methods: i) c_{API}^{inst}; ii) $c_{API}^{theoretical}$; iii) assessed on the powder mixture that enters the tablet press feed frame using near-infrared (NIR) spectroscopy; iv) assessed on the final tablets using a Bruker Tandem IIIA FT-NIR instrument (Bruker Corp. Germany); v) assessed on the final tablets using HPLC.*

2.3. Materials and methods

2.3.1 Process description and experimental data collection

The experimental data for designing and validating the proposed monitoring methodology were collected in a continuous direct compression pilot plant for a molecule in development (García-Muñoz et al., 2017; Figure 2.2). Five ($F = 5$) loss-in-weight feeders (Coperion K-Tron KT20) dispense the API and the other four ingredients of the formulation to the downstream continuous horizontal mixer. The powder is then discharged into a vertical hopper and, after passing

through a transition piece, enters the feed frame of the rotary tablet press. All operations occurring in between the feeders and the press feed frame form the blending section of the process. Powder blending is necessary for mixing the formulation ingredients, and allows partially compensating for feeder disturbances. Powder blending is known to occur not only in the horizontal mixer but in the tablet press feed frame. The specific details about the formulation are irrelevant for the study and are withheld for confidentiality.

The loss-in-weight feeders (Figure 2.3) are made up of a motor-driven screw system, an agitator, a hopper where the material is stored, and a load cell installed below the unit. The measurements provided by the feeder and used by the proposed monitoring system as if they were available in real time are: *i*) the net weight m_i of material in feeder, and *ii*) the motor drive command DC_i, ranging from 0 to 100%.

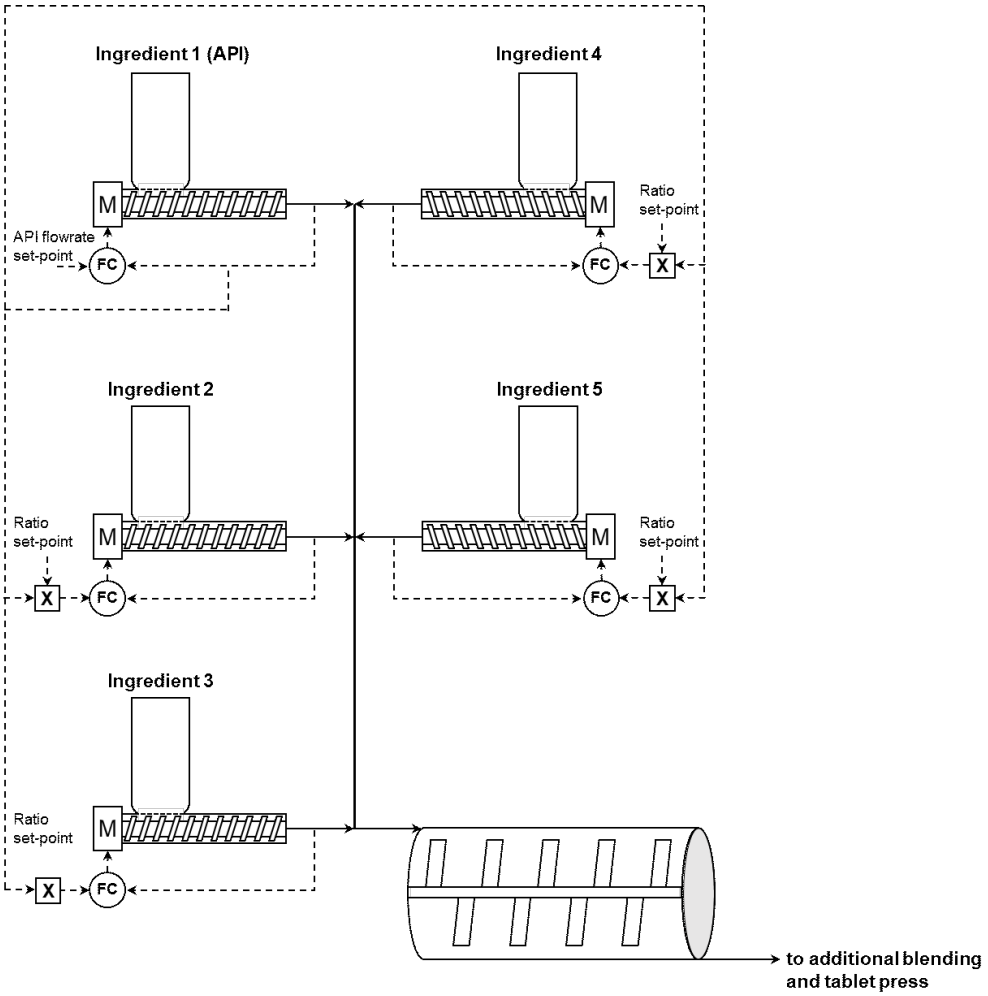


Figure 2.2. *Schematic of the direct compression line used for data collection. Five loss-in-weight feeders operate in parallel, sending the powder formulation to the horizontal mixer. Additional powder blending occurs in the feed frame of the tablet press, before the powder is fed in the die. The blending model in this work encompasses all the mixing occurring before the powder enters the die.*

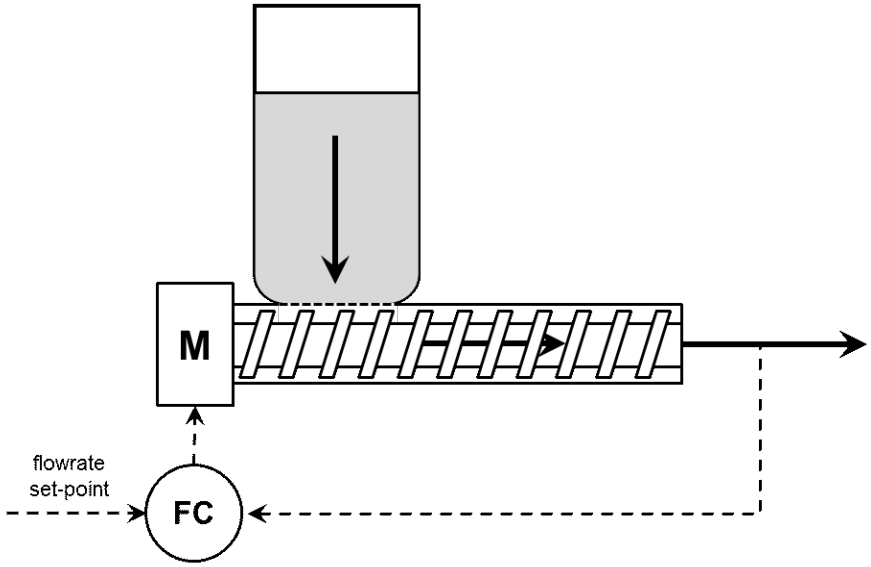
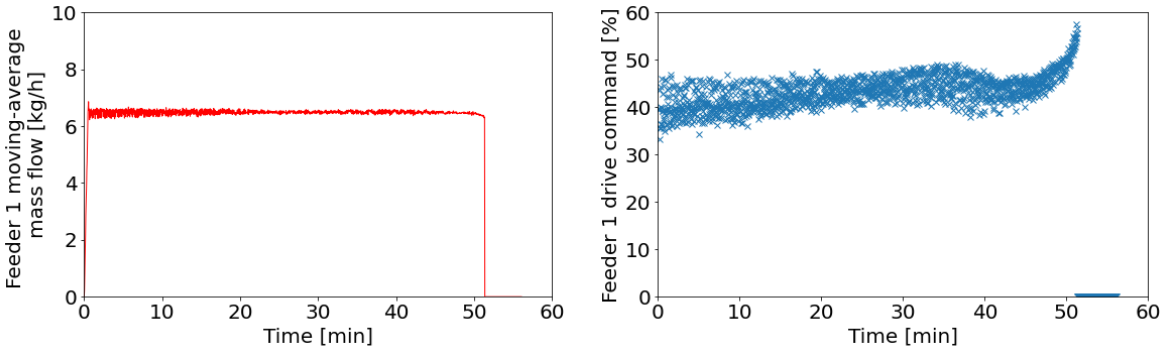


Figure 2.3. *Schematic of a loss-in-weight feeder. The powder stored in the hopper, aided by an agitator (not shown), flows into the twin-screw system, from where it is pushed out of the feeder. A load cell (not shown) measures the mass of powder in the unit. From the mass measurements, a powder mass flow measurement is inferred and used by the flow controller for adjusting the motor (M) speed, to track the reference mass flow setpoint.*

The feeders measurements DC_i and m_i and the three aforementioned (Section 2.2) mass flow estimations $\dot{m}_i^{inst}$, $\dot{m}_i^{avg}$ and $\dot{m}_i^{smooth}$ are updated by the internal feeder software through high-frequency readings, but they are all made available to the user only every 2 s. The feeders can operate at open loop (volumetric mode) or at closed loop (gravimetric mode). When operating in closed-loop mode, the internal mass flow controller uses the high-frequency measurements (not accessible to the user) to keep the mass flow at the relevant setpoint, by adjusting DC_i. A ratio controller (Hanson, 2018) continuously receives the high-frequency API mass flow estimation and, based on that, updates the set points of the other flow controllers, in such a way as to enforce the target concentration for each ingredient.

The API concentration in the powder mixture in the feed frame was estimated through near-infrared (NIR) spectroscopy. Extended iterative optimization technology (EIOT), a recently-developed methodology (Shi et al., 2019) was used to correlate the NIR signal to the API concentration, here denoted as c_{API}^{NIR}, using optimization to enforce a modified Beer-Lambert law to the spectra. Discrete API concentration measurements were also obtained on tablets sampled at the press outlet. HPLC and concentration estimations derived from data collected with a Bruker Tandem IIIA FT-NIR instrument (Bruker Corp. Germany) were used to this purpose, and the relevant measurements are indicated with c_{API}^{HPLC} and c_{API}^{tandem}, respectively. While c_{API}^{NIR} was made available to the user every 2 s, only few measurements of c_{API}^{HPLC} and c_{API}^{tandem} were collected during the operation.

In this study, three datasets (Table 2.1) were collected during closed-loop continuous runs (without hopper refills) for a molecule in development in a pilot plant facility. This data was used for calibration and validation of the monitoring system. The profiles of $\dot{m}_1^{avg}$ and DC_1 for Dataset A in feeder 1 are shown in Figure 2.4 to illustrate how a typical closed-loop operation evolves. As time progresses, powder densification in the screws decreases. To maintain $\dot{m}_1^{avg}$ close to the set-point of 6.5 kg/h (Figure 2.4a), the controller progressively increases DC_1 (Figure 2.4b), thus increasing the number of revolutions per minute of the feeder screw.



(*a*) (*b*)

Figure 2.4. *Dataset A, feeder 1: time profiles of (a) $\dot{m}_{1,avg}$ and (b) DC_1. The built-in feeder flow controller manipulates DC_1 to keep the mass flow at its set-point. While time progresses during a run, powder densification in the screw decreases, due to the smaller hold-up in the hopper. The controller responds by increasing DC_1 (hence the number of revolutions per minute of the screw).*

Table 2.1. *List of datasets collected during continuous runs of the direct compression line.*

	Final tablet mass [mg]	**API target concentration in tablet [%]**	**Experiment duration [min]**
Dataset A	50	40	55
Dataset B	150	40	35
Dataset C	200	40	35

All datasets in Table 2.1 refer to the same tablet formulation. Although the final mass of the tablets obtained varies across datasets, the feeding and blending sections always operate in the same way. The set of time-dependent variables characterizing each dataset is summarized in Table 2.2. As will be detailed later, DC_i and m_i are the only measurements required by the proposed monitoring system; the other variables are used for validating and benchmarking it. Additional real-time measurements available through the feeder software interface, but not meaningful for process monitoring, are not discussed here.

Table 2.2. *List of time-dependent variables characterizing each dataset. A comprehensive list of symbols is reported as a separate section.*

Symbol	Type	Variable	Measurement unit
c_{API}^{HPLC}	Measured	API concentration measured with HPLC on tablets	[kg/kg]
c_{API}^{NIR}	Measured	API concentration measured with NIR on powder mixture entering the feed frame	[kg/kg]
c_{API}^{tandem}	Measured	API concentration estimated with spectra collected from tablets with a Bruker Tandem IIIA FT-NIR instrument	[kg/kg]
c_{API}^{inst}	Calculated	Fed concentration of API calculated from $\dot{m}_i^{inst}$ (Equation 2.2)	[kg/kg]
$c_{API}^{theoretical}$	Calculated	Theoretical fed concentration of API (Equation 2.1)	[kg/kg]
DC_i	Measured	Motor drive command of feeder *i*	[%]
m_i	Measured	Powder mass in feeder *i*	[kg]
$\dot{m}_i^{inst}$	Calculated	Mass flow from feeder *i* obtained as rate of change of measured mass (raw instantaneous mass flow)	[kg/h]
$\dot{m}_i^{avg}$	Calculated	Mass flow from feeder *i* obtained as thirty-seconds moving-average of $\dot{m}_i^{inst}$	[kg/h]
$\dot{m}_i^{smooth}$	Calculated	Mass flow from feeder *i* obtained by smoothing $\dot{m}_i^{inst}$ with a statistical filter	[kg/h]

2.3.2 State estimation: moving-horizon estimation of state variables

Let us consider the mathematical model of a process. The model predicts the time evolution of the system states according to:

$$\frac{\mathrm{d}\tilde{\mathbf{x}}(t)}{\mathrm{d}t} = \mathbf{F}(\tilde{\mathbf{x}}(t), \mathbf{u}(t)) \,, \tag{2.3}$$

where $\tilde{\mathbf{x}} \in \mathbb{R}^{n_x}$ is the vector of the n_x predicted states, $\mathbf{u} \in \mathbb{R}^{n_u}$ is the vector of the n_u system inputs, $\mathbf{F}(\cdot)$: $\mathbb{R}^{n_x+n_u} \mapsto \mathbb{R}^{n_x}$ is the vector function representing the model, and t is the time. The system states may be measured or not. Note that in the remainder of the manuscript, the tilde symbol "$\tilde{}$" on top of a variable denotes that the variable has been calculated by means of a model prediction.

The continuous-time model (3) can be expressed in discrete-time form as:

$$\tilde{\mathbf{x}}(t_{k+1}) = \mathbf{f}(\tilde{\mathbf{x}}(t_k), \mathbf{u}(t_k)) \,, \tag{2.4}$$

in which t_{k+1} is the time step subsequent to t_k, and $\mathbf{f}(\cdot)$: $\mathbb{R}^{n_x+n_u} \mapsto \mathbb{R}^{n_x}$ is the discrete-time form of the model.

Assume that n_y variables are measured from the process at a given time step t_k, and that they are collected in vector $\mathbf{y} \in \mathbb{R}^{n_y}$. A prediction of the measured variables can be obtained by:

$$\tilde{\mathbf{y}}(t_k) = \mathbf{g}(\tilde{\mathbf{x}}(t_k)) \,, \tag{2.5}$$

where $\tilde{\mathbf{y}} \in \mathbb{R}^{n_y}$ is the vector of the measured values as predicted by the model, and $\mathbf{g}(\cdot)$: $\mathbb{R}^{n_x} \mapsto \mathbb{R}^{n_y}$ is the vector function describing the measurement model.

Even though a mathematical model can provide real-time predictions at any time step as a soft-sensor, unfortunately this piece of information cannot be used on its own for process monitoring due to the effect of the accumulation of error. Modeling errors (often referred to as plant/model mismatch) appear since the state variables predicted by Equation 2.4 are different from the actual (and unknown) ones, as evidenced by the fact that $\tilde{\mathbf{y}}(t_k) \neq \mathbf{y}(t_k)$. On the other hand, also process monitoring based on real-time measurements only may be ineffective, due to both lack of measurements and measurement noise.

State estimation (Ray, 1981) is a family of techniques that combines the use of the mathematical model of a system and of real-time measurements from the system to improve the prediction of the system states over the one provided by the model on its own. Namely, at each time step the state estimator calculates an estimation of the states in such a way as to reconcile the values of the measurements predicted by the model at that time step (or within a given time window including it) to the actual measurements at the same time step (or time window). State estimation adjusts the model at each sampling instant using the available measurements, thus improving the model predictions by providing accurate estimates for the states, namely the unmeasured quantities of interest.

In this study, we make use of moving-horizon state estimation (Rao, 2000; Rao and Rawlings, 2002). An intuitive representation of the rationale behind an MHE is illustrated in Figure 2.5 and discussed in the following text, while a formal mathematical background for MHE is provided at the end of this subsection. At the current time step t_k (Figure 2.5a), the MHE estimates the states of the system by finding a trade-off between the predictions from the model and the measurements available in a receding horizon $T(t_k) = \{t | t_{k-H} \leq t \leq t_k\}$ that includes the past *H* time steps (*H* is called window size). Instead, older measurements and estimations are retained in the so-called archived horizon (where the estimated states are not updated anymore). This is done practically by solving an optimization problem, namely by finding the set of estimated states in $T(t_k)$ that minimize an objective function that is basically made by the sum of the squared model errors and measurement errors within the receding horizon. At time step t_l inside the receding horizon, the vectors of model error $\mathbf{w} \in \mathbb{R}^{n_x}$ and measurement error $\mathbf{v} \in \mathbb{R}^{n_y}$ are defined as follows:

$$\mathbf{w}(t_l) = \hat{\mathbf{x}}(t_{l+1}) - \tilde{\mathbf{x}}(t_{l+1}), \quad \text{for } l = k-H, k-H+1, \dots, k-1 \tag{2.6}$$

$$\mathbf{v}(t_l) = \mathbf{y}(t_l) - \mathbf{g}\big(\hat{\mathbf{x}}(t_l)\big), \quad \text{for } l = k-H, k-H+1, \dots, k \quad , \tag{2.7}$$

where $\hat{\mathbf{x}} \in \mathbb{R}^{n_x}$ is the vector of the estimated states (note that in the remainder of the manuscript, the caret symbol "$\widehat{\ }$" on top of a variable denotes that the variable has been calculated by means of state estimation). At the left boundary of the window, instead of the model error, the arrival cost (which depends on the estimation obtained at time step t_{k-1} instead that on a model prediction) is considered in the objective function of the optimization problem, accounting for

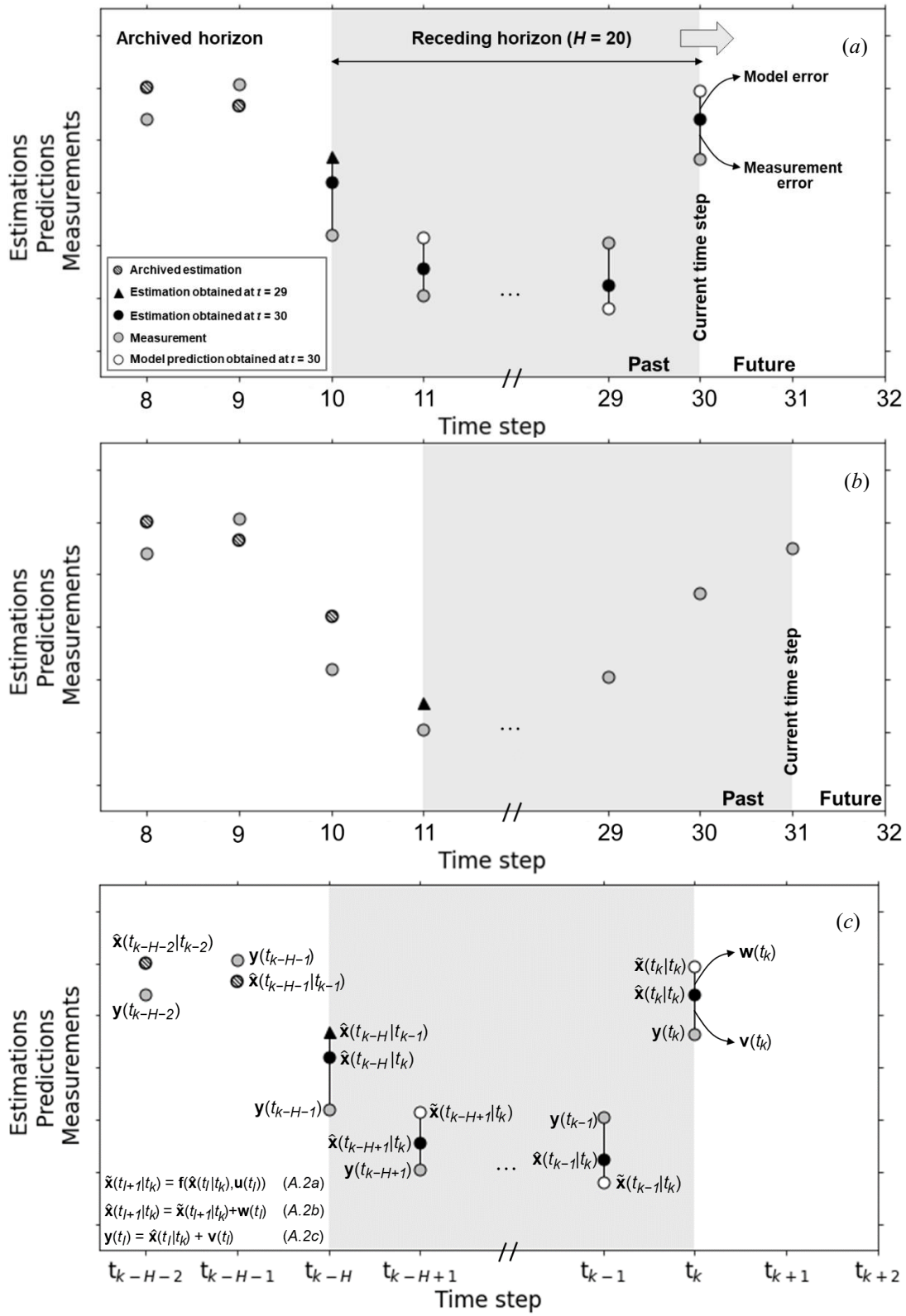


Figure 2.5. *Schematic view of an MHE. For illustrative purposes, we assume that the measured variables coincide with the system states ($\boldsymbol{g}(\hat{\boldsymbol{x}}(t_l)) = \hat{\boldsymbol{x}}(t_l)$ in Equation 2.7). (a): State estimation at current time step ($t_k = 30$). (b): Window slide at next time step ($t_k = 31$), before of performing the state estimation procedure again. (c): Schematic view of the MHE of Figure 5a, proposed again using the advanced notation for estimations, predictions and measurements introduced in Equations 2.8-2.10.*

all the discarded measurements now laying inside the archived horizon. At the next time step t_{k+1} (Figure 2.5b), when new measurements become available, the window slides one time step forward. The new measurement set is included in the receding horizon, while the oldest one is discarded and enters the archived horizon. The estimation procedure is then iterated. The arrival cost at t_{k+1} is calculated using the estimation at the new left boundary of the window obtained at t_k.

A formal mathematical definition of the mathematical background of MHE is further provided. Let us consider a time step t_l lying inside the receding horizon $T(t_k)$. We further detail the introduced notation by denoting with $\hat{\mathbf{x}}(t_l|t_k)$ the estimation of the system states at time t_l when the system is at time t_k (i.e., the estimation is based on all measurements within the receding horizon). At the current time step t_k, the optimal estimation of the states inside $T(t_k)$, namely $\{\hat{\mathbf{x}}(t_{k-H}|t_k), \dots, \hat{\mathbf{x}}(t_k|t_k)\}$, is obtained by solving the following nonlinear programming problem:

$$\min_{\hat{\mathbf{x}}(t_{k-H}|t_k),\mathbf{w}(t_{k-H}),\dots,\mathbf{w}(t_{k-1})} \{\Theta\big(\hat{\mathbf{x}}(t_{k-H}|t_k)\big) + \frac{1}{2}\sum_{t=t_{k-H}}^{t_k} \mathbf{v}(t)^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{v}(t) + + \frac{1}{2}\sum_{t=t_{k-H}}^{t_{k-1}} \mathbf{w}(t)^{\mathrm{T}}\mathbf{Q}^{-1}\mathbf{w}(t)\} \quad (2.8)$$

subject to:

$$\tilde{\mathbf{x}}(t_{l+1}|t_k) = \mathbf{f}(\hat{\mathbf{x}}(t_l|t_k), \mathbf{u}(t_l)) \quad (2.9a)$$

$$\hat{\mathbf{x}}(t_{l+1}|t_k) = \tilde{\mathbf{x}}(t_{l+1}|t_k) + \mathbf{w}(t_l) \quad (2.9b)$$

$$\mathbf{y}(t_l) = \mathbf{g}(\hat{\mathbf{x}}(t_l|t_k)) + \mathbf{v}(t_l) \quad (2.9c)$$

$$\mathbf{w}(t_l) \in \mathbf{W}_l, \quad (2.9d)$$

$$\hat{\mathbf{x}}(t_{l+1}|t_k) \in \mathbf{X}_l, \quad (2.9e)$$

where, $\Theta(\tilde{\mathbf{x}}(t_{k-H}))$ is referred to as the arrival cost, $\mathbf{R}\in\mathbb{R}^{n_y \times n_y}$ is the measurement error variance matrix, $\mathbf{Q}\in\mathbb{R}^{n_x \times n_x}$ is the model error variance matrix, $\tilde{\mathbf{x}}(t_{l+1}|t_k)$ is the vector of predicted states at time step t_{l+1}, obtained at t_k from initial conditions $\hat{\mathbf{x}}(t_l|t_k)$, and $\mathbf{W}_l$ and $\mathbf{X}_l$, (which define the boundaries of, respectively, $\mathbf{w}(t_l)$ and $\hat{\mathbf{x}}(t_l)$, are taken as polyhedral convex sets.

The objective function in Equation 2.8 represents in mathematical form the request of finding the optimal set of estimated states that minimize the sum of squared model errors and measurement errors within the receding horizon. Note that Equations 2.9b and 2.9c are, respectively, equivalent to Equations 2.6 and 2.7. The arrival cost accounts for all the data collected in the archived horizon ($0 \le t < t_{k-H}$), not directly considered in the objective function (Equation 2.8). It is defined as:

$$\Theta(\hat{\mathbf{x}}(t_{k-H}|t_k)) = \frac{1}{2}\Big((\hat{\mathbf{x}}(t_{k-H}|t_{k-1}) - \hat{\mathbf{x}}(t_{k-H}|t_k))^{\mathrm{T}}\,\mathbf{P}^{-1}(t_{k-H}|t_{k-1})\big(\hat{\mathbf{x}}(t_{k-H}|t_{k-1}) - \hat{\mathbf{x}}(t_{k-H}|t_k)\big)\Big) \quad (2.10)$$

in which $\mathbf{P}(t_{k-H}|t_{k-1})$ is the covariance of the estimation of $\hat{\mathbf{x}}(t_{k-H}|t_{k-1})$.

The goodness of the MHE estimation increases with both *H* and the arrival cost computation accuracy. Hence, the computational burden for solving an MHE problem can be reduced by using a smaller *H*, and compensating with an improved approximation of the arrival cost. In this study, we follow a recently proposed approach for the calculation of the arrival cost (López-Negrete and Biegler, 2012), which computes $\mathbf{P}(t_{k-H}|t_{k-1})$ from the reduced Hessian of Problem 2.8-2.10 at the solution. The mathematical notation introduced to formalize the MHE problem is summarized in Figure 2.5c, to be compared with the qualitative schematic view of the MHE algorithm presented in Figure 5.

In the MHE formulation of Equations 2.8-2.10, $\mathbf{v}(t_l)$ and $\mathbf{w}(t_l)$ are assumed to be zero-mean Gaussian processes of variances **R** and **Q**, respectively. In the presence of gross errors, robust implementations of MHE can be resorted to (Nicholson et al., 2014). Matrices **Q** and **R** are regarded as tuning parameters, respectively representing the confidence in the model and in the measurements. If the model is trusted more than the measurements, due to significant noise, larger elements for the matrix **R** should be selected than for **Q**, and vice versa.

2.3.3 Feeder model description

Many data-driven and hybrid first-principles/data-driven models for loss-in-weight feeders are described in the literature. In this study, we use a model recently proposed by Bascone et al. (2020). The model is made up of two differential equations:

$$\frac{\mathrm{d}\tilde{m}_i(t)}{\mathrm{d}t} = \dot{m}_i^{refill}(t) - \tilde{\dot{m}}_i(t) \tag{2.11}$$

$$\tau_i \frac{\mathrm{d}\tilde{\dot{m}}_i(t)}{\mathrm{d}t} + \tilde{\dot{m}}_i(t) = n\, PA\, \eta\, \rho_i(t)\, N_i(t), \tag{2.12}$$

where $\tilde{m}_i$ is the predicted mass of powder in the feeder, $\dot{m}_i^{refill}$ is the powder refill mass flow, $\tilde{\dot{m}}_i$ is the predicted powder mass flow at the feeder outlet, τ_i is the feeder time constant (calibration parameter), n is the number of starts of the screw thread, *P* is the pitch of the screw, *A* is the cross-sectional area available for powder flow in the screw, η is the volumetric efficiency of the feeder (ratio between conveyed powder volume and screw volume available during one revolution, calculated from the material and feeder geometrical properties as in Bascone et al., 2020), N_i is the screw rotation speed, and ρ_i is the time-variable effective density of the powder in the screws.

The effective density is given by:

$$\rho_i(t) = \rho_i^0 + k_i \ln(\sigma_V(t)) \quad , \tag{2.13}$$

where ρ_i^0 is the density in the absence of vertical stress due to the powder in the hopper (calibration parameter), k_i is an additional calibration parameter, and σ_V is the vertical stress in the hopper, expressed in [Pa]. The vertical stress σ_V varies with time, and is a nonlinear function of $\tilde{m}_i$; in this study, we calculate σ_V with the set of equations proposed by Bascone et al. (2020) for Coperion K-Tron KT20 feeders.

The screw rotation speed N_i is obtained from DC_i and the maximum screws rotation speed N_{max} as:

$$N_i(t) = \frac{DC_i}{100} N_{max} \quad . \tag{2.14}$$

Overall, the dynamic model described by Equations 2.11-2.14 has (for each feeder) two states ($\tilde{m}_i$ and $\dot{\tilde{m}}_i$), two inputs (DC_i and $\dot{m}_i^{refill}$), and three calibration parameters (k_i, ρ_i^0 and τ_i). Materials physical properties and quantitative geometrical information on the twin-screw system configuration used in the process of interest are not disclosed due to confidentiality reasons. The main geometrical properties of KT20 feeders are reported in Bascone et al. (2020).

2.4 Implementation of the powder composition monitoring system

In this section, we present the general workflow required for implementing the powder composition monitoring system in a continuous direct compression line (Figure 2.6). The steps to be carried out are: *i*) feeder model calibration, *ii*) blending model calibration, and *iii*) MHE design. A general overview of the monitoring system is given in Subsection 2.4.1, together with remarks on the role of the blending model and the procedure for its calibration. For the feeding section, model calibration and state estimator design will be described in detail in Sections 2.4.2 and 2.4.3, respectively. We remind that the caret symbol $\hat{}$ on top of a variable means that the variable has been obtained through state estimation, while variables with the tilde symbol $\tilde{}$ on top have been obtained through model predictions. Concentrations and mass flows without any symbol on top are instead measurements, either direct or indirect (Table 2.2).

$$\hat{c}_{API}^{pre-blending}(t) = \frac{\hat{\dot{m}}_1(t)}{\sum_{i=1}^{5} \hat{\dot{m}}_i(t)} \tag{2.15}$$

The blending effect throughout the line up to the point where the powder enters the die in the tablet press is modeled through a first-order-plus-dead-time dynamics (FOPDT; Seborg et al., 2017), accounting for the powder axial mixing and the residence time in this section of the process:

$$\tau_{blending} \frac{d\hat{c}_{API}^{blended}(t)}{dt} + \hat{c}_{API}^{blended}(t) = \hat{c}_{API}^{pre-blending}(t - \theta_{blending}) \, , \tag{2.16}$$

where $\hat{c}_{API}^{blended}$ is the estimated API concentration at the blender outlet, and the calibration parameters $\theta_{blending}$ and $\tau_{blending}$ are, respectively, the blending dead time and time constant. As a remark, $\hat{c}_{API}^{blended}$ is the variable of interest for monitoring purposes, representing the API concentration that will also characterize the final tablets.

The blending model is calibrated offline with data from a step experiment (Figure 2.7) taken from the literature (García-Muñoz et al., 2018) and carried out in the same blending units used for the experimental activities of this study. The estimated values of the blending parameters are: $\theta_{blending}$ = 60 s and $\tau_{blending}$ = 35 s. For simplicity, the blending model is used in a purely predictive fashion, instead of being embedded in the state estimation framework.

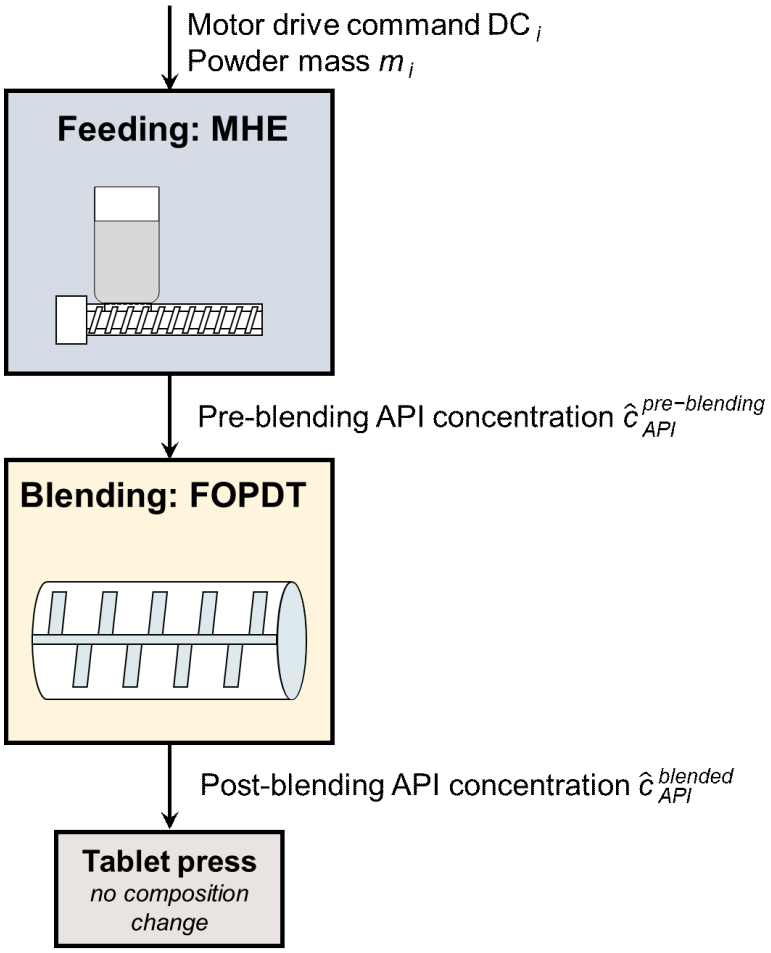


Figure 2.6. *Scheme of the proposed monitoring system for a direct compression line. The moving-horizon estimator uses real-time measurements of motor drive command and of powder mass from each feeder, and estimates the pre-blending API concentration. Powder blending occurring in the horizontal mixer and (to a lesser extent) in the vertical surge hopper and in the transition piece is accounted for through a first-order-plus-dead-time (FOPDT) dynamics. The API concentration in the blended powder mixture entering the feed frame is eventually calculated.*

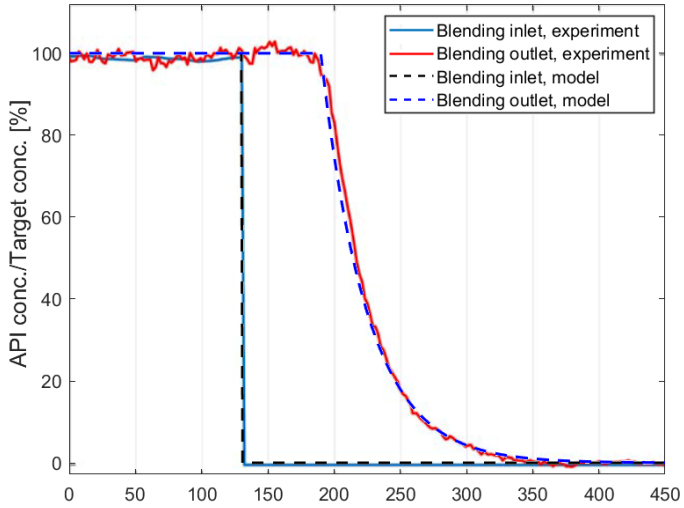


Figure 2.7. *Step change experiment for blending model calibration. Experimental data are reported as solid lines and are taken from Figure 5 in García-Muñoz et al. (2017).*

2.4.1 Feeder model calibration

Dataset A (Table 2.2) is used for calibrating the feeder model (Section 2.3.3) to each of the five feeders supplying the ingredients of the formulation. Maximum-likelihood estimation (MLE) under the assumptions discussed by Rawlings et al. (1993) is used for model parameter estimation, minimizing the objective function:

$$\Phi_{\text{MLE}}(\boldsymbol{\theta}_i) = \frac{n_{exp}}{2} \ln\left(\sum_{t_k=0}^{t_f} (\tilde{m}_i(t_k, \boldsymbol{\theta}_i) - m_i(t_k))\right), \tag{2.17}$$

where n_{exp} is the number of experimental points in the dataset, t_f is the final time step of the experiment, and $\boldsymbol{\theta}_i$ is the set vector of estimated parameters for feeder *i*. Note that, for a given feeder *i*, $\boldsymbol{\theta}_i$ needs to be estimated again if the ingredient, the type of feeder, or the feeder characteristics (e.g., hopper or screws) are changed during the process life cycle. Bascone et al. (2020) indicate a linear dependence on N_i for all calibration parameters. From preliminary parameter estimation results, we conclude that, for the process of interest, the calibration parameters are constant and do not depend on N_i, except for ρ_i^0, which is calculated according to:

$$\rho_i^0 = \rho_i^{0,intercept} + \rho_i^{0,slope} N_i, \tag{2.18}$$

where $\rho_i^{0,intercept}$ and $\rho_i^{0,slope}$ respectively are the intercept and the slope of the equation representing the linear dependency of ρ_i^0 on N_i. Hence, θ_i defined as:

$$\boldsymbol{\theta}_i = [k_i \; \rho_i^{0,intercept} \rho_i^{0,slope} \; \tau_i] \tag{2.19}$$

We solve the optimization problem for parameter estimation in the Pyomo environment, based on Python (Hart et al., 2017). The feeders model is first coded into the Pyomo environment in the time-continuous form of Equations 2.11-2.14. Then, the set of differential-algebraic equations are transformed in time-discrete form using a simultaneous strategy (Biegler, 2007), exploiting the Pyomo.dae toolbox (Nicholson et al., 2018).The differential equations are approximated with polynomials using Radau orthogonal collocation (Biegler, 2010), with 1500 finite elements (corresponding to n_{exp}) and 3 collocation points. The obtained time-discrete feeder model equations are used as constraints in the optimization problem for minimizing Φ_{MLE} (Equation 2.17). The solution is reached using the large-scale nonlinear solver IPOPT (Wächter and Biegler, 2006), providing the optimal estimation of parameters $\widehat{\boldsymbol{\theta}}_i$ for i = 1, 2, …, 5 (Table 2.3). The corresponding estimation uncertainty is given by the variance-covariance matrix (whose diagonal is reported in standard deviation form in Table 2.3), approximated as the inverse of the Hessian of the objective function (Equation 2.17) at the solution (Bard, 1974). For computing the Hessian at the solution, we use sIPOPT (Pirnay et al., 2012), an extension of IPOPT. All parameters are estimated with sufficiently small uncertainty, except $\rho_4^{0,slope}$. However, we verified that this parameter has a low sensitivity on the API concentration in the tablet.

Table 2.3. *Feeder model: parameter estimation results. For each parameter, the estimation uncertainty is reported as standard deviation from the estimated value.*

Parameter	Feeder 1		Feeder 2		Feeder 3		Feeder 4		Feeder 5	
	Estimation	Std. dev.	Estimation	Std. dev.	Estimation	Std. dev.	Estimation	Std. dev.	Estimation	Std. dev.
k_i [kg/m^3]	-3.16	1.34E-2	6.66	2.40E-2	-1.85	3.44E-2	10.46	1.81	-2.21	1.64E-1
$\rho_i^{0,intercept}$ [kg/m^3]	403.12	8.52E-1	265.16	1.64	274.61	1.08	133.19	1.44E1	249.36	2.01
$\rho_i^{0,slope}$[kg/m^3/rpm]	-3.28	1.11E-1	-1.25	8.54E-2	-4.71	2.18E-1	-4.34	1.14E1	-10.05	3.50E-1
τ_i [s]	0.72	1.28E-1	2.10	5.09E-7	0.52	6.42E-2	0	1.10E-4	0.98	1.27E-1

The model performance (no state estimation) is assessed in Figure 2.8 for Dataset A (calibration) and Dataset B (validation). The plots in the figure refer to Ingredients 1 (API) and 2 of the formulation, making up for more than 80% of the tablet. Figure 2.8a, and its zoomed-in version Figure 2.8b, show that, in Dataset A, the experimental powder mass profile (i.e., the only measurement used in the objective function of Equation 2.17) is fitted well by the model for material 1. The model predicts API mass flow values ($\tilde{\dot{m}}_1$) that agree well with $\dot{m}_1^{avg}$ (Figure 2.8c), resulting in a composition signal that is much less noisy than $\dot{m}_1^{inst}$. Analogous considerations can be drawn for Ingredient 2 in Dataset A (Figure 2.8d), and Ingredients 1-2 in Dataset B (Figure 2.8e-f).

2.4.2 MHE design

Based on the calibrated feeder model, the MHE is implemented in the Pyomo environment, starting from the continuous form of the feeders model already coded for parameter estimation. A time step of 2 s is used (corresponding to the sampling time of the actual measurements), and a value of $H = 30$ is selected, since it is found to be a good compromise between computational burden and estimation accuracy. The MHE framework is defined by the optimization problem of Equations 2.8-2.10. The discrete-time model $\mathbf{f}(\hat{\mathbf{x}}(t_{l+1}|t_k), \mathbf{u}(t_k))$ of Equation 2.9a is obtained using the same approach adopted for parameter estimation, namely through discretization with Radau orthogonal collocation of Equations 2.8-2.11. Also in this case, 3 collocation points are used, but the number of finite elements is set to 30, corresponding to the selected value of *H*. The measurement model $\mathbf{g}(\hat{\mathbf{x}}(t_l|t_k))$ appearing in Equation 2.9c is defined as:

$$\mathbf{g}\big(\hat{\mathbf{x}}(t_l|t_k)\big):\ \hat{\mathbf{x}}(t_l|t_k) \ \mapsto\ [\hat{m}_1(t_l|t_k)\ \ \hat{m}_2(t_l|t_k)\ \ \ldots\ \ \hat{m}_5(t_l|t_k)] \tag{2.20}$$

At each t_k, IPOPT (Wächter and Biegler, 2006) is used for solving the MHE optimization problem, and the estimated hold-ups and mass flows at the current time step (respectively, $\hat{m}_i(t_k)$ and $\hat{\dot{m}}_i(t_k)$ for *i* = 1, …, 5) are stored. Then, sIPOPT (Pirnay et al., 2012) is resorted to for calculating the reduced Hessian of Problem 2.8-2.10 at the solution with respect to $\hat{\mathbf{x}}(t_{k-H+1}|t_k)$. Following the approach proposed by López-Negrete and Biegler (2012), from the reduced Hessian we calculate $\mathbf{P}(t_{k-H+1}|t_k)$, which is used at t_{k+1} for updating the arrival cost (Equation 2.10). At each time step, the estimated state vector includes 10 elements, while the input and measurement vectors have 5 elements each, namely:

$$\hat{\mathbf{x}} = [\hat{m}_1\ \ \hat{\dot{m}}_1\ \ \hat{m}_2\ \ \hat{\dot{m}}_2\ \ \ldots\ \ \hat{m}_5\ \ \hat{\dot{m}}_5\ ] \in \mathbb{R}^{10} \tag{2.21}$$

$$\mathbf{u} = [\mathrm{DC}_1\ \ \mathrm{DC}_2\ \ \ldots\ \ \mathrm{DC}_5] \in \mathbb{R}^5 \tag{2.22}$$

$$\mathbf{y} = [m_1\ \ m_2\ \ \ldots\ \ m_5] \in \mathbb{R}^5 \quad . \tag{2.23}$$

Implementation of the MHE requires designing the tuning matrices **R** (measurement noise variance matrix) and **Q** (model error variance matrix. Matrix **R** is designed as a diagonal matrix:

$$\mathbf{R}=\mathrm{diag}([\sigma^2_{m_1}\ \ \sigma^2_{m_2}\ \ \ldots\ \ \sigma^2_{m_5}])\ , \tag{2.24}$$

where $\sigma^2_{m_i}$ (variance of the mass measurement of feeder *i*) is assigned the value 1×10^{-6} for *i* = 1, …, 5. Matrix **Q** is tuned as a diagonal matrix, too:

$$\mathbf{Q} = \mathrm{diag}([\sigma^2_{\hat{m}_1}\ \ \sigma^2_{\hat{\dot{m}}_1}\ \ \sigma^2_{\hat{m}_2}\ \ \sigma^2_{\hat{\dot{m}}_2}\ \ \ldots\ \ \sigma^2_{\hat{m}_5}\ \ \sigma^2_{\hat{\dot{m}}_5}])\ , \tag{2.25}$$

in which, for each feeder *i*, $\sigma^2_{\hat{m}_i}$ and $\sigma^2_{\hat{\dot{m}}_1}$ are the variances of the element of $\mathbf{w}(t_l)$ corresponding

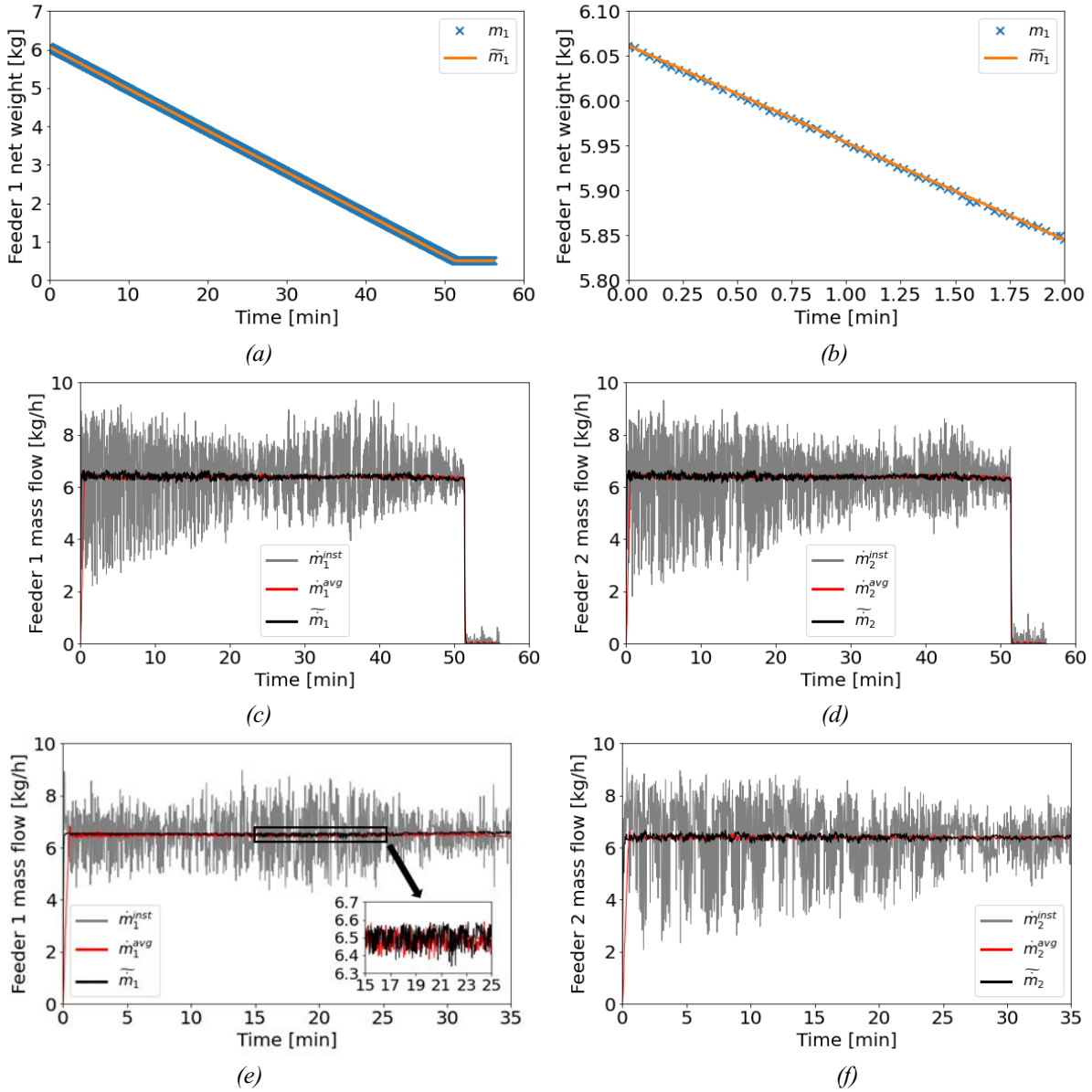


Figure 2.8. *Experimental measurements compared against the predictions of the calibrated model. (a) Dataset A, m_1 vs $\widetilde{m}_1$, (b) zoom-in of Figure 2.8a, (c) Dataset A, $\dot{m}_1^{inst}$ vs $\dot{m}_1^{avg}$ vs $\widetilde{m}_1$, (d) Dataset A, $\dot{m}_2^{inst}$ vs $\dot{m}_2^{avg}$ vs $\widetilde{m}_2$, (e) Dataset B, $\dot{m}_1^{inst}$ vs $\dot{m}_1^{avg}$ vs $\widetilde{m}_1$, (f) Dataset B, $\dot{m}_2^{inst}$ vs. $\dot{m}_2^{avg}$ vs $\widetilde{m}_2$. Figure 2.8e also reports a zoom-in of $\dot{m}_1^{avg}$ vs $\widetilde{m}_1$.*

to the discretized version of (respectively) Equation 2.10 and Equation 2.11. Note that Equation 2.10 has no modeling error, as it represents the material balance of powder in the feeder. Hence, the elements of $\mathbf{w}(t_l)$ corresponding to Equation 2.10 are set to 0. For the same reason, $\sigma^2_{\widetilde{m}_i}$ is set to 0, for i = 1, …, 5. On the other hand, Equation 2.11 is subject to error, and the variances $\sigma^2_{\widetilde{m}_i}$ are assigned by trial and error. Namely, we used Dataset A and adjusted the diagonal elements until the variability interval observed for $\hat{c}_{API}^{blended}$ was roughly comparable to the one observed experimentally for c_{API}^{NIR} (i.e., for the only concentration measurement available at sufficiently high rate from downstream). The tuning parameters that were found to have the strongest impact on the state estimation are $\sigma^2_{\widetilde{m}_1}$ and $\sigma^2_{\widetilde{m}_2}$, which is not surprising because

Ingredients 1 and 2 are the main components of the tablet formulation.
The $\hat{c}_{API}^{blended}$ values obtained with four different tunings of the MHE (Table 2.4) are illustrated in Figure 2.9, with increasing values of $\sigma_{\tilde{m}_i}^2$ from Figure 2.9a to Figure 2.9d. As expected, the lowest $\hat{c}_{API}^{blended}$ fluctuations are registered in Figure 2.9a, where the MHE response is practically equivalent to the model prediction. By increasing $\sigma_{\tilde{m}_i}^2$, the MHE is tuned to trust the model less and less, and to give increasingly more importance to the measurements (see Equation 2.8). When measurements are weighted very strongly (Figure 2.9d), the $\hat{c}_{API}^{blended}$ profile becomes qualitatively similar to the very noisy API concentration calculated from the raw instantaneous mass flows (Figure 2.1a). The parameters used for the estimation shown in Figure 2.9c are eventually selected, as they lead to a $\hat{c}_{API}^{blended}$ profile that more closely resembles the c_{API}^{NIR} one.

Table 2.4. *MHE tuning comparison (Dataset A). Values of the tuning parameters of* $\boldsymbol{Q}$ *(Equation 2.25) used for obtaining the estimations reported in Figure 2.9.*

	$\sigma_{\tilde{m}_1}^2$	$\sigma_{\tilde{m}_2}^2$	$\sigma_{\tilde{m}_3}^2$	$\sigma_{\tilde{m}_4}^2$	$\sigma_{\tilde{m}_5}^2$
Tuning Figure 2.9a	0	0	0	0	0
Tuning Figure 2.9b	7.5	7.5	1.5	7.5	15
Tuning Figure 2.9c	300	7.5	1.5	7.5	15
Tuning Figure 2.9d	3000	7.5	1.5	7.5	15

2.5. Real-time monitoring: proof of concept

We test the proposed monitoring system on Datasets A-C. Real-time monitoring is mimicked by using the data stored in the datasets as if they were coming from the sensors of the running plant. For Dataset A, the obtained profiles of the estimated mass flows $\widehat{m}_1$ and $\widehat{m}_2$ at the feeder outlet (Figures 2.10a and 2.10b, respectively) show much greater variability than the corresponding $\dot{m}_i^{avg}$, and slightly smaller than $\dot{m}_i^{inst}$. The same considerations apply to $\hat{c}_{API}^{pre-blending}$, which is compared to c_{API}^{inst} and $c_{API}^{theoretical}$ in Figure 2.10c.
The resulting $\hat{c}_{API}^{blended}$ profile (downstream the blending section) is validated against c_{API}^{HPLC}, c_{API}^{NIR} and c_{API}^{tandem} measurements in Figure 2.11a (in the industrial practice, c_{API}^{HPLC} is considered the most reliable API concentration measurement across the whole process line). We remark that none of these validation measurements are used by the monitoring system for estimating $\hat{c}_{API}^{blended}$. Visual inspection shows that $\hat{c}_{API}^{blended}$ is aligned well with the HPLC and NIR-derived measurements in the final tablet. Quantitively speaking, the sum of squared errors (SSE) between c_{API}^{HPLC} and $\hat{c}_{API}^{blended}$ is ~20% smaller than the SSE between c_{API}^{HPLC} and the blending outlet concentration calculated applying the blending dynamics (Equation 2.16) to $c_{API}^{theoretical}$ (Figure 2.11b). The SSE reduction is even greater (~40%) when c_{API}^{tandem} is considered, instead of c_{API}^{HPLC}. When $\hat{c}_{API}^{blended}$ is benchmarked against the blending outlet concentration calculated

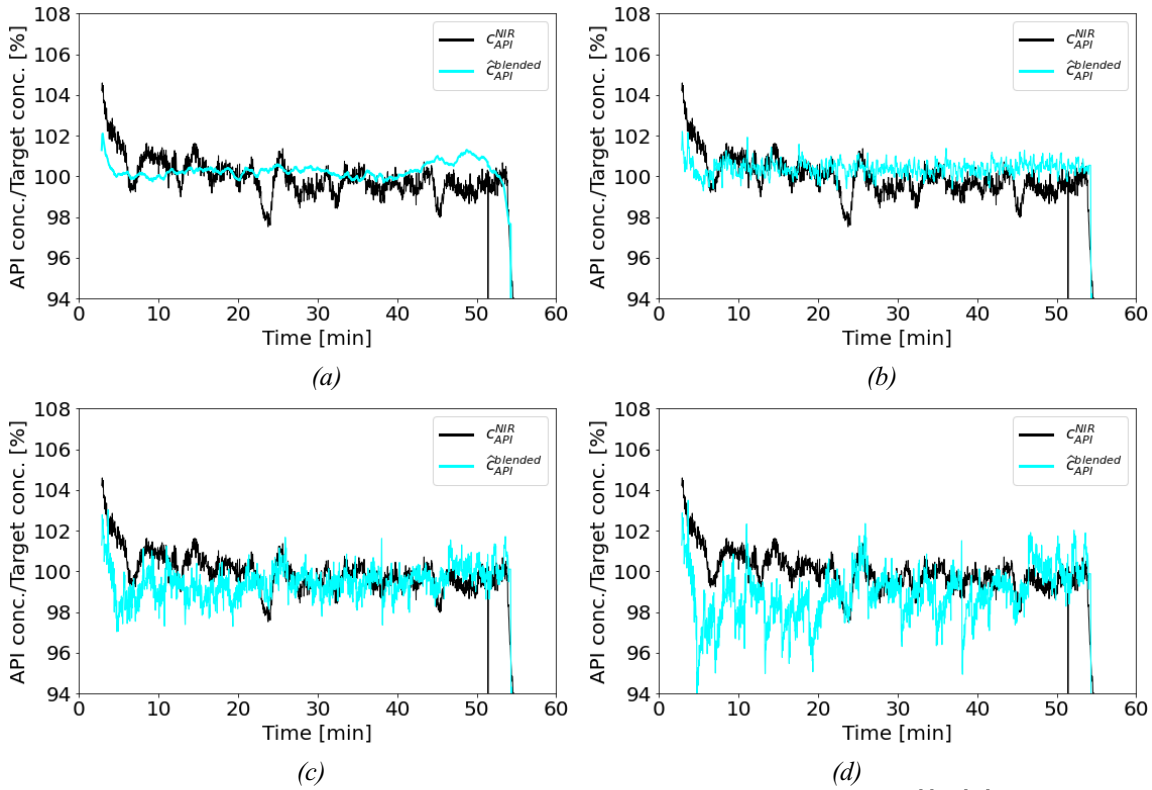


Figure 2.9. *MHE tuning comparison (Dataset A). Time profile of* $\hat{c}_{API}^{blended}$ *obtained with decreasing confidence in the model from (a) to (d). The values of the different tuning parameters are reported in Table 2.4. The time profile of* c_{API}^{NIR} *is reported for comparison.*

applying the blending dynamics to c_{API}^{inst} (Figure 2.11c), the SSE reduction with respect to c_{API}^{HPLC} is greater than 40%, while the one with respect to c_{API}^{tandem} is greater than 60%.

Although $\hat{c}_{API}^{blended}$ and c_{API}^{NIR} do not match perfectly in Figure 2.11a, it appears from the same figure that $\hat{c}_{API}^{blended}$ captures the underlying low-frequency dynamics that is apparent in the c_{API}^{NIR} signal. Instead, the blending outlet concentration calculated from $c_{API}^{theoretical}$ (Figure 2.11b), essentially, does not display any dynamics, due to the low signal-to-noise ratio of $c_{API}^{theoretical}$ (Figure 2.1). Application of fast Fourier transform (a mathematical technique capable of identifying the frequencies of the underlying dynamics of a signal) to $\hat{c}_{API}^{blended}$, c_{API}^{NIR} and $c_{API}^{theoretical}$ confirms these observations. Figure 2.12 shows the Fourier transform-derived power spectra (not to be confused with the spectra obtained through PAT in the plant) of the $\hat{c}_{API}^{blended}$, c_{API}^{NIR} and $c_{API}^{theoretical}$signals. Both $\hat{c}_{API}^{blended}$ and c_{API}^{NIR} have their main components laying in the low frequencies band of 0.001-0.005 Hz, corresponding to periods of about 3-15 min (coherently with the plots displayed in Figure 2.11a). Instead, $c_{API}^{theoretical}$ mostly presents higher-frequency components that are associated with noise, and, after the blending effect is applied to $c_{API}^{theoretical}$, almost no low-frequency dynamics is left (Figure 2.11b).

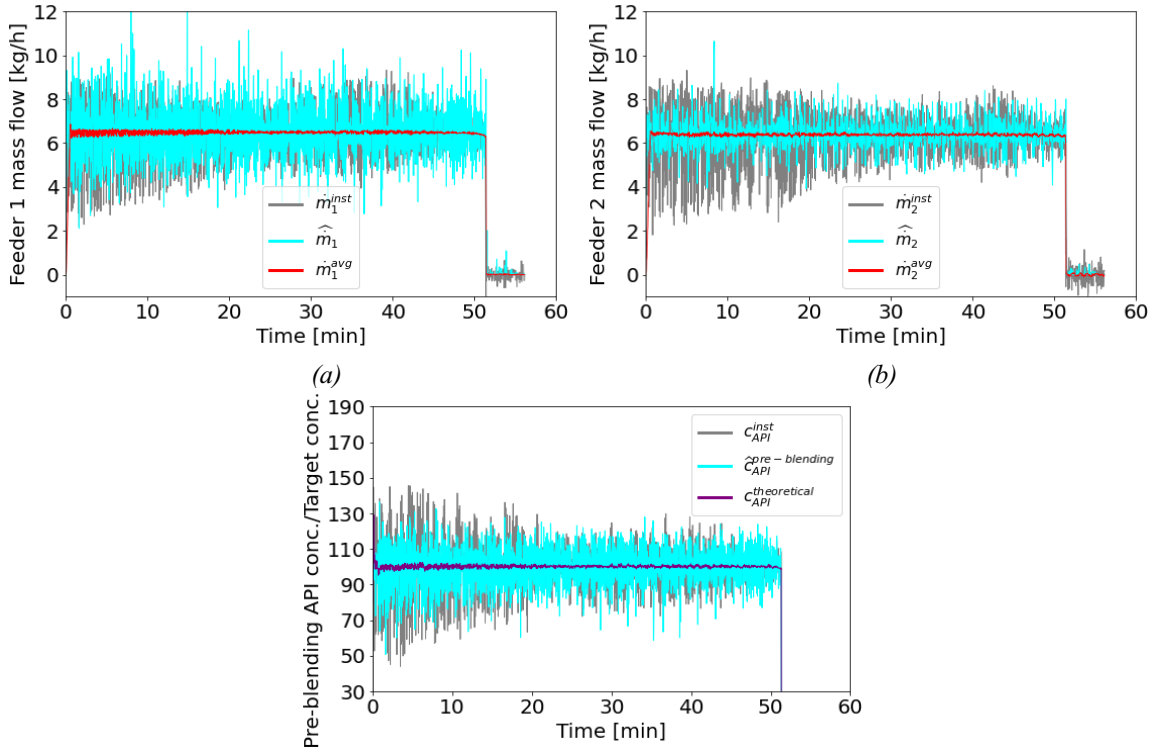


Figure 2.10. *Comparison of pre-blending estimations and measurements (Dataset A). (a) $\dot{m}_1^{inst}$ vs $\widehat{\dot{m}}_1$ vs $\dot{m}_1^{avg}$; (b) $\dot{m}_2^{inst}$ vs $\widehat{\dot{m}}_2$ vs $\dot{m}_2^{avg}$; (c) c_{API}^{inst} vs $\hat{c}_{API}^{pre-blending}$ vs $c_{API}^{theoretical}$.*

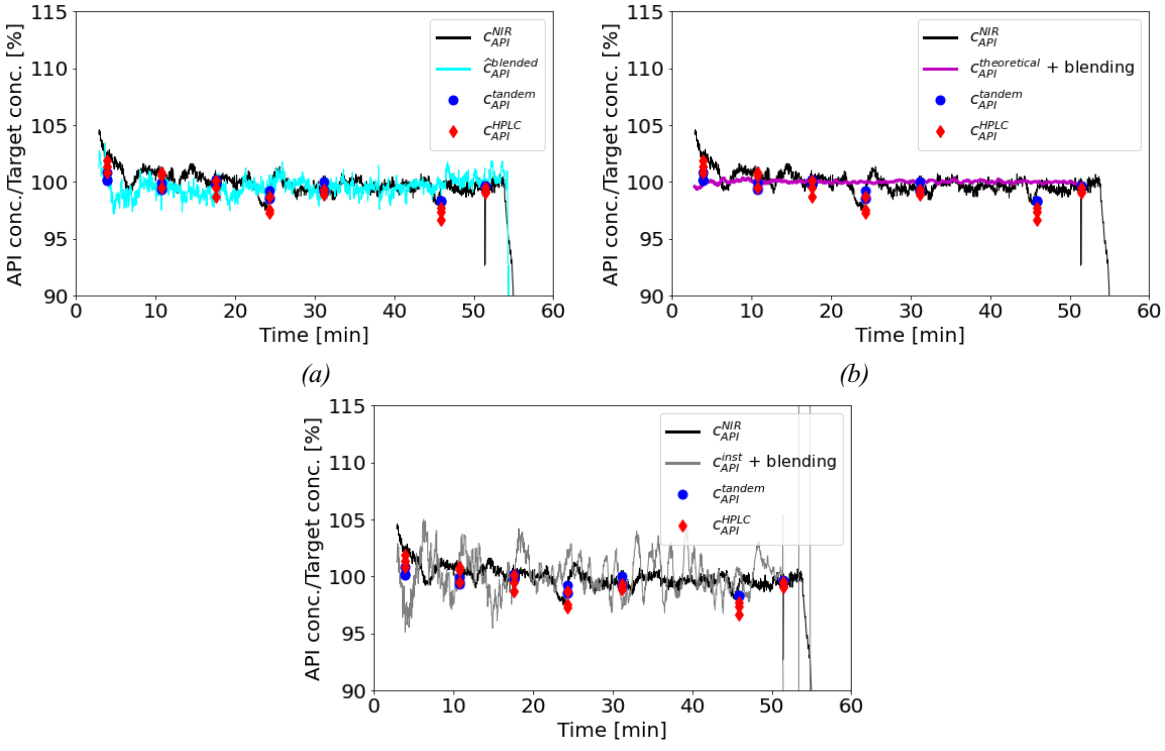


Figure 2.11. *Comparison of post-blending estimations and measurements (Dataset A). (a) $\hat{c}_{API}^{blended}$, (b) $c_{API}^{theoretical}$ + blending effect (Equation 2.16) and (c) c_{API}^{inst} + blending effect (Equation 2.16). We include c_{API}^{HPLC}, c_{API}^{NIR} and c_{API}^{tandem} for reference.*

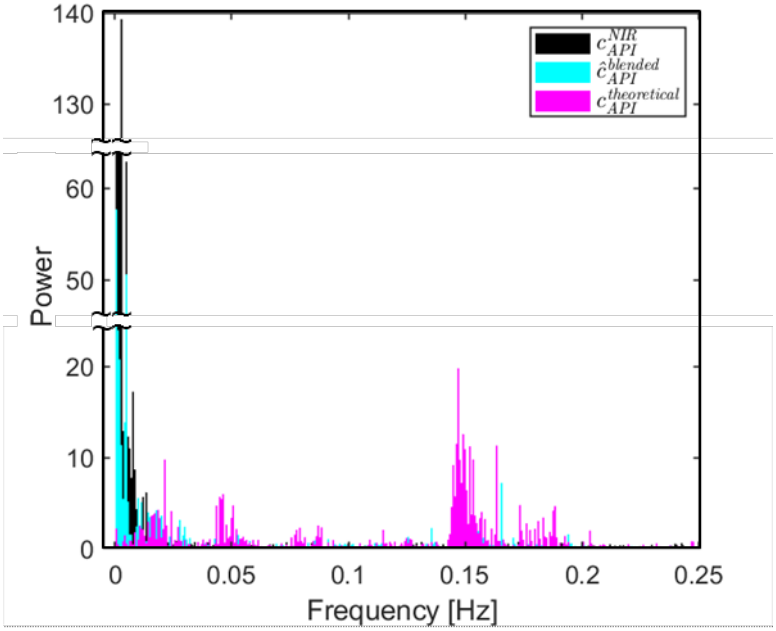


Figure 2.12. *Single-sided Fourier transform derived power spectrum of* c_{API}^{NIR}, $\hat{c}_{API}^{blended}$ *and* $c_{API}^{theoretical}$ *(Dataset A). Note that the sampling time of the datasets is 2 s, corresponding to a sampling frequency of 0.50 Hz.*

The partial misalignment between $\hat{c}_{API}^{blended}$ and c_{API}^{NIR} in Figure 2.11a might be due to unmodelled stochastic phenomena occurring in the blending section, or to the need of improving the feeder model (on the structural side and/or on the parametric one). In particular, the datasets used in this study come from a pilot-plant experimental campaign on a drug product that was not analyzed in further experiments. Hence, limited data were available for calibrating the chemometric model relating the NIR spectra to the API concentration, and complete trust in c_{API}^{NIR} cannot be argued.

The results demonstrate that the proposed monitoring framework provides better estimations of the API concentration across the process than done by a conventional approach based on monitoring the API concentration by $c_{API}^{theoretical}$. Notice that the variability of the API concentration upstream the blending section as estimated by the MHE is greater than the one represented by $c_{API}^{theoretical}$ (see Figure 2.10c). This is expected given that the latter is calculated with a moving average, which is a dampened signal by definition.

To further appreciate the benefits of including the state estimator in the monitoring system, consider again Figure 2.11c. It can be seen that the API concentration at the blending outlet as obtained by applying the blending dynamics of Equation 2.16 to c_{API}^{inst} still displays high-frequency fluctuations (e.g. from ~30 to ~45 min), not resembling any of the measured c_{API}^{HPLC}, c_{API}^{NIR} and c_{API}^{tandem} profiles. No practical improvement is achieved by either using a discrete three-point derivative for calculating $\dot{m}_i^{inst}$, or applying a low-pass filter to c_{API}^{inst}. This means that, despite $\hat{c}_{API}^{pre-blending}$ variability is similar to the one of c_{API}^{inst} (Figure 2.10c), the fluctuations observed in the MHE estimations reflect the underlying process dynamics, instead

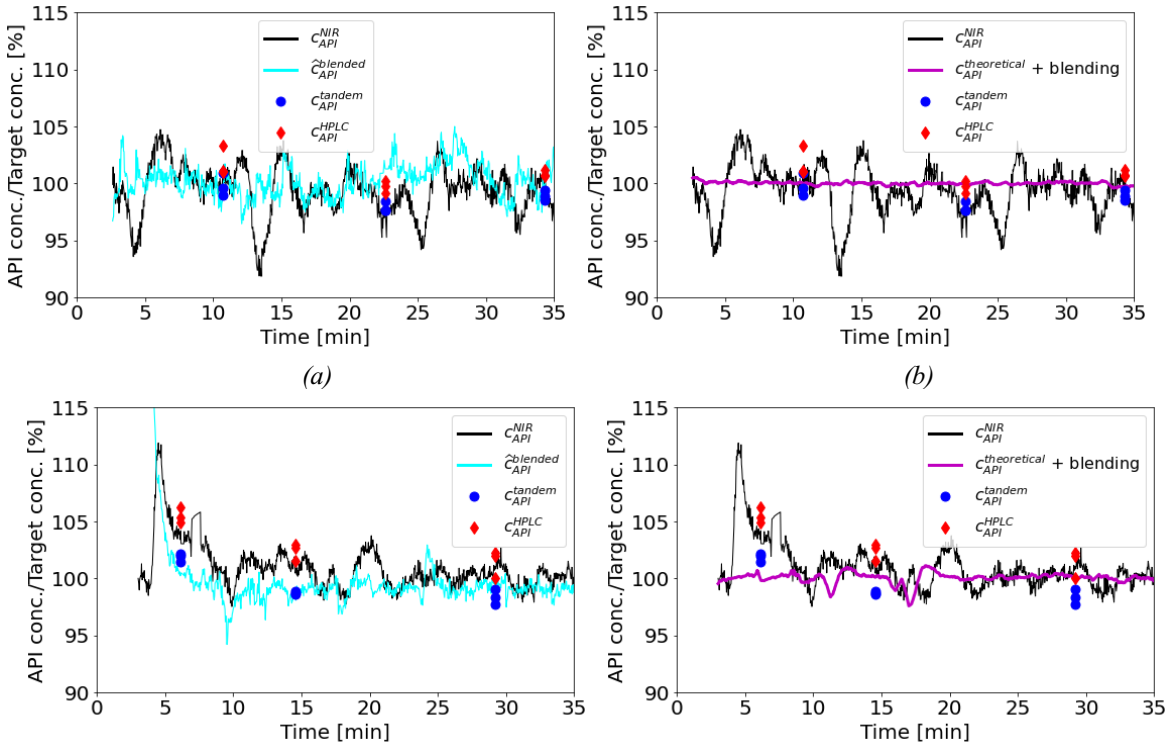


Figure 2.13. *Comparison of post-blending estimations and measurements (Datasets B and C). (a) $\hat{c}_{API}^{blended}$, Dataset B, (b) $c_{API}^{theoretical}$ + blending effect (Equation 2.16), Dataset B, (c) $\hat{c}_{API}^{blended}$, Dataset C, (d) $c_{API}^{theoretical}$ + blending effect (Equation 2.16), Dataset C. c_{API}^{HPLC}, c_{API}^{NIR} and c_{API}^{tandem} are reported in all plots for comparison.*

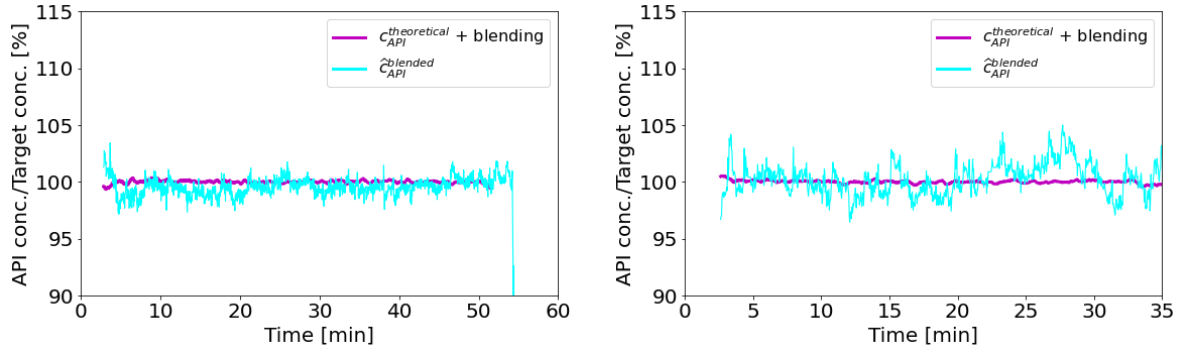


Figure 2.14. *$\hat{c}_{API}^{blended}$ compared against $c_{API}^{theoretical}$ + blending effect (Equation 2.16) for (a) Dataset A and (b) Dataset B.*

of noise as in c_{API}^{inst}; re-configuring the number of points of the moving-average for calculating $\dot{m}_i^{avg}$ would not suffice for achieving the good performance reached by the MHE.

We test the MHE-based monitoring system also for Datasets B and C of Table 2.1. The manufactured tablets have a different mass than those related to Dataset A, and this would suggest to consider retuning both the model and the MHE. Nevertheless, for simplicity we consider a worst-case scenario where the tuning is not adjusted. The results shown in Figure

2.13 are still satisfactory. In fact, $\hat{c}_{API}^{blended}$ is coherent with c_{API}^{HPLC} and c_{API}^{tandem}, while the blending outlet concentration calculated by applying the blending dynamics to $c_{API}^{theoretical}$ essentially does not display any dynamics.

By overlapping $\hat{c}_{API}^{blended}$ with $c_{API}^{theoretical}$ after blending (Figure 2.14), it is once again confirmed that $c_{API}^{theoretical}$ is an over-smoothed signal mainly including high-frequency noise components, while $\hat{c}_{API}^{blended}$ includes lower-frequency components representing the underlying dynamics of the process. The misalignment between $\hat{c}_{API}^{blended}$ and c_{API}^{NIR} in Figures 2.13a and 2.13c might be due to the aforementioned reasons, such as disturbances in the blending or to the need to improve the feeders model. However, for these datasets an even more important role in the mismatch might be played by the unreliability of c_{API}^{NIR}, possibly due to poor sensor readings. For instance, referring to Figure 2.13e, the occurrence of 10% variations of API concentrations in less than 2 min (e.g., from ~12 to ~14 min) with no likewise variations in the feeders mass measurements seems unlikely. Finally, since the tablets produced in the experiments of Datasets B and C have increasingly larger mass, another reason for the mismatch might be related to the need of tuning the MHE for experiments producing tablets of a fixed mass. These possibilities will be further inquired in future work.

As a concluding remark, the computational burden of the monitoring framework is compatible with real-time implementation. The state estimation results were obtained in approximately 50% of the process time with an Intel® Core™ i7-8565U CPU @1.80 GHz processor and total memory of 16.0 GB RAM. The computational time can be further reduced by using smaller values for *H* as in Liu et al. (2018) or with fast MHE approaches (Zavala et al., 2008).

2.6. Conclusions

We presented a novel approach to monitoring powder feeding in continuous solid-dosage forms manufacturing, and successfully validated it against experimental datasets collected on a direct compression line. The monitoring system is based on a state estimator (MHE), which effectively reconciles the mass measurements coming from loss-in-weight feeders with downstream measurements potentially coming from a PAT instrument, and estimates the delivered powder mass flows by means of a model-based optimization strategy. The monitoring system exploits a detailed mathematical model of the process for state estimation purposes, meeting the Quality-by-Design framework invitation to develop control strategies rooted on enhanced process understanding. The powder mass flows estimated with the proposed monitoring approach (and, in turn, the estimated API concentration in the final dosage forms) are highly consistent with downstream HPLC and spectroscopic measurements, when compared with traditional approaches to feeding monitoring. Additionally, powder concentration estimations are provided in a practically continuous way, which is a significant

advantage over sampled measurements. We showed that using statistical filters instead of the state estimator for calculating the powder mass flows from the mass measurements, as done by default by loss-in-weight feeder software, can lead to inaccuracies. Finally, the required computational time is compatible with real time implementation.

Future work on a different molecule will involve additional validation activities on extended datasets, for which a reliable chemometric model relating the NIR spectra to the API concentration in the powder mixture is available. The integration of the state estimator with latent-variable modeling for fault detection and diagnosis purposes is also envisioned, following a hybrid monitoring approach outlined in Chapter 3.

We believe state estimation technology is the appropriate way to obtain maximum synergy from the resources invested in the development of a deterministic model for a process, and the resources dedicated to the implementation of PAT solutions. State estimation puts the data from PAT solutions in the context of a the wholistic view of the process that is represented in a deterministic model. This is, in the opinion of the authors, the first step in the journey to adopt model-based closed loop control in the pharmaceutical industry.

Chapter 3

Hybrid data-driven/knowledge-driven process monitoring*

In this Chapter traditional standalone data-driven and knowledge-driven process monitoring approaches are bridged by proposing a novel hybrid framework that exploits the advantages of both simultaneously. Namely, a novel process monitoring approach is proposed, based on a data-driven model that includes two different data types: i) "actual" data coming from sensor measurements, and ii) "virtual" data coming from a state estimator, based on a first-principles model of the system under investigation. The proposed approach is tested on three simulated case studies: a continuous polycondensation process for the synthesis of poly-ethylene terephthalate, a fed-batch fermentation process for the manufacturing of penicillin, and a pharmaceutical segmented fluid bed dryer. The hybrid monitoring approach shows superior fault detection and diagnosis performances with respect to conventional monitoring techniques, even when the first-principles model is relatively simple and process/model mismatch exists. The hybrid monitoring system is particularly relevant within the Quality-by-Design context, as the available physical knowledge on the process is directly factored into the control strategy.

3.1 Introduction

Process monitoring is a key task in the process industry, as detecting a fault and assessing its cause before the production is compromised can save valuable assets. Several data-driven (DD) methodologies for fault detection and diagnosis have been proposed in the last decades (Jiang et al., 2019; Qin, 2012). Among them, latent-variables models (LVMs; Jackson, 1991) are a powerful class of DD multivariate approaches that proved very effective for fault detection and

*Destro, F., P. Facco, S. García-Muñoz, F. Bezzo, M. Barolo (2020). A hybrid framework for process monitoring: Enhancing data-driven methodologies with state and parameter estimation. *J. Process Cont.*, **92**, 333-351
Destro, F., A.J. Salmon, P. Facco, C.C. Pantelides, F. Bezzo, M. Barolo (2020). Monitoring a segmented fluid bed dryer by hybrid data-driven/knowledge-driven modeling. *IFAC-PapersOnLine*, **53**, 11638-11643.

diagnosis (Kresta et al., 1991), and gained increased relevance with the Industry 4.0 “big data” era. LVMs detect a fault when new measurements coming from the plant sensors are unknown to the correlation structure of the training data, which define the normal operating conditions (NOC) for the process. Faults are typically detected using multivariate control charts (Nomikos and Macgregor, 1995). Contribution plots (Miller et al., 1998) can then be exploited to pinpoint the measurements most related to the faulty conditions. However, when the number of available measurements is relatively small, detection of a fault through an LVM may be delayed, because the fault must propagate into the system until the few measured variables are affected. On the other side, diagnosing the root-cause of a fault may be challenging due to the smearing-out effect (Qin, 2012). In fact, since LVMs are not cause-effect models, it may be difficult to identify causality patterns between measurements under abnormal process conditions. This issue is particularly relevant if the variables embodying the root-cause of the fault are not measured, and therefore cannot be included in the LVM.

To overcome this limitation, monitoring methodologies exploiting first-principles knowledge about the process under investigation may be considered. Process monitoring methodologies based on knowledge-driven (KD) models have been thoroughly reviewed elsewhere (Gao et al., 2015; Venkatasubramanian et al., 2003a). The most popular KD approaches are based on parity relations (Gertler, 1998) or on state estimators (Blanke et al., 2006; Caccavale et al., 2009; Deshpande et al., 2009; Mohd et al., 2015; Rusinov et al., 2013), possibly implemented for simultaneous state and parameter estimation (Varshney et al., 2019). Generally speaking, KD models have the advantage of embedding the available understanding on the mechanisms driving the process under investigation. This piece of information can help fault detection and diagnosis, and is missing in DD monitoring approaches. However, KD models are generally more complex to develop than their DD counterparts and, when used for monitoring, the performances can be severely affected by process-model mismatch. In addition, the fault models have typically to be known a priori (Caccavale et al., 2009; Rusinov et al., 2013).

Hybrid models (von Stosch et al., 2013; Zendehboudi et al., 2018) combine DD methods with the information available from first-principles knowledge about the process, and are promising techniques for overcoming the limitations of DD and KD monitoring (He and Wang, 2018; Reis et al., 2019). Hybrid models for process monitoring usually consist of a KD soft-sensing framework in which a DD component is added to make up for missing deterministic information (Bonvin et al., 2016; Jia et al., 2011). Other hybrid approaches (Tidriri et al., 2016) use DD techniques to monitor the residuals of a KD model, or they develop complex schemes with subsequent DD and KD steps, tailored to specific applications (Ghosh et al., 2011). Recent contributions (Baklouti et al., 2019, 2018) showed the benefits of building advanced control charts to monitor the states estimated by a state estimator. However, these approaches are typically univariate, and therefore suffer from the well-known limitations of univariate monitoring with respect to its multivariate counterpart (Jackson, 1991; Kresta et al., 1991;

Seborg et al., 2017). Even though all these methodologies are gaining interest, state-of-the-art hybrid monitoring methods still lack a general framework to combine results coming from DD and KD modeling approaches (Tidriri et al., 2016).

In this study, we couple the easy-to-design features of DD process monitoring approaches to the descriptive capability of KD models, in order to develop a novel multivariate monitoring methodology with improved fault detection and diagnostic capabilities. Namely, we propose a latent-variable-based monitoring model that uses an augmented data matrix including two different data types: *i*) "actual" data coming from sensor measurements, and *ii*) "virtual" data coming from a dynamic KD model able to capture the main features of the system under investigation. Process-model mismatch for the KD model is (at least partially) compensated for by using a state estimator, which returns a set of virtual data consisting of estimated system states, adapted parameters, and reconstructed measurements to be included in the DD model. We test the proposed methodology on three simulated processes: a fed-batch fermentation process for the production of penicillin, a pharmaceutical segmented fluid-bed dryer, and a continuous poly-ethylene terephthalate (PET) polymerization process. The latter case study is here included to demonstrate how the proposed hybrid approach is effective also for non-pharmaceutical applications.

The remainder of this article is organized as follows. In Section 3.2 the mathematical methodologies later applied are briefly summarized. The proposed hybrid monitoring framework is outlined in Section 3.3. The case studies are presented in Section 3.4, and the results are discussed in Sections 3.5 and 3.6. Conclusions to the study are finally reported in Section 7.

3.2 Methods

3.2.1 Process monitoring by extended Kalman filtering

Let the first-principles model (FPM) of a dynamic system be expressed as a set of ordinary differential equations:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t) + \mathbf{w}(t) \tag{3.1}$$

where $\mathbf{f}$ is a nonlinear function, $\mathbf{x}(t)$ denotes the system state vector at time t, $\mathbf{u}(t)$ is the input vector, and $\mathbf{w}(t)$ is the process noise vector, which is assumed to be a white Gaussian process with mean $\mathbf{0}$ and variance $\mathbf{Q}(t)$. Assuming that measurements are available at discrete time steps t_k from the plant, they can be related to the system states through an appropriate measurement model:

$$\mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t), \mathbf{u}(t), t) + \mathbf{v}(t) \quad (3.2)$$

where $\mathbf{y}(t)$ is the measurement vector, and $\mathbf{v}(t)$ is the measurement noise vector, which is assumed to be a white Gaussian process with mean $\mathbf{0}$ and variance $\mathbf{R}(t)$.

The discrete time data extended Kalman filter (EKF; Ray, 1981) provides the estimated state vector $\hat{\mathbf{x}}(t)$ and the state covariance $\mathbf{P}(t)$, given the initial estimation of the states $\hat{\mathbf{x}}_0$ and the initial state covariance $\mathbf{P}_0$. The algorithm includes two steps, prediction and update, which are alternatively performed at each time point $k = 1, 2, \ldots, K$. The predictions of the states and of the state covariance at time t_k (t_k^-), obtained *before* the measurements are available, are respectively referred to as $\hat{\mathbf{x}}(t_k^-)$ and $\mathbf{P}(t_k^-)$, whereas the corrected estimations *after* the sampling time (t_k^+) are denoted as $\hat{\mathbf{x}}(t_k^+)$ and $\mathbf{P}(t_k^+)$.

During the prediction step, $\hat{\mathbf{x}}(t_{k-1}^+)$ and $\mathbf{P}(t_{k-1}^+)$ are propagated, with the integration of Equations 3.3-3.4, to obtain (respectively) $\hat{\mathbf{x}}(t_k^-)$ and $\mathbf{P}(t_k^-)$:

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{f}(\hat{\mathbf{x}}(t), \mathbf{u}(t), t) \quad (3.3)$$

$$\dot{\mathbf{P}}(t) = \mathbf{F}\mathbf{P} + \mathbf{P}\mathbf{F}^{\mathrm{T}} + \mathbf{Q}(t) \quad (3.4)$$

where $\mathbf{F}$ is the Jacobian matrix:

$$\mathbf{F} = \left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right)_{\hat{\mathbf{x}}(t), u(t), t} \quad (3.5)$$

At each sampling point k, the predicted estimations are corrected with the update equations:

$$\hat{\mathbf{x}}(t_k^+) = \hat{\mathbf{x}}(t_k^-) + \mathbf{K}(t_k)\, \boldsymbol{\gamma}(t_k) \quad (3.6)$$

$$\mathbf{P}(t_k^+) = \mathbf{P}(t_k^-) - \mathbf{K}(t_k)\mathbf{H}_k\mathbf{P}(t_k^-) \quad (3.7)$$

where the Kalman gain $\mathbf{K}(t_k)$, the innovation $\boldsymbol{\gamma}(t_k)$, and the Jacobian matrix $\mathbf{H}_k$ are respectively calculated with:

$$\mathbf{K}(t_k) = \mathbf{P}(t_k^-)\, \mathbf{H}_k^{\mathrm{T}} \mathbf{V}(t_k)^{-1} \quad (3.8)$$

$$\boldsymbol{\gamma}(t_k) = \mathbf{y}(t_k) - \mathbf{h}(\hat{\mathbf{x}}(t_k^-), \mathbf{u}(t_k), t_k) \quad (3.9)$$

$$\mathbf{H}_k = \left(\frac{\partial \mathbf{h}}{\partial \mathbf{x}}\right)_{\hat{\mathbf{x}}(t_k^-), \mathbf{u}(t_k), t_k} \quad (3.10)$$

The matrix $\mathbf{V}(t_k)$ in Equation 3.8 is calculated with:

$$\mathbf{V}(t_k) = \mathbf{H}_k\, \mathbf{P}(t_k^-)\, \mathbf{H}_k^{\mathrm{T}} + \mathbf{R}(t_k) \quad (3.11)$$

In Equation 3.8, for the inversion of matrix $\mathbf{V}(t_k)$ factorization is resorted to instead of direct inversion, for computational efficiency and robustness. Rigorous methods to design $\mathbf{P}_0$, $\mathbf{Q}$ and $\mathbf{R}$ exist (Schneider and Georgakis, 2013). In this study, we take a simpler approach that proved

effective also in several other studies (Delgado-Aguiñaga et al., 2016; Pérez et al., 2015; Ricker and Lee, 1995). Namely, we design $\mathbf{P}_0$ as a diagonal matrix, based on the expected uncertainty on the initial estimation for each state. Additionally, we set $\mathbf{Q}$ as a time-invariant diagonal matrix, and we tune it by trial and error in such a way as to obtain robust convergence. Finally, we design $\mathbf{R}$ as:

$$\mathbf{R} = \mathrm{diag}(\sigma_i) \ , \tag{3.12}$$

where σ_i is the variance of the noise of the *i*-th measurement sensor.

The EKF provides reconstructed values for $\mathbf{y}(t_k)$, tackling the measurement noise. Such reconstructed values are useful for better monitoring the process evolution. Specific implementations of the EKF can be set up for tackling colored noise or measurement bias (Simon, 2006), if that emerges useful for monitoring.

The EKF can also perform online parameter adaptation, upon augmentation of the state vector $\hat{\mathbf{x}}$ with the subset $\hat{\mathbf{p}}$ of the FPM parameters that one seeks to adapt in real time. The nominal values $\mathbf{p}_0$ of the parameters are taken as initial conditions for $\hat{\mathbf{p}}$, and negligible dynamics with additive process noise of small variance can be assumed. State augmentation can be resorted to for improving the estimation of unmeasured states through bias estimation (Liotta et al., 1997), too. Bias estimation consists in the insertion of an additive term (bias) at the right-hand side of Equation 3.3 for selected states. Biases are assigned null initial values and a random walk model of small variance. To meet the observability conditions, the augmented states cannot exceed in number the available measurements.

Conventional KD fault detection approaches (Blanke et al., 2006; Deshpande et al., 2009) are based on the assumption that the innovation sequence $\boldsymbol{\gamma}(t)$ of the EKF without augmented states follows a white $\mathbf{0}$ mean Gaussian distribution with variance $\mathbf{V}(t)$. Deviations from this behavior indicate a fault and can be detected by monitoring for each time instant the test statistics in Equation 3.13, which follows the central χ^2 distribution with *R* degrees of freedom (where *R* is the number of measurements in $\mathbf{y}$):

$$\epsilon(t_k) = \boldsymbol{\gamma}(t_k)^{\mathrm{T}}\mathbf{V}(t_k)^{-1}\boldsymbol{\gamma}(t_k) \quad . \tag{3.13}$$

Upon rejection of the null hypothesis for $\epsilon(t_k)$ at time t_k, the innovation sequence over a time window of size *S* (tuning parameter) is used for confirming the fault condition, with the following test statistic:

$$\epsilon(t_k, S) = \sum_{i=k}^{k+S} \boldsymbol{\gamma}(t_i)^{\mathrm{T}}\mathbf{V}(t_i)^{-1}\boldsymbol{\gamma}(t_i) \ . \tag{3.14}$$

If $\epsilon(t_k, S)$ violates the confidence limit for the central χ^2 distribution with $R{\cdot}S$ degrees of freedom, the fault is confirmed. Fault isolation is then typically carried out by online adaptation of all the parameters related to possible fault conditions, and by identifying which parameter is

drifting from its nominal value (Che Mid and Dua, 2017; Ku et al., 1994).

3.2.2 Multivariate process monitoring by principal component analysis

Principal component analysis (PCA) is a dimensionality reduction technique aimed at extracting the few underlying factors (called principal components, PCs) that explain most of the variability from a NOC dataset $\mathbf{Z}$ [N×M] including N observations on M variables (Jackson, 1991). The data matrix $\mathbf{Z}$ is auto-scaled to zero mean and unit variance, and the PCs are extracted by decomposing $\mathbf{Z}$ as (Wise and Gallagher, 1996):

$$\mathbf{Z} = \sum_{a=1}^{A} \mathbf{t}_a \mathbf{p}_a^{\mathrm{T}} + \mathbf{E}\,, \tag{3.15}$$

where $\mathbf{t}_a$ are [N×1] score vector of the a-th PC, $\mathbf{p}_a$ [M×1] is the loading vector for the same PC, and $\mathbf{E}$ is the matrix of residuals, which are random noise if an appropriate number A of PCs is selected. For each observation n, two monitoring statistics can be calculated from the PCA model, namely the Hotelling T^2 and the squared prediction error (SPE), according to:

$$T_n^2 = \sum_{a=1}^{A} t_{a,n}\, \lambda_a^{-1}\, t_{a,n} \tag{3.16}$$

$$\mathrm{SPE}_n = \boldsymbol{e}_n \boldsymbol{e}_n^{\mathrm{T}} \tag{3.17}$$

where $t_{a,n}$ is the element in $\mathbf{t}_a$ corresponding to observation n, λ_a is the eigenvalue associated to the a-th PC, and $\boldsymbol{e}_n$ is the residual vector. Confidence limits can be obtained for both statistics from the available set of NOC. In this study, we obtained the confidence limits $\mathrm{SPE}_{\mathrm{lim}}$ (on SPE) and T_{lim}^2 (on T^2) by means (respectively) of the Jackson-Mudholkar equation (Wise and Gallagher, 1996), and of the F-distribution confidence limit equation (Wise and Gallagher, 1996). We calculated both limits at 99% confidence.

When a new observation $\mathbf{z}_{\mathrm{new}}$ [1×M] becomes available from the process at time t_{new}, it is first normalized on the mean and variance of the NOC dataset, then it is projected onto the model space, and finally $\mathrm{SPE}_{\mathrm{new}}$ and T_{new}^2 are calculated. If at least one of the two statistics exceeds its relevant confidence limit for some (e.g., three) consecutive observations, a fault is alarmed. The fault can be diagnosed using contribution plots (Miller et al., 1998), which point to the variables included in $\mathbf{Z}$ that most contribute to the confidence limit violation. The [1×M] contribution vectors for $\mathbf{z}_{\mathrm{new}}$ can be built calculating the contributions for each variable m as:

$$\mathrm{c}_{\mathrm{new},m}^{\mathrm{SPE}} = e_{\mathrm{new},m}\,, \tag{3.18}$$

$$\mathrm{c}_{\mathrm{new},m}^{T^2} = \sum_{a=1}^{A} t_{a,new}\, \lambda_a^{-0.5}\, p_{a,m} \quad , \tag{3.19}$$

where $e_{\mathrm{new},m}$ and $p_{a,m}$ are the elements corresponding to variable m respectively in $\boldsymbol{e}_{\mathrm{new}}$ and $\mathbf{p}_a$. We use the residuals as contributions to SPE in order to preserve the sign of the error, which proves useful for fault diagnosis. The vector of residuals $\mathbf{c}_{\mathrm{new}}^{\mathrm{SPE}}$ is known to follow a normal

distribution if A is selected appropriately (Wise and Gallagher, 1996). For the case studies under investigation, we found that also the $\mathbf{c}_{\mathrm{NEW}}^{T^2}$ contributions are normally distributed, an occurrence that has been noted also by other investigators (Ündey et al., 2003; Westerhuis et al., 2000). Therefore, we calculated Gaussian control limits (at 99% confidence) for both contributions. Although these limits should not be considered to have statistical significance, they are nevertheless helpful for comparing contributions presenting significantly different magnitudes also with respect to the NOC.

When the observations in $\mathbf{Z}$ are auto-correlated, the standard PCA approach should be modified to account for the effect of time. For example, dynamic PCA (DPCA; Ku et al., 1995) can be used, where L lagged measurements are included in $\mathbf{Z}$ to obtain the NOC matrix $\mathbf{Z}_{dyn}$:

$$\mathbf{Z}_{dyn} = \begin{bmatrix} \mathbf{z}^{\mathrm{T}}(t_1) & \mathbf{z}^{\mathrm{T}}(t_0) & \dots & \mathbf{z}^{\mathrm{T}}(t_1 - L) \\ \mathbf{z}^{\mathrm{T}}(t_2) & \mathbf{z}^{\mathrm{T}}(t_1) & \dots & \mathbf{z}^{\mathrm{T}}(t_2 - L) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{z}^{\mathrm{T}}(t_M) & \mathbf{z}^{\mathrm{T}}(t_M - 1) & \dots & \mathbf{z}^{\mathrm{T}}(t_M - L) \end{bmatrix} \quad . \tag{3.20}$$

Alternatively, a multi-model moving-window PCA (MW-PCA, (Camacho et al., 2008)) approach can be resorted to, with a PCA model calibrated at each time point t_k on the data for the previous W time points, where W is the time window width. The NOC matrix $\mathbf{Z}_{MW}(t_{\mathrm{k}})$ for the PCA model at time $t_k$is defined as:

$$\mathbf{Z}_{MW}(t_{\mathrm{k}}) = [\mathbf{z}^{\mathrm{T}}(t_k) \quad \mathbf{z}^{\mathrm{T}}(t_k - 1) \quad \dots \quad \mathbf{z}^{\mathrm{T}}(t_k - W)] \quad . \tag{3.21}$$

3.3. Proposed hybrid monitoring framework

The proposed hybrid monitoring framework is based on the approach sketched in Figure 3.1. A process with measured inputs **u** and unknown disturbances **d** produces a set of measured outputs **y**. The KD block exploits an FPM and the online measurements **u** and **y** to perform state estimation, online adaptation of the FPM parameters, and measurement reconstruction. In this study, an EKF (Ray, 1981) has been used to carry out these tasks, though other estimators might be used (Mohd et al., 2015). The DD block is based on an LVM that extracts operation-relevant information from the available set of field measurements (**u** and **y**) as well as from the virtual data (estimated states $\hat{\mathbf{x}}$, adapted model parameters $\hat{\mathbf{p}}$, and reconstructed measurements $\hat{\mathbf{y}}$) returned by the KD block. The augmented overall data matrix becomes:

$$\mathbf{Z} = [\hat{\mathbf{x}} \quad \hat{\mathbf{p}} \quad \hat{\mathbf{y}} \quad \mathbf{u} \quad \mathbf{y}] \tag{3.22}$$

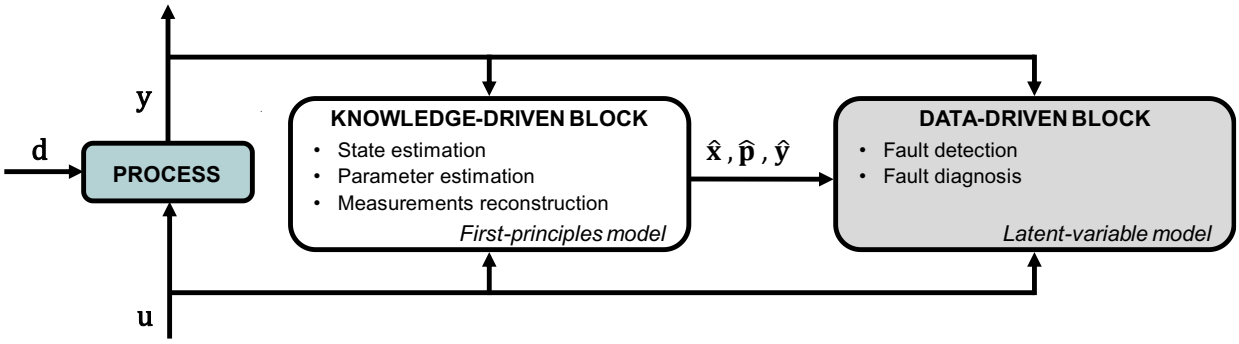


Figure 3.1. *Proposed hybrid monitoring framework.*

The advantage of this architecture is that, through the state estimator, the DD multivariate monitoring model can receive information related also to the inner mechanisms driving the system (in the form of states and FPM parameters), which can be very useful to monitor the process, but would otherwise not be accessible in the absence of a KD block. Therefore, the proposed hybrid approach is multivariate in nature, a feature that is known to offer significant advantages over single-variable methods (Jackson, 1991; Kresta et al., 1991; Seborg et al., 2017). The main features of the hybrid monitoring framework are the following:

- the estimated states (and possibly adapted parameters) provide meaningful indications about the phenomena involved in faults, which can facilitate fault detection and diagnosis. Large contributions from an estimated state or an adapted parameter indicate that the fault might be related to a physical phenomenon linked to that state or parameter. For this reason, the parameters to be selected for online adaptation (if any) are those mostly related to specific faults to be monitored. Although parameter estimation for fault detection and diagnosis has already been discussed in the literature (Che Mid and Dua, 2017; Ku et al., 1994), the distinctive advantage of the proposed framework is that the overall *co-variation* of states, parameters, and measurements is assessed by the hybrid framework, which can improve the monitoring performance;
- a subset of the measured inputs and outputs (y and u) may be not modeled by the FPM. Yet, due to their inclusion in the Z matrix, the LVM can exploit the deterministic information they embed, by assessing how they correlate not only with the other measurements, but also with the estimated variables;
- measurements are considered twice by the LVM: once in terms of y and once in terms of $\hat{\mathbf{y}}$. Hence, the filter innovations $\boldsymbol{\gamma}(t)$, which are sometimes monitored for fault detection in KD monitoring approaches (Blanke et al., 2006; Deshpande et al., 2009; Venkatasubramanian et al., 2003a), are (indirectly) fed to the monitoring system and are analyzed in a multivariate fashion.

3.4. Case studies

In this section, the three simulated case studies used to test the proposed methodology are presented. For each case study, two models are employed, namely:

- a detailed model is used to represent the true plant behavior; this model will be referred to as "the process";
- a simplified model is used to design the state estimator; this model will be referred to as "the FPM" or, more simply, "the model".

The main features of the case studies are summarized in Table 3.1. Case studies 1 and 2 involve a process of average complexity and size (<10 differential equations), while for Case study 3 a detailed process (~200 differential equations) validated on an industrial pharmaceutical segmented fluid bed dryer is resorted to. Parametric and structural process-model mismatch

Table 3.1. *Comparison of the main features of the two case studies investigated.*

Feature	Case study 1	Case study 2	Case study 3
process name	PET manufacturing	penicillin manufacturing	Pharmaceutical segmented fluid bed dryer
process type	continuous	fed-batch	continuous
number of differential states of process	4	9	223
process-model mismatch	parametric and (mild) structural	parametric and structural	parametric
unmodeled measurements	pressures	temperature	none
knowledge-driven block	state estimation; measurement reconstruction	state estimation; measurement reconstruction; parameter adaptation	state estimation; measurement reconstruction
data-driven block	dynamic PCA	moving-window PCA	moving-window PCA

exist in Case studies 1 and 2, although in Case study 2 the structural mismatch is more significant. Case study 3, instead, involves only parametric mismatch. In Case studies 1 and 3, only state estimation and measurement reconstruction are carried out by the EKF in the KD block, whereas in Case study 2 the EKF is also exploited for online parameter adaptation. DPCA is used in the DD block for Case studies 1 and 3, whereas MW-PCA is used in Case study 2.

3.4.1 Case study 1: PET manufacturing

PET synthesis occurs through three main steps: transesterification/esterification, pre-polymerization, and polycondensation. In this study, we refer to the polycondensation step. The

process is constituted by a series of three CSTRs (Figure 3.2) described by the following set of equations (Ling and Kravaris, 2016):

$$\frac{dc_{\mathrm{EG},i}}{dt} = \frac{1}{\tau}\left(c_{\mathrm{EG},i-1} - c_{\mathrm{EG},i}\right) - h_i\left(c_{\mathrm{EG},i} - c^*_{\mathrm{EG},i}\right) + 0.5\, R_{\mathrm{pol},i} \tag{3.23}$$

$$\frac{dc_{\mathrm{OH},i}}{dt} = \frac{1}{\tau}\left(c_{\mathrm{OH},i-1} - c_{\mathrm{OH},i}\right) - R_{\mathrm{pol},i} \tag{3.24}$$

$$\frac{dc_{\mathrm{COOH},i}}{dt} = \frac{1}{\tau}\left(c_{\mathrm{COOH},i-1} - c_{\mathrm{COOH},i}\right) + R_{\mathrm{degr},i} \tag{3.25}$$

$$\frac{dc_{\mathrm{ESTER},i}}{dt} = \frac{1}{\tau}\left(c_{\mathrm{ESTER},i-1} - c_{\mathrm{ESTER},i}\right) + 0.5\, R_{\mathrm{pol},i} - R_{\mathrm{degr},i} \tag{3.26}$$

$$R_{\mathrm{pol},i} = k_{\mathrm{pol},i}(c^2_{\mathrm{OH},i} - 8\, c_{\mathrm{EG},i} c_{\mathrm{ESTER},i}) \tag{3.27}$$

$$R_{\mathrm{degr},i} = k_{\mathrm{degr},i} c_{\mathrm{ESTER},i} \tag{3.28}$$

$$c^*_{\mathrm{EG},i} = \frac{P_i}{P^{sat}_{\mathrm{EG},i}(T)\, \hat{v}_{\mathrm{EG}}\, \exp(1+\chi)} \tag{3.29}$$

where i identifies a given reactor (i = 1, 2, 3).

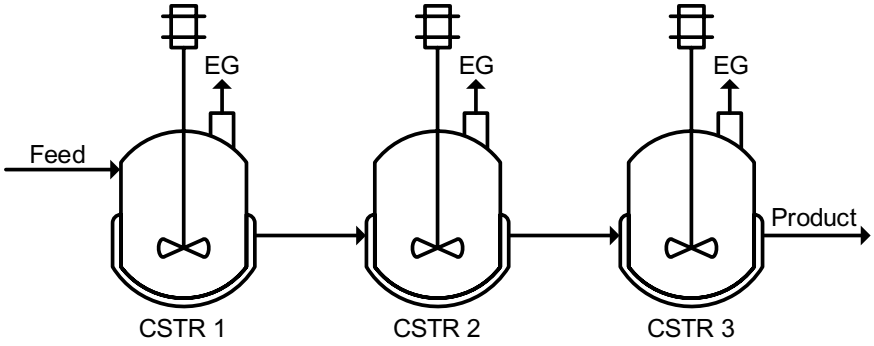


Figure 3.2. *Case study 1: scheme of the PET polycondensation process with three CSTRs in series (adapted from* (Ling and Kravaris, 2016)*).*

In the above model equations, EG, OH, COOH and ESTER respectively denote ethylene glycol, hydroxyl end-groups, acid end-groups, and ester groups; $c_{j,i}$ is the concentration of species j in reactor i ($c_{j,0}$ being the inlet concentration of species j to the first reactor); τ is the residence time in a reactor, and is the same for all reactors; $R_{\mathrm{pol},i}$ and $R_{\mathrm{degr},i}$ respectively refer to the rates for the polycondensation and degradation in reactor i. In the Flory-Huggins equation (3.29), $c^*_{\mathrm{EG},i}$ is the equilibrium concentration of ethylene glycol (the only species existing in the vapor phase) in reactor i, P_i is the pressure in reactor i, $P^{sat}_{\mathrm{EG},i}$ is the EG vapor pressure, $\hat{v}_{\mathrm{EG}}$ is the EG molar volume in the liquid phase, and χ is the polymer-solvent interaction parameter. The values of the parameters of the Flory-Huggins equation are set as in (Ling and Kravaris, 2016). The process is assumed to be isothermal, so the energy balance is neglected. The meaning of

the rest of the symbols is reported in Table 3.2, together with the values of all parameters and feed conditions. Their nominal values are taken from (Ling and Kravaris, 2016), but fluctuations are added as smoothed pseudo-random binary signals in order to more closely mimic a real situation where process noise increases normal process variability.

The key performance indicator is the degree of polymerization in the third reactor (DP_3), which cannot be measured online and is calculated as:

$$DP_3 = 1 + \frac{2\, c_{\mathrm{ESTER},3}}{c_{\mathrm{OH},3} + c_{\mathrm{COOH},3} + c_{\mathrm{Eg},3}} \tag{3.30}$$

where $c_{\mathrm{Eg},3}$ is the concentration of a byproduct, which is assumed to be equal to $c_{\mathrm{COOH},3}$ (Ling and Kravaris, 2016).

The 9 measurements available online from the process are:

- $y_{c_{\mathrm{OH},i}}$[mol L^{-1}], concentration of hydroxyl end-groups in reactor *i*;
- $y_{c_{\mathrm{COOH},i}}$[mol L^{-1}], concentration of acid end-groups in reactor *i*;
- y_{P_i} [Pa], pressure in reactor *i*.

White noise with standard deviation of typical industrial sensors is added to the detailed model outputs; namely, the standard deviations are 7E-3 mol L^{-1} for $y_{c_{\mathrm{OH},i}}$, 3E-4 mol L^{-1} for $y_{c_{\mathrm{COOH},i}}$, and 1 Pa for y_{P_i} (notice that the types of sensors for OH and COOH concentration measurements are different (Ling and Kravaris, 2016)). The measurement intervals for $y_{c_{\mathrm{OH},i}}$ and $y_{c_{\mathrm{COOH},i}}$ are set to 1 min and 10 min, respectively (Ling and Kravaris, 2016). Measurements for pressure are recorded every 10 min, because they are not needed at greater frequency for process monitoring.

The NOC dataset includes data from 5300 min of steady-state operation. We generate faulty datasets by running the process as under NOC, but applying the fault after 300 min from the

Table 3.2. *Case study 1: parameters and feed conditions in the detailed and simplified models. In the detailed model, fluctuations are added to the nominal values as smoothed-pseudo random binary signals with the indicated maximum/minimum amplitudes. In the simplified model, constant (nominal) values are used.*

Parameter or feed condition	Symbol	Units	Nominal value	Max/min amplitude
Residence time (all reactors)	τ	min	60	0
Mass transfer coefficient reactor 1	h_1	min^{-1}	2.70	0.80
Mass transfer coefficient reactor 2	h_2	min^{-1}	2.03	0.61
Mass transfer coefficient reactor 3	h_3	min^{-1}	1.35	0.41
Kinetic constant, reaction 1	$k_{\mathrm{pol},i}$	L mol^{-1} min^{-1}	6.66E-02	0.017
Kinetic constant, reaction 2	$k_{\mathrm{degr},i}$	min^{-1}	8.34E-06	1.25E-06
Pressure in reactor *i*	P_i	Pa	130	6.5
Feed concentration of ethylene glycol	$c_{\mathrm{EG},0}$	mol L^{-1}	6.50E-03	9.75E-04
Feed concentration of OH end-groups	$c_{\mathrm{OH},0}$	mol L^{-1}	0.40	0.080
Feed concentration of COOH end-groups	$c_{\mathrm{COOH},0}$	mol L^{-1}	2.57E-03	2.57E-04
Feed concentration of ester groups	$c_{\mathrm{ESTER},0}$	mol L^{-1}	11.20	0.022

start of the NOC sequence. Simulations are run for 1300 min. We consider four faulty sequences, all of which eventually result in an impactful decrease of DP_3:

- Fault #1: slowly decreasing ester concentration in the feed ($c_{\mathrm{ESTER},0}$ decreases by 0.01% per min). As a consequence, the concentration of ester in the first reactor starts decreasing, eventually reducing DP_3 (Equation 3.30).
- Fault #2: minor fault in the agitation system of the second reactor (h_2 decreases by 0.05% per min). The smaller mixer speed hinders the mass transfer of ethylene glycol from the liquid to the vapor phase. Ethylene glycol starts accumulating in the second reactor, thus reducing the rate of polycondensation (Equation 3.27), and eventually affecting DP_3 (Equation 3.30).
- Fault #3: significant fault in the agitation system of the second reactor (h_2 decreases by 0.1% per min). The consequences are as for Fault #2, but with greater magnitude.
- Fault #4: increasing pressure in the second reactor (P_2 increases by 3% per min). The pressure increase inhibits ethylene glycol mass transfer (Equations 3.23 and 3.29), with effects on DP_3 similar to Faults #2 and #3.

To assess reproducibility of the results for different patterns of measurement and process noise, we consider 10 different realizations of each fault scenario.

The simplified FPM model employs Equations 3.23-28, but with constant parameters and feed conditions. In addition, the effect of pressure on $c^*_{\mathrm{EG},i}$ in each reactor i (Equation 3.29) is neglected and $c^*_{\mathrm{EG},i}$ is assumed constant. As a result, the FPM presents parametric and (mild) structural mismatch.

3.4.2 Case study 2: penicillin manufacturing

The manufacturing of penicillin by biomass fermentation is modeled by Birol et al. (Birol et al., 2002). The process is carried out in a reactor operating batchwise for the first 50 h (growth phase). During this period, the concentration of biomass grows, and no penicillin is produced. Then, the substrate feed is turned on, and in this fed-batch phase the biomass concentration grows slowly and the penicillin concentration increases.

For ease of reading, the set of equations defining the detailed model (Birol et al., 2002) is reported in Appendix A, together with the values of all inputs, parameters, and process noise characteristics. The simplified FPM model includes significant parametric and structural mismatch, and is also reported in Appendix A.

The 6 measurements available online from the process are pH (y_{pH}), temperature (y_T), oxygen concentration (y_{O_2}), volume (y_V), CO_2 concentration (y_{CO_2}), and feed flow rate (y_F). A measurement interval of 3 min is considered for process monitoring. Note that y_T and y_{pH} are not accounted for by the FPM, y_F is the only input of the FPM, and the other available measurements correspond to states of the FPM. White noise is added to all measurements,

consistently with the typical precision of industrial instrumentation. The standard deviations of the noise signals are 0.05 for y_{pH}, 0.05 K for y_T, 0.0025 g_{O_2} L^{-1} for y_{O_2}, 0.01 L for y_V, 0.06 mol_{CO_2} L^{-1} for y_{CO_2}, and 0.0002 L h^{-1} for y_F.

The NOC dataset includes 35 batches, each one lasting 300 h. Four faulty batches are considered, with the same length as the normal ones. The faulty batch characteristics are as follows:

- Fault #1: slow ramp decrease in the aeration rate (unmeasured variable), which causes a drop in the oxygen mass transfer coefficient and in the oxygen concentration in the reactor;
- Fault #2: slow ramp decrease of the substrate concentration in the feed (unmeasured variable), which inhibits the biomass growth;
- Fault #3: slow ramp decrease of the maximum growth rate kinetic parameter, which reduces the biomass concentration in the reactor;
- Fault #4: high cooling water temperature (unmeasured variable), which causes the reactor temperature to rise.

Numerical details on the faulty sequences are reported in Appendix A. We implement ten different realizations of each fault scenario.

3.4.3 Case study 3: pharmaceutical segmented fluid bed drying

The segmented fluid bed dryer model available in gPROMS FormulatedProducts is used for synthetic data generation for Case Study 3. The model (223 differential equations and 6256 algebraic equations) represents the physical phenomena occurring in real fluid bed dryers (Burgschweiger and Tsotsas, 2002). We use the dryer model to simulate a pharmaceutical process (Figure 3.3) in which the moisture content of wet granules fed to the unit is reduced by flowing hot air. The dryer receives a continuous feed of wet granules, with each of the six segments behaving as a fluidized bed that cycles through four phases: loading, drying, discharging and waiting. We refer to the sequence of loading, drying and discharging phases in a given segment as a "batch"; the waiting phase is not considered here as the segment is empty in that period of time. During the dryer operation, a batch is processed in each of the segments. In Figure 3.3, segment #2 is being loaded with wet granules, thus starting a new batch for that segment. When segment #2 is fully loaded, the loading of segment #3 (which was in the waiting phase until that moment) starts. Meanwhile, segment #4 is discharging, while segments #1, #5 and #6 are in the drying phase. Hot air is continuously fed to the dryer, and its flow is distributed between the six segments through a distributor plate (not shown in Figure 3.3). Details about the model assumptions and equations can be found in Burgschweiger and Tsotsas (2002). We assume that, as in typical industrial settings, measurement sensors are available for five inputs

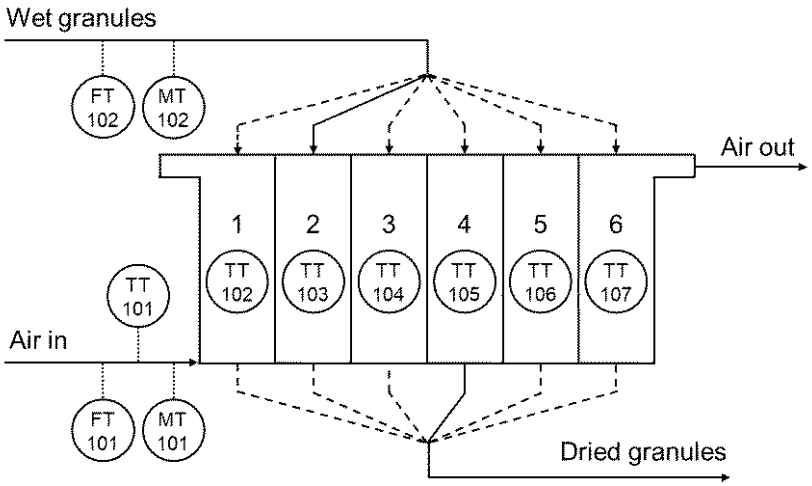


Figure 3.3. *Case study 3: sketch of a six-segmented fluid bed dryer. The full lines represent the active streams in the current phase of the process, while the dashed lines are streams not currently under operation. FT, MT and TT represent flowrate, moisture/humidity and temperature transmitters. The air distributor plate is not shown.*

(total flowrate, temperature and relative humidity of the inlet air; total flowrate and moisture of the wet granules) and six outputs (temperature in each segment). The air flows to individual segments are not measured. Note that, for a given batch, only one output measurement is available (namely, the temperature of the segment wherein the batch is being carried out). We simulate 100 NOC batches, corresponding to a total of about 6 h of dryer operation. The duration of each batch is fixed and equal to 800 s. To mimic inter-batch variability under NOC, all the five measured inputs are varied according to smoothed pseudo-random binary sequences (PRBSs) of maximum amplitude ± 0.5% around the set-point. Also the split ratios of the air among the six segments follow a smoothed PRBS pattern of maximum amplitude ±1% around their nominal values. Thus, the air is not distributed evenly between the segments (not even under NOC), a situation that may arise in practice due to the distributor plate design. We generate data for 3 faulty batches involving disturbances in the air flow to the segments, something which could be caused by partial blockage of the distributor plate. All abnormal batches start from NOC; at t = 300 s, for a given faulty batch we introduce a step (–5% for Fault #1, –10% for Fault #2) or a ramp (–0.025%/s for Fault #3) decrease in the air flow to the relevant segment. The segment air flow changes are simulated by changing the air split ratios. For each fault, Figure 3.4 compares the inlet air flow to the segment to the flow variability (expressed as 99% confidence limits) induced by the application of the smoothed PRBSs on both the total inlet air inflow and the split ratios. The input and output measurements are affected by white noise with standard deviations of 1% of the set-point for the inputs, and 0.05 °C for the outputs.

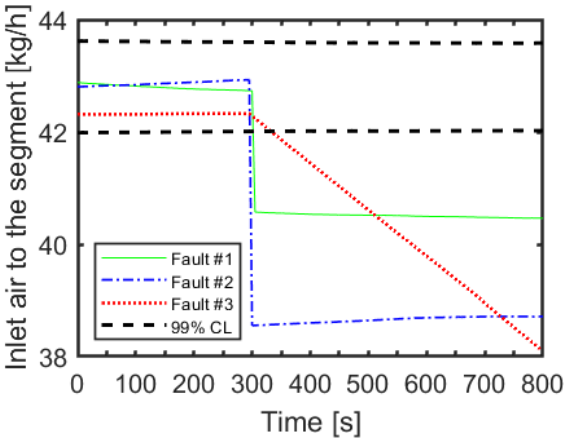


Figure 3.4. *Case study 3: profiles of segment inlet air flow (unmeasured) for the three faulty batches.*

3.5. Results and discussion for Case Study 1

3.5.1 Design of the hybrid monitoring model

The 21 variables selected for inclusion in the augmented data matrix of the hybrid monitoring model are listed in Table 3. They comprise the 9 available field measurements and the 12 states estimated by the EKF using the FPM and the available measurements. No model parameters are included in the augmented matrix.

The difference in the concentration measurement intervals (1 min vs. 10 min, see §3.4.1) is dealt with using a two time-scale EKF (Figure 3.5), which we borrow from (Scali et al., 1997). EKF-1 receives the frequent measurements and performs the prediction and correction steps for all states except $c_{\mathrm{COOH},i}$, which is predicted at open loop due to the observability conditions. EKF-2 receives the infrequent measurements and provides corrections for $c_{\mathrm{COOH},i}$. The accuracy of estimation of the unmeasured states $c_{\mathrm{EG},i}$ and $c_{\mathrm{ESTER},i}$ is improved through bias estimation (Liotta et al., 1997). Following the approach adopted in (Ling and Kravaris, 2016), to satisfy the observability conditions three biases are updated in EKF-1 ($b_{\mathrm{EG},1}$, $b_{\mathrm{EG},2}$ and $b_{\mathrm{EG},3}$), whereas the other three biases are updated in EKF-2 ($b_{\mathrm{ESTER},1}$, $b_{\mathrm{ESTER},2}$ and $b_{\mathrm{ESTER},3}$).

We design the initial state variance matrices ($\mathbf{P}_{0,\,\mathrm{EKF\text{-}1}}$ and $\mathbf{P}_{0,\,\mathrm{EKF\text{-}2}}$) as diagonal matrices with zero variance for the biases, and the same variance (equal to 1E-6) for all the other states. We tune $\mathbf{Q}_{\mathrm{EKF\text{-}1}}$ and $\mathbf{Q}_{\mathrm{EKF\text{-}2}}$ heuristically to achieve quick and robust convergence, resulting in:

$$\mathbf{Q}_{\mathrm{EKF\text{-}1}} = \mathrm{diag}(4\mathrm{E}\text{-}6 \quad 0 \quad 0 \quad 0 \quad 1\mathrm{E}\text{-}6 \quad 1\mathrm{E}\text{-}8 \quad 0 \quad 0 \quad 1\mathrm{E}\text{-}6 \quad 2.5\mathrm{E}\text{-}9 \quad 0 \quad 0 \quad 4\mathrm{E}\text{-}10 \quad 0 \quad 4\mathrm{E}\text{-}10 \quad 0 \quad 4\mathrm{E}\text{-}10 \quad 0) \tag{3.31}$$

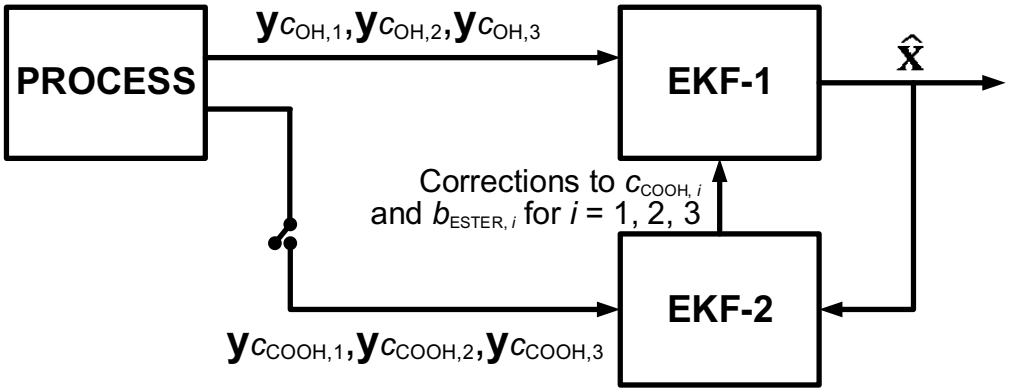


Figure 3.5. *Architecture of the knowledge-driven block (two time-scale EKF) implemented for Case Study 1 (adapted from Ling and Kravaris, 2016).*

Table 3.3. *Case study 1: list of variables included in the augmented data matrix of the hybrid monitoring model.*

Var. no.	Variable name	Reactor no.	Symbol	Units	Variable type
1	pressure	1	y_{P_1}	Pa	Measurement
2	hydroxyl end-groups concentration	1	$y_{c_{OH,1}}$	mol L^{-1}	Measurement
3	acid end-group concentration	1	$y_{c_{COOH},1}$	mol L^{-1}	Measurement
4	pressure	2	y_{P_2}	Pa	Measurement
5	hydroxyl end-group concentration	2	$y_{c_{OH,2}}$	mol L^{-1}	Measurement
6	acid end-group concentration	2	$y_{c_{COOH,2}}$	mol L^{-1}	Measurement
7	pressure	3	y_{P_3}	Pa	Measurement
8	hydroxyl end-group concentration	3	$y_{c_{OH,3}}$	mol L^{-1}	Measurement
9	acid end-group concentration	3	$y_{c_{COOH,3}}$	mol L^{-1}	Measurement
10	ethylene glycol concentration	1	$c_{EG,1}$	mol L^{-1}	Estimated state
11	hydroxyl end-group concentration	1	$c_{OH,1}$	mol L^{-1}	Estimated state/Reconstructed measurement
12	acid end-group concentration	1	$c_{COOH,1}$	mol L^{-1}	Estimated state/Reconstructed measurement
13	ester end-group concentration	1	$c_{ESTER,1}$	mol L^{-1}	Estimated state
14	ethylene glycol concentration	2	$c_{EG,2}$	mol L^{-1}	Estimated state
15	hydroxyl end-group concentration	2	$c_{OH,2}$	mol L^{-1}	Estimated state/Reconstructed measurement
16	acid end-group concentration	2	$c_{COOH,2}$	mol L^{-1}	Estimated state/Reconstructed measurement
17	ester end-group concentration	2	$c_{ESTER,2}$	mol L^{-1}	Estimated state
18	ethylene glycol concentration	3	$c_{EG,3}$	mol L^{-1}	Estimated state
19	hydroxyl end-group concentration	3	$c_{OH,3}$	mol L^{-1}	Estimated state/Reconstructed measurement
20	acid end-group concentration	3	$c_{COOH,3}$	mol L^{-1}	Estimated state/Reconstructed measurement
21	ester end-group concentration	3	$c_{ESTER,3}$	mol L^{-1}	Estimated state

$$\mathbf{Q}_{\text{EKF-2}} = \text{diag}(0 \quad 0 \quad 2.5\text{E-}9 \quad 0 \quad 0 \quad 0 \quad 2.5\text{E-}9 \quad 0 \quad 0 \quad 0 \quad 2.5\text{E-}9 \quad 0 \quad 0 \quad 3.6\text{E-}7 \quad 0 \quad 3.6\text{E-}7 \quad 0 \quad 3.6\text{E-}7) \tag{3.32}$$

In Equations 3.31 and 3.32, the first twelve elements along the diagonal correspond to the estimated states (variable nos. 10-21 in Table 3), whereas the last six refer to the bias vector $[b_{\text{EG},1} \; b_{\text{ESTER},1} \; b_{\text{EG},2} \; b_{\text{ESTER},2} \; b_{\text{EG},3} \; b_{\text{ESTER},3}]$. Finally, $\mathbf{R}_{\text{EKF-1}}$ and $\mathbf{R}_{\text{EKF-2}}$ are built according to Equation 3.12, with $\sigma_{y_{\text{OH},i}}$= 4.9E-5 and $\sigma_{y_{\text{COOH},i}}$= 9E-8. The typical estimation performance during a transient is shown in Figure 3.6 for one unmeasured state in the second reactor. The estimation accuracy is satisfactory.

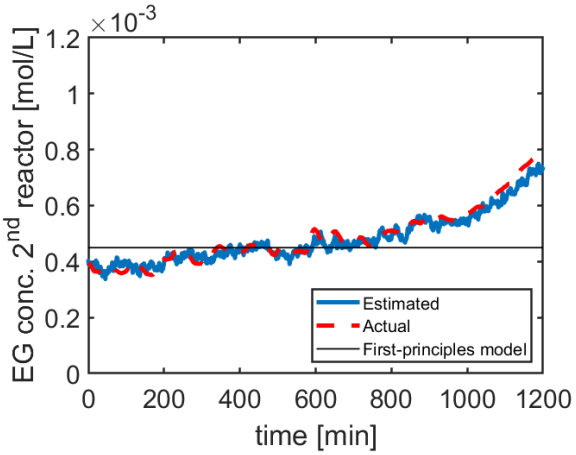


Figure 3.6. *Case study 1: EKF estimation performance for one unmeasurable state in the second reactor during a transient.*

We design a DPCA model on the 21 variables listed in Table 3 under NOC, sampling all signals every 10 min. The number L of lagged measurements to be included in $\mathbf{Z}_{dyn}$ (Equation 3.20) is heuristically derived from the overall residence time in the three reactors (3 × 60 = 180 min), resulting in L = 18. The resulting size of $\mathbf{Z}_{dyn}$ is [512×399]. Eight PCs are used (this number being found by cross-validation), explaining 54% of data variability on T^2, with the remaining 46% being explained by SPE. The resulting control charts under NOC are shown in Figure 3.7: no false alarms are issued.

3.5.2 Fault detection and diagnosis

We compare the monitoring performances of the hybrid model to those of a standard DD monitoring model and of a standard KD monitoring approach. The standalone DD model uses only the 9 measurements available from the process and is based on a DPCA model designed with the same characteristics of the hybrid model (18 lagged measurements spaced by 10 min). Cross-validation suggests using 8 PCs for this model, too. Monitoring through a standalone KD model is carried out using an EKF and χ^2 tests on $\epsilon(t_k)$ and $\epsilon(t_k, S)$ (Deshpande et al., 2009).

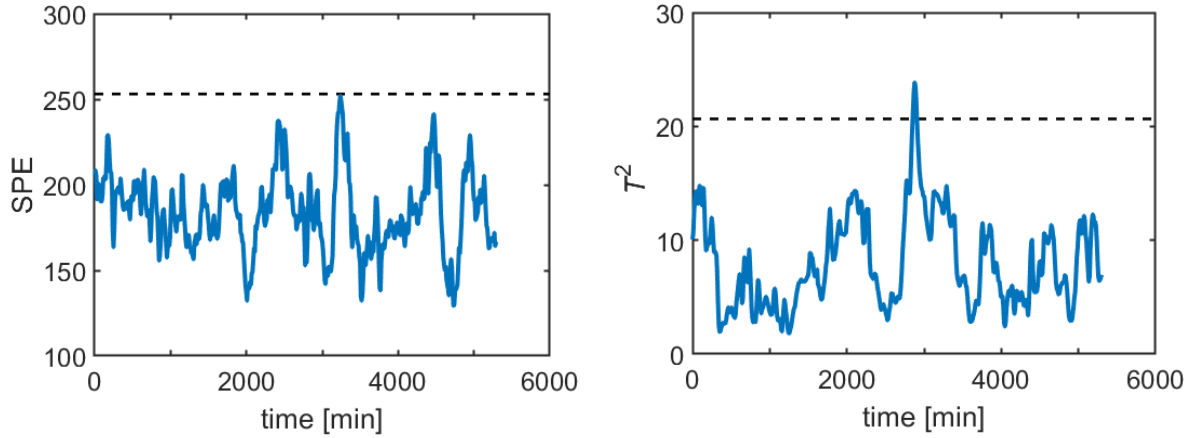


Figure 3.7. *Case study 1: control charts under normal operating conditions: (a) SPE chart, (b) Hotelling T^2 chart, (c) contributions to SPE. The dashed lines represent the 99% confidence limits.*

The confirmation test is implemented with the same moving window of the hybrid and DD models; no state augmentation is used.

For each fault scenario, the fault detection time is reported in Table 3.4 as an average across the relevant fault realizations. Whereas the detection performances for Fault #4 result the same for the hybrid and the DD models, the hybrid model detects Fault #1 and Fault #3 much more promptly than the DD one (170 and 320 min earlier, respectively). Additionally, Fault #2 (a subtle one) goes undetected by the DD model, whereas it is correctly detected by the hybrid model. The KD method leads to the worst detection performance, as it is severely compromised by the process-model mismatch. Fault #2 is not detected, and the average detection times for the other fault scenarios are greater than those of the hybrid and DD models (changing the confirmation test window size does not lead to any substantial improvement). For this reason, the KD model will not be investigated further for this case study. Incidentally, if the process is monitored by univariate control charts on measurements or estimated states, unsatisfactory detection performances for all fault scenarios are obtained (results are expected and are not reported for conciseness).

The reason why the hybrid model performs better than the DD model in fault detection can be explained as follows. Thanks to the presence of the KD block (Figure 3.1), the hybrid model embeds more information about the inner working of the process, namely on how the measured variables *and* the states are expected to co-vary under NOC. Hence, mutual deviations of measured variables and states from the relevant reference trajectories (such as those occurring after the onset of a fault) can be detected effectively. For example, Fault #3 starts impacting on the (estimated) states earlier than it does on the measurements, as can be seen from Figure 3.8 for the representative profiles of the ethylene glycol concentration (an estimated state in the second reactor) and of the hydroxyl end-group concentration (a measured output in the same

Table 3.4 *Case study 1: fault detection time using the hybrid monitoring model, the data-driven monitoring model and the knowledge-driven monitoring model. The detection time runs from the fault onset, and the reported values are averaged across the 10 realizations of the relevant fault. For the hybrid and data-driven models, the monitoring statistics alarming the fault is also indicated.*

Fault number and type	Primarily affected variable or parameter	Hybrid model detection time (min)	Data-driven model detection time (min)	Knowledge-driven model detection time (min)
#1: ester feed concentration decrease	$c_{ESTER,0}$	550 (SPE)	720 (T^2)	880
#2: minor agitation fault	h_2	590 (SPE)	(undetected)	(undetected)
#3: major agitation fault	h_2	380 (SPE)	700 (SPE)	930
#4: pressure increase	P_2	40 (SPE)	40 (SPE)	890

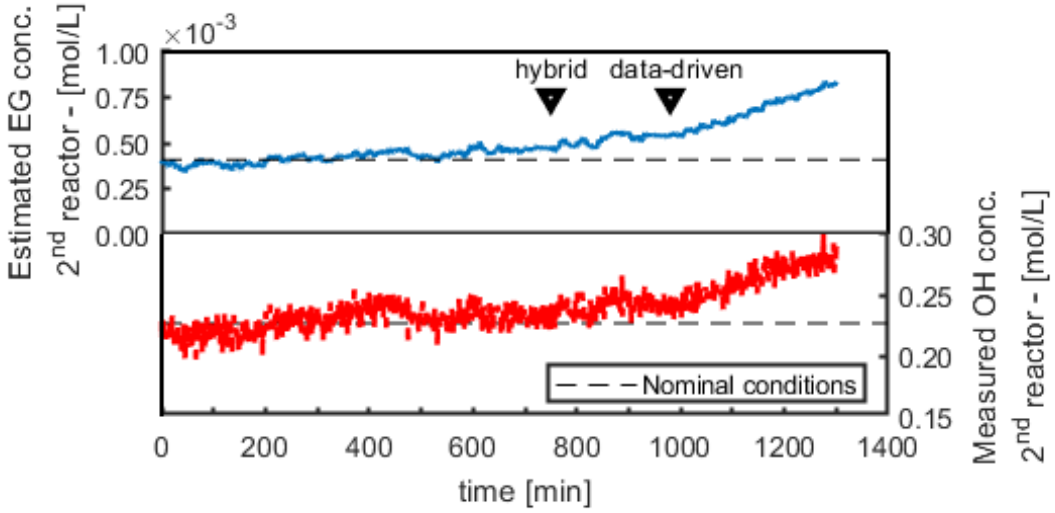


Figure 3.8. *Case study 1: time profiles of the ethylene glycol concentration (estimated state) and of the ethylene glycol concentration (measured variable) before and after the onset of Fault #3 (time = 300 min). The triangles indicate the detection instant of the hybrid model and of the data-driven model.*

reactor). This piece of information is captured by the DD block, and this allows anticipating the fault detection. On the other hand, if the fault impacts on a measured variable directly, rather than through (or after) the change in one or more states (as occurs for reactor 2 pressure in Fault #4), it is unlikely that the information provided by the KD block to the hybrid model can lead to a significant improvement in the detection performance.

Contribution plots of the hybrid and of the DD models at the first out-of-control signal are used for fault diagnosis. For Fault #1, Figure 3.9a shows that the hybrid monitoring model points to the concentration of ester in the first reactor (variable no. 13 in Table 3.3, an estimated state) as most related to the deviation from the NOC, which would straightforwardly suggest an abnormal feed concentration change as a possible root-cause of the fault. On the other hand, the information provided by the DD monitoring model for fault diagnosis is more ambiguous: Figure 3.10a misleadingly draws the attention to the concentration of acid end-groups in the third reactor (variable no. 9), a measurement that is not related to the root-cause of the fault

directly. In ~50% of the realizations of this fault, the DD model provides T^2 contributions outside the control limits for acid end-groups also in the first and in the second reactors. Indeed, due to the reduced ester concentration in the reactors caused by the fault, also the acid end-group concentrations are expected to decrease because of the ester degradation kinetics (Equation 3.28), but this is only a secondary (slower) effect. With respect to Faults #2 and #3, the hybrid model (Figure 3.9b and 3.9c) clearly identifies an abnormal ethylene glycol concentration in the second reactor (variable no. 14, an estimated state) as the variable most directly related to the fault, thus correctly pointing the attention to a possibly abnormal mass transfer of EG in that reactor. On the other hand, as already mentioned, Fault #2 is not detected by the DD monitoring model because it is too small in magnitude, whereas for Fault #3 the DD model points to measured variable no. 5, i.e. to the hydroxyl end-group concentration in reactor

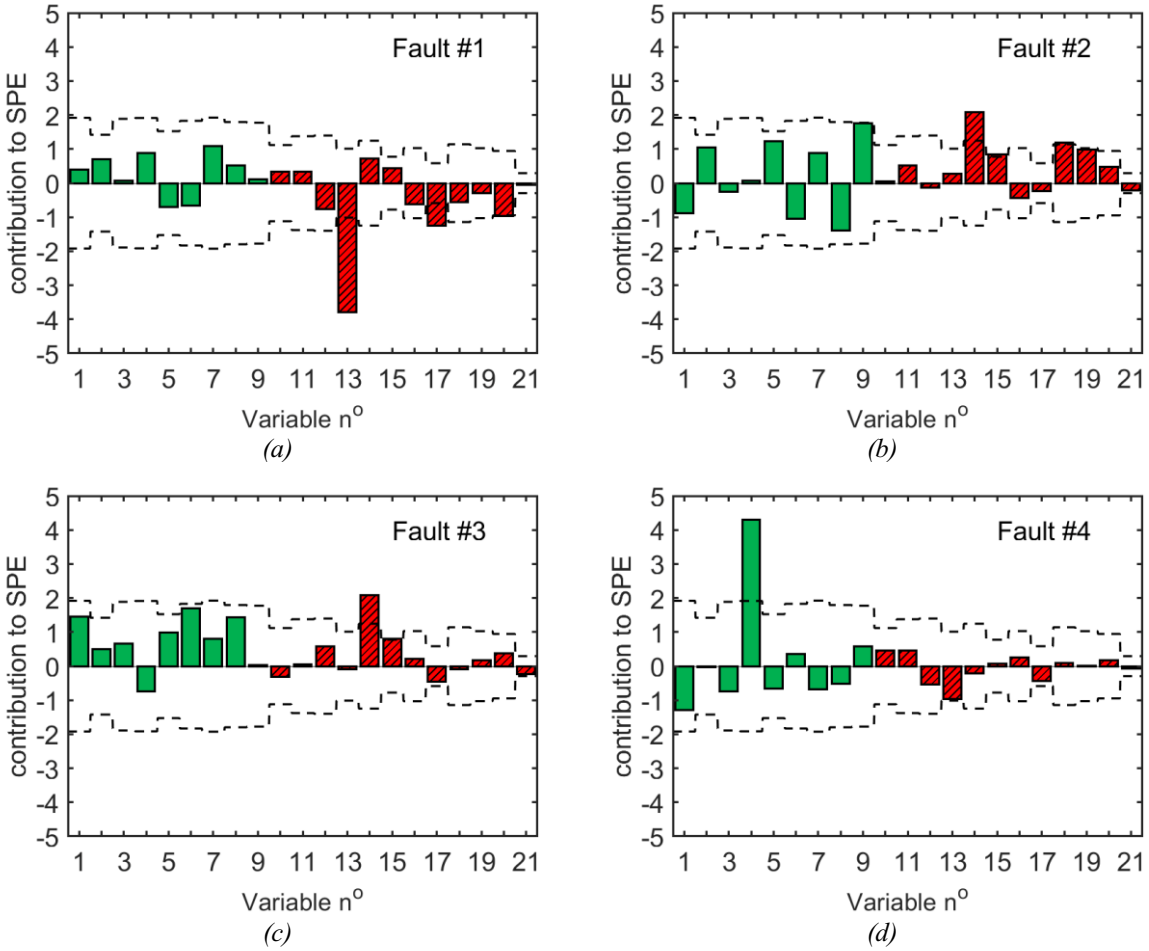


Figure 3.9. *Case study 1: representative contribution plots for the hybrid monitoring model at the first out-of-control observation for (a) Fault #1, (b) Fault #2, (c) Fault #3, and (d) Fault #4. Variables are numbered as in Table 3. In all plots, the contributions of field measurements are in green, while those of estimated/reconstructed variables are in red with diagonal lines. Control limits at 99% confidence are shown as dashed lines.*

2 (Figure 3.10b). In fact, accumulation of ethylene glycol due to the fault hinders the polycondensation reaction (Equation 3.27), leading to an increase in $y_{c_{\mathrm{OH},2}}$. However, attributing a reduction in the polycondensation rate of reaction to a high ethylene glycol concentration may not be straightforward.

Finally, Fault #4 can be diagnosed very easily both by the hybrid model (Figure 3.d) and by the DD model (Figure 3.c), because this fault impacts on reactor 2 pressure (measured variable no. 4) directly. Note that very accurate state estimation is not required for the hybrid monitoring system to perform well. In fact, the main task of the KD block in Figure 3.1 is only to provide information on how the estimated states (together with reconstructed measurements

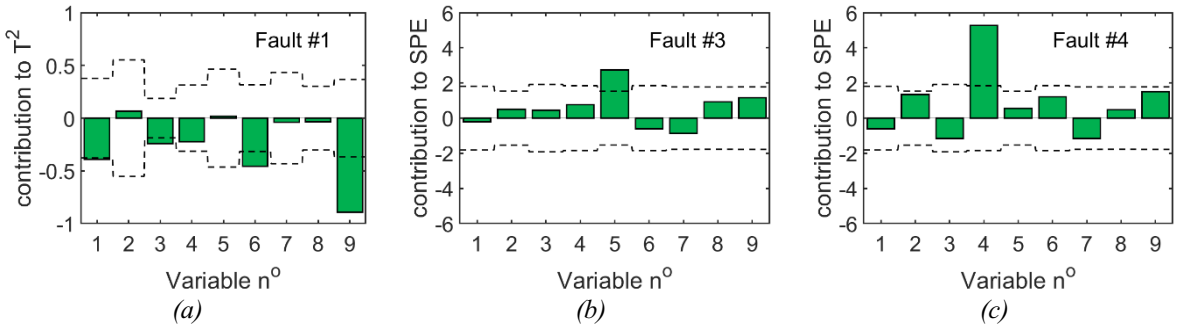


Figure 3.10. *Case study 1: representative contribution plots for the purely data-driven model at the first out-of-control observation for (a) Fault #1, (b)Fault #3, and (c) Fault #4. Variables are numbered as is in Table 3. Control limits at 99% confidence are shown as dashed lines.*

and possibly adapted parameters) co-vary during the process operation, regardless of the fact that the actual values of the states may be somewhat different from the actual (and unknown) ones. To clarify this point, we consider a stronger process/model mismatch by altering the FPM dynamics through summation of a constant term (equal to –2E-3) to the right-hand side of Equation 3.24, even though the process dynamics remains the same. Additionally, we degrade the filter performance by tuning $\mathbf{Q}_{\text{EKF-2}}$ as:

$$\mathbf{Q}_{\text{EKF-2}} = \text{diag}(0 \quad 0 \quad 2.5\text{E-}9 \quad 0 \quad 0 \quad 0 \quad 2.5\text{E-}9 \quad 0 \quad 0 \quad 0 \quad 2.5\text{E-}9 \quad 0 \quad 0 \quad 4\text{E-}8 \quad 0 \quad 4\text{E-}8 \quad 0 \quad 4\text{E-}8) \quad . \tag{3.33}$$

This results in unsatisfactory state estimation, as shown in Figure 3.11 for the same state and transient considered in Figure 3.6. Nevertheless, the fault detection and fault diagnosis performance of the hybrid monitoring model are almost the same as those reported in Table 3.4 and shown in Figure 3.9 (results are not reported for conciseness).

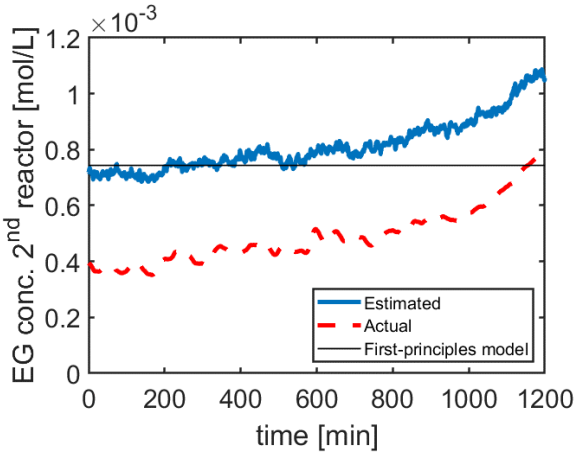


Figure 3.11. *Case study 1: EKF estimation performance on the unmeasurable state and under the same dynamic conditions of Figure 3.4, in the presence of significant process-model mismatch and bad filter tuning.*

3.6. Results and discussion for Case Study 2

3.6.1 Design of the hybrid monitoring model

The augmented data matrix of the hybrid monitoring model includes the 15 variables listed in Table 3.4. To improve the monitoring performance, in addition to state estimation the EKF performs online adaptation of three FPM parameters: $\mu_{X,\max}$ (maximum growth parameter), K_{la} (mass transfer coefficient) and s_F (substrate concentration in the feed). We select these parameters because they can provide useful insights for typical potential faults that may affect the reactor, namely changes in kinetics, mass transfer, or feed composition.
The modeled measurements (y_{O_2}, y_V, y_{CO_2}) and measured input (y_F) are supplied to the EKF every 3 min. The $\mathbf{P}_0$, $\mathbf{Q}$ and $\mathbf{R}$ matrices used are designed using the same criteria as in Case study #1, resulting in (states and measurements are ordered as is in Table 3.4):

$$\mathbf{P}_0 = \text{diag}(1.00\text{E-}4 \quad 1.00\text{E-}4 \quad 1.00\text{E-}4 \quad 1.00\text{E-}4 \quad 1.00\text{E-}4 \quad 1.00\text{E-}4 \quad 2.25\text{E-}4 \quad 1.00\ 0.00) \tag{3.34}$$

$$\mathbf{Q} = \text{diag}(2.50\text{E-}8 \quad 1.00\text{E-}9 \quad 9.00\text{E-}7 \quad 2.50\text{E-}8 \quad 9.00\text{E-}5 \quad 2.50\text{E-}8 \quad 6.50\text{E-}8 \quad 2.50\text{E-}3\ 8.00\text{E-}2) \tag{3.35}$$

$$\mathbf{R} = \text{diag}(6.25\text{E-}06\ 1.00\text{E-}04\ 3.60\text{E-}03) \tag{3.36}$$

The resulting state estimation performance is satisfactory (Figure 3.12). Nevertheless, as illustrated in Figure 3.13, the adapted parameters drift away from their "true" values even under NOC, because the EKF adjusts the model parameters in the attempt to compensate for the detected process-model mismatch. This does not represent an issue for the hybrid system,

Table 3.4 *Case study 2: list of variables included in the augmented data matrix of the hybrid monitoring model.*

#	Symbol	Variable	Unit	Type of variable
1	y_{pH}	pH	-	Measurement
2	y_T	temperature	K	Measurement
3	y_{O_2}	oxygen concentration	g_{O_2} L^{-1}	Measurement
4	y_V	volume	L	Measurement
5	y_{CO_2}	CO_2 concentration	$mol_{CO_2}L^{-1}$	Measurement
6	y_F	feed flow rate	L h^{-1}	Measurement
7	X	biomass concentration	g_X L^{-1}	Estimated state
8	P	penicillin concentration	g_P L^{-1}	Estimated state
9	S	substrate concentration	g_S L^{-1}	Estimated state
10	c_{O_2}	oxygen concentration	g_{O_2} L^{-1}	Estimated state/Reconstructed measurement
11	V	volume	L	Estimated state/Reconstructed measurement
12	c_{CO_2}	CO_2 concentration	$mol_{CO_2}L^{-1}$	Estimated state/Reconstructed measurement
13	$\mu_{X,max}$	maximum growth parameter	h^{-1}	Adapted parameter
14	K_{la}	mass transfer coefficient	h^{-1}	Adapted parameter
15	s_F	feed concentration	g_S L^{-1}	Adapted parameter

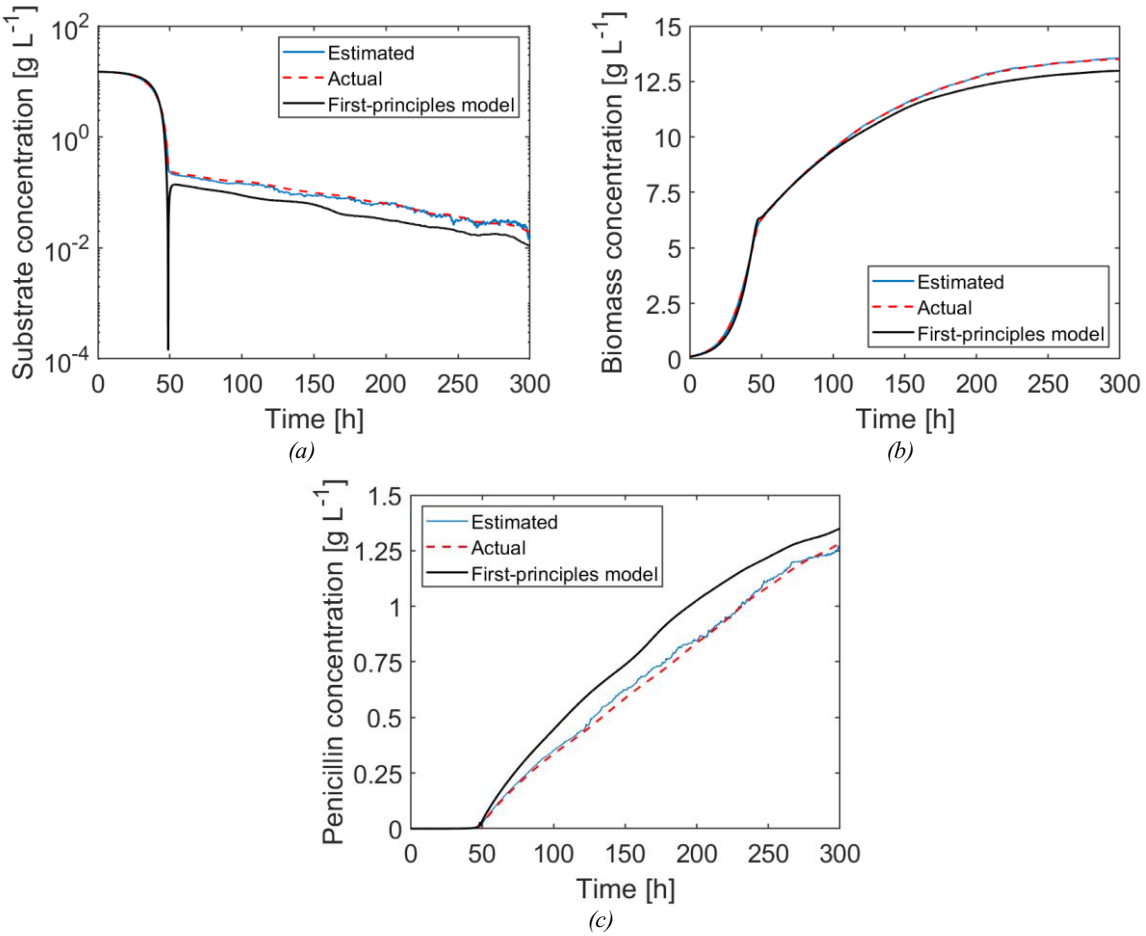


Figure 3.12. *Case study 2: EKF state estimation performance for three unmeasured states during a representative NOC batch: (a) substrate concentration, (b) biomass concentration, and (c) penicillin concentration*

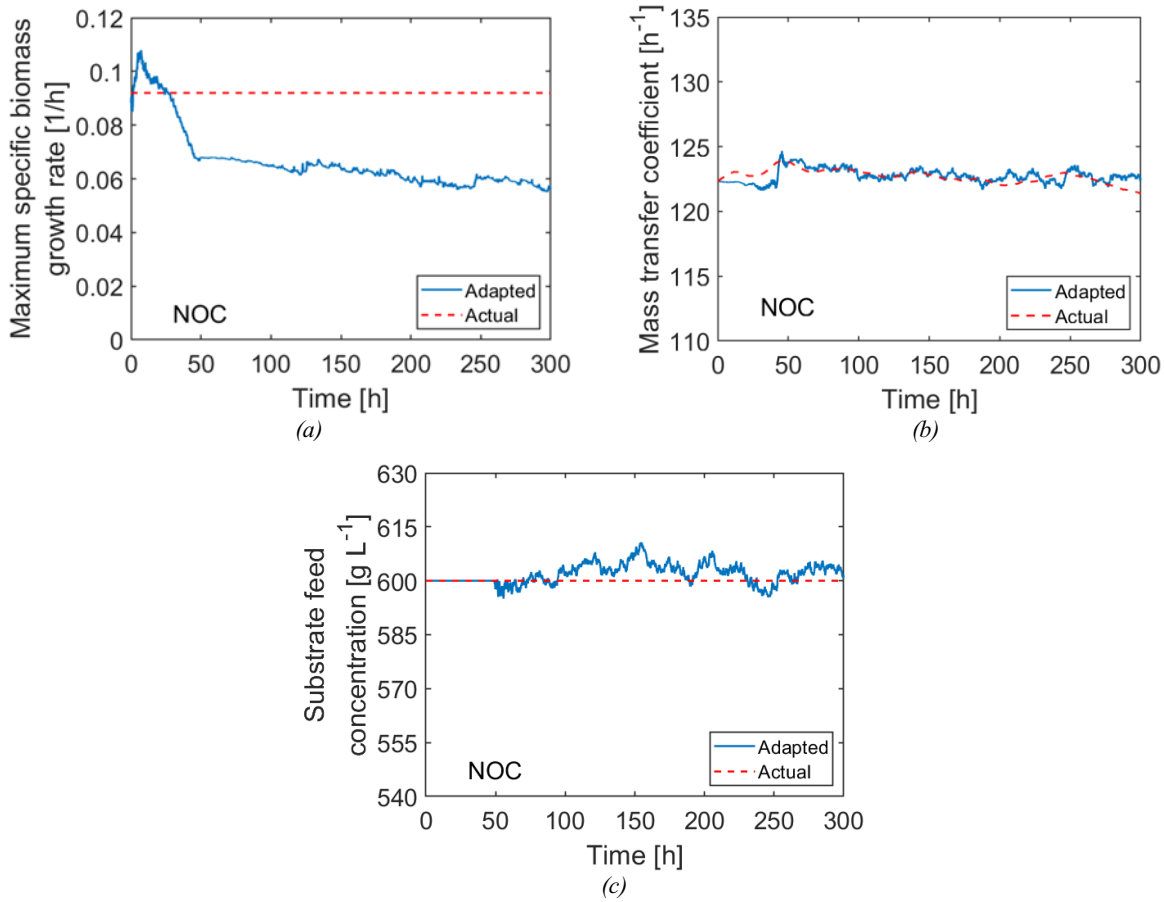


Figure 3.13. *Case study 2: EKF parameter adaptation performance during a representative NOC batch of (a) the maximum specific biomass growth rate $\mu_{X,max}$, (b) the mass transfer coefficient K_{la}, and (c) the substrate feed concentration s_F.*

because it is the pattern of change of the adjusted parameter profiles together with the profiles of all other measured and estimated variables that matters for process monitoring, regardless of the fact that each single parameter is estimated accurately or not. Incidentally, note that, in the presence of significant structural mismatch, a comparison between the "process" and the "model" parameters might even not be entirely appropriate, because the meaning of a parameter within the "process" may be different from the one the same parameter has in the "model". This is especially true when a model parameter (e.g., the maximum growth rate kinetic parameter; see Appendix A) is used to compactly represent a set of physical mechanisms that are expected to occur in the process, but are not described accurately by the model equations.

We design an MW-PCA model for the DD block of the hybrid monitoring system. Considering that the total batch duration is 300 h, measurements are retained every 1 h. After exploring several windows widths, we select W = 50 h as the width leading to satisfactory monitoring performances. The size of $\mathbf{Z}_{MW}$ is [35×15] after 1 h, and grows of 15 columns per hour until 50 h. Then, a fully developed $\mathbf{Z}_{MW}$ of size [35×750] is used until the end of the batch. Three PCs

are used (found by cross-validation), with an explained variance profile of ~45% in the production phase, after reaching a minimum of ~25% at the switch between the growth and the production phases, due to the high variability of the switching instant across the NOC batches. The loadings of the first two PCs over one window width during the production phase (Figure1) allow assessing the auto- and cross-correlation of the variables included in the augmented data matrix. MW-PCA provides a correlation model in which estimated states, adapted parameters and measurements (including those not accounted for by the FPM) are all intimately linked. For example, from the analysis of the loadings on the first PC (Figure 3.14a), it emerges that the reactor temperature (variable no.2, which is measured but not modeled by the FPM) is not only strongly auto-correlated, but also cross-correlated to several variables modeled by the FPM, such as for example the reactor volume (variable no.11). As will be shown in the next section, exploitation of the cross-correlation between states, measurements and parameters enhances the monitoring capability of the hybrid model.

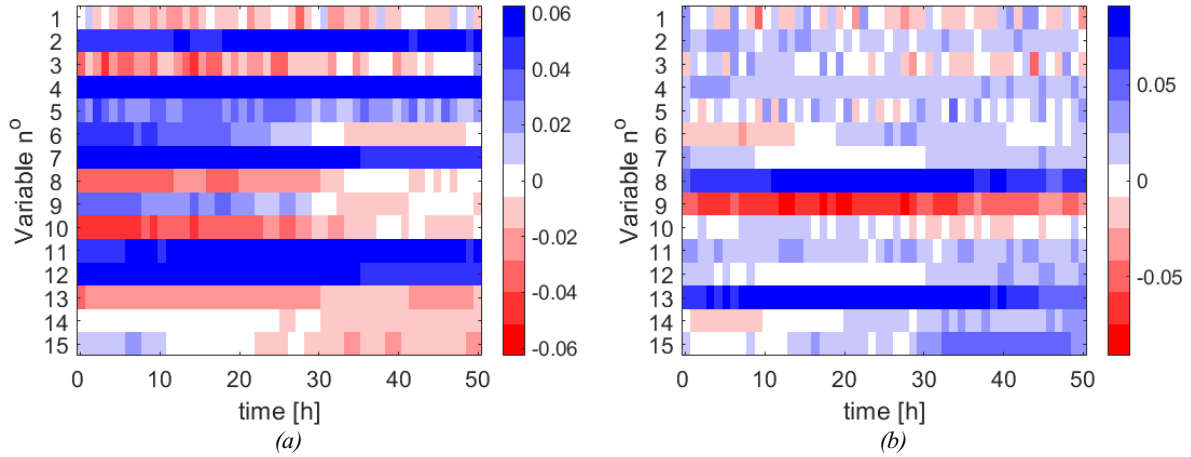


Figure 3.14. *Case study 2: loadings for (a) the first PC and (b) the second PC of the MW-PCA model over one window width during the production phase. The variables are numbered as in Table 3.4.*

3.6.2 Fault detection and diagnosis

We compare the fault detection and diagnosis performances of the hybrid monitoring model to those of a standalone DD monitoring model that uses only the 6 measurements available from the process (3.4.2 and first six entries in Table 3.5). We use a window width of 50 h also for the DD model, and cross-validation suggests retaining 3 PCs.

We also implement the KD fault detection and confirmation tests on the innovation sequence of the EKF without parameter estimation for comparison. However, as shown in Figure 3.15, this approach performs poorly: due to process-model mismatch, false alarms are issued under NOC. Consequently, we build a benchmark KD approach by monitoring drifts of the adjusted

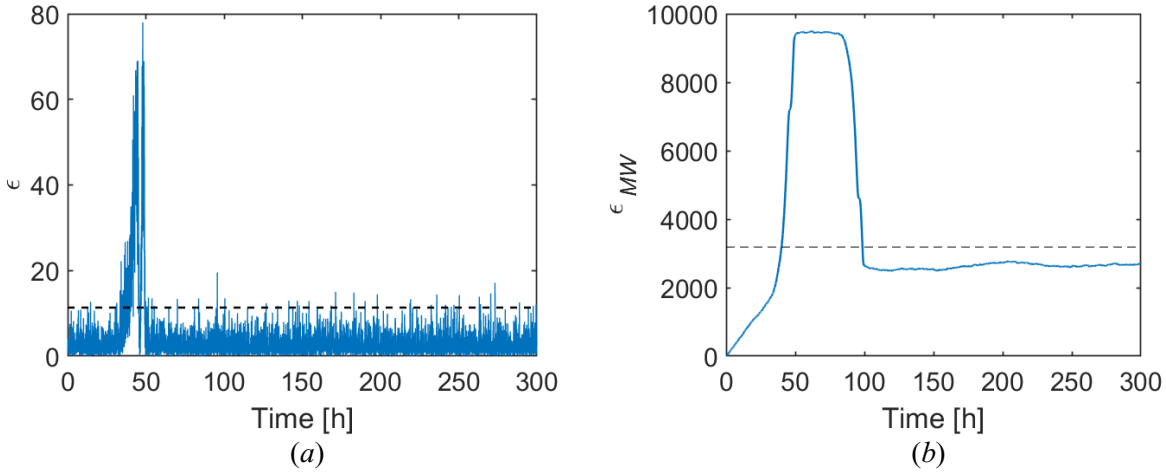


Figure 3.15. *Case study 2: control charts under normal operating conditions for the knowledge-driven approach. (a) χ^2 test on the single-point innovation (fault detection), (b) χ^2 test on the innovations on a 50 h moving window (fault confirmation). The dashed lines represent the 99% confidence limits.*

parameters with respect to their nominal values (Ku et al., 1994). Typical profiles of the adjusted parameters under faulty conditions are shown in Figures 3.16-3.18.

For each fault scenario, the fault detection time is reported in Table 3.6 as an average across the relevant fault realizations for the three monitoring approaches. The KD monitoring approach struggles with the high variability of the adjusted parameter profiles under NOC, which masks faulty parametric drifts. Only Fault #1 and Fault #2 are correctly detected and diagnosed (see also Figure 3.17a and Figure 3.17b), although much later than with the hybrid model. On the other hand, neither Fault #3 nor Fault #4 can be detected. This approach also suffers from an additional limitation. Due to the fact that EKF adjusts the model parameters to compensate for the process-model mismatch, the effect of a fault gets smeared into simultaneous variations of all parameters. This limits the possibility to monitor each single parameter in order to detect and isolate the faults. For example, looking at the very fluctuating substrate feed concentration estimation during certain realizations of Fault #3 (Figure 3.18c), one might wrongly ascribe the faulty condition to this parameter.

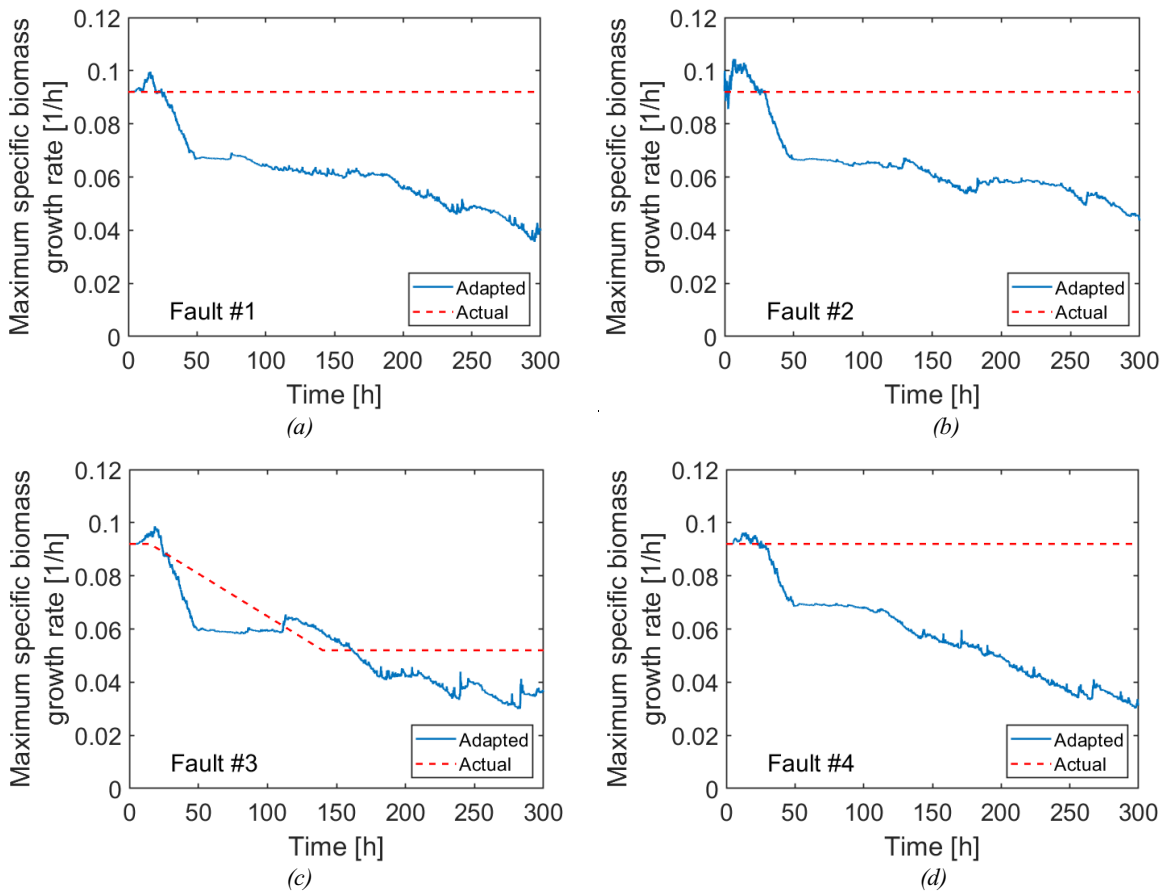


Figure 3.16. *Case study 2: EKF parameter adaptation performance for the maximum specific biomass growth rate $\mu_{X,max}$ during a representative batch for (a) Fault #1, (b) Fault #2, (c) Fault #3, and (d) Fault #4. Adapted = value of the parameter estimated by the EKF. Actual = actual value of the parameter in the plant.*

Table 3.5. *Case study 2: fault detection time using the hybrid monitoring model, the data-driven monitoring model and the knowledge-driven monitoring model. The detection time runs from the fault onset, and the reported values are averaged across the 10 realizations of the relevant fault. For the hybrid and data-driven models, the monitoring statistics alarming the fault is also indicated.*

Fault number type	Primarily affected variable or parameter	Hybrid model detection time (h)	Data-driven model detection time (h)	Knowledge-driven model detection time (h)
#1: aeration rate decrease	K_{la}	45 (SPE)	110 (SPE)	70
#2: substrate feed concentration decrease	s_F	35 (SPE)	70 (SPE)	62
#3: growth rate decrease	$\mu_{X,\max}$	14 (SPE)	28 (SPE)	(undetected)
#4: cooling water temperature rise	T	35 (SPE)	35 (SPE)	(undetected)

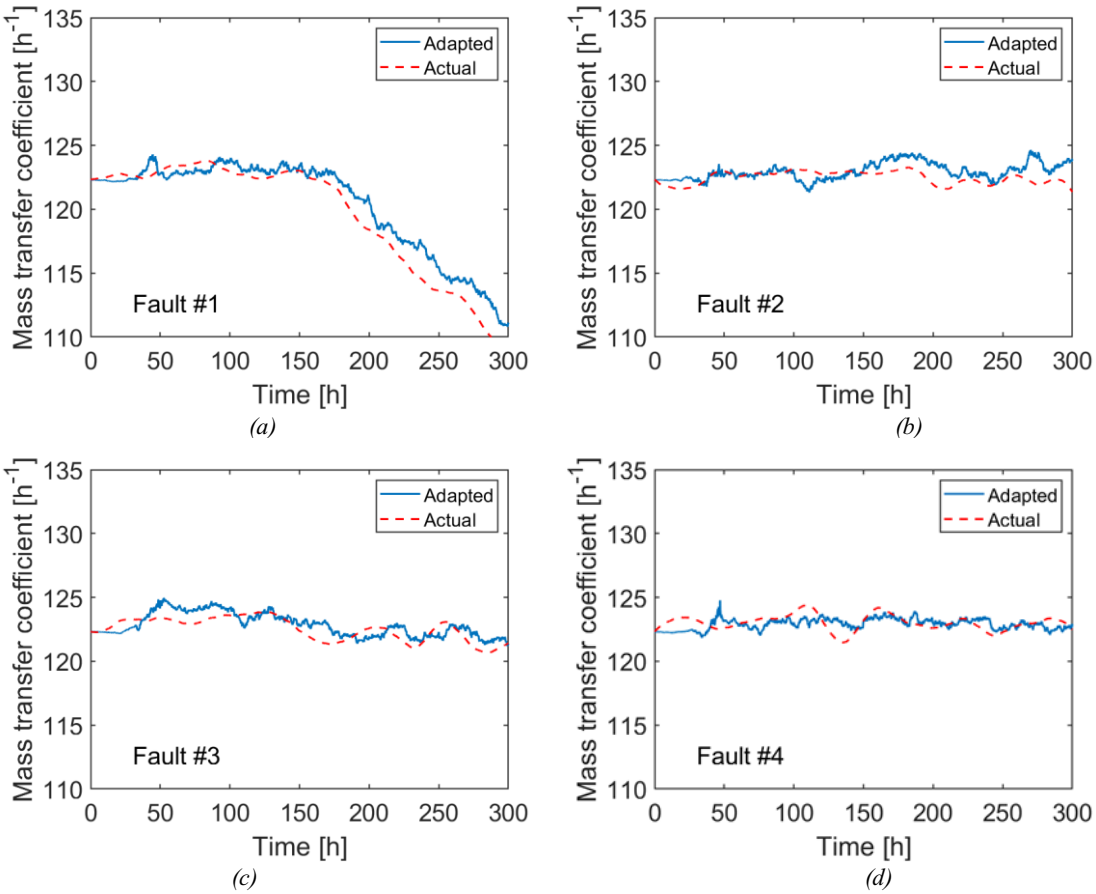


Figure 3.17 *Case study 2: EKF parameter adaptation performance for the mass transfer coefficient K_{la} during a representative batch for (a) Fault #1, (b) Fault #2, (c) Fault #3, and (d) Fault #4. Adapted = value of the parameter estimated by the EKF. Actual = actual value of the parameter in the plant.*

On the other hand, all fault scenarios are detected with hybrid and DD monitoring models for all fault realizations. Still, the hybrid monitoring model alarms the faults much earlier than the DD one for all faults except Fault #4. For this fault, the reactor temperature (a measured variable) is affected by the fault at the same time as the states, and therefore the information the KD block passes to the DD block does not contribute to improve the detection performance.

The DD monitoring model does not allow to diagnose the faults unambiguously, because different faults generate qualitatively similar contribution plots. For example, both Fault #1 (Figure 3.19a) and Fault #3 (Figure 3.19c) point to the oxygen concentration (variable no.3), which is indeed a measurement strongly affected by both faults, but is not useful to clearly discriminate the root-causes of the faults. Reactor temperature (measured variable no.2) is pinpointed as the suspected variable for both Fault #2 (Figure 3.19b) and Fault #4 (Figure 3.19d), but this leaves the diagnosis problem open. The main difficulty with the DD monitoring model is that in this system there are too few measurements on which a fault can manifest. This implies that different faults become visible only through the same measurements, which makes fault diagnosis harder. On the other hand, in the hybrid monitoring system the estimated states

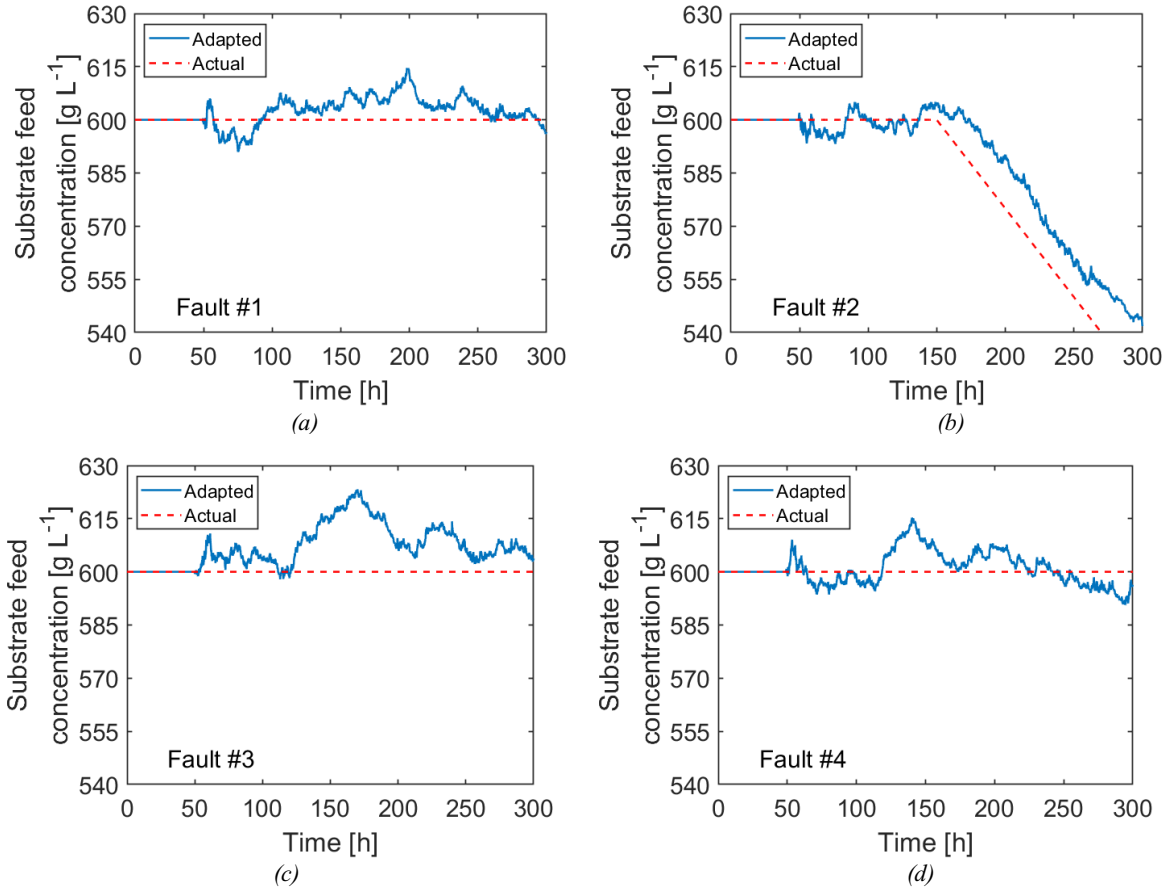


Figure 3.18. *Case study 2: EKF parameter adaptation performance for the substrate feed concentration s_F during a representative batch for (a) Fault #1, (b) Fault #2, (c) Fault #3, and (d) Fault #4. Adapted = value of the parameter estimated by the EKF. Actual = actual value of the parameter in the plant.*

and parameters provide a set of additional "virtual measurements" that can capture a qualitative signature of the fault. Not only does this allow to anticipate fault detection, but it can also point to the root-cause of the fault in a more straightforward way, because – by design – the virtual measurements represent the underlying mechanisms through which the fault propagates into the system. In fact, the contribution plots derived from the hybrid model (3.20) provide information that is very helpful for fault diagnosis. The aeration problem (Fault #1, 3.20a) is marked by anomalously small contributions for the mass transfer coefficient (variable no. 14) and the oxygen concentration (variable no. 10), whereas in Fault #2 the abnormal feed concentration is clearly spotted (Figure 3.4b, variable no.15). For the biomass growth rate decrease problem (Fault #3; 4c), the most significant contributions refer to variables related to the biomass reaction, including the online adapted kinetic parameter (variable no. 13); from this piece of information, a biomass growth reaction problem can be diagnosed in a relatively easy way. The 3.20d), but temperature is a measurement where an abnormal reactor cooling can leave a footprint directly. Hence, the hybrid model contribution plot does no better than the DD model one.

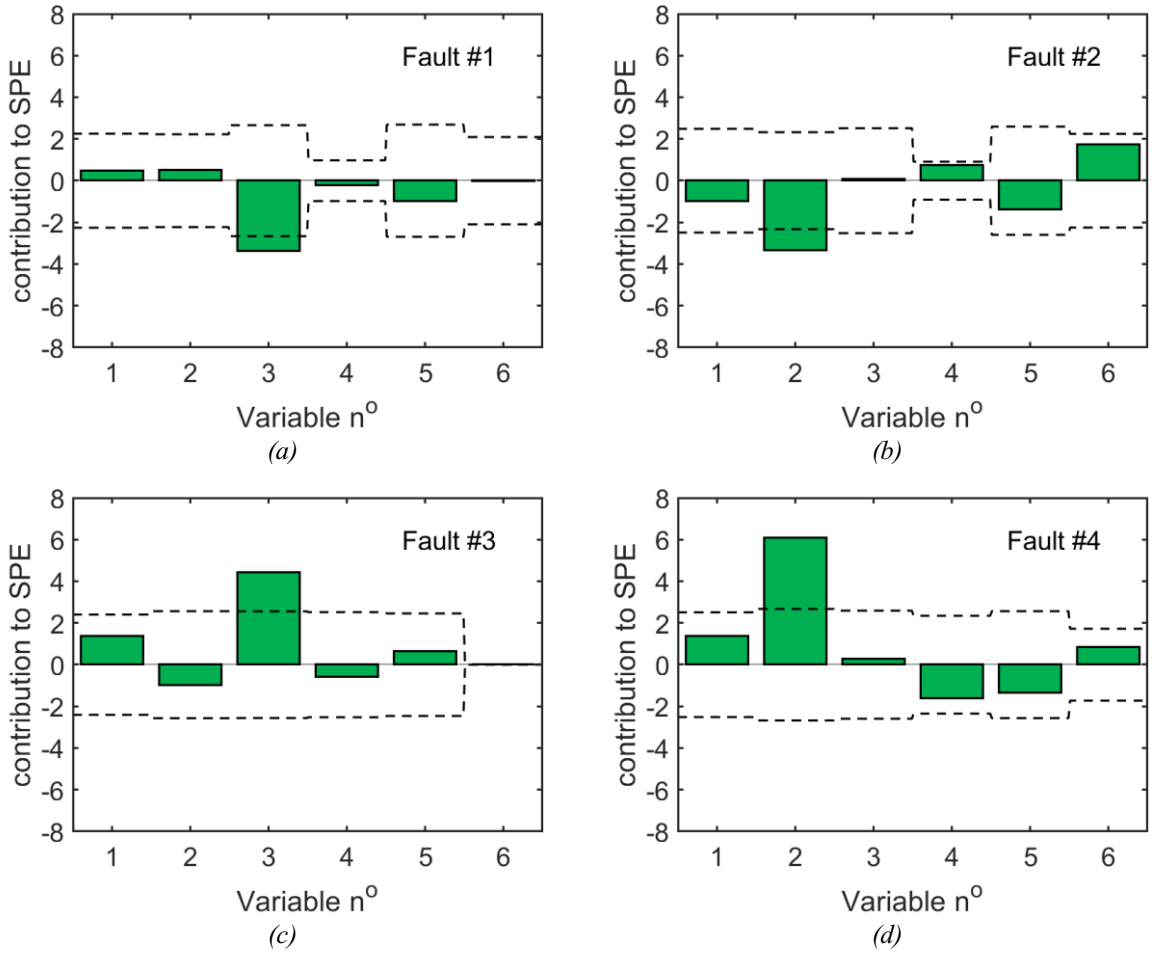


Figure 3.19. *Case study 2: representative contribution plots for the purely data-driven model at the first out-of-control observation for (a) Fault #1, (b) Fault #2, (c) Fault #3, and (d) Fault #4. Variables are numbered as is in Table 3.5. Control limits at 99% confidence are shown as dashed lines.*

3.7. Results and discussion for Case Study 3

3.7.1 Design of the hybrid monitoring model

The hybrid system of Figure 3.1 is implemented to monitor the batches operated in the dryer segments. The measured inputs **u** and outputs **y** for a batch are summarized in Table 3.7 (Variables #1-6).

The EKF available within the gPROMS platform is employed as the state estimator in the knowledge-driven block. Process-model mismatch arises from the fact that the EKF is not aware of the disturbances in the split ratios. The update step of the EKF is performed every 5 s. To improve the EKF robustness, we filter its inputs with a moving average approach. The state estimator is initialized with the true initial states $\mathbf{x}_0$ under NOC and with a null $\mathbf{P}_0$ matrix,

because at the beginning of the process the dryer is empty and there is no uncertainty on this condition. The only non-null entries of the **Q** matrix are the diagonal elements corresponding to the vapor phase enthalpy of each segment, which are set equal to 1. **R** is a diagonal matrix, whose i^{th} element is the variance of the sensor noise for measurement y_i.

We apply the EKF to each of the 100 NOC batches to reconstruct the differential states and the algebraic variables. We arrange dataset **Z** as in Equation 3.22, thus augmenting the set of 5 input measurements **u** and 1 output measurement **y** (Table 3.7, Variables #1-6) with 9 estimated

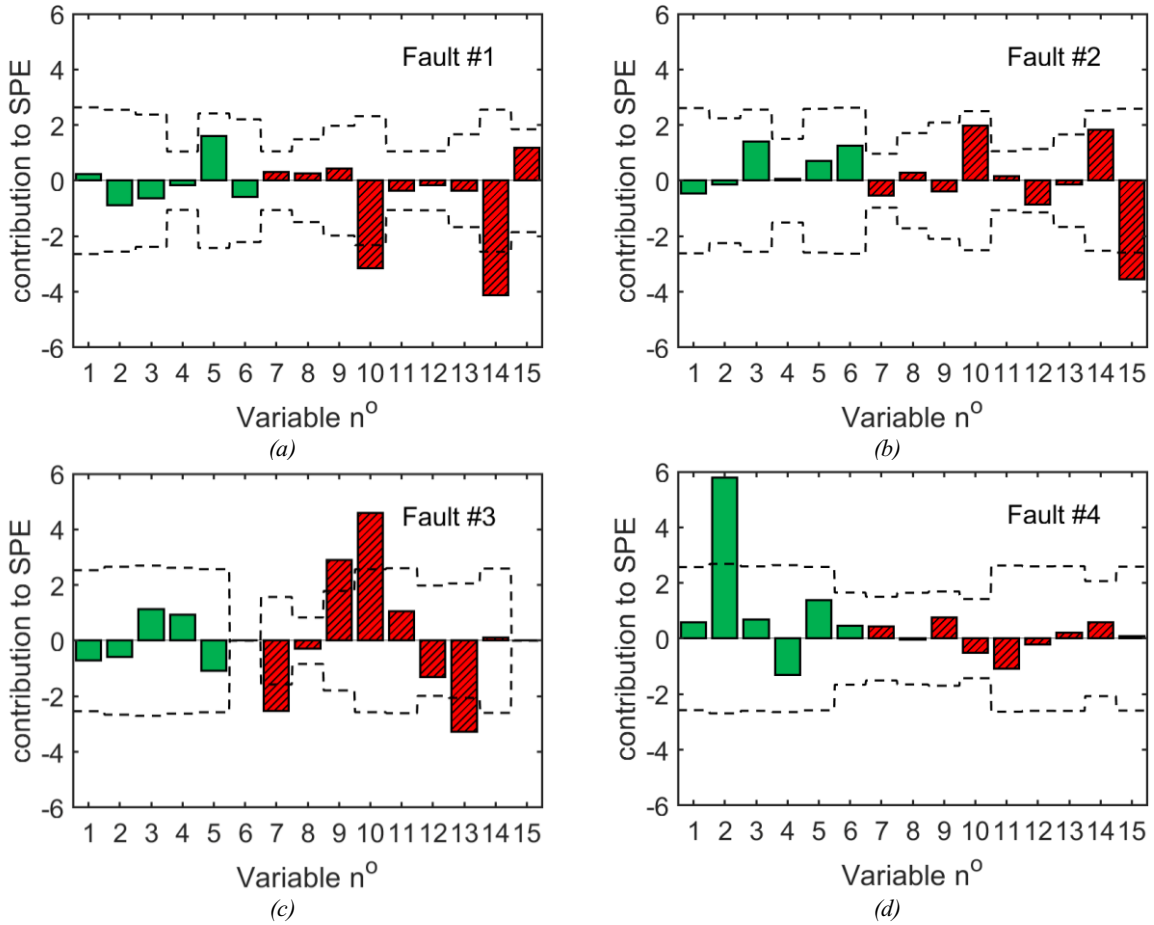


Figure 3.20. *Case study 1: representative contribution plots for the hybrid monitoring model at the first out-of-control observation for (a) Fault #1, (b) Fault #2, (c) Fault #3, and (d) Fault #4. Variables are numbered as in Table 3.5. In all plots, the contributions of field measurements are in green, while those of estimated/reconstructed variables are in red with diagonal lines. Control limits at 99% confidence are shown as dashed lines.*

Table 3.7. *List of variables included in the augmented data matrix of the hybrid monitoring model*

#	Variable	Unit	Variable type
1	Total flowrate of inlet air to the dryer	kg/h	Input ($\in \mathbf{u}$)
2	Relative humidity of inlet air to the dryer	%	Input ($\in \mathbf{u}$)
3	Temperature of inlet air to the dryer	°C	Input ($\in \mathbf{u}$)
4	Total flowrate of inlet granules	kg/h	Input ($\in \mathbf{u}$)
5	Moisture of inlet granules	kg/kg	Input ($\in \mathbf{u}$)
6	Segment temperature	°C	Output ($\in \mathbf{y}$)
7	Heat loss rate	J/s	Estimated ($\in \tilde{\mathbf{x}}$)
8	Mass of air in the segment	kg	Estimated ($\in \tilde{\mathbf{x}}$)
9	Mass of granules in the segment	kg	Estimated ($\in \tilde{\mathbf{x}}$)
10	Temperature of air in the segment	°C	Estimated ($\in \tilde{\mathbf{x}}$)
11	Temperature of granules in the segment	°C	Estimated ($\in \tilde{\mathbf{x}}$)
12	Drying rate	kg/s	Estimated ($\in \tilde{\mathbf{x}}$)
13	Moisture of granules in the segment	kg/kg	Estimated ($\in \tilde{\mathbf{x}}$)
14	Absolute humidity of air in the segment	g/kg	Estimated ($\in \tilde{\mathbf{x}}$)
15	Relative humidity of air in the segment	%	Estimated ($\in \tilde{\mathbf{x}}$)

variables $\tilde{\mathbf{x}}$ (Table 3.7, Variables #7-15) selected to provide additional information on the unmeasured phenomena occurring in the dryer.

We divide each batch of 800 s into time intervals of 10 s, and construct a separate dynamic PCA model at each of the corresponding 81 time points. By trial and error, we select a number of lags $l = 4$ (cfr. Equation 3.21), resulting in a [100×75] $\mathbf{Z}_{dyn}$ (t_k) matrix for each time point. We retain 10 PCs in each of the 81 PCA models, with the explained variance ranging from 75% to 85%.

The hybrid monitoring system can be implemented for real-time applications as the EKF, i.e., the most demanding component of the framework, requires a computational time smaller than the sampling time.

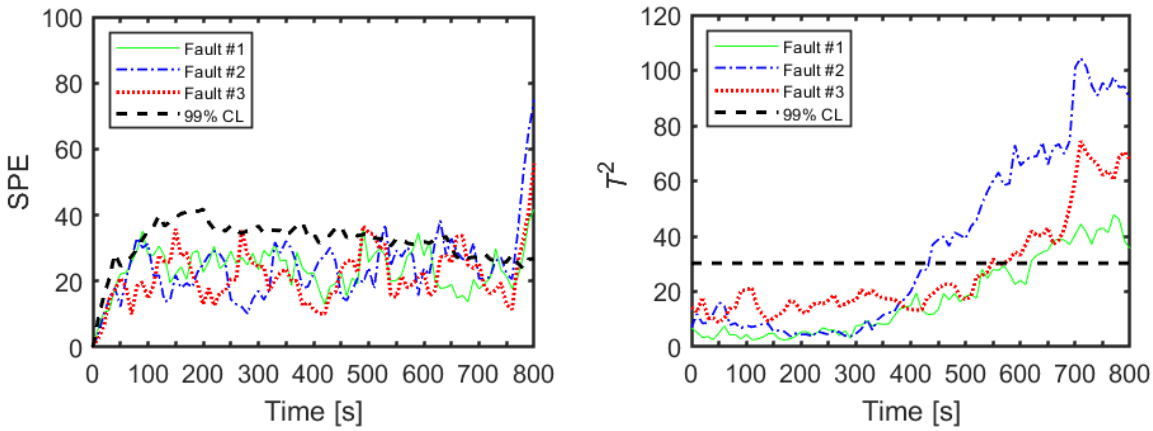
3.7.2 Fault detection and diagnosis

The hybrid model proves capable of detecting all faults (Figure 3.21), with the first out-of-control signal always coming from the T^2 chart. The larger step decrease (–10%, Fault #2) in the air flow to a segment is detected earlier than the ramp fault (Fault #3). The hardest fault to detect is the smaller step (–5%, Fault #1). The SPE does increase sharply at the beginning of the discharging phase ($t = 750$ s), but the faults are already detected well before that on the basis of the T^2 criterion.

Contributions to T^2 shortly after fault detection are similar for all faults, and Figure 3.22 shows an example for Fault #2. Note that most of the out-of-limit contributions to T^2 result not from measurements (green bars), but from estimated values (red bars). Figure 3.22 shows that, for the batch under investigation, the segment temperature is smaller than normal, both for the raw measurement (Variable #6) and for its value as reconstructed by the EKF (Variable #10). In addition, the drying rate (Variable #12) is smaller than normal. Taken together, these two results suggest than the fault is probably due to a reduced energy exchange (low temperature) in the

segment, which is causing a reduction of the drying rate. Since the source of energy for the process is the total hot air feed, one may diagnose the fault as a problem in the air feed received by the segment. This diagnosis is corroborated by the fact that the relative air humidity in the segment (Variable #15) is abnormally high despite the lower drying rate (Variable #12). This indicates that the flux of water being vaporized is picked up by a lower air flow.

To compare the monitoring performance of the proposed hybrid system to the one of a standard PCA approach, we also performed a PCA on a reduced dataset including sensor measurements only (Variables #1-6). Results are not shown for conciseness, but they nevertheless deserve discussion. Though the faults are still detected (with minor delay), the contributions can only point to the measured temperature in the segment (Variable #6) as responsible for the fault. Without the additional information generated by the EKF, diagnosing the fault is much harder and further investigation would therefore be required. Note that the multivariate analysis introduced by the PCA is essential for fault detection. Figure 3.23 shows univariate charts for three of the five variables exhibiting out-of-control contributions in Figure 3.22; the corresponding charts for the other two variables are very similar. We note that Fault #1 does not result in significant deviations from the confidence limits established under NOC. Faults #2 and #3 do result in some deviations, but their magnitude is very small: the strongest deviations, those in the measured temperature (Variable #6), are only ~0.2°C for Fault #2 and ~0.5°C for Fault #3.



(a) *(b)*

Figure 3.21. *Fault detection: (a) SPE monitoring chart and (b) T^2 monitoring chart for the three fault scenarios.*

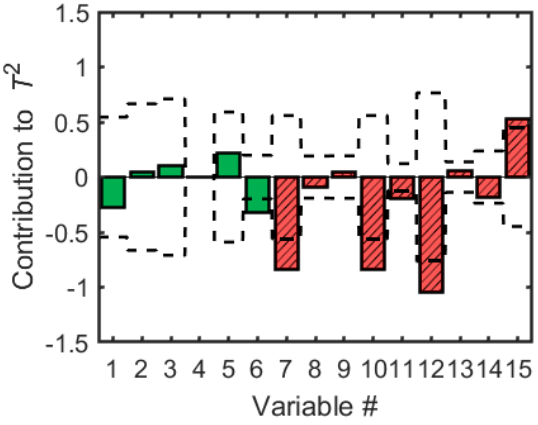


Figure 3.22. *Contributions to T^2 few instants after fault detection for Fault #2. Variables are numbered as in Table 3.7. Green bars refer to measured variables, red bars to estimated variables. Confidence limits at 99% for NOC are shown as black dashed lines.*

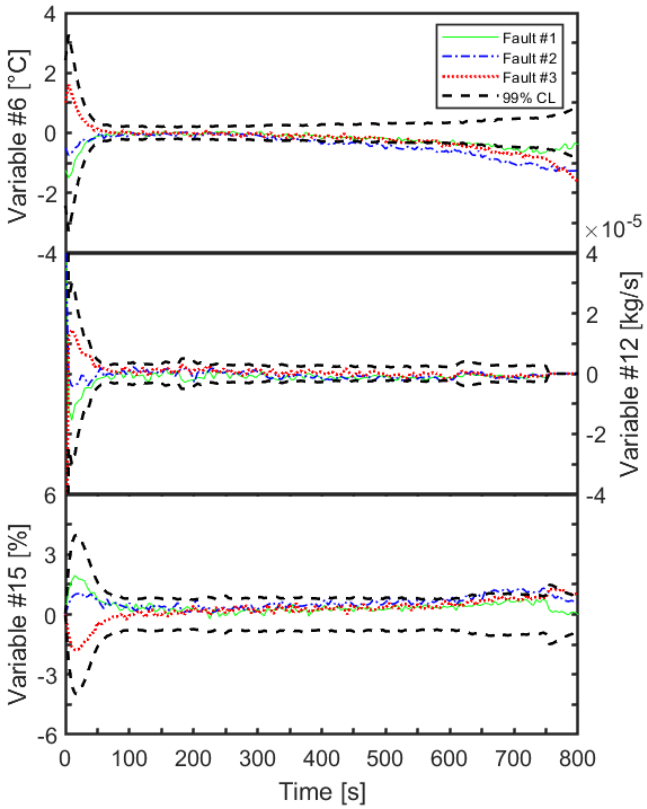


Figure 3.23 *Mean-centered univariate monitoring charts of selected variables displaying high contributions in Fig. 3.22. Variables are numbered as in Table 3.7.*

3.8. Conclusions

In this study, we proposed a novel framework for multivariate process monitoring based on a hybrid modeling approach. Real-time deterministic information about the process is first obtained in a knowledge-driven block from a state estimator in the form of estimated states, reconstructed measurements, and possibly adapted parameters. The information is then passed to a data-driven block, where it is exploited, in conjunction with the available field measurements, by a latent-variable model that accomplishes multivariate fault detection and diagnosis. The design of the two blocks is largely independent, which makes implementation of the proposed methodology easier. The proposed hybrid monitoring framework perfectly responds to the request of the Quality-by-Design initiative to factor the available knowledge on the physics of the process into the control strategy.

We tested the hybrid methodology on three simulated case studies, namely two continuous process and a fed-batch one. It typically allowed for earlier fault detection than standard data-driven and knowledge-driven approaches taken in isolation, even when the state estimator did not perform entirely satisfactorily. In addition, using the hybrid approach significantly facilitated fault diagnosis.

The very satisfactory fault detection performance of the hybrid approach derives from the fact that the estimated states (and possibly the adapted parameters) provide a set of additional variables a fault can leave a footprint on. In most cases, these variables respond to the fault earlier than the measurements, causing an anticipated shift or break of the normal correlation structure of the data, which can be promptly captured and alarmed by the latent-variable model. With respect to fault diagnosis, if a fault manifests itself as an abnormal change in one or more states (or parameters), diagnosing that fault with the hybrid model is generally easier, because the states and parameters straightforwardly point to the inner mechanism that is being impacted by the fault. This enables one to disclose the root-cause of the fault with less ambiguity than can be done using field measurements alone.

The successful performance of the hybrid monitoring system is due to the inclusion of the estimated states (and possibly of the adapted parameters) within a multivariate framework together with the measurements, rather than to the mere implementation of a knowledge-driven component. In fact, traditional knowledge-driven approaches (like innovation sequence monitoring or parametric drift detection) lack the well-known advantages of process monitoring by latent-variable modeling, and they were found not to be able to cope with even mild process-model mismatch.

As for all data-driven methodologies, a word of caution must be mentioned for the hybrid approach in relation to the region wherein the process is run. In fact, the monitoring performance may be compromised if the process is operated away from the region over which the data-driven component was calibrated.

Chapter 4

Intensified continuous filtration-drying of pharmaceuticals: mathematical modeling and design space description*

This Chapter introduces a mathematical model of a novel integrated filter-dryer carousel system, designed for continuously filtering, washing and drying a slurry stream into a crystals cake. For a set of feed conditions and control inputs, the model allows tracking the solvents and impurities content in the cake (critical quality attributes) throughout the whole process. For the isolation of paracetamol from a crystallization slurry, the filtration and drying model parameters are identified through experiments, respectively, on a Nutsche filter and on a thermogravimetric analyzer. The calibrated model is then used for operation design for the isolation of paracetamol from a multi-component slurry, containing a non-volatile impurity, too. The probabilistic design space and the maximum throughput are identified.

4.1 Introduction

Over the past decade, industry and academia have dedicated considerable effort in aiding the shift of pharmaceutical manufacturing from the traditional batch production towards a more continuous operating mode (Burcham et al., 2018; Ierapetritou et al., 2016). Continuous manufacturing allows for greater product consistency and easier scale-up compared to batch processing, at typically lower costs and production times (Fisher et al., 2016). This change of paradigm has also been promoted by regulatory agencies (Lee et al., 2015) within the Quality-by-Design (QbD) initiative (Food & Drug Administration, 2004), aimed at modernizing pharmaceutical manufacturing. Even though many pharmaceutical unit operations are already continuous, few examples of end-to-end continuous pharmaceutical processes are documented in the literature (Domokos et al., 2020; Mascia et al., 2013). It is accepted that active pharmaceutical ingredient (API) purification (also known as isolation), namely, the sequence of API crystallization, filtration, washing and drying operations before of the

*Destro, F., I. Hur, V. Wang, M. Abdi, X. Feng, E. Wood, S. Coleman, P. Firth, A. Barton, M. Barolo, Z. K. Nagy (2021). Mathematical modeling and digital design of an intensified filtration-washing-drying unit for pharmaceutical continuous manufacturing. *Chem. Eng. Sci.*, **244**, 116803.

downstream section, is a main bottleneck for the transition to end-to-end continuous processing (Rantanen and Khinast, 2015). API purification stages are of fundamental importance, since impurities herein retained within the crystals will inevitably be included in the final drug product, subject to tight impurity content requirements (International Council for Harmonisation, 2016). Significant progress has recently been made in the digital design and control of continuous crystallization systems, including experimental validation and implementation (Nagy et al., 2013; Wood et al., 2019b). Nonetheless, significant advancement is still needed in the subsequent solid/liquid separation, namely in continuous filtration, washing and drying at a scale suitable for API manufacturing (Burcham et al., 2018; McWilliams et al., 2018).

Traditional solid/liquid separation design, thoroughly reviewed by Tarleton and Wakeman (2006), is based on empirical knowledge and shortcut approaches. These methods do not rely on detailed process understanding, recommended by the QbD guidelines, nor they consider the strong interactions existing among design and operating parameters of the different purification unit operations. Examples of these coupled dynamics are (but are not limited to) the impact of the size and shape distribution of the crystals produced in the crystallizer on the following steps (Acevedo et al., 2016) and the effect of the washing operating conditions on drying (Ottoboni et al., 2020b). All this considered, the steps of API purification should be considered together for the purposes of process design, optimization and control, with the support of detailed mathematical modeling. However, both experimental and computational studies on integrated continuous API purification are scarce in the literature, and little progress has been made from traditional short-cut design of the individual unit operations. Wibowo and coauthors (2001) use short-cut models for evaluating the effect of the crystal size distribution (CSD) attained during crystallization on the subsequent solid-liquid separation, and for comparing different design alternatives. Empirical correlations and short-cut models are also exploited by Cheng and coauthors (2010) for proposing a workflow of experiments and modeling for managing the impurity content in the entire train of isolation operations. More detailed purification models are developed by Benyahia et al. (2012) for their plant-wide dynamic simulator of a continuous pharmaceutical process. The authors analyze the effect of the washing conditions on the impurity content of the final product, but results are quite limited to their specific application, and the interactions among the operating conditions of the purification steps are not examined in depth. Neither pioneering experimental works on end-to-end continuous pharmaceutical manufacturing (Domokos et al., 2020; Mascia et al., 2013) specifically tackle the current issues in drug isolation, albeit they represent important steps forward in continuous purification. Sen et al. (2013) provide more insights on the integration of purification unit operations and on their effect on downstream operation through hybrid population balances/discrete element method modeling. They develop an integrated flowsheet model combining continuous crystallization, filtration, drying and powder mixing. Subsequent work (Sen et al., 2013b) uses the model for *in-silico* optimization of one process parameter per unit operation (crystallization cooling schedule, filtration pressure gradient, drying gas temperature and mixer rotation speed). However, the authors do not

investigate important phenomena occurring in real plants (for example, cake washing or the dependence of cake porosity on the CSD), nor they discuss the effect of model uncertainty. A study relying on comprehensive mathematical modeling for knowledge-based design of continuous integrated API filtration, washing and drying is still missing in the literature.

In this work, we develop a detailed mechanistic model of a novel intensified continuous filtration-washing carousel, particularly suited for API purification (Liu et al., 2019), which has recently been upgraded to include a drying component. This work follows up on recent experimental and modeling progress that successfully integrates the carousel operation (in filtration-only mode) with upstream continuous crystallizers (Acevedo et al., 2016; Liu et al., 2019; Nagy et al., 2021). The carousel presents a main cylindrical body, composed by multiple stations. In each station, one or more processing steps are carried out batchwise in a cylindrical port. Every fixed time interval (cycle duration), the main body rotates, moving each port to the next station and enabling continuous operation. Traditional short-cut design is particularly challenging for the unit, as the set cycle duration fixes the residence time for each processing step, further increasing the aforementioned coupling among purification operations. Following a QbD approach, we enhance the process understanding by developing standalone dynamic models for the different processing steps, using multi-component macroscopic and microscopic mass, energy and momentum balances (Bird et al., 2006). Then, we assemble together the models into a carousel simulator. We account for the effect of the upstream crystallization by calculating the physical properties of the filtered cake based on the size and shape distributions of the crystals, if they are not available from experimental data. For this purpose, we tailor literature models for mixtures of non-spherical particles, such as porosity models (Yu et al., 1996; Zou and Yu, 1996) and specific cake resistance models (Bourcier et al., 2016), to applications for API purification.

After the modeling activities, we develop a case study on paracetamol (PCM) isolation from a multi-component slurry (also including a non-volatile impurity). We calibrate the filtration and drying components of the carousel model with an experimental campaign. To have more flexibility on the operating conditions, we perform the experiments in separate standalone equipment for filtration (Nutsche filter) and drying (thermogravimetric analyzer, TGA). Then, we use the calibrated model for identifying the probabilistic design space (DS) with a risk-based approach (García-Muñoz et al., 2015), using Monte Carlo simulations to account for model uncertainty. The calculated DS represents a step forward in the design of continuous integrated API purification, as, for every investigated combination of critical process parameters (CPPs) and feed conditions (critical material attributes, CMAs), the DS provides the probability of satisfying the target critical quality attributes (CQAs) of the product.

The rest of the manuscript is organized as follows. The carousel technology is described in Section 4.2, while the developed mathematical models are discussed in Section 4.3. Experimental results and model calibration are presented in Section 4.4, before of DS identification results (Section 4.5) and final conclusions follow.

4.2 The continuous carousel for integrated filtration-washing-drying

The carousel manufacturer (Alconbury Weston Ltd, UK) produces different types of continuous carousels for filtration, washing and drying, varying in size and number of ports. The aim of the carousel is to transform the crystallization slurry into a high purity dry cake of API crystals, containing a concentration of residual solvents and impurities (CQAs) below a certain target. The carousel setup that is the object of this study is a prototype unit (Figure 4.1) with five ports and a maximum hold-up of 10 mL per port, installed in the Crystallization Systems Engineering laboratory at Purdue University. The five ports are anchored to the main cylindrical body in correspondence of five processing stations, all at atmospheric pressure in the top part. The bottom of the first four stations is closed with a filter mesh, connected to a vacuum pump providing the pressure drop ΔP acting as driving force for the processes occurring in each port. The fifth station is instead open at the bottom, to allow for product discharge. Continuous carousel operation is made possible by alternating a processing cycle, during which every port operates batchwise, to a carousel rotation, which moves every port (and the material in it) into the following processing station. At the beginning of every cycle, the slurry from the crystallizer is charged into the first port, and filtration subsequently starts. The crystals in the slurry settle on top of the filter mesh, leading to cake formation, while the liquid is filtered out. At the end of filtration, the liquid is not present anymore on top of the cake, but only inside of the pores. If filtration finishes before the end of the cycle, deliquoring starts, consisting in the mechanical removal of the liquid in the porosity by the effect of ΔP. Meanwhile, cake washing is carried out in the second station, where a controlled volume of wash solvent is released on top of the cake at the beginning of every cycle. Cake washing consists in replacing the mother liquor with a wash solvent, aiming: *i)* to reduce the content of impurities, especially the non-volatile ones that cannot be eliminated through thermal drying and/or *ii)* to enhance cake drying by using a wash solvent more volatile than the mother liquor. After the end of cake washing, which is a fast process, cake deliquoring occurs for the rest of the cycle duration in the second station. Deliquoring is carried out also in the third station, for all the cycle duration. A final drying step with hot air is performed on the cake in the fourth station. Thermal drying is necessary to meet the target CQA since, due to capillarity and mass transfer limitations, cake deliquoring cannot remove the liquid in the cake pores below a certain equilibrium concentration. Non-volatile impurities are not removed through thermal drying, hence their content must be reduced accordingly to quality thresholds earlier in the processing stream. Finally, the cake is discharged from the carousel in the fifth station, through a pneumatic piston. In the unit in Figure 4.1, stations 1-4 do not actually operate all simultaneously, but one cycle of operation in stations 1-3 is alternated to a drying step in station 4, since the unit is a prototype for data collection. For the modeling work in this study, we refer to the operation mode of the production scale carousels, in which all ports operate simultaneously.

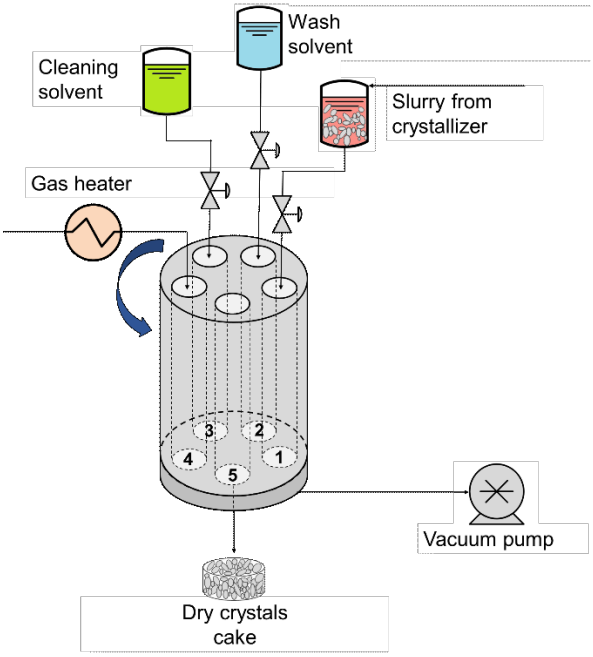


Figure 4.1. *Schematic diagram of the continuous carousel for integrated filtration, washing and drying installed at Purdue University. Stations 1-4 present a filter mesh at the bottom, and are connected to the vacuum pump, while station 5 is open for cake discharge. A cleaning-in-place procedure is automatically triggered when significant mesh fouling is detected, allowing the cleaning solvent stored above Station 3 into the carousel.*

4.3 Carousel mathematical model development

4.3.1 Carousel mathematical model overview

For the purposes of this work, it is of interest to predict the residual solvents and impurities content in the discharged product for a set of given inputs, rather than simulating the processing of different cakes simultaneously in the carousel stations. Hence, the carousel model was developed with the input/output structure sketched in Figure 4.2. A part from the values of CMAs, CPPs and control variables (CVs) and the thermophysical properties of the pure components of the solid, liquid and gas phases, the model needs as inputs the physical properties of the cake (porosity ε and specific resistance α), which can be either measured or calculated from the size and shape distribution of the crystals in the feed slurry (§4.3.2). The outputs of the model are the CQAs, expressed as mass fractions of residual solvents and impurities in the discharged cake.

The carousel model is obtained by assembling together standalone dynamic models of the four processes occurring in the carousel, namely i) slurry filtration, ii) cake deliquoring, iii) cake washing and iv) thermal drying. All the models are coded as MATLAB functions, except for the deliquoring and drying models, implemented in a C environment and interfaced with MATLAB through a C-MEX function to increase the computational speed. An external wrapper calls the standalone models

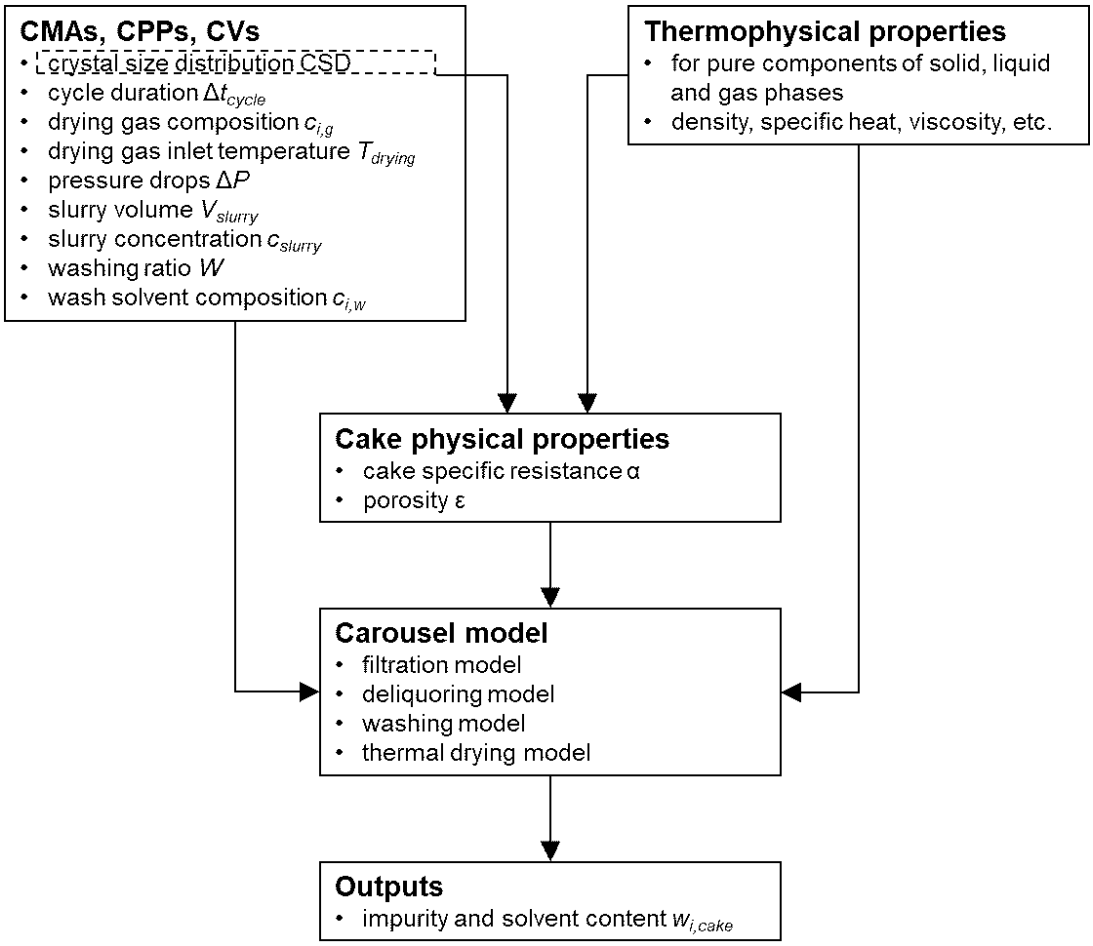


Figure 4.2. *Input/output structure of the carousel mathematical model.*

of the processing steps with the order in which they occur in the carousel (Figure 4.3), as described in Section 4.2. To increase the code robustness, the simulator is implemented for being able to handle the limiting cases in which, due to low cake filterability or adverse combinations of CPPs, CMAs and CVs, filtration and/or washing last for more than one cycle duration. Furthermore, if the cake entering the drying step has the pores still filled with liquid above 20% of the equilibrium deliquoring content, the simulator performs a deliquoring step instead than a drying one, until the liquid content in the pores is low enough to actually allow the gas to flow in the cake.

The models of the processing steps are all based on differential multi-component balances for the reference system in Figure 4.4, to track the dynamic axial profiles of cake composition through the process. The hypotheses common to all the models are: *i)* multiphase system (solid and liquid phases always present, gas phase present in deliquoring and drying), *ii*) multicomponent system (formed by API, one or more solvents, and possibly additional impurities), *iii)* absence of radial gradients of any variable and property (axial symmetry hypothesis), *iv)* homogeneity and isotropy of cake physical properties and of crystals size and shape distribution, and *v)* ideality of liquid and gas phases. All processes are also assumed to be isothermal, except for the thermal drying. The input/output structure

of the standalone models of the processing steps is summarized in Table 4.1. Beside the relevant physical properties, all models need as input ΔP and the initial profiles of cake saturation (S, ratio between volume of liquid in the pores and pores volume) and of components concentration in the liquid phase ($c_{i,l}$ [kg/m^3] for species i), and they provide the final values of the same profiles as outputs, which become inputs to the following model. The only exception is filtration, for which the volume and crystals concentration of the slurry fed to the first station of the carousel has to be provided, instead of the initial cake saturation profile. Additional inputs and outputs specific to only certain models are listed in Table 4.1, and their meanings are discussed in detail upon presentation of the models in the next subsections. From the S and $c_{i,l}$ profiles calculated by the models, the solvents and impurities content (CQAs) in every point of the cake can be monitored all across carousel processing.

In the next subsections, we first briefly introduce the models that, using the crystals size and shape distribution, calculate ϵ and α, which are properties needed by all the other models. Then, we outline the models of the different processing steps.

4.3.2 Cake porosity and specific resistance models

We provide the carousel simulator with a cake properties module, for predicting ϵ and α from the crystals size and shape distribution of the inlet slurry. Although experimental measurements of ϵ and α lead to higher accuracy, the module is useful for connection with an upstream crystallizer model, for integrated crystallization-filtration-washing-drying simulation. The porosity model is based on the work of Yu *et al*. (1996), who proposed (and validated) a modified linear packing model for predicting the porosity of non-spherical particles mixtures. Given the CSD expressed as percentage volume distribution f(d_i), with respect to particle size d_i, the modified linear packing model assumes that one bin of size d_i is the controlling component, determining the porosity of the whole mixture. Under this assumption, the cake specific volume V (with ϵ =1-1/V) is given by:

$$V = \max\{ V_i^C \} , \tag{4.1}$$

where V_i^C is the specific volume calculated assuming that the i-th component of the CSD is the controlling one. Arranging the bins in decreasing size order, V_i^C is obtained with the following set of equations, for every i-th component:

$$V_i^C = V_i + \sum_{j=1}^{i-1}[V_j - (V_j - 1)\, g(r) - Vi]\, x_{v,j} + \sum_{j=i+1}^{n}[V_j - V_j\, f(r) - V_i]\, x_{v,j} \tag{4.2}$$

$$r = \begin{cases} \frac{d_{p,i}}{d_{p,j}} \text{ if } j < i \\ \frac{d_{p,j}}{d_{p,i}} \text{ if } j > i \end{cases} \tag{4.3}$$

$$f(r) = (1-r)^{3.3} + 2.8\, r\, (1-r)^{2.7} \tag{4.4}$$

$$g(r) = (1-r)^2 + 0.4\,r\,(1-r)^{3.7} \tag{4.5}$$

Table 4.1. *Summary of inputs and outputs required by the models of the carousel processing steps. In addition to the listed inputs, the physical properties relevant to each model have to be provided. The symbols t_i and t_f respectively denote the initial and final time instant of a processing step.*

Processing step	I/O	Symbol	Variable
Filtration	Inputs	ΔP	pressure drop
		V_{slurry}	slurry volume
		c_{slurry}	slurry concentration
		$c_{i,l}(t_i, z)$	initial concentration profile of liquid phase components
	Outputs	$\Delta t_{filtration}$	filtration duration
		H_{cake}	cake height
		$V_{filt}(t)$	filtrate volume dynamic profile
		$S(t_f, z)$	final cake saturation profile
		$c_{i,l}(t_f, z)$	final concentration profile of liquid phase components
Deliquoring	Inputs	ΔP	pressure drop
		Δt_{deliq}	deliquoring duration
		$S(t_i, z)$	initial cake saturation profile
		$c_{i,l}(t_f, z)$	initial concentration profile of liquid phase components
	Outputs	$S(t_f, z)$	final cake saturation profile
		$c_{i,l}(t_f, z)$	final concentration profile of liquid phase components
Washing	Inputs	ΔP	pressure drop
		$c_{i,w}$	wash solvent composition
		W	wash ratio
		$S(t_i, z)$	initial cake saturation profile
		$c_{i,l}(t_i, z)$	initial concentrations of liquid phase components
	Outputs	$\Delta t_{washing}$	washing duration
		$S(t_f, z)$	final cake saturation profile
		$c_{i,l}(t_f, z)$	final concentration profile of liquid phase components
Thermal drying	Inputs	ΔP	pressure drop
		Δt_{drying}	drying duration
		T_{drying}	drying gas inlet temperature
		$c_{i,g}$	drying gas inlet composition
		$S(t_i, z)$	initial cake saturation profile
		$c_{i,l}(t_i, z)$	initial concentration profile of liquid phase components
	Outputs	$S(t_f, z)$	final cake saturation profile
		$c_{i,l}(t_f, z)$	final concentration profile of liquid phase components

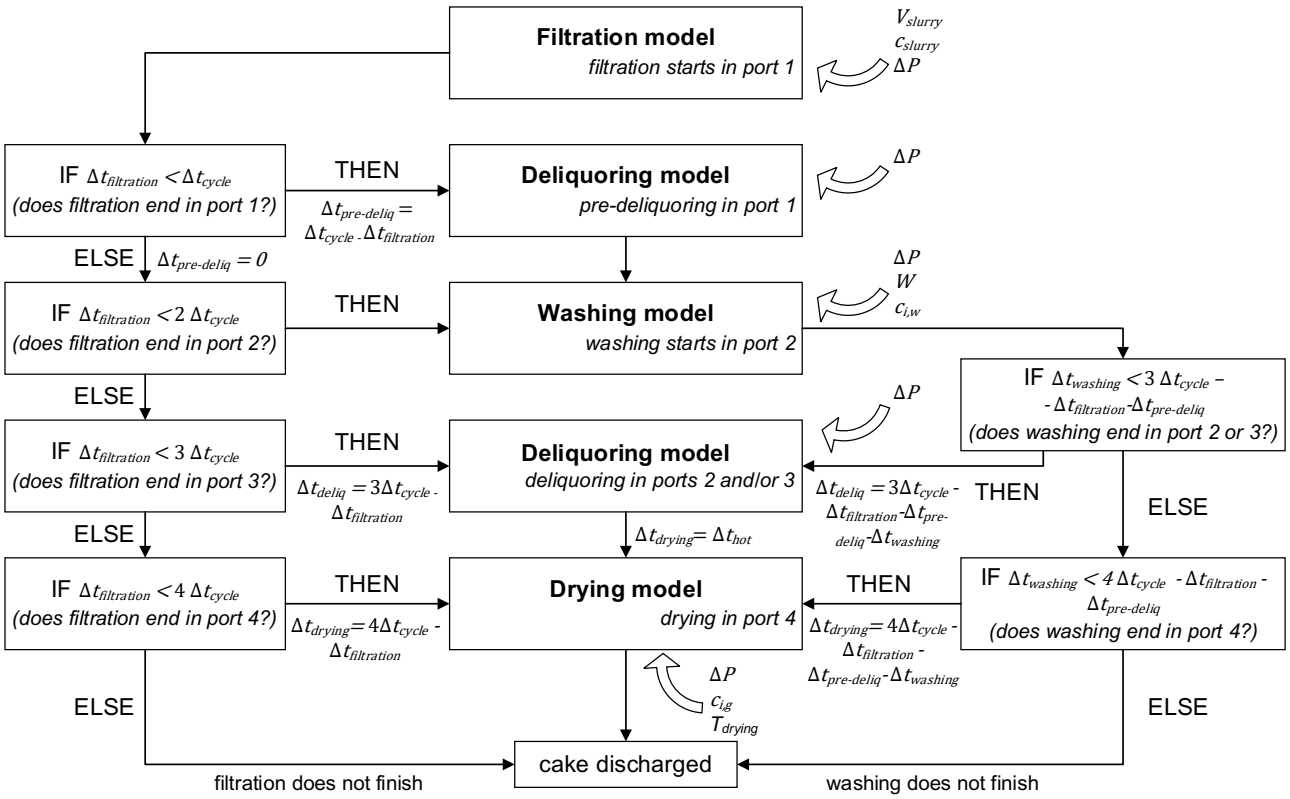


Figure 4.3. *Structure of the carousel model, showing how the models of the processing steps are assembled together to mimic the operation of the physical unit. For conciseness, the only reported inputs are the CMAs, CPPs and CVs, to remark which step they directly affect. The complete inputs and outputs list is in Table 4.1.*

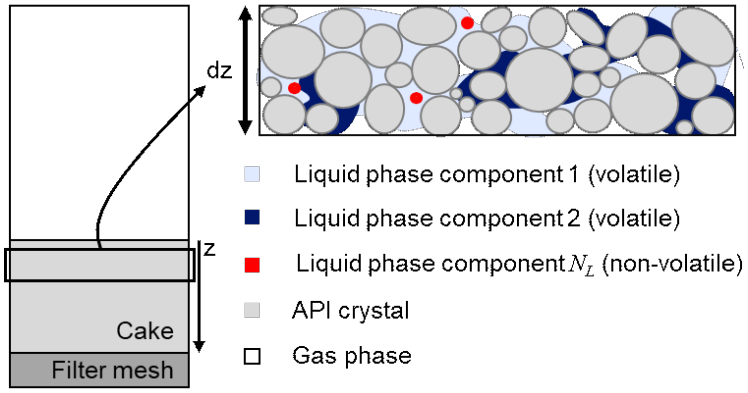


Figure 4.4. *Reference system of the models of the carousel processing steps. In the most general case, solid, liquid and gas phases are all present in a differential volume. The components of the liquid phase can be volatile (vaporized by thermal drying) or not.*

where V_i is the specific volume of a cake composed only by particles of size d_i, $g(r)$ and $f(r)$ are interaction functions between two components of size ratio r, $x_{v,i}$ is the volumetric fraction of the i-th component in the mixture (calculated from $f(d_j)$), and $d_{p,i}$ is the equivalent packing diameter of the i-th component (calculated following Yu *et al.*, 1996). The effect of the shape distribution on ϵ is accounted for by calculating V_i with the relations proposed by Zou and Yu (1996).
We calculate α following a resistance additivity hypothesis (Bourcier et al., 2016), considering the contribution of every component of the particles mixture:

$$\alpha = \sum f(d_i) 180 \frac{1-\epsilon}{\epsilon^3} \frac{1}{\phi_i^2 d_i^2 \rho_s} \tag{4.6}$$

where ρ_s is the solid mass density, and Φ_i is the sphericity of the i-th component of the particle mixture, calculated as:

$$\Phi_i = \frac{\pi^{\frac{1}{3}} (6\, k_V(d_i))^{\frac{2}{3}}}{k_S(d_i)} \tag{4.7}$$

in which $k_V(d_i)$ and $k_S(d_i)$ are, respectively, the surface and the volume shape factors of the crystals of size d_i. We neglect the cake compressibility, as the cakes processed in the carousel are not particularly subject to this phenomenon, due to small cake size and relatively low applied ΔP. Suitable power law relations (Huggins et al., 2019) can be introduced for compressible systems. The cake permeability k is then defined from α as:

$$k = \frac{1}{\alpha_m \rho_s (1-\epsilon)} \tag{4.8}$$

4.3.3 Filtration model

At the end of filtration, the cake is fully saturated (S equal to one in every point of the cake), and the liquid composition is the same as in the fed slurry. The remaining outputs of the filtration model (Table 4.1) to be calculated are the filtration duration $\Delta t_{filtration}$, the final cake height H_{cake} and the dynamic profile of filtrate V_{filt}, which are obtained accordingly to the following discussion.
The driving force for filtration ΔP is equal to the sum of the pressure drops through the cake ΔP_{cake} and of the pressure drops through the filter mesh ΔP_{filter}:

$$\Delta P = \Delta P_{cake}(t) + \Delta P_{filter}(t) \tag{4.9}$$

Factoring the Darcy law (Muskat and Meres, 1936) into Equation 4.9 and rearranging, the instantaneous filtrate flowrate is:

$$\frac{dV_{filt}}{dt} = \frac{\Delta P}{\frac{\alpha \mu_l V_{slurry} c_{slurry}}{A^2 V_{filt,final}} V + \frac{R_m \mu_l}{A}} \tag{4.10}$$

where μ_l is the liquid viscosity, A is the filter cross-section, V_{slurry} is the slurry volume loaded in the carousel at every cycle, c_{slurry} is the crystals concentration in the slurry, R_m is the filter mesh resistance and $V_{filt,\,final}$ is the volume of filtrate at the end of filtration, which is calculated with a mass balance:

$$V_{filt,\,final} = V_{slurry}\left(1-\frac{c_{slurry}}{\rho_s}\left(1+\frac{\epsilon}{1+\epsilon}\right)\right) \tag{4.11}$$

The integration of Equation 4.10 assuming constant ΔP yields the quadratic law for $V_{filt}(t)$:

$$\frac{\alpha\,\mu_l}{2\,A^2}\frac{V_{slurry}\,c_{slurry}}{V_{filt,final}}V_{filt}^2(t)\;+\frac{R_m\mu_l}{A}V_{filt}(t)-\;\Delta P\;t\;=\;0 \tag{4.12}$$

Imposing V_{filt} equal to $V_{filt,final}$ in Equation 4.12, $\Delta t_{filtration}$ is obtained as:

$$\Delta t_{filtration} = \frac{\mu_l\,\alpha\,V_{slurry}\,c_{slurry}\,V_{filt,final}}{2\,A^2\Delta P}+\frac{\mu_l R_m V_{filtrate,final}}{A\,\Delta P} \tag{4.13}$$

At the end of filtration and during all the subsequent carousel processing, H_{cake} and ΔP_{cake} correspond to, respectively:

$$H_{cake}=\frac{V_{slurry}\,c_{slurry}}{\rho_s(1-\epsilon)\,A} \tag{4.14}$$

$$\Delta P_{cake}\left(t\;\geq\;\Delta t_{filtration}\right)=\Delta P\left(1-\frac{R_m}{\alpha\,H_{cake}\,\rho_s(1-\epsilon)+R_m}\right) \tag{4.15}$$

4.3.4 Deliquoring model

Differently from the filtration model, the deliquoring model has to account for the presence of a gas phase, flowing into the cake and replacing the liquid under the action of ΔP. Due to capillary forces, there is a minimum pressure threshold P_b to be applied to the cake to set in motion the liquid in the pores. Because of capillarity, there is also an equilibrium saturation of the cake S_∞, at which deliquoring stops, and further cake desaturation can be carried out only through thermal drying. P_b and S_∞ can be obtained for a system of interest with experimental measurements, or they can be predicted based on the cake physical properties and on ΔP_{cake}. We calculate P_b and S_∞ with literature equations (Wakeman, 1976) for cakes of mono-sized particles, to which we introduce an additive hypothesis to account for the CSD in the cake as follow:

$$P_b\;=\;\sum \mathrm{f}(d_i)\;\frac{4.6\,(1-\epsilon)\,\sigma}{\epsilon\,d_i} \tag{4.16}$$

$$S_\infty = \sum \mathrm{f}(d_i)\;0.155\;(1+0.031\;N_{cap}^{-0.49}) \tag{4.17}$$

where σ is the liquid surface tension and N_{cap} is the capillary number, calculated as:

$$N_{cap} = \frac{\epsilon^3 d_i^2 (\rho_l\, g\, H_{cake} + \Delta P_{cake})}{(1-\epsilon)^2\, H_{cake}\, \sigma} \tag{4.18}$$

The local velocity of the liquid u_l is given by the Darcy law for multiphase flow (Muskat and Meres, 1936):

$$u_l = -\frac{k\, k_{rl}}{\mu_l} \frac{\mathrm{d}P_l}{\mathrm{d}z} \tag{4.19}$$

in which P_l is the local pressure of the liquid and k_{rl} is the liquid relative cake permeability. We calculate k_{rl} with the following relations (Wakeman, 1979):

$$k_{rl} = k\, S_R^{2+3\lambda} \tag{4.20}$$

$$S_R = \frac{S - S_\infty}{1 - S_\infty} = \left(\frac{P_b}{P_g - P_l}\right)^{\lambda} \tag{4.21}$$

where λ is the pore size distribution parameter (usually assumed equal to 5), S_R is the local reduced saturation and P_g is the local gas pressure. In Equation 4.19, we calculate P_l through Equation 4.21, assuming linear and constant gas pressure gradient during the process, with total gas pressure drop through the cake equal to ΔP_{cake} (Equation 4.15). With this assumption, we neglect the initial deliquoring transient for a fully saturated cake, during which there is no gas at the outlet of the bed, to avoid solving the gas mass balance and the Darcy law for the gas phase. This initial transient is very quick and it has little impact on the process.

In the academic and industrial practice, deliquoring is modeled and designed through design charts (Tarleton and Wakeman, 2006). However, design charts provide only the average values of cake saturation and liquid composition within the cake, neglecting axial gradients. Considering the axial gradients in the model is important not only for calculating the cake composition accurately, but also for monitoring regions of excessively low saturation in the cake, which might lead to cracking. For these reasons, in this work we develop a detailed deliquoring model based on partial differential equations, starting our derivation from Wakeman's early results (Wakeman, 1979).

Under the hypotheses stated in §4.3.1 and in absence of mass transfer between liquid and gas, the liquid phase mass balance reads:

$$\frac{\partial S}{\partial t} = -\frac{1}{\epsilon}\frac{\partial u_l}{\partial z} \tag{4.22}$$

We account for initial gradients in the liquid composition by including in the model the liquid phase species mass balances, assuming absence of species diffusion in the liquid:

$$\frac{\partial c_{i,l}}{\partial t} = -\frac{u_l}{\epsilon S}\frac{\partial c_{i,l}}{\partial z}, \text{for } i = 1, \dots, N_L \tag{4.23}$$

The initial conditions are inputs of the model (Table 4.1). The boundary conditions are:

$$\begin{cases} S_R(t, z=0) = 0, \ \forall t > 0 \\ \frac{\partial c_{i,l}\,(t, z=0)}{\partial z} = 0, \ \ \forall t > 0 \end{cases} \tag{4.24}$$

The model of Equations 4.16-24 presents $1+N_L$ partial differential equations (PDEs), which we semi-discretize with a high-resolution finite volume approach (Van Leer, 1974) along z. The resulting ODEs are integrated with MATLAB's ode23s solver. To increase numerical stability, we reformulate the model in dimensionless form, as suggested by Wakeman (1979).

4.3.5 Washing model

During the washing of a cake that did not undergo a pre-deliquoring step, full saturation is always maintained. In this situation, the dynamic profiles of $c_{i,l}$ are given by the following species mass balances:

$$\frac{\partial c_{i,l}}{\partial t} = -\frac{u_l}{\epsilon}\frac{\partial c_{i,l}}{\partial z} - \frac{\partial}{\partial z}\left(D_{i,ax,l}\frac{\partial\, c_{i,l}}{\partial z}\right) \ \text{for } i = 1, \dots, N_L \tag{4.25}$$

where $D_{i,ax,l}$ is the axial diffusion coefficient for species i in the liquid, and u_l is calculated with Equation 4.19, assuming constant liquid pressure gradient with total liquid pressure drops through the cake equal to ΔP_{cake} (Equation 4.15). For all the N_L components, $D_{i,ax,l}$ can be obtained from experimental data, or it can be calculated with (Wakeman and Tarleton, 2005):

$$\begin{cases} D_{i,ax,l}/D_{i,l} = \frac{1}{\sqrt{2}} & if\ \text{Re Sc} < 1 \\ D_{i,ax,l}/D_{i,l} = \frac{1}{\sqrt{2}} + 55.5\ (\text{Re Sc})^{0.96} & if\ \text{Re Sc} > 1\ and\ L < 10\ cm \\ D_{i,ax,l}/D_{i,l} = \frac{1}{\sqrt{2}} + 1.75\ \text{Re Sc} & if\ \text{Re Sc} > 1\ and\ L > 10\ cm \end{cases} \tag{4.26}$$

where $D_{i,l}$ is the molecular diffusion coefficient of species i in the liquid and $ReSc$ is the product of the Reynolds and of the Schmidt numbers, which we compute using the additive hypothesis to account for the CSD:

$$ReSc \ = \ \sum \mathrm{f}(d_i)\ u_l \frac{d_i}{D_{i,l}} \tag{4.27}$$

The initial profiles of $c_{i,l}$ are inputs of the model. The boundary conditions are:

$$c_{i,l}\,(t, z=0) = c_{i,w}, \ \forall t > 0, \ \text{for } i = 1, \dots, N_L \tag{4.28}$$

where $c_{i,w}$ is the concentration of species i in the wash liquid.

Popular short-cut methods employ pre-computed solutions of Equation 4.25 (Tarleton and Wakeman, 2006) to calculate the average cake composition and/or the composition at the bottom of the cake. These washing design charts are available only for certain values of $D_{i,ax,l}$, and interpolation is resorted to if the solution at intermediate values is needed. In this work, we calculate the whole

dynamic axial profiles of solvents and impurities concentrations in the liquid phase of the cake during washing through the analytical solution of Equations 4.25 (Lapidus and Amundson, 1952), which does not require any interpolation exercise. As usually done for washing, instead of time we use as independent variable the washing ratio W (ratio between the used volume of wash solvent and the volume of pores of the cake):

$$\phi_i(W,\mathrm{z}) = \frac{c_{i,l}(W,\ z) - c_{i,l}(W=0,\ z)}{c_{i,w} - c_{i,l}(W=0,\ z)} = 0.5\left\{\mathrm{erfc}\left(\frac{z/H_{cake} - \lambda_{ads}W}{2\sqrt{\lambda_{ads}W}}\sqrt{\frac{u_l H_{cake}}{\epsilon D_{i,ax,l}}}\right) + \right.$$

$$\left.\exp\left(\frac{u_l z}{\epsilon D_{i,ax,l}}\right)\ \mathrm{erfc}\left(\frac{z/H_{cake} + \lambda_{ads}W}{2\sqrt{\lambda_{ads}W}}\sqrt{\frac{u_l H_{cake}}{\epsilon D_{i,ax,l}}}\right)\right\} \tag{4.29}$$

in which $\lambda_{i,ads}$ accounts for sorption effects. Note that the washing duration $\Delta t_{washing}$ is immediately obtained from the definition of W (input of the carousel model), using the (known) volume of cake pores and Equation 4.19.

Simulating the washing of a cake that underwent pre-deliquoring requires solving additional equations, accounting for the coupling between cake resaturation and dispersion phenomena. A comprehensive model exist in literature for this process (Wakeman and Attwood, 1990), from which Wakeman and Tarleton (2005) derived the following relation, used in this work to compute the effect of pre-deliquoring on washing:

$$W_{corr} = W + 1.51\left(1 - S_{avg}\right)\exp\left(-1.56\ \phi_i(W,\mathrm{z}=L)\right) -$$

$$7.4\left(1 - S_{avg}^2\right)\exp(-1.72\ \phi_i(W,\mathrm{z}=L)) \tag{4.30}$$

where W_{corr} is a corrected washing ratio, accounting for pre-deliquoring, W is the washing ratio for the saturated cake, S_{avg} is the average initial saturation, and ϕ_i is calculated with Equation 4.29.

The outlined washing model neglects changes of dry cake mass and CSD due to partial dissolution of the crystals into the wash solvent. This simplifying assumption holds true for most applications, since the wash solvent flow through the cake is typically fast, and the volume of wash solvent employed in common industrial operation is a fraction of the volume of the cake. Dissolution during washing should instead be accounted for in systems where the crystals are highly soluble in the wash solvent, and in which large volumes of wash solvent have to be employed due to difficult impurities removal. This scenario is out of the purposes of this work.

4.3.6 Thermal drying model

The model for convective cake drying is based on the following equations. The local drying rate $\dot{m}_i^{\mathrm{L}\rightarrow\mathrm{G}}$ [kg/(m^3 s)] for the $N_{L,vol}$ volatile species still present in the liquid phase is (Burgschweiger and Tsotsas, 2002):

$$\dot{m}_i^{\mathrm{L}\to\mathrm{G}} = \begin{cases} h_{M,i}\, a\left(P_{i,sat} - P_{i,g}\right)\eta_i & if\ P_{i,sat} > P_{i,g} \\ 0 & if\ P_{i,sat} \le P_{i,g} \end{cases}, \text{ for } i = 1, \dots, N_{L,vol} \tag{4.31}$$

where $h_{M,i}$ is the mass transfer coefficient, calculated with correlations or from experimental data, a is the cake specific surface, either computed as $a = 6/d_p$ (where d_p is the Sauter diameter from the CSD) or measured, and, for species i, $P_{i,sat}$ is the saturation pressure, $P_{i,g}$ is the partial pressure, and η_i is a factor accounting for mass transfer limitations occurring, mostly because of capillarity, when the mass fraction of i in the cake ($w_{i,\,cake}$) becomes lower than a critical value $w_{i,\,cake}^{crit}$ (falling rate period). η_i is instead equal to one when $w_{i,\,cake}$ is greater than $w_{i,\,cake}^{crit}$ (constant rate period). In the falling rate period, η_i is typically linearly or quadratically dependent on $w_{i,\,cake}$, and it should be estimated with experiments. More than one falling rate period (and the corresponding critical solvent content) can be identified for certain systems. When the equilibrium mass fraction of i ($w_{i,\,cake}^{eq}$) is reached, η_i drops to zero. The species mass balance with respect to the concentration of i in the cake $c_{i,cake}$ is:

$$\frac{\partial}{\partial t} c_{i,cake} = -\dot{m}_i^{\mathrm{L}\to\mathrm{G}}, \text{for } i = 1, \dots, N_{L,vol} \tag{4.32}$$

Equation 4.32 is not solved for non-volatile species, as their concentration in the cake does not vary during drying. In addition, in the equation both $c_{i,cake}$ and $\dot{m}_i^{\mathrm{L}\to\mathrm{G}}$ are function of time and z. The local cake saturation is related to the species concentrations in the cake with:

$$S = \frac{\sum_{i=1}^{N_L} c_{i,cake}/\rho_{i,l}}{\epsilon}, \text{for } i = 1, \dots, N_L \tag{4.33}$$

where $\rho_{i,l}$ is the liquid density of pure i. The species mass balances in the gas phase for the volatile solvents and impurities read, in terms of mass fraction $w_{i,g}$:

$$\rho_g \epsilon (1 - S) \frac{\partial}{\partial t} w_{i,g} = -\rho_g u_g \frac{\partial}{\partial z} w_{i,g} + \dot{m}_i^{L\to G}, \qquad \text{for } i = 1, \dots, N_{L,vol} \tag{4.34}$$

where ρ_g is the density of the gas, and u_g is the gas velocity, calculated with the Darcy law for mono-phase gas flow in a porous medium (Muskat and Meres, 1936).

Inter-phase energy transfer is very fast for the process, as found from experiments on the carousel and literature correlations (Bird et al., 2006). Hence, we assume local thermal equilibrium among phases for the development of the differential energy balance:

$$\left(\rho_s c_{p,s}(1-\epsilon) + \rho_l c_{p,l} \epsilon S + \rho_g c_{p,g} \epsilon (1-S)\right) \frac{\partial T}{\partial t} = -\sum_i \left(\dot{m}_i^{\mathrm{L}\to\mathrm{G}} \lambda_i\right) - u_g\, c_{p,g} \rho_g \frac{\partial T}{\partial z} + \dot{Q} \tag{4.35}$$

where ρ_l is the liquid phase density, $c_{p,g}$ is the gas phase specific heat, $c_{p,l}$ is the liquid phase specific heat, $c_{p,s}$ is the solid phase specific heat, T is the local temperature, λ_i is the latent heat of vaporization of species i and $\dot{Q}$ is the heat exchange with the environment, namely the heat loss through the dryer

walls (assumed equal to zero for the purposes of this study).
The saturation and composition initial conditions are inputs of the model. The initial temperature corresponds to the room value for every z. The boundary conditions are given by the drying gas inlet composition and the drying gas inlet temperature T_{drying}, both inputs of the carousel model. The drying model of Equations 4.31-35 presents $1+N_{L,vol}$ PDEs, which we semi-discretize along z with a first-order upwind scheme. The resulting ODEs are integrated with MATLAB's ode15s solver.

4.4 Paracetamol case study: experimental results and model calibration

We develop a case study on API isolation from a slurry containing PCM (the API), isopropyl alcohol (IPA, the mother liquor), and a non-volatile impurity. Pure ethanol (EtOH) is selected as wash solvent. The case study involves two steps: *i)* carousel model calibration, and *ii)* probabilistic DS identification through the calibrated carousel model. The former step is discussed in this section, the latter in the following one.
Table 4.2 reports the thermophysical properties used in the case study for the solid phase and for the pure components of the liquid phase. The transport properties of pure components of the liquid phase are obtained from temperature-dependent relations: molecular diffusion coefficients are assumed to be the same as in water, and calculated from Yaws (2009), while viscosities are calculated according to Yaws and Le Xuan Dang (2009). Following modeling assumption *v)* (§3.1), ideal solution and ideal gas mixing rules are used, respectively, for the liquid and gas phases. Thermophysical properties of the pure components of the gas phase and the saturation pressures were calculated with temperature-dependent relations from Perry's Chemical Engineer's Handbook (Green and Southard, 2019). Note that in the simulation studies we consider the presence of a non-volatile impurity as a generic component in the slurry to demonstrate the carousel simulator capabilities to deal with multi-component systems and non-volatile impurities. However, in the experimental campaign carried out for this study, no impurities were present in the slurry. Physical and transport properties of the non-volatile impurity in the simulations are assumed to be the same as IPA's (Table 4.2), the mother liquor in which it is dissolved, except for the saturation pressure, set to zero at all temperatures. It is also assumed that the non-volatile impurity does not adsorb in the solid (λ_{ads} equal to one).
We calibrate the carousel model on the system through filtration and drying experiments. Experimental validation of the washing component of the carousel model is not carried out in this study, as washing is not a critical step under the introduced assumption of absence of adsorption of the non-volatile impurity. Hence, the dispersion coefficients are directly calculated through Equations 4.26-27.
We use standard equipment for filtration and drying experiments: a Nutsche filter (a pocket filter produced by Alconbury Weston Ltd) and a TGA (Perkin Elmer TGA 4000). This experimental procedure allows for greater flexibility and easier data collection, compared to experiments directly

Table 4.2. *Paracetamol case study: thermophysical properties of the solid phase and of the pure components of the liquid phase.*

Symbol	Parameter	UOM	Value
$c_{p,EtOH,l}$	Pure liquid specific heat - EtOH	[kJ/kg K]	2.570
$c_{p,impurity,l}$	Pure liquid specific heat – impurity	[kJ/kg K]	2.667
$c_{p,IPA,l}$	Pure liquid specific heat - IPA	[kJ/kg K]	2.667
$c_{p,s}$	Solid specific heat	[kJ/kg K]	1.300
Φ_i	Crystals sphericity	[-]	1
λ_{EtOH}	Latent heat of vaporization – EtOH	[kJ/kg]	846
λ_{IPA}	Latent heat of vaporization – IPA	[kJ/kg]	664
$\lambda_{EtOH,ads}$	Washing sorption coefficient – EtOH	[-]	1
$\lambda_{impurity,ads}$	Washing sorption coefficient – non-volatile impurity	[-]	1
$\lambda_{IPA,ads}$	Washing sorption coefficient – IPA	[-]	1
$\rho_{EtOH,l}$	Pure liquid density – EtOH	[kg/m^3]	842
$\rho_{impurity,l}$	Pure liquid density – non-volatile impurity	[kg/m^3]	786
$\rho_{IPA,l}$	Pure liquid density – IPA	[kg/m^3]	786
ρ_s	Solid mass density	[kg/m^3]	1290
σ	Liquid surface tension	[N/m]	22.39

performed on the carousel. On the other hand, filtration occurs in the Nutsche filter and in the carousel ports in the same way, and the drying kinetics in the TGA and in the carousel are the same. Actually, a slightly conservative value is obtained with the TGA, as the higher gas velocity during carousel drying can enhance the kinetics. The CSD of the PCM crystals used for the experiments, measured with a Mastersizer 3000 by Malvern Panalytical (UK), is reported in Figure 4.5.

A PCM/EtOH slurry with crystals concentration of 25%w is used for the filtration experiments, to estimate α and R_m (note that the parameters are not influenced by the type of mother liquor). The Nutsche filter presents a mesh of the same material and pores size (20 μm) of the meshes installed in the carousel, but with a larger diameter (5 cm). An overhead pressure provides the ΔP necessary for filtration. We carried out two experiments, each one consisting of two filtration runs. The first run of each experiment (Run #1 and #3) is performed at ΔP=0.1 bar, while the second one (Run #2 and #4) at ΔP=0.7 bar. The filter mesh is cleaned with pure EtOH after Run #2, but in between the two runs of each experiments (Runs #1-2 and Runs #3-4) the cake is unloaded without further cleaning. This choice aims to reproduce carousel operation, where the meshes are cleaned with an automatic cleaning-in-place procedure (Liu et al., 2019) only when significant fouling is detected, instead than at the end of each cycle. The regressed parameters (Table 4.3), obtained through maximum likelihood estimation from the measured filtrate volume profiles, display small estimation uncertainty. Overall, the cake demonstrates high filterability properties. The good fitting of the runs (Figure 4.6) at different ΔP using the same α verifies the low compressibility assumption used for the modeling. Since the degree of fouling in each run is different, independent filter mesh resistance parameters are considered: $R_{m,1}$, $R_{m,2}$, $R_{m,3}$ and $R_{m,4}$. As expected, the regressed $R_{m,1}$ and $R_{m,3}$ are lower than $R_{m,2}$ and $R_{m,4}$, proving that part of the fine crystals in the CSD (Figure 4.5) remain trapped in the mesh,

increasing its resistance. The obtained cakes present an average ε of 0.36. Using the PCM CSD, α and ε predicted by the models reported in §4.3.2 show good agreement with the experimental findings (Φ_i equal to one is used in the calculations), as reported in Table 4.3.

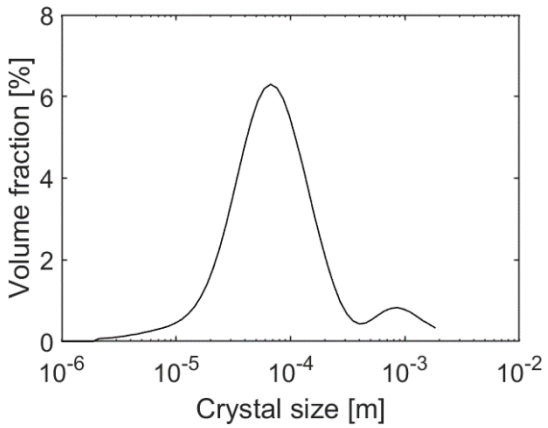


Figure 4.5. *Measured CSD of the PCM used for the filtration and drying experiments.*

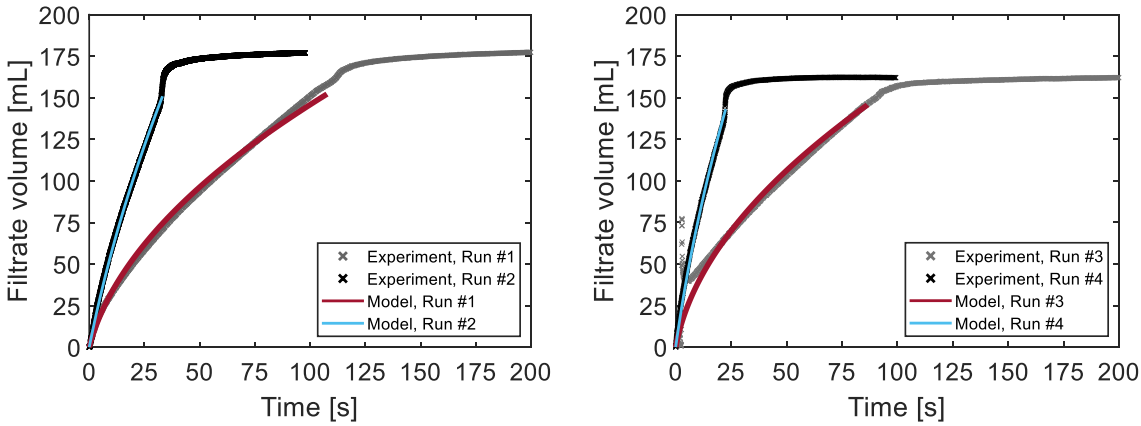


Figure 4.6. *Filtration experiments: measured filtrate volume profile vs model fitting. Left: Runs #1-2. Right: Runs #3-4. ΔP=0.1 bar for Runs #1 and #3, while ΔP=0.7 bar for Runs #2 and #4.*

Table 4.3. *Filtration experiments: estimated parameters and relevant standard deviations. The values of α and ε predicted by the models in §4.3.2 are reported for comparison.*

Symbol	Parameter	UOM	Estimated value	Estimation std. dev.	Model prediction
$R_{m,1}$	Filter mesh resistance, run #1	[1/m]	4.18E9	1.45E6	-
$R_{m,2}$	Filter mesh resistance, run #2	[1/m]	22.48E9	3.30E7	-
$R_{m,3}$	Filter mesh resistance, run #3	[1/m]	2.98E9	3.04E6	-
$R_{m,4}$	Filter mesh resistance, run #4	[1/m]	15.58E9	1.03E7	-

α	Specific cake resistance	[m/kg]	1.04E9	5.13E5	3.52E9
ε	Cake porosity	[-]	3.6E-1	1.0E-2	3.5E-1

Due to the quick deliquoring transient reported in Figure 4.6, deliquoring parameters are not estimated, as additional specific experiments are required for this purpose (Ripperger, 2013) and deliquoring is not a critical step for this system.

The only parameters to be estimated for the drying model are those appearing in Equation 4.31, the drying rate equation: $h_{M,i}$ and the relation $\eta_i(w_{i,cake})$, for *i* corresponding to EtOH and IPA. We only estimate the parameters for EtOH, assuming that they are equal to those for IPA, considering that *i)* they are similar solvents, *ii)* the parameters mainly depend on the cake properties, rather than on the solvent type, and *iii)* in the case study, the liquid phase of the cakes entering the dryer is mainly composed by EtOH. The wet cakes obtained at the end of the filtration experiments are used for the drying experiments, after resaturation with pure EtOH. Five TGA experiments were carried out, each one at a different temperature (30 °C, 40 °C, 50 °C, 60 °C and 70 °C), to assess the temperature dependence of the parameters. For each experiment, a pan containing a small (~20-40 mg) piece of wet cake is inserted in the TGA chamber, which is subsequently heated up. When the designed temperature for a given experiment is reached, the TGA starts recording the sample weight, keeping at the same time the chamber temperature (approximately) constant and uniform. The drying rates for the experiments (Figure 4.7) are obtained as time derivate of the measured weight, normalized by the sample volume. Two critical solvent contents are identified for all the experiments: $w_{EtOH,\,cake}^{crit,1}$ equal to 0.11, at which the drying rate switches from the constant period to a linearly falling period, and $w_{EtOH,\,cake}^{crit,2}$ equal to 0.018, below which a higher order decay is followed. The dry cakes obtained at the end of the experiments were left to dry in an oven at 40°C for 48 h. No additional change of weight was registered, leading to the conclusion that $w_{EtOH,\,cake}^{eq}$ is equal to zero.

For parameter estimation, we assume that the volatilized EtOH immediately leaves the cake during the experiment, due to the small sample size. Hence, $\dot{m}_{\mathrm{EtOH}}$ is uniform (but not constant) in the sample, and during the constant rate period (from Equation 4.31):

$$h_{M,\mathrm{EtOH}}\, a = \frac{\dot{m}_{\mathrm{EtOH}}}{P_{\mathrm{EtOH},sat}}^{\mathrm{L}\rightarrow\mathrm{G}} \tag{4.36}$$

As first attempt, we assume that $h_{M,\mathrm{EtOH}}$ and the relation $\eta_{\mathrm{EtOH}}(w_{\mathrm{EtOH},cake})$ do not depend on temperature, as found for the critical solvent contents. Using the collected data in the constant rate periods of all the experiments and Equation 4.36 in a maximum likelihood estimation framework, the specific mass transfer coefficient $h_{M,EtOH}\, a$ results equal to 2.25E-5 kg m^{-3} s^{-1} Pa^{-1} (standard deviation = 5.89E-8 kg m^{-3} s^{-1} Pa^{-1}). Independent values of $h_{M,EtOH}$ and a are not identified, since they are not required by any of the models. Then, the following $\eta_{\mathrm{EtOH}}(w_{\mathrm{EtOH},cake})$ polynomials are regressed for the falling rate periods:

$$\begin{cases} \eta_{EtOH}=3.37\ w_{EtOH,cake}+0.63 & \text{if } 0.0118< w_{EtOH,cake}<0.110 \\ \eta_{EtOH}=-3.57E7\ w_{EtOH,cake}^{4}+1.71E6\ w_{EtOH,cake}^{3} & \\ \quad -3.02\ w_{EtOH,cake}^{2}+2.34E2\ w_{EtOH,cake}+8.60E-3 & \text{if } 0<w_{EtOH,cake}<0.018 \end{cases} \quad (4.37)$$

The drying rate calculated with Equation 4.31 using the regressed $h_{M,EtOH}\,a$ and Equation 4.37 shows good agreement with experimental results (Figure 4.7). Hence, the increasing drying rate at higher temperatures is mainly due to larger saturation pressures, instead than to changes in the mass transfer phenomena. The experiment at 70 °C shows a slightly worse fitting when compared to the other ones, especially in the constant rate period. This may be due to that fact that the small temperature gradients developing within the TGA chamber due to the proximity with the EtOH boiling temperature are not considered in the calculation of $P_{EtOH,sat}$ (very steep at 70 °C) for the model prediction in Figure 4.7. This issue does not occur when the drying model is used for carousel operation simulation, for which temperature gradients are not neglected.

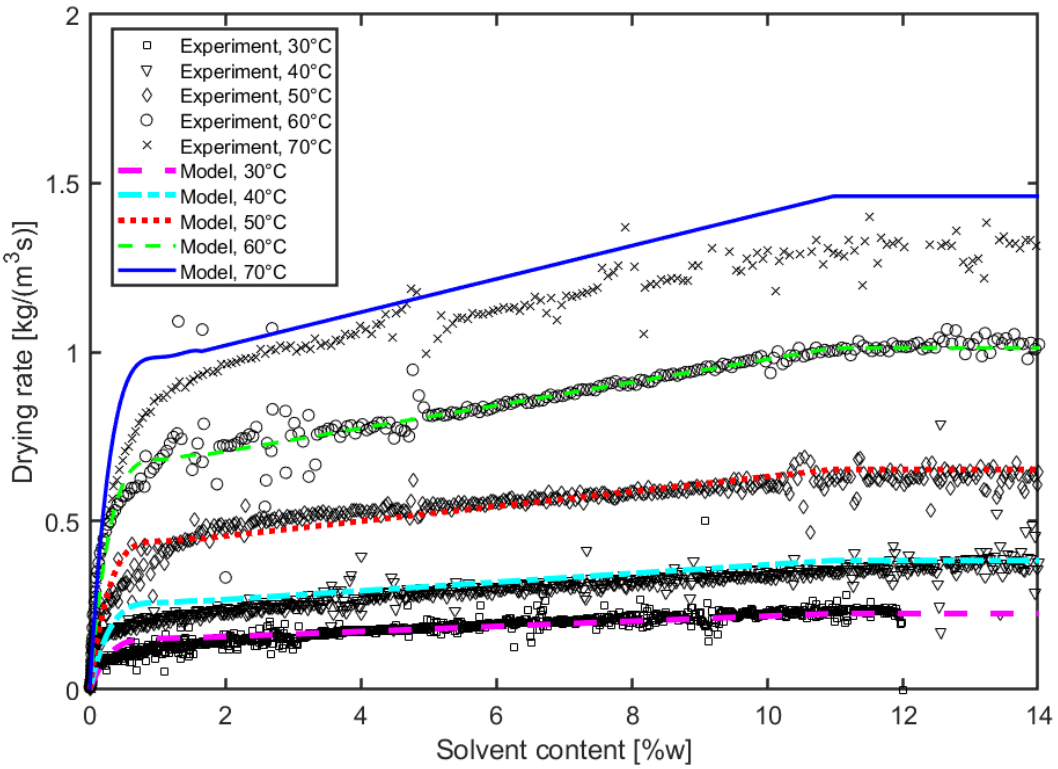


Figure 4.7. *Drying experiments: drying rate obtained from the experimental weight measurements vs drying model prediction. The specific drying rate from the experiments is smoothed with a low-pass filter.*

4.5 Paracetamol case study: design space description and throughput maximization

We calculate the DS for PCM isolation on the carousel prototype presented in §4.2, making use of the calibrated carousel model (§4.3-4). The considered carousel has a port diameter equal to 1 cm. The liquid phase of the slurry is constituted at 95%w by IPA, and at 5%w by a non-volatile impurity. Cake washing is carried out with pure EtOH, to reduce the impurity content. For all solvents and impurities, the maximum acceptable content in the discharged cake is 0.5%.

The CPPs are identified as Δt_{cycle} and V_{slurry}, while c_{slurry} is the CMA. We consider W as a CV instead of a CPP, since, under the assumption of absence of adsorption of impurities in the solid, the CQAs are not impacted by even significant variations of *W*. In the case study, we fix W to a value of one, which proved to be enough to meet the target impurity content. The other CVs are T_{drying} and ΔP, which we set to the maximum admissible values for the process (respectively, 70 °C and 50 kPa), to maximize their contribution to meeting the CQAs target values. The DS is identified within a probabilistic framework (García-Muñoz et al., 2015), accounting for uncertainty on six model parameters (Table 4.4). The uncertain model parameters are those estimated from experimental data (h_M a, R_m , α and ϵ), and two additional ones, $c_{slurry,dist}$ and $V_{slurry,dist}$, expressing the small fluctuations that c_{slurry} and V_{slurry} present around their set-points during operation. In the simulations for DS identification, c_{slurry} and V_{slurry} are calculated adding to the values set as input of the carousel model by, respectively, $c_{slurry,dist}$ and $V_{slurry,dist}$ (normally distributed with zero mean and standard deviation equal to 3% of the corresponding set-point). The parameters h_M a, α and ϵ are considered normally distributed, with mean equal to the respective estimated values. The standard deviations of the distributions are set to 5% of the mean, instead than using the (smaller) standard deviation obtained in §4.3 with parameter estimation, to conservatively account for additional process variability and disturbances that can occur during operation (e.g.: small changes in the CSD of the fed slurry). A uniform distribution is resorted to for R_m, aiming to reproduce the increase of fouling during operation assessed with the filtration experiments. The lower bound of the distribution corresponds to the resistance values found for the clean meshes (Runs #1 and #3), while the upper bound is representative of the threshold at which the cleaning-in-place procedure is triggered in the carousel. For the purposes of this work, it is assumed that all meshes foul at the same rate.

For DS identification, a CPPs and CMAs grid is first built. We vary Δt_{cycle} from 25 s to 395 s, with step 10 s, while $V_{slurry}$is varied from 1 mL to 10 mL (maximum port capacity), with step 0.5 mL (smaller steps would be of the same magnitude of $V_{slurry,dist}$). The CMAs vector contains only c_{slurry}, which is varied between 50 kg/m^3 and 200 kg/m^3, with step 25 kg/m^3. For every grid point, a Monte Carlo simulation with 400 realizations is carried out. For each realization, different values for the uncertain model parameters are sampled from their probability distributions, and the associated CQAs of the discharged cake are calculated with the carousel model. The percentage of

realizations satisfying the CQAs requirements in a grid point corresponds to the probability of meeting the quality target for the associated combination of CPPs and CMA.

The probabilities calculated through the Monte Carlo simulations are reported in Figure 4.8, as planes at constant c_{slurry}. We set the minimum acceptable probability for a grid point to be in the DS to 90%. As expected, the minimum Δt_{cycle} for obtaining the target CQAs with an acceptable probability increases with V_{slurry} and c_{slurry}. We obtain the following equation for the curve of separation between the DS and the region of non-acceptable probability, through multi-linear regression:

$$\Delta t_{cycle} = a_1 V_{slurry} + a_2 c_{slurry} + a_3 c_{slurry} V_{slurry} + a_4 \quad (4.38)$$

with a_1=-2.78E6 s/m^3, a_2=-1.05E-1 s m^3/kg, a_3=3.06E5 s/kg and a_4=5.14E1 s. The values of a_1, a_2, a_3 and a_4 are calibrated on the grid points at c_{slurry} = 50, 100, 150 and 200 kg/m^3, while the grid points at c_{slurry} = 75, 125 and 175 kg/m^3 are used for validation. The curve obtained with Equation 4.38 effectively delimits the DS on both the calibration and validation datasets (Figure 4.8). We also validate it for $\Delta t_{cycle} > 395$ s for c_{slurry}=150, 175 and 200 kg/m^3 (conditions in Figures 4.8e-g), to obtain the DS boundary for all the explored values of V_{slurry} and c_{slurry} in the grid. For this purpose, we carry out additional Monte Carlo simulations in a less refined grid (the plots are not reported here for conciseness). Hence, the DS is expressed by:

$$\text{DS} = \left\{ \left(V_{slurry}, \Delta t_{cycle}, c_{slurry} \right) \middle| \Delta t_{cycle} \geq a_1 V_{slurry} + a_2 c_{slurry} + a_3 c_{slurry} V_{slurry} + a_4 \right\} \quad (4.39)$$

with $V_{slurry} \in \left[V_{slurry,min}, V_{slurry,max} \right]$ and $c_{slurry} \in \left[c_{slurry,min}, c_{slurry,max} \right]$ ($V_{slurry,min}$, $V_{slurry,max}$, $c_{slurry,min}$ and $c_{slurry,max}$ are the boundaries of the grid used for the Monte Carlo simulations). As a remark, for cakes presenting larger α, higher order terms might be needed for fitting the DS boundary beside the linear dependencies on V_{slurry} and c_{slurry} and the interaction term used in Equation 4.38.

Table 4.4. *Uncertain parameters for probabilistic DS calculation.*

Uncertain parameter	Unit	Probability distribution
Cake porosity ϵ	[-]	N(0.36, 3.24E-4)
Filter medium resistance R_m	[1/m]	U(3E9, 3E10)
Specific mass transfer coefficient $h_M a$	[kg/(m^3 s Pa)]	N(2.25E-5, 1.27E-12)
Slurry concentration disturbance $c_{slurry,dist}$	[-]	N(0, 9.00E-4 c^2_{slurry})
Slurry volume disturbance $V_{slurry,dist}$	[kg/kg]	N(0, 9.00E-4 V^2_{slurry})
Specific cake resistance α	[m/kg]	N(1E9, 2.50E15)

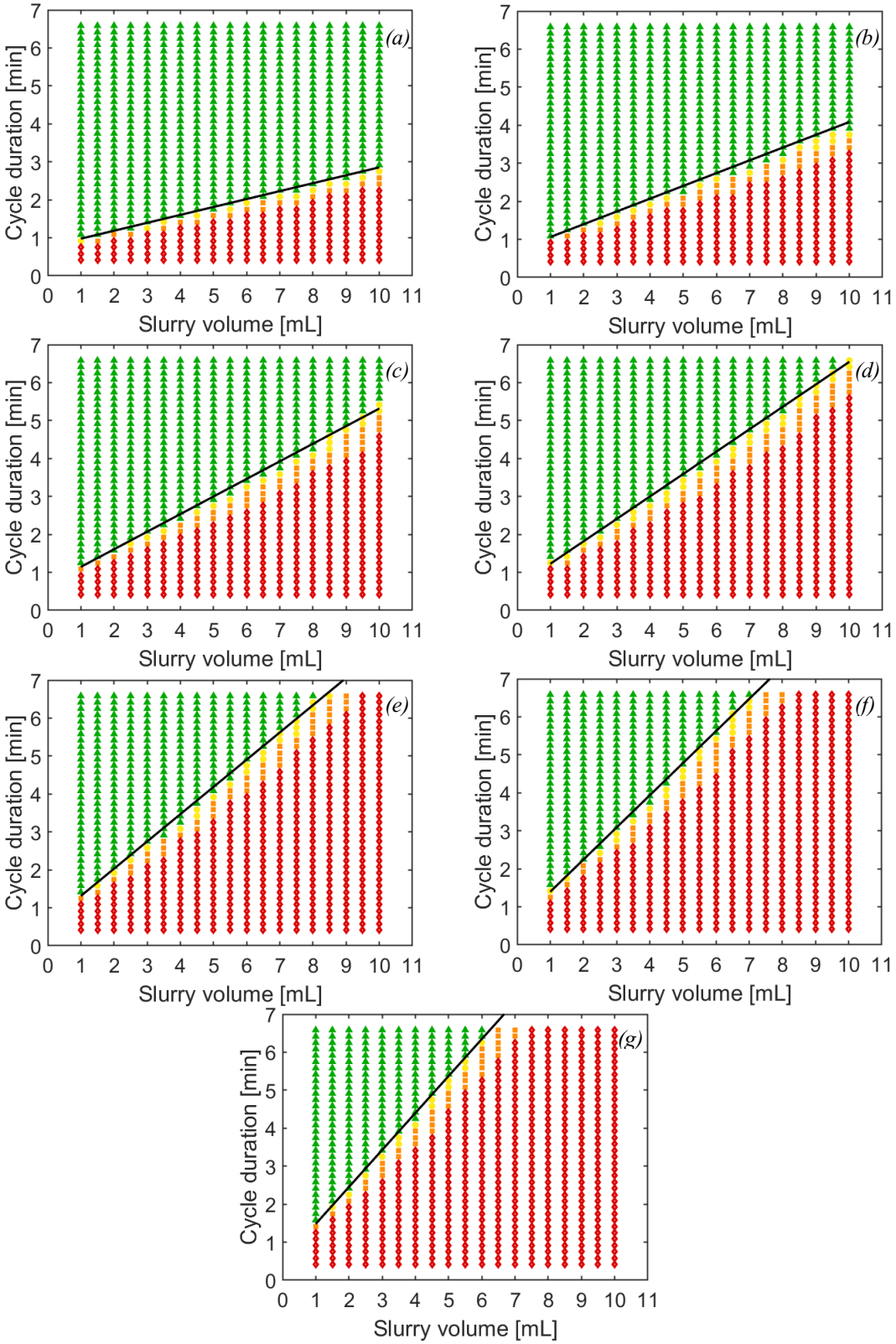


Figure 4.8. *Paracetamol case study: probabilistic DS, representing the probability of meeting the target CQAs. Green triangles: probability ≥ 90%, yellow circles: 80% ≤ probability < 90%, orange squares: 60% ≤ probability < 80% and red diamonds: probability < 60%. In the figures, c_{slurry} is (a) 50 kg/m^3, (b) 75 kg/m^3, (c) 100 kg/m^3, (d) 125 kg/m^3, (e) 150 kg/m^3, (f) 175 kg/m^3 and (g) 200 kg/m^3. The solid line represents the DS boundary calculated with Equation 4.38.*

The maximum throughput T_{max} that can be processed within the DS is obtained solving the following optimization problem:

$$T_{max} = \max_{V_{slurry}, \Delta t_{cycle}, c_{slurry}} T \tag{4.40}$$

subject to:

$$T = \frac{c_{slurry} V_{slurry}}{\Delta t_{cycle}} \tag{4.41a}$$

$$\Delta t_{cycle} \geq a_1 V_{slurry} + a_2 c_{slurry} + a_3 c_{slurry} V_{slurry} + a_4 \tag{4.41b}$$

$$c_{slurry} \in [c_{slurry,min}, c_{slurry,max}] \tag{4.41c}$$

$$V_{slurry} \in [V_{slurry,min}, V_{slurry,max}] \tag{4.41d}$$

where T is the throughput processed in the carousel. It can be easily assessed that the optimum of Problem 4.40-41 lays on the DS boundary, hence Equations 4.40 and 4.41a can be reformulated as:

$$T_{max} = \max_{V_{slurry}, c_{slurry}} T \tag{4.42}$$

subject to:

$$T = \frac{c_{slurry} V_{slurry}}{a_1 V_{slurry} + a_2 c_{slurry} + a_3 c_{slurry} V_{slurry} + a_4} \tag{4.43}$$

Overall, the reformulated optimization problem is defined by Equations 4.42, 4.43 and 4.41c-d. Δt_{cycle} does not explicitly appear anymore, although it is immediately available from Equation 4.38. The gradients of the objective function T with respect to c_{slurry} and V_{slurry} are always positive under the domain defined by the constraints, hence T_{max} (195 mg of crystals per minute) is reached for $c_{slurry} = c_{slurry,max}$= 200 kg/m^3 and $V_{slurry} = V_{slurry,max}$ = 10 mL. The plots of the throughput at the DS boundary for different values of c_{slurry} and V_{slurry} (Figure 4.9) support the derived analytical solution of the optimization problem. In other applications, T_{max} may be found in between the grid boundaries, rather than at the boundaries themselves. For this case study, the low nonlinearity of the DS boundary possibly prevents achieving a stationary point of maximum, although from Figure 4.9 it can be noted that such condition would eventually have been reached if the maximum capacity of the ports were larger. In any case, for the purposes of pharmaceutical manufacturing, guaranteeing the quality of the product assumes a greater importance than maximizing the throughput. Hence, it is suggested to operate in proximity of T_{max}, but more inside the DS, instead than at the boundary.

To conclude this section, we demonstrate an additional application of the carousel model, regarding the selection of the number of ports during the design of a new carousel. Thanks to the flexibility of arrangement of the processing steps models in the carousel simulator, we repeat the DS space calculation with the same procedure employed for obtaining the results in Figure 4.8d, but simulating a carousel with an additional drying station (six stations overall). From the comparison (Figure 4.10), it clearly emerges that, for a given V_{slurry}, acceptable CQAs can be reached with two drying stations in less than half of the time needed with only one drying station. Actually, for the carousel with only five stations, simulations show that drying is the rate determining step.

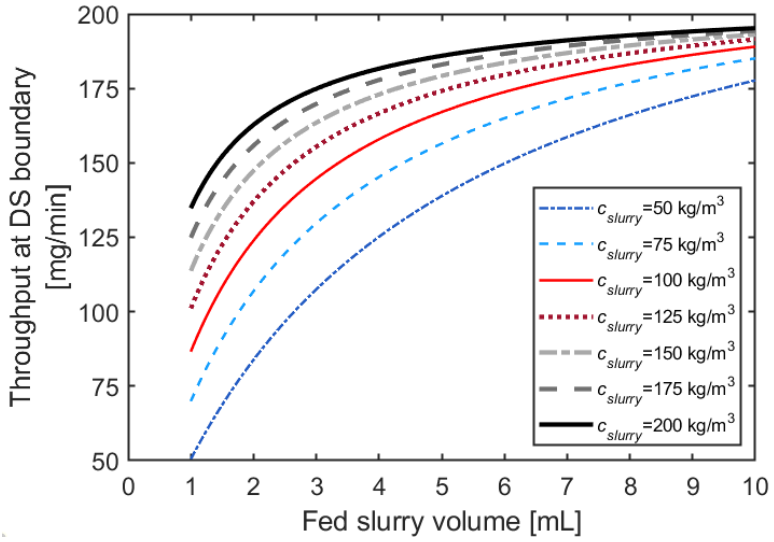


Figure 4.9. *Paracetamol case study: throughput T at the design space boundary (Equation 4.38) for varying V_{slurry} and c_{slurry}.*

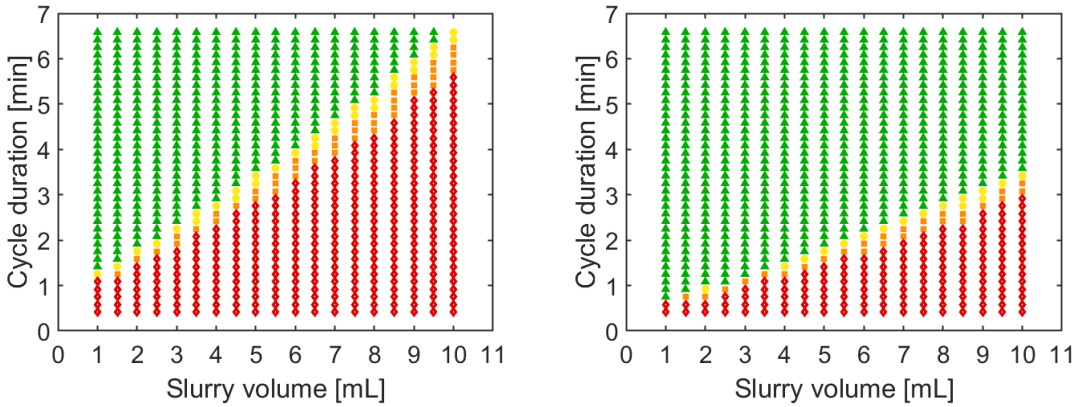


Figure 4.10. *Paracetamol case study: DS comparison of the carousel configuration with one drying station (left, equivalent to Figure 4.8d) and two drying stations (right). The markers legend is as in Figure 4.8, and c_{slurry} is 125 kg/m^3.*

4.6 Conclusions

The paper presents a mathematical model of a novel continuous carousel for the isolation of crystals from a slurry. The model is developed by assembling together independent filtration, deliquoring, washing and drying modeling components, that can also be arranged to mimic different continuous isolation setups. A cake properties module is implemented for calculating the cake porosity and specific resistance from the slurry crystal size and shape distribution. Dynamic profiles of cake composition are tracked throughout all carousel processing by the model, that overcomes the short-cut methods more commonly used for designing this type of process. In particular, a differential

deliquoring model is used, instead of widely employed design charts. The developed cake drying model presents novel features, as it accounts for the presence of a flow of hot gas, not usually present in other equipment (and models) for cake drying.

A case study on the digital design of a process for isolating PCM from a crystallization slurry, is presented to demonstrate the carousel model capabilities. After successfully calibrating the model with filtration and drying experiments, the probabilistic design space of the process is identified. Model uncertainty is considered through Monte Carlo simulations, sampling the uncertain model parameters from their probability distributions. The maximum throughput that can be processed in the carousel with acceptable probability of meeting the product quality specifications was also evaluated.

The proposed model and approach to the digital design of continuous integrated crystals isolation represents a step forward in the field and in end-to-end continuous pharmaceutical manufacturing. The calculated design space responds to the need of considering the critical process parameters, the variability in the feed conditions and the model uncertainty altogether in a probabilistic framework when designing the filtration, washing and drying sections of a pharmaceutical process. Future work will involve the digital design of other crystals isolation processes and the integration of an upstream crystallizer model for carrying out integrated crystallization-filtration-drying digital design. Experimental validation of the washing component of the model for systems where washing is a critical step will be considered, and the effect of partial dissolution of the crystals into the wash solvent will be incorporated in the modeling framework. Process control applications are also envisioned, by using the developed model as digital twin of the process for testing control strategies (including active control), and for model-based process control.

Chapter 5

Intensified continuous filtration-drying of pharmaceuticals: Quality-by-Control on a novel real time simulator*

In this Chapter, the mathematical modeling framework of Chapter 4 is developed into *ContCarSim*, a real time simulator of a continuous carousel for filtration-drying of active pharmaceutical ingredients. Differently from the carousel model of Chapter 4, the simulator developed in this Chapter includes computational routines simulating sensors and actuators present in physical carousels (Ottoboni et al., 2020a), and supports the implementation of control loops. The simulator is calibrated with novel filtration and drying experiments on a pilot scale carousel for a paracetamol/ethanol slurry system. A closed-loop control strategy for the unit, based on the Quality-by-Control paradigm, is then conceived and tested on the simulator, under a set of disturbances known to affect the unit operation. The designed Quality-by-Control framework leads to improved performance compared to early stage Quality-by-Design control strategies, which rely on open-loop quality control. The developed simulator will be made freely available at the link: https://github.com/francescodestro/ContCarSim.

5.1 Introduction

In the latest years, pharmaceutical development and manufacturing have been undergoing a modernization trend, aimed at increasing the economic efficiency and the capability to attain and consistently maintain product quality. These efforts have been promoted by regulators such as the United States Food & Drug Administration (FDA) and the International Council for Harmonisation (ICH), through the process analytical technology (FDA, 2004b) and quality-by-design (QbD; ICH, 2009) initiatives. Under a QbD approach to pharmaceutical development and manufacturing, quality is inherently built into the product, by designing and conducting the manufacturing process with a

*Destro, F., Z. K. Nagy, and M. Barolo. A benchmark simulator for quality-by-design and quality-by-control studies in continuous pharmaceutical manufacturing – Intensified filtration-drying of paracetamol/ethanol slurries. *In preparation*
Destro, F., Z. K. Nagy, and M. Barolo. Quality-by-Control of intensified continuous filtration-drying of active pharmaceutical ingredients. *In preparation*

science- and risk-based approach. QbD represents a major improvement with respect to the traditional pharmaceutical approach of Quality-by-Testing (QbT), based on extensive testing on the end-product. The guidelines provided by regulators (ICH, 2012b, 2009, 2008, 2005) present different QbD elements that can be adopted by practitioners for process development and manufacturing. Following a typical QbD pharmaceutical development workflow (ICH, 2009), the product critical quality attributes (CQAs) are first identified from the quality target product profile. Subsequently, the critical process parameters (CPPs) and the raw materials properties (critical material attributes, CMAs) that are critical for product CQAs attainment are established through a quality risk assessment procedure. Then, the design space (DS) is determined, as the multivariable region of CMAs and CPPs that allow obtaining the target product CQAs. Finally, a control strategy is designed, as a set of actions, routines and controls to maintain the process in a state of control and guarantee the product quality. A document by FDA regulators (Yu et al., 2014) distinguished among three levels of control strategies. A Level 3 control strategy, corresponding to the QbT approach, consists in operating the process under tight intervals of CMAs and CPPs, which have been proven to yield the target product CQAs in the regulatory approval application. In a more advanced Level 2 control strategy, a DS is established, and the process is operated therein. Level 2 control strategies guarantee more flexibility on the operating conditions than Level 3 strategies, and allow reducing the reliance on end-product testing. However, in a Level 2 control strategy the product CQAs are still controlled at open-loop, by moving the CPPs within the DS in response to registered changes in the CMAs. In a Level 1 control strategy, instead, the product quality is actively controlled making use of process control techniques. The CPPs are automatically adjusted in response to measured changes of the CMAs (feedforward control) or of the CQAs (feedback control). This is achieved through PID control or more advanced techniques, such as model predictive control and real time optimization (RTO; Seborg et al., 2017). A Level 1 control strategy facilitates real time release testing, namely the ability to evaluate and ensure the quality of intermediates and/or final product during process operation based on process data, including a combination of CMAs, CPPs and CQAs (ICH, 2009). Real time release testing can potentially replace end-product testing. Soft sensors (e.g., those based on state estimation; Ray, 1981) can replace direct measurements of CMAs, CPPs and CQAs to support real time release testing (ICH, 2011). Although both Level 1 and Level 2 control strategies include QbD elements, the ultimate aim of the QbD initiative is to reach widespread establishment of Level 1 control strategies, from which both pharmaceutical companies and patients can benefit (Collins, 2018; Fisher et al., 2019). Active process control on quality can reduce the occurrence of shortages and recalls and can provide higher quality attainment. Moreover, McKinsey (McKinsey & Company, 2021) estimated that introducing an advanced quality control system for process development and manufacturing could have a paramount impact on profit of pharmaceutical companies, by reducing the product launch time by even 30% and increasing manufacturing and supply chain capacity by up to 30%, in addition to preventing major product shortages and recalls. Even though the adoption of Level 1 control strategies is still in embryonic phase, many recent academic publications (Mesbah et al., 2017; Rehrl et al.,

2016; Sen et al., 2014) demonstrated the advantages of closed-loop quality control in different pharmaceutical processes, including plant-wide applications (Lakerveld et al., 2013). The recent interest towards control strategies comprising elements of active process control has led to the establishment of a novel paradigm in pharmaceutical development and manufacturing, which has been named Quality-by-Control (QbC; Su et al., 2019b). QbC is an evolved form of QbD, rather than a QbD-independent initiative, in which active process control represents the core feature of the control strategy. A hierarchical structure, following the ISA-95 Enterprise-Control System Integration Standard, has recently been proposed (Su et al., 2019b, 2017) to support the development of QbC control systems. The hierarchical structure includes three layers (Figure 1.3): *i)* Layer 0, consisting of the built-in control systems of the equipment, *ii)* Layer 1, at which the loop on quality is closed through PID control, and *iii)* Layer 2, which features advanced model-based process control and process monitoring techniques.

In addition from active quality control, the transition to a more continuous production mode is another pharmaceutical emerging technology promoted by the QbD initiative (ICH, 2021; Lee et al., 2015). Continuous processing, when feasible, is preferred to the traditional batch production mode, due to the many advantages it offers (Burcham et al., 2018; Fisher et al., 2019; Ierapetritou et al., 2016). As for active process control, the benefits of continuous processing are for both manufacturers and patients, and include reduced manufacturing time and cost, greater product quality consistency, and potential to reduce recalls and shortages. Furthermore, process control is usually easier in a continuous plant and has been explored more in the literature for continuous processes rather than for their batch counterparts.

As a matter of fact, a tight interconnection exists between the two most important pharmaceutical emerging technologies: QbC is gaining much relevance under the increasing popularity of continuous processes, while the transition to more continuous operation modes is boosted by the QbC tools for systematic control system design. Many literature contributions on continuous processing and QbC have been published around the topics of reacting systems (Bana et al., 2017; A E Cervera-Padrell et al., 2012; Albert E. Cervera-Padrell et al., 2012; Nikolakopoulou et al., 2020), crystallization (Acevedo et al., 2016; Nagy et al., 2013; Nagy and Braatz, 2012; Wood et al., 2019a), and of solid-dosage form manufacturing lines (García-Muñoz et al., 2010; Hanson, 2018; Singh et al., 2015; Su et al., 2019a, 2019b, 2017). However, end-to-end continuous pharmaceutical processes with active quality control are scarcely explored in the literature (Lakerveld et al., 2013; Mesbah et al., 2017). Filtration, washing and drying of crystallization slurries are pivotal unit operations for connecting the drug substance and drug product manufacturing sections of a pharmaceutical process into an end-to-end continuous integrated system. However, although they can constitute a bottleneck in the implementation of plant-wide continuous automated pharmaceutical processes, they have been scarcely studied from a continuous integrated processing and QbC perspective.

In this study, we make a step forward toward the implementation of end-to-end continuous pharmaceutical manufacturing with closed-loop quality control, by presenting a QbC framework for

a novel continuous carousel for integrated filtration and drying of crystallization slurries. Cake washing can also be carried out within the unit, but this processing step is not considered in this chapter. The carousel presents multiple processing stations embedded in a main cylindrical body. During carousel operation, cylindrical ports containing the material being processed are aligned with the stations, and processing steps are carried out simultaneously in each port in a batch mode. Continuous operation is enabled by carousel rotations, which transfer each port to the following station at each time interval (cycle duration), as defined by the user. The slurry from the crystallizer is loaded into the first station, and is eventually discharged as dry crystals cake from the last station, after having been processed in all the stations in between. Previous studies on similar carousels involved experimental design of the unit and its operation (Ottoboni et al., 2020b, 2020a, 2019), and experimental and modeling results that successfully integrated the carousel, in filtration-only mode, with upstream continuous crystallization systems (Acevedo et al., 2016; Liu et al., 2019; Nagy et al., 2021). In Chapter 4, a comprehensive mathematical model of the unit has been developed and used for designing the carousel operation, in filtration-washing-drying mode, in a QbD perspective, namely by describing the probabilistic DS of the unit for a paracetamol crystals isolation process. The model was calibrated and validated through filtration experiments in a Nutsche filter, and drying experiments on a thermogravimetric analyzer.

In this study, we elaborate on the mathematical model of the carousel of Chapter 4, to develop *ContCarSim*, a real-time simulator of the unit, calibrated and validated through experiments carried out on an actual prototype carousel, for an ethanol/paracetamol slurry. Differently from the carousel model of Chapter 4, the simulator developed in this Chapter includes computational routines simulating sensors and actuators present in physical carousels (Ottoboni et al., 2020a), and supports the implementation of control loops. Improvements to the energy balance of the model drying component are proposed to enhance the match with novel experimental findings. A three-layer QbC control strategy for the carousel is then conceived and tested on the simulator. Layer 0 consists in the built-in controllers of the equipment, such as those needed for manipulating the CPPs and handling the carousel rotation/feeding routines. In Layer 1, the control strategy features an end-point controller for automatically triggering a carousel rotation when the target product CQAs are met in the final station of the carousel. In Layer 2, an RTO algorithm is used for optimizing the operating conditions at each cycle. A state estimator is also put in place for monitoring the product CQAs in a probabilistic framework, to support the end-point controller at Layer 1 and to enable real time release testing. The proposed QbC control strategy is benchmarked against a traditional QbD approach under a set of disturbances occurring in normal operating conditions, such as mesh fouling and concentration fluctuations in the fed slurry.

The remainder of the manuscript is organized as follows. In Section 5.2, the pilot scale carousel used for the experimental campaign is presented. In Section 5.3, the simulator is described, including the underlying mathematical modeling framework. The calibration of the simulator through an experimental campaign is reported and discussed in Section 5.4. Section 5.5 presents QbD and QbC

design and control challenges that can be addressed with *ContCarSim*. Section 5.6 proposes a QbC framework for the carousel. The response of the proposed QbC control strategy to a set of disturbances is discussed and benchmarked with traditional QbD approaches in Section 5.7. The concluding section follows.

The simulator will be made publicly available at the link: https://github.com/francescodestro/ContCarSim.

5.2 Experimental setup: the pilot scale carousel

In this section, we complement the description of the carousel technology given in §4.2 with technical information on the pilot scale carousel used for the experimental campaign reported in §5.4. The unit (Figure 4.1) can continuously process an inlet slurry stream into a dry crystals cake. A schematic P&ID of the process is provided in Figure 5.1, with the legend of equipment, sensors and controllers reported in Table 5.1.

The carousel features five cylindrical ports, each one of 15 mm diameter, which allow a maximum hold-up of 10 mL. The ports are embedded in a main cylindrical body, aligned to five processing stations (Stations 1-5). For illustrative purposes, in Figure 5.1 the stations are represented as vessels in series (V102-V106), although the actual layout of the carousel is as in Figure 4.1. In the carousel

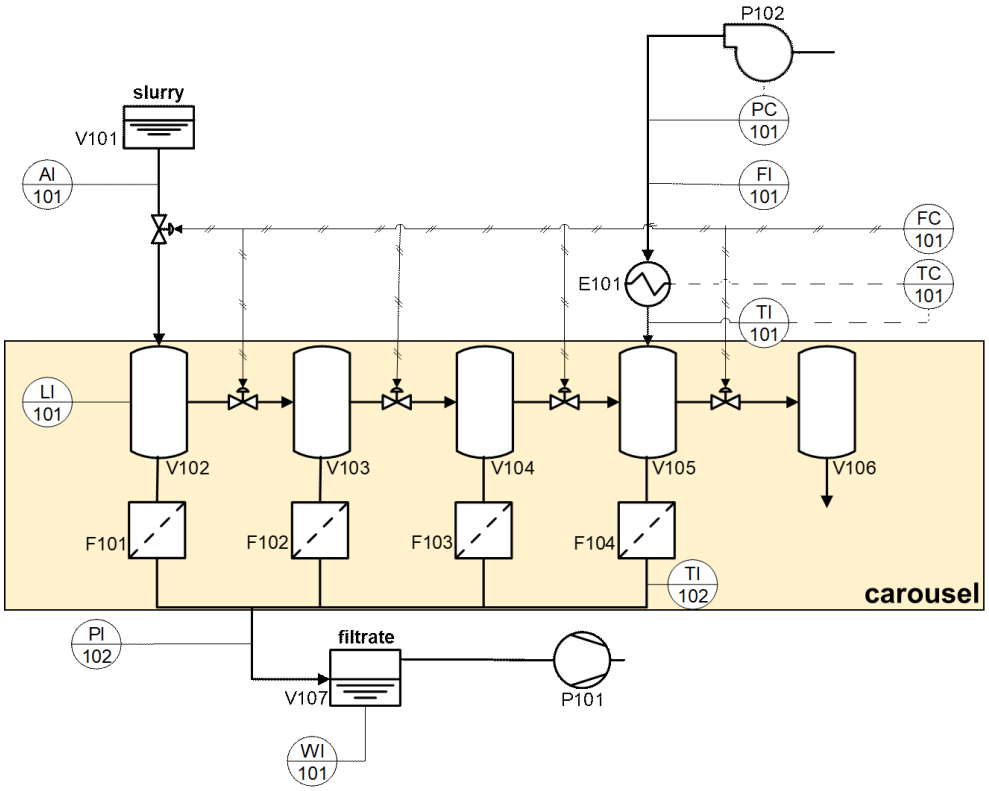


Figure 5.1. *P&ID of the prototype continuous carousel setup for integrated filtration, washing and drying. The equipment legend is reported in Table 5.1.*

prototype, Stations 1-4 present a filter mesh at the bottom (F101-104), and are connected to a vacuum pump (P101) that provides the pressure gradient necessary for filtration and drying. Instead, Station 5 is open at the bottom for cake discharge, which is enabled by the action of a pneumatic piston. In commercial scale carousels and in the developed simulator (Section 5.3), the pressure gradient is increased with a compressor (P102), whereas all ports are maintained at atmospheric pressure on the top part.

The carousel operates in cyclic mode: processing cycles, during which every port processes batch-wise the material therein contained, are alternated to carousel rotations, during which the ports containing the material being processed are moved to the following station. Carousel rotations are represented in the P&ID as material streams, controlled by FC101. The alternating processing cycles and carousels rotations are interrupted when significant mesh fouling is detected: a cleaning-in-place cycle is triggered, and all meshes are automatically cleaned by sending a wash solvent into the carousel (Figure 4.1). Stations 1-3 are dedicated to filtration and deliquoring, while in Station 4 thermal drying is carried out. In Station 2 cake washing can also be carried out through a wash solvent

Table 5.1. *Legend of the P&ID diagram of Figure 2, including unit operations and ancillary equipment. Some pieces of equipment are present both in the carousel prototype used for data collection and in the implemented carousel simulator (Section 5.3), while other ones are present in only one of the two.*

Name	Description	Present in prototype?	Present in simulator?
Unit ID			
F101-F104	Filter mesh below Stations 1-4 (respectively)	Yes	Yes
P101	Vacuum pump	Yes	No
P102	Compressor	No	Yes
E101	Drying gas heater	Yes	Yes
VI101	Slurry tank	Yes	Yes
VI102	Carousel Station 1	Yes	Yes
VI103	Carousel Station 2	Yes	Yes
VI104	Carousel Station 3	Yes	Yes
VI105	Carousel Station 4	Yes	Yes
VI106	Carousel Station 5	Yes	Yes
VI107	Filtrate collector	Yes	Yes
Controllers and sensors			
AI101	Slurry concentration sensor (ultrasonic probe; Bamberger and Greenwood, 2004a, 2004b)	No	Yes
FC101	Fictitious flowrate controller representing carousel rotation and slurry feeding routines enabled by PLC	Yes	Yes
FI101	Flowmeter for gas entering carousel ports	Yes	Yes
LI101	Camera system (Ottoboni et al., 2020a) measuring volume of fed slurry and cake height	No	Yes
PC101	Pressure controller		
PI102	Pressure indicator	Yes	Yes
TC101	Drying gas inlet temperature controller	Yes	Yes
TI101	Thermocouple for inlet drying gas temperature	Yes	Yes
TI102	Thermocouple for outlet drying gas temperature	Yes	Yes
WI101	Scale for inferring filtrate flowrate	Yes	Yes

(Figure 4.1), but this feature was not considered in this study (the washing equipment is not reported in the P&ID of Figure 5.2). In Station 5, only cake discharge occurs.

Slurry processing in the prototype occurs as follows. The crystallization slurry is fed to Station 1 at the beginning of each cycle, by keeping the valve between the slurry tank (V101) and Station 1 open for a time interval (feed duration) assigned by the user (a linear correlation between feed duration and fed slurry volume for a given slurry system can be quickly regressed with suitable experiments). After slurry feeding, a subsequent filtration step starts in Station 1, and it continues until filtration ends, or throughout the whole cycle duration, until the following carousel rotation. In the latter situation, filtration will continue in Station 2. During filtration, the liquid contained in the slurry is filtered out of the port by the action of the vacuum pump and stored in filtrate collector V107, while the crystals are retained on top of the filter mesh, leading to cake formation. We distinguish between actual filtration, when there is a slurry hold-up on top of the cake being formed, and the subsequent deliquoring, during which the only remaining liquid is the one retained inside the cake pores. Upon deliquoring, the liquid in the pores of the cake is mechanically displaced out of the cake by the action of the vacuum pump, until a certain pore saturation equilibrium is achieved. Filtration duration depends on the cake properties and on the pressure drop delivered by the vacuum pump. Depending on filtration duration, the cake can also be partially deliquored in Stations 1-3, or it might even enter Station 4 with some slurry hold-up (drying cannot be properly conducted in this situation, which should be avoided). Thermal drying is performed in Station 4 by flowing a hot gas stream through the cake. As the carousel is a prototype for process development and scale-up studies, the filtration and drying stations are not activated simultaneously as in commercial carousels, but in subsequent filtration and drying cycles, followed by a carousel rotation. This feature has been exploited in this study to decouple filtration and drying experiments, which is more informative for model calibration purposes.

The sensor and controller network installed in the carousel and exploited in the experimental campaign is made up of: *i)* FC101, a fictitious flowrate controller that represents the rotation and slurry feeding routines controlled by the carousel programmable logic controller (PLC), *ii)* FI101, a flowmeter measuring the gas flowrate entering the carousel ports, *iii)* PI102, measuring the pressure at the bottom of the filter meshes, *iv)* TC101, controlling the temperature of the gas entering the drying port by acting on heater E101, *v)* thermocouple TI101, measuring the drying gas inlet temperature, i*v)* thermocouple TI102, measuring the drying gas outlet temperature, and *vii)* WI101, a scale placed under filtrate collector V107 for inferring the filtrate flowrate. All measurements are sampled every 0.15 s. Level indicator L101 (camera vision system; Ottoboni et al., 2020a) and slurry composition sensor AI101 (ultrasonic probe; Bamberger and Greenwood, 2004a, 2004b) are present in commercial scale carousels and are implemented in the simulator (§5.3), but they are not installed in the available carousel prototype.

5.3 Carousel simulator description

The simulator is made up by assembling together filtration, deliquoring and drying models (detailed in §5.3.2) with the same sequence with which they occur during carousel operation (§5.2). The models are dynamic, and are developed for the one-dimensional system of Figure 5.2, with the axial coordinate of the cake denoted by the symbol z. Filtration and deliquoring models are used for simulating V102-V104, while the drying model is used for V105. Implemented routines automatically detect when filtration finishes for a given batch, and simulate deliquoring for the remaining time spent by the material in V102-V104. If a cake entering V105 is not deliquored enough for proper air flow, deliquoring is simulated in place of drying, until drying can actually start. The simulator also properly handles the limiting case in which filtration has not finished in the batch entering V105, yet.
The carousel simulator reproduces the operation of the carousel setup described by Figure 5.1. As detailed in Table 5.1, in the simulator the pressure gradient is provided by compressor P102, and vacuum pump P101 is not implemented.

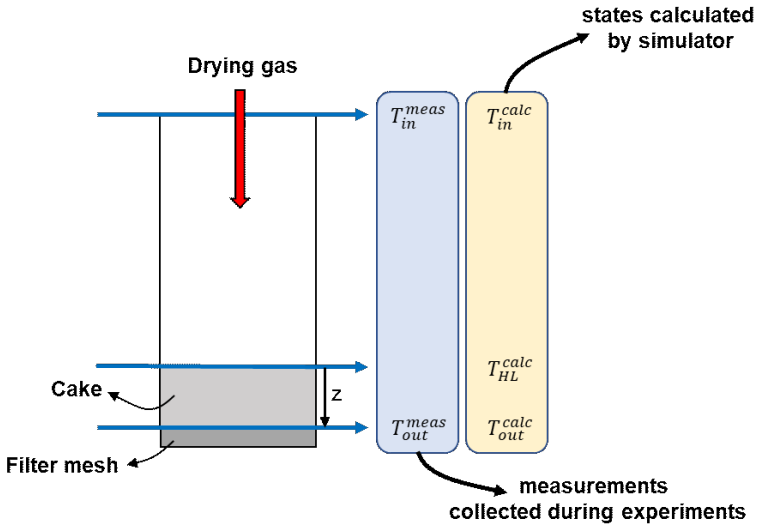


Figure 5.2. *Reference system for the filtration, deliquoring and drying models. The different notation for experimental gas temperature measurements and gas temperatures calculated by the simulator is elucidated. Note that the gas outlet temperature T_{out}^{meas} is measured with a thermocouple placed in the outlet pipe placed below the filter mesh, however it can fairly be approximated to the temperature of the gas at the bottom of the cake.*

5.3.1 Simulator input/output structure

The input/output structure of the simulator is given in Table 5.2. The inputs to be provided to the simulator are: *i)* parameters to be calibrated from experiments, *ii)* the thermophysical properties of the pure components of the system and of the cake, *iii)* set-points for the controllers implemented in the simulator, and *iv)* suitable profiles for the system disturbances.

Table 5.2. *Input/output structure of the carousel simulator.*

	Symbol	Variable	Unit	Relevant equipment
Inputs				
Calibration parameters	$\alpha_{HL,1}$	Air-wall heat transfer coefficient - intercept	-	V105
	$\alpha_{HL,2}$	Air-wall heat transfer coefficient - slope	s/m	V105
	η_i	Drying effectiveness factor for species i	-	V105
	ξ	Tunneling parameter	-	V105
	$\tau_{HL,1}$	Air-wall heat transfer time constant	s	V105
Disturbances	c_{slurry}	Slurry concentration	kg/m^3	V102, AI101
	h_M^{dist}	Drying kinetic constant disturbance	-	V105
	h_T^{dist}	Cake/gas heat transfer coefficient in dryer disturbance	-	V105
	$R_{m,1}$	Mesh resistance – Station 1	1/m	V102
	$R_{m,2}$	Mesh resistance – Station 2	1/m	V103
	$R_{m,3}$	Mesh resistance – Station 3	1/m	V104
	$R_{m,4}$	Mesh resistance – Station 4	1/m	V105
	V_{slurry}^{dist}	Fed slurry volume disturbance	-	V102
	$\mathbf{w}_{in,g}$	Inlet mass fraction gas phase components	-	V105
	$\mathbf{w}_{l,slurry}$	Mass fraction of slurry liquid phase components	-	V102
	ϵ^{dist}	Cake porosity disturbance	-	V102-V105
Thermophysical properties		Cake physical properties	[…]	V102-V105
	S_∞	Cake deliquoring equilibrium saturation	-	
	α_{cake}	Specific cake resistance	m/kg	
	ϵ_{cake}	Cake porosity	-	
		Properties of pure components of solid, liquid and gas phases	[…]	All
Controlled variables set-points	$T_{in,g}^{sp}$	Drying gas inlet temperature set-point	°C	V106, TC101
	V_{slurry}^{sp}	Fed slurry volume set-point	m^3	V102, FC101
	ΔP^{sp}	Pressure gradient set-point	bar	V102-V105, PC101
	Δt_{cycle}^{sp}	Cycle duration set-point	s	V102-V105, FC101
Outputs				
Simulated measurements	$y_{c_{slurry}}$	Slurry concentration measurement	kg/m^3	AI101
	$y_{H_{cake}}$	Cake height measurement	m	LI101
	$y_{M_{filt}}$	Filtrate mass measurement	kg	WI101
	y_P	Pressure at meshes bottom	bar	PI101
	$y_{T_{in}}$	Drying inlet gas temperature measurement	°C	TC101
	$y_{T_{out}}$	Drying outlet gas temperature measurement	°C	TI101
	$y_{\dot{V}_{dryer}}$	Gas flowrate measurement	NL/min	FI101
	$y_{V_{slurry}}$	Fed slurry volume measurement measurement	m^3	LI101
States	S_1	Cake saturation – station 1	-	V102
	S_2	Cake saturation – station 2	-	V103
	S_3	Cake saturation – station 3	-	V104
	S_4	Cake saturation – station 4	-	V105
	$w_{i,cake,1}$	Mass fraction of solvent/impurity i in cake – Station 1	-	V102

$w_{i,cake,2}$	Mass fraction of solvent/impurity *i* in cake – Station 2	-	V103
$w_{i,cake,3}$	Mass fraction of solvent/impurity *i* in cake – Station 3	-	V104
$w_{i,cake,4}$	Mass fraction of solvent/impurity *i* in cake – Station 4	-	V105

5.3.1.1 Inputs

The calibration parameters are thoroughly described in §5.3.2, when introducing Equations 5.10-5.12 and 5.18-5.21. Thermophysical properties of pure components or of the cake should be taken from cake porosity ϵ_{cake} and specific resistance α_{cake}). For running the simulator, the following controller set-points have to be assigned: cycle duration Δt_{cycle}^{sp} and fed slurry volume V_{slurry}^{sp}, needed by PLC FC101, drying gas inlet temperature $T_{in,g}^{sp}$ for TC101, and pressure gradient ΔP^{sp} provided by compressor P102. The cycle duration, pressure gradient and drying gas inlet temperature are assumed to be perfectly controlled with no dynamics; therefore, their calculated values are always equal to the relevant set-points. On the other hand, the fed slurry volume V_{slurry} is not necessarily equal to V_{slurry}^{sp}, because a set of multiplicative factors V_{slurry}^{dist}, each one associated to a specific cycle, can be set as disturbances, so that, for every cycle, $V_{slurry} = V_{slurry}^{sp} V_{slurry}^{dist}$. Analogous multiplicative disturbances, whose values vary at every cycle, are set for the specific cake resistance α, the cake porosity ϵ, the drying kinetic constant $h_{M,i}$, for species *i,* and the heat transfer coefficient between cake and air during drying h_T, and are, respectively, denoted as: α^{dist}, ϵ^{dist}, $h_{M,i}^{dist}$, and h_T^{dist}. Hence, in the equations of the simulator: $\alpha = \alpha_{cake}\ \alpha^{dist}$, $\epsilon = \epsilon_{cake}\ \epsilon^{dist}$, $h_{M,i} = h_{M,i,cake}\ h_{M,i}^{dist}$, and $h_T = h_{T,cake}\ h_T^{dist}$ ($h_{M,i,cake}$ and $h_{T,cake}$ are, respectively, defined in Equations 5.11-12). While ϵ^{dist} is an input of the simulator (Table 5.2), α^{dist} is calculated according to the Kozeny-Carman equation (Ripperger, 2013): $\alpha^{dist} = 1 + (1 - \epsilon^{dist})/(\epsilon^{dist})^3$. The crystals concentration of the fed slurry c_{slurry}, the vector $\mathbf{w}_{l,slurry}$ of the mass fractions of every component i of the liquid phase of the slurry (for i = 1, 2, …, N_L; with N_L total number of components of the liquid phase of the system), and the vector $\mathbf{w}_{in,g}$ of the inlet mass fraction of every component *i* of the drying gas are additional disturbances in the simulator (note that the solid phase of the slurry is constituted by pure crystals). Profiles for the resistances of the meshes of Stations 1-4 ($R_{m,1}$, $R_{m,2}$, $R_{m,3}$ and $R_{m,4}$), representing the degree of mesh fouling, also have to be provided.

5.3.1.2 Outputs

The outputs generated by the simulator are *i)* variables that are not measured in the real system (e.g., states) and *ii)* variables that are also measured in the real system (further referred to as *simulated measurements* and denoted by the symbol y). The former group is made up of solvents and impurities content ($w_{i,cake,j}$, for i = 1, 2, …, N_L, for every processing station j =1, 2, …, 4) and saturation (S_j, j =1, 2, …, 4; ratio between volume of liquid in the cake and pores volume) of the cakes being

processed in each of the stations. Variables $w_{i,cake,j}$ and S_j are provided as dynamic one-dimensional profiles along the axial coordinate of the cake, following the reference system of Figure 5.2. The simulated measurements and the corresponding sensors in Figure 5.1 are: the slurry concentration measurement $y_{c_{slurry}}$ (from AI101), the cake height $y_{H_{cake}}$ and slurry volume $y_{V_{slurry}}$ (from LI101), the measured inlet (from TI101) and outlet drying gas temperatures (from TI102), the measurement of pressure below the filter meshes y_P (from PI101; y_P always corresponds to atmospheric pressure in the simulator, due to the absence of vacuum pump P101), the gas flowrate measurement $y_{\dot{V}_{dryer}}$ (from FI101), and the measurements of weight of filtrate $y_{M_{filt}}$ (from WI101). All simulated measurements are generated from the corresponding calculated variables in the simulator, either by adding white noise, or by rounding the calculated variable up to the readability of industrial sensors. White noise with standard deviation equal to 2 kg/m^3, as for sensors described in the literature (Bamberger and Greenwood, 2004a, 2004b) is assigned to $y_{c_{slurry}}$. Variables $y_{H_{cake}}$ and $y_{V_{slurry}}$ are considered noise free, due to the high accuracy of camera systems installed in commercial scale carousels (Ottoboni et al., 2020a). Variables $y_{T_{in}}$ and $y_{T_{out}}$ are generated by rounding the actual temperatures to a 0.1°C readability, as for standard commercial thermocouples. The actual gas flowrate used in the simulator is rounded to 0.1 NL/min readability to obtain $y_{\dot{V}_{dryer}}$, to meet the accuracy of the flowmeter installed in the prototype carousel. Filtrate weight measurements have white noise of standard deviation of 0.05 g, as in the scale installed in the prototype carousel used for data collection in this work (§5.2). The sampling time of the measurements and of the states made available to the user can be freely adjusted, and is fixed to 0.25 s in this study.

5.3.2 Mathematical modeling

The filtration and deliquoring models implemented in the simulator are basically the same of those described in Chapter 4, respectively in §4.3.3 and §4.3.4. The only change regards the inclusion in the filtration model of an equation for calculating the dynamic profile of filtrate mass, a required output of the carousel simulator (Table 5.2), but not directly available from the equations in §4.3.3:

$$M_{filt}(t) = \frac{-b_{filt}+\sqrt{b_{filt}^2-4a_{filt}c_{filt}}}{2\,a_{filt}\,\rho_L}\,, \tag{5.1}$$

where $M_{filt}(t)$ is the filtrate mass collected from the filtration onset up to time t, ρ_L is the filtrate density, and:

$$a_{filt} = \frac{\alpha\,\mu_l}{2\,A^2}\frac{V_{slurry}\,c_{slurry}}{V_{filt,final}} \tag{5.2}$$

$$b_{filt} = \frac{R_m\mu_l}{A} \tag{5.3}$$

$$c_{filt} = -\,\Delta P\;t\,, \tag{5.4}$$

in which α is the specific cake resistance, μ_l is the liquid viscosity, V_{slurry} is the fed slurry volume,

c_{slurry} is the crystals concentration in the fed slurry, A is the filter cross-section, $V_{filt,\,final}$ is the volume of filtrate at the end of filtration (Equation 4.11), R_m is the filter mesh resistance and ΔP is the pressure gradient across the port.

The filtration model is implemented as a MATLAB function, whereas the deliquoring and drying models are coded in a C environment and interfaced with MATLAB through a C-MEX function, to enhance the computational speed.

The drying component is derived from the model in §4.3.6, but it has been updated based on the experimental results of §5.4.3 to better reproduce the behavior of the carousel, as outlined in the remainder of this section. The drying model is based on the material balance on the cake of the $N_{L,vol}$ volatile species (Equation 5.5), on the material balance in the gas phase of the $N_{L,vol}$ volatile species (Equation 5.6), on the energy balance of the cake (considered as a pseudo-homogeneous phase combining liquid and solid; Equation 5.7) and on the energy balance of the gas phase (Equation 5.8), all dynamic one-dimensional partial differential equations:

$$\frac{\partial}{\partial t} w_{i,cake} = -\frac{\dot{m}_i^{\mathrm{L\to G}}}{\rho_{cake}}, \quad \text{for } i = 1, \dots, N_{L,vol} \tag{5.5}$$

$$\rho_g \epsilon (1-S) \frac{\partial}{\partial t} w_{i,g} = -\rho_g u_g \frac{\partial}{\partial z} w_{i,g} + \dot{m}_i^{L\to G}, \quad \text{for } i = 1, \dots, N_{L,vol} \tag{5.6}$$

$$\left(\rho_s c_{p,s}(1-\epsilon) + \rho_l c_{p,l} \epsilon S\right) \frac{\partial T_{cake}}{\partial t} = h_T a \left(T_g - T_{cake}\right) + \sum_i^{N_{L,vol}} (\dot{m}_i^{L\to G} \lambda_i) + \dot{Q}_s \tag{5.7}$$

$$\left(\rho_g c_{p,g} \epsilon (1-S)\right) \frac{\partial T_g}{\partial t} = -h_T a (T_g - T_{cake}) - u_g c_{p,g} \rho_g \frac{\partial T_g}{\partial z}, \tag{5.8}$$

where $w_{i,cake}$ is the mass fraction of i in the cake, ρ_{cake} is the cake density ($\rho_{cake} = \rho_s(1-\epsilon) + \rho_l S \epsilon$), $\dot{m}_i^{\mathrm{L\to G}}$ is the drying rate [kg/m^3], ρ_g is the gas density, $w_{i,g}$ is the mass fraction of i in the gas phase, u_g is the superficial gas velocity, $c_{p,s}$ is the solid phase specific heat, $c_{p,l}$ is the liquid phase specific heat, T_{cake} is the temperature of the cake, a is the cake specific surface, T_g is the gas temperature, λ_i is the latent heat of vaporization of species i, $\dot{Q}_s$ is the heat exchange between the cake and the environment (namely the heat loss through the dryer walls, assumed equal to zero in this work), and $c_{p,g}$ is the gas phase specific heat.

For simulation purposes, u_g is calculated from through the Darcy law for mono-phase gas flow in a porous medium (Muskat and Meres, 1936):

$$u_g = -\frac{1}{\alpha\, \rho_s\, (1-\epsilon)\, \mu_g} \frac{\Delta P}{H_{cake}}, \tag{5.9}$$

As a remark, all the variables of Equations 5.5-5.8 are local variables, that vary with both time and z. Note that the material balances in the cake for the non-volatile components of the cake are not solved, since their concentration in the cake is constant during drying. For the same reason, the species material balance is solved only for the volatile components of the cake.

The local drying rate $\dot{m}_i^{\mathrm{L\to G}}$ [kg/(m^3 s)] for $i = 1, \dots, N_{L,vol}$ is computed as (Burgschweiger and Tsotsas, 2002):

$$\dot{m}_i^{\mathrm{L}\to\mathrm{G}} = \begin{cases} h_{M,I}\, a\left(P_{i,sat} - P_{i,g}\right)\eta_i & if\ P_{i,sat} > P_{i,g} \\ 0 & if\ P_{i,sat} \le P_{i,g} \end{cases}, \text{ for } i = 1, \dots, N_{L,vol} \tag{5.10}$$

where $P_{i,sat}$ is the saturation pressure, $P_{i,g}$ is the partial pressure of i in the gas phase, and η_i is the effectiveness factor, accounting for internal mass transfer limitations, varying from zero and one. Function $\eta_i(w_{i,cake,4})$ is equal to one when $w_{i,cake,4}$ is larger than a certain critical value $w_{i,cake}^{crit}$, and starts decaying when $w_{i,cake,4}$ drops below $w_{i,cake}^{crit}$, eventually reaching a null value when $w_{i,cake,4}$ approaches an equilibrium value $w_{i,cake}^{eq}$. Suitable experiments (Burgschweiger and Tsotsas, 2002) should be conducted for identifying $\eta_i(w_{i,cake,4})$, $w_{i,cake}^{crit}$ and $w_{i,cake}^{eq}$ for $i = 1, 2, \dots, N_L$.

The main difference between the model of Equations 5.5-5.8 and the model in §4.3.6 is that here the assumption of fast heat transfer (i.e., temperature equilibrium) between the gas phase and the cake is dropped. Following a more general approach the gas and the cake temperature , are separately computed through two distinct energy balances (Equations 5.7 and 5.8), connected by the heat transfer term.

Conventional correlations for heat and mass transfer in packed beds (Bird et al., 2006; Treybal, 1980), lead to overestimation of heat and mass transfer coefficients when applied to systems with small particles, such as crystals cakes, in which the Reynolds number (proportional to d_{32}, the Sauter mean diameter of the crystal size distribution) can assume abnormally low values. Different reasons have been proposed to explain this phenomenon (Cornish, 1965; Gabor et al., 1985; Glicksman and Joos, 1980). In the simulator, we use correlations that describe the reduction of heat and mass transfer in packed beds composed of fine particles by considering channeling phenomena, affecting the fluid flow (Kunii and Suzuki, 1967):

$$h_{M,i,cake} = \frac{\rho_g u_g}{a\ \xi\, d_{32}\, P} \tag{5.11}$$

$$h_{T,cake} = \frac{c_{p,g}\rho_g u_g}{a\ \xi\, d_{32}}\ , \tag{5.12}$$

where ξ is the channeling parameter, which should be regressed from experimental data. In Equations 5.7-5.9, $h_{M,i}$ and h_T are, respectively, calculated by multiplying the mass ($h_{M,i,cake}$) and heat ($h_{T,cake}$) transfer coefficients found from Equations 5.11-5.12 by the corresponding multiplicative disturbances ($h_{M,i}^{dist}$ and h_T^{dist}).

The drying model of Equations 5.5-5.12 presents $1+N_{L,vol}$ partial differential equations, which are integrated with MATLAB's ode15s solver upon semi-discretization along z (grid spacing equal to 0.3 mm) with a first-order upwind scheme. The simulator outputs related to drying in Table 5.2 are directly obtained from the integration, or indirectly through trivial mass balances. The initial conditions of the drying model are the outputs of the filtration or of the deliquoring model. The initial T_g and T_{cake} correspond to room temperature for every z. The boundary conditions are given by the drying gas inlet composition, input of the simulator (Table 5.2), and the temperature of the drying gas entering the cake $T_g(t, z = 0)$.

Preliminary experiments drying experiments on the carousel showed that, due to the heat transfer from the air flowing into the dryer to the walls, the gas temperature T_{HL} on top of the cake is lower than T_{in}^{meas}, the gas temperature measured by TI101 (Figure 5.2). To reproduce this phenomenon in the simulator, we introduce an equation for calculating T_{HL}^{calc}, the calculated temperature of the drying gas on top of the cake (Figure 5.2), used as boundary condition in the drying model by imposing $T_g(t, z = 0) = T_{HL}^{calc}$. The dynamic macroscopic energy balance for the gas flowing in the part of the dryer above the cake reads, under the simplifying assumption of perfect mixing:

$$\frac{d(\rho_g c_{p,g} V_{empty}\, T_{HL}^{calc})}{dt} = \rho_g u_g c_{p,g} A_{dryer}^{cross}\left(T_{in}^{calc} - T_{HL}^{calc}\right) - h_{HL} A_{empty}^{lat}(T_{in}^{calc} - T_{room})\,, \tag{5.13}$$

where V_{empty} is the volume of the empty portion of the dryer, A_{dryer}^{cross} is the cross-section of the dryer, h_{HL} is the heat transfer coefficient between the drying gas and the dryer walls, A_{empty}^{lat} is the surface of the empty part of the dryer, and T_{room} is the room temperature during the experiment. In Equation 5.13, V_{empty} and A_{empty}^{lat} depend on H_{cake}. Since, even with highly concentrated slurries, H_{cake} is much smaller (<20%) than the dryer height, we simplify Equation 5.13 by assuming that, for every H_{cake}, $V_{empty} = V_{dryer}$ (the dryer volume), and $A_{empty}^{lat} = A_{dryer}$ (the dryer surface). Practically speaking, this assumption means that the temperature decrease from the dryer inlet to the top of the cake does not depend on H_{cake}. Rearranging Equation 5.13 with the introduced simplifying assumption:

$$\frac{V_{dryer}}{u_g\, A_{dryer}^{cross}} \frac{dT_{HL}^{calc}}{dt} = T_{in}^{calc} - T_{HL}^{calc} - \frac{h_{HL} A_{dryer}^{lat}}{\rho_g u_g c_{p,g} A_{dryer}^{cross}}\left(T_{in}^{calc} - T_{room}\right) \tag{5.14}$$

Introducing the time constant τ_{HL} and the heat loss coefficient α_{HL}, Equation 5.14 is reformulated as:

$$\tau_{HL} \frac{dT_{HL}^{calc}}{dt} = T_{in}^{calc} - T_{HL}^{calc} - \alpha_{HL}\left(T_{in}^{calc} - T_{room}\right), \tag{5.15}$$

with:

$$\tau_{HL} = \frac{V_{dryer}}{u_g\, A_{dryer}^{cross}} \tag{5.16}$$

$$\alpha_{HL} = \frac{h_{HL} A_{dryer}^{lat}}{\rho_g u_g c_{p,g} A_{dryer}^{cross}} \tag{5.17}$$

Discretizing Equation 5.15 with respect to time, at time step t_k:

$$T_{HL}^{calc}(t_k) = \left(1 - \alpha(t_k)\right) T_{HL}^{calc}(t_{k-1}) + \alpha(t_k)\left(T_{in}^{calc} - \beta_{HL}\left(T_{in}^{calc} - T_{room}\right)\right), \tag{5.18}$$

where:

$$\alpha(t_k) = \frac{t_k - t_{k-1}}{\tau_{HL} + t_k - t_{k-1}} \tag{5.19}$$

Equation 5.18 is initialized by setting at $T_{HL}^{calc} = T_{room}$ at $t_k = 0$ (beginning of each carousel cycle). From Equations 5.16-5.17, and neglecting the weak dependence of α_{HL} on temperature and pressure, u_g is the only variable on which τ_{HL} and α_{HL} depend. We introduce an additional simplifying assumption, by reformulating Equations 5.16-5.17 in linear terms with respect to u_g, and introducing calibration parameters $\tau_{HL,1}$, $\tau_{HL,2}$, $\alpha_{HL,1}$ and $\alpha_{HL,2}$. Equations 11 and 12, respectively, become:

$$\tau_{HL} = \tau_{HL,1} + \tau_{HL,2} u_g \quad (5.20)$$

$$\alpha_{HL} = \alpha_{HL,1} + \alpha_{HL,2} u_g \quad (5.21)$$

From a sensitivity analysis, $\tau_{HL,2}$ does not significantly impact the model predictions, and is set equal to zero for reducing the number of calibration parameters. Overall, in the simulator, T_{HL}^{calc} is calculated through Equations 5.18-21, which present three calibration parameters ($\tau_{HL,1}$, $\alpha_{HL,1}$ and $\alpha_{HL,2}$), and depend on u_g and T_{in}^{calc}.

5.4 Carousel simulator calibration

5.4.1 Overview

The simulator is calibrated to a process for isolating paracetamol from a paracetamol/ethanol slurry, with 25%w crystals concentration. The considered paracetamol, used for the experiments reported in this section, comes from the same batch of the one used for the experiments of Chapter 4. The size distribution of the crystals used for the experiments is characterized with a Mastersizer 3000 by Malvern Panalytical (Malvern, UK), yielding d_{32} equal to 47.2 μm and a specific surface a equal to 100.9 m^2/kg. The crystal size distribution has been reported in Figure 4.5. No impurities are present in the slurry; hence the liquid phase is solely composed by ethanol, and the solid phase is constituted by pure paracetamol crystals.

Simulator calibration consists in assigning the thermophysical properties and the calibration parameters of Table 5.2, together with providing suitable profiles for the disturbances.

The thermophysical properties of liquid ethanol are obtained from Perry's Chemical Engineers' Handbook (Green and Southard, 2019). While paracetamol crystals density is taken from the literature (Haynes, 2014), we consider the specific heat as a calibration parameter, to improve the agreement of the simulator predictions with drying experiments results (Section §5.4.4). We obtain the cake properties α_{cake}, ϵ_{cake} and S_∞ (deliquoring equilibrium saturation §4.3.4) from experimental data (§5.4.1), too. Air is selected as the drying gas. Thermophysical properties of the pure components of the gas phase and the saturation pressures are also taken from Perry's (Green and Southard, 2019), resorting to temperature-dependent relations. Ideal gas mixing rules are used for calculating the gas phase properties from the pure components properties.

Table 5.3 shows the values of the parameters estimated from experiments in this study, including both the default calibration parameters of the simulator (Table 5.2), the thermophysical properties treated as calibration parameters following the above discussion ($c_{p,s}$, S_∞, α_{cake}, and ϵ_{cake}), and additional parameters estimated from the experiments, although not directly needed by the simulator. Note that all the estimated parameters pass the *t*-test for statistical significance. The next subsections detail the experimental and the model calibration procedures followed for reaching the parameter estimation results reported in Table 5.3. A final subsection describes the disturbances implemented in the simulator.

Table 5.3. *Simulator calibration: parameter estimation results from filtration, drying and heat loss experiments.*

Symbol	Parameter	Unit	Estimated value	Estimated from	Estimation technique	*t*-value (95% CL)*
$c_{p,s}$	Paracetamol crystals specific heat	J/(kg K)	2267	Drying experiments	MLE	3.69
$R_{m,1}^{filt1}$	Filter mesh resistance – filtration experiment #1	1/m	7.6E9	Filtration experiments	MLE	63.58
$R_{m,1}^{filt2}$	Filter mesh resistance – filtration experiment #2	1/m	4.7E9	Filtration experiments	MLE	11.67
$R_{m,1}^{filt3}$	Filter mesh resistance – filtration experiment #3	1/m	4.9E9	Filtration experiments	MLE	26.59
$R_{m,1}^{filt4}$	Filter mesh resistance – filtration experiment #4	1/m	3.8E9	Filtration experiments	MLE	77.70
S_0^{dry1}	Initial ethanol content – drying experiment #1	-	0.089	Drying experiments	MLE	2.36
S_0^{dry2}	Initial ethanol content – drying experiment #2	-	0.132	Drying experiments	MLE	2.10
S_∞	Cake deliquoring equilibrium saturation	-	0.085	Drying experiments	Qualitative estimation	-
α_{cake}	Specific cake resistance	m/kg	2.6E9	Filtration experiments	MLE	97.59
$\alpha_{HL,1}$	Air-wall heat transfer coefficient - intercept	-	0.70	Heat loss experiments	MLE	254.68
$\alpha_{HL,2}$	Air-wall heat transfer coefficient - slope	s/m	-0.72	Heat loss experiments	MLE	60.52
ϵ_{cake}	Cake porosity	-	3.5E-1	Filtration and drying experiments	Measurement (§2.2)	-
η_{EtOH}	Drying effectiveness factor for ethanol	-	Equation 18	Chapter 4	-	-
ξ	Tunneling parameter	-	199.8	Drying experiments	MLE	3.66
$\tau_{HL,1}$	Air-wall heat transfer time constant	s	68.9	Heat loss experiments	MLE	50.34

*Reference *t*-value at 95% confidence = 1.65.

5.4.2 Filtration: experiments and model calibration

Four filtration experiments (#F1-4) are carried out by loading the crystallization slurry into the first port of the carousel setup (§5.2), and activating vacuum pump P101. Filtrate time profiles (Figure

5.3) are recorded through WI101 (Mettler Toledo ML6002T precision balance, of readability 0.01 g). The filtration pressure gradient is also measured (through PI101) and recorded. In each experiment, a different volume of slurry in the range 4-10 mL is fed to the carousel, to better characterize the filtration behavior across the whole span of slurry volumes that are fed to the carousel during normal operation. Before every filtration experiment, pure ethanol is loaded into the first port of the carousel and filtered out through the vacuum pump, to clean up the mesh and restore similar initial mesh

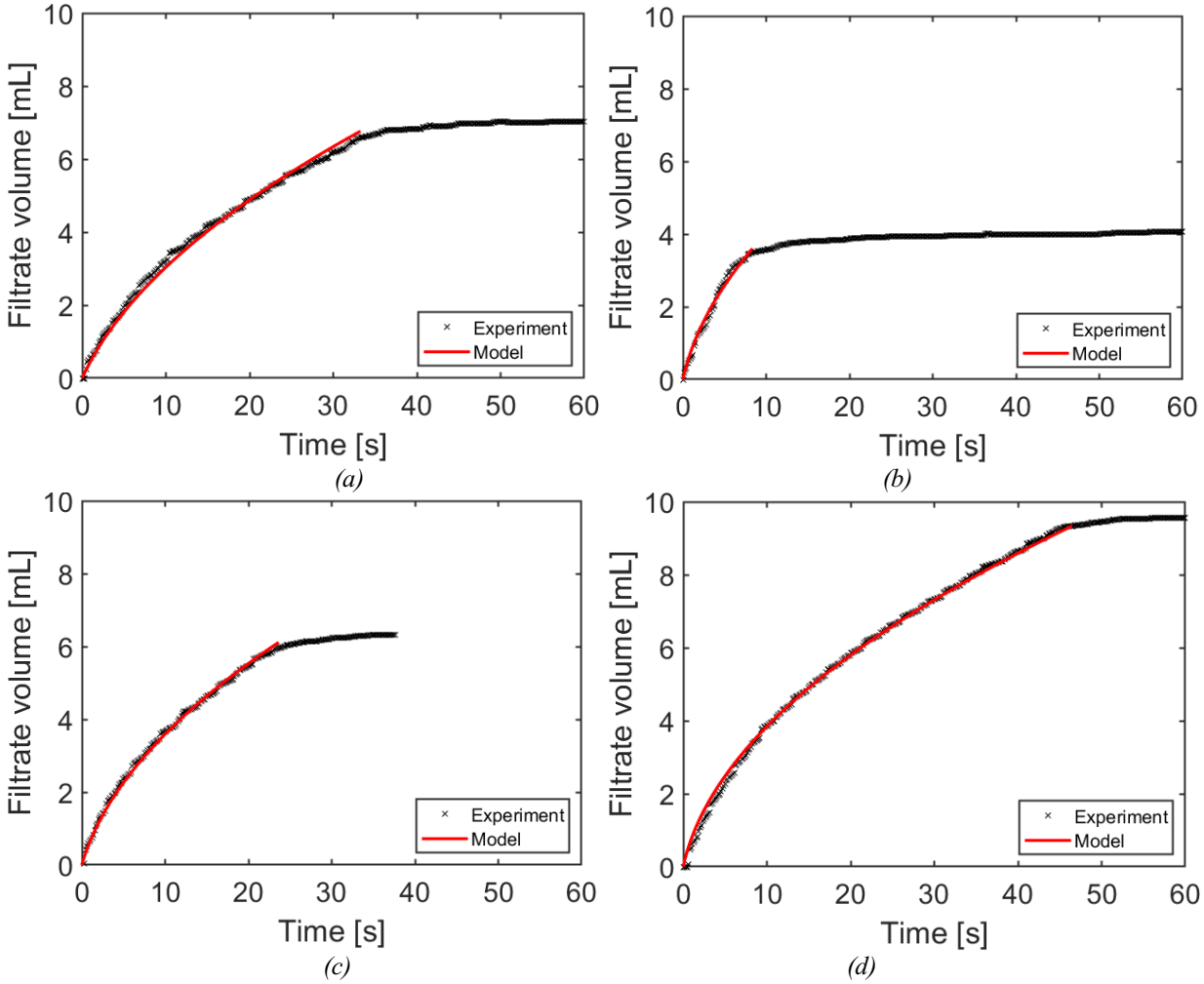


Figure 5.3. *Filtration experiments: measured filtrate volume vs. model prediction for: (a) #F1, (b) #F2, (c) #F3, and (d) #F4. Experimental values are reported also for deliquoring, while model predictions stop at the end of filtration.*

conditions for every run.

Filtration experiments #F1-4 are used for estimating α_{cake} (Table 5.3), fitting (Figure 5.3) the filtrate weight profiles with the filtration model (§5.3.2) through maximum likelihood estimation (MLE). The pressure gradient profiles measured during the experiments with PI101 are used as model inputs. Even though the filter mesh is cleaned after every run, independent filter mesh resistance parameters ($R_{m,1}^{filt1}$, $R_{m,1}^{filt2}$, $R_{m,1}^{filt3}$ and $R_{m,1}^{filt4}$) are regressed for the four experiments. A non-negligible variability of the fouling conditions after mesh cleaning is registered (Table 5.3). The estimated α_{cake}, equal to

2.6E9 m/kg, is aligned to specific cake resistances previously found on the analogous paracetamol/ethanol slurry analyzed in Chapter 4 (§4.4).

5.4.3 Heat loss in dryer: experiments and model calibration

Heat loss experiments are carried out to calibrate the heat loss model of Equations 5.18-5.21, used in the simulator for calculating the gas temperature on top of the cake T_{HL}^{calc}. Four (#H1-4) "heat loss experiments" (Figure 5.4) are carried out, at different air inlet temperature and air flowrate conditions, to study their impact on the heat transfer phenomena. The experiments are carried out by flowing hot air in the empty dryer, with no material in the port, and recording the air inlet T_{in}^{meas} and outlet temperatures T_{out}^{meas}, and the air flowrate $\dot{V}_{in}^{meas}$. The air inlet temperature set-point T_{sp}^{meas} is changed stepwise (Figure 5.4b,c,d), to cover the temperature interval at which paracetamol is

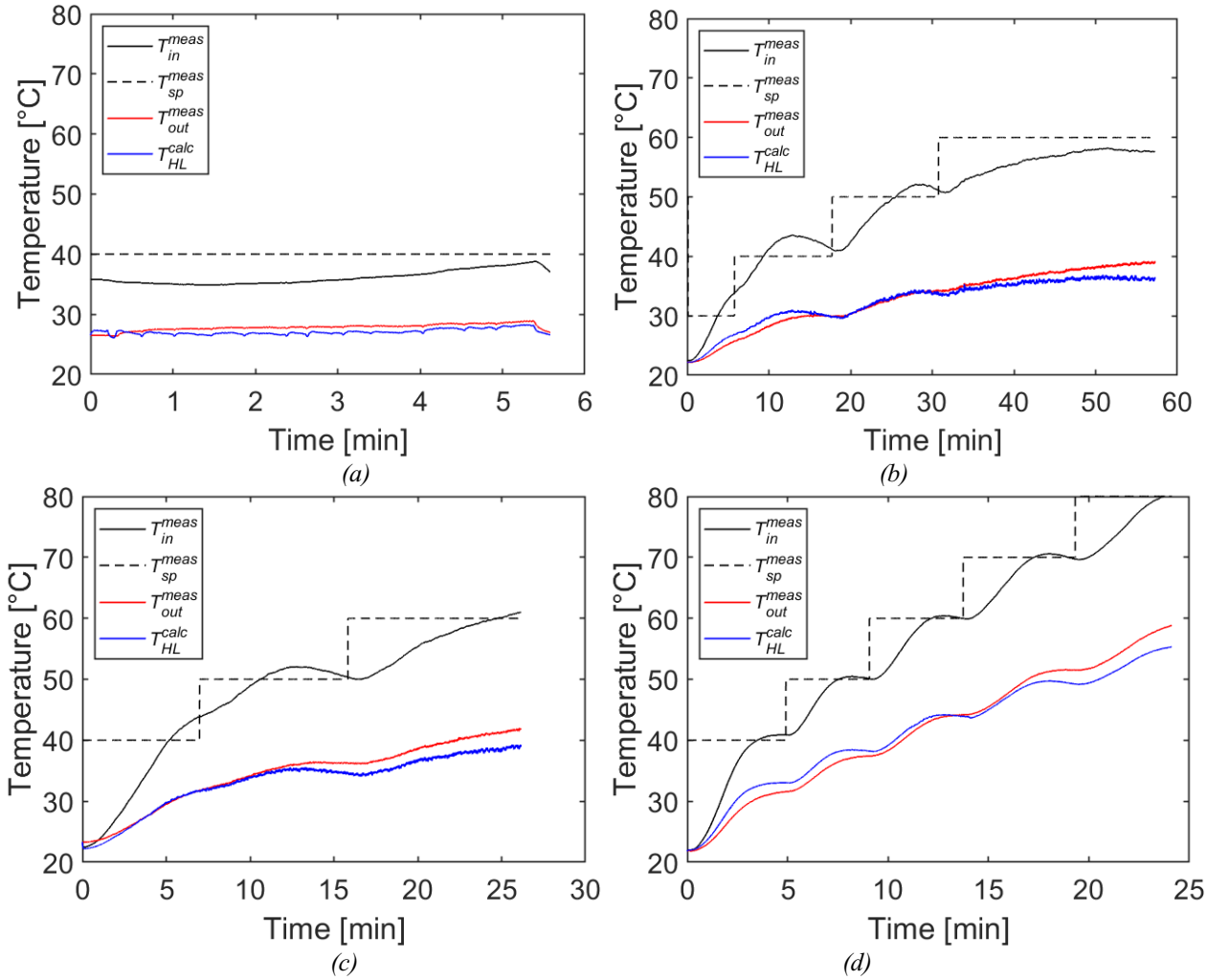


Figure 5.4. *Heat loss experiments. measured inlet drying air temperature (T_{in}^{meas}), inlet drying air temperature set-point (T_{sp}^{meas}), and of outlet drying air temperature (T_{out}^{meas})vs model predictions of outlet drying air temperature ($T_{HL}^{calc} = T_{out}^{meas}$) for: (a) #H1, (b) #H2, (c) #H3 and (d): #H4.*

typically, dried (30-80°C). Note that, in the carousel prototype, the air flowrate is not a controlled variable, and it depends on the local mesh fouling conditions and on the cake height. The air flowrate measured during heat loss experiments (not plotted here for conciseness) fluctuates within small ranges, as in the drying experiments of Figures 5.5c,d. For heat loss experiments #H1-4, the ranges are, respectively: 0.8-1.4 NL/min, 1.3-1.8 NL/min, 1.9-2.5 NL/min, and 4.1-4.4 NL/min. The air flowrate ranges of experiments #H1-4 cover all the range of variability encountered during carousel operation. The air flowrate range of #H1 corresponds to severe mesh fouling conditions, when very little gas flows. On the other hand, the air flowrate registered during #H4 corresponds to the maximum flowrate that can be obtained for the current carousel/vacuum pump setup, and corresponds to completely cleaned mesh conditions. Mesh fouling conditions of #H2 and #H3 are intermediate between those of #H1 and #H4.

Experiments #H1-4 are used for estimating $\tau_{HL,1}$, $\alpha_{HL,1}$ and $\alpha_{HL,2}$ (Equations 5.18-5.21), through MLE, by fitting T_{out}^{meas} with T_{HL}^{calc}. T_{in}^{meas} and $\dot{V}_{in}^{meas}$ profiles are used as inputs to the model ($\dot{V}_{in}^{meas}$ is used for calculating the actual u_g, instead of resorting to Equation 5.9). A good fitting is achieved (Figure 5.4), and all estimated parameters pass the *t*-test (Table 5.3). Note that in the drying experiments of Figures 5.5a and 5.5b, at the end of drying, T_{out}^{meas} approaches T_{HL}^{calc}, validating the calibrated heat loss model.

5.4.4 Drying

Two drying experiments are performed using the same slurry employed for filtration, for calibrating the drying model. Air is used as drying gas. Drying experiments are carried out as follow. A slurry volume of 8 mL is fed to the carousel, filtered for 40 s in the first station, and then filtered for other 40 s in the second one. Immediately afterwards, the wet cake is directly transferred to Station 4 through carousel rotations, for the drying step. The measurements of air inlet temperature (T_{in}^{meas}), air outlet temperature (T_{out}^{meas}) and air inlet flowrate ($\dot{V}_{in}^{meas}$) during the drying are recorded (Figure 5.5). Note that, as for heat loss experiments, the air flowrate is not a manipulated variable in the carousel prototype. The set-point of the air inlet temperature T_{sp}^{meas} is set to 50 °C for experiment #D1, and to 70 °C for experiment #D2. However, the heating system of the prototype carousel cannot quickly heat the inlet air up to the set-point: a slowly increasing T_{in}^{meas} profile, starting from ambient temperature, is registered during the experiments (Figures 5a and 5b). At the beginning of drying, the outlet temperature T_{out}^{meas} drops (Figure 5a,b), due to the latent heat of vaporization of ethanol. As the drying rate slows down, T_{out}^{meas} starts rising again. The cakes collected at the end of the drying step are discharged from the carousel, weighed and placed in a convective oven at 30 °C. Subsequent measurements after 24 h revealed no change of weight, meaning that no residual ethanol was present in the cake at the end of drying in the carousel. The cakes obtained from filtration and drying experiments present a porosity of 0.35, consistently with results of Chapter 4 (§4.4).

Parameters $c_{p,s}$ and ξ are estimated by making use of experiments #D1-2. Since the actual initial cake

saturations in the experiments (S_0^{dry1} and S_0^{dry2}) are unknown, they are considered additional parameters to be estimated. The drying effectiveness factor η_{EtOH} for ethanol and the relevant critical ethanol contents in the cake are taken from §4.4, as derived (Equation 4.37) from thermogravimetric analysis experiments on an analogous paracetamol/ethanol drying cake. $\dot{V}_{in}^{meas}$ and T_{in}^{meas} profiles are used as inputs to the model. Variable $\dot{V}_{in}^{meas}$ is used to directly calculate u_g, instead than resorting to Equation 5.9. The heat loss model (Equations 5.18-5.21) is used to calculate T_{HL}^{calc} from T_{in}^{meas} and u_g.

The measured air outlet temperature profiles T_{out}^{meas} are fitted with the calculated outlet air temperature profiles $T_{out}^{calc}(= T_g(t, z = H_{cake}))$ through MLE (Figures 5.5a and 5.5b). The estimated set of parameters leads to predicting the temperatures profiles of the cake at the bottom $T_{cake,bottom}^{calc}$, reported in Figures 5a and 5b. However, since in the experiments the cake temperature is not measured in any point, $T_{cake,bottom}^{calc}$ cannot be validated, and the estimated ξ, $c_{p,s}$, S_0^{dry1} and S_0^{dry2}

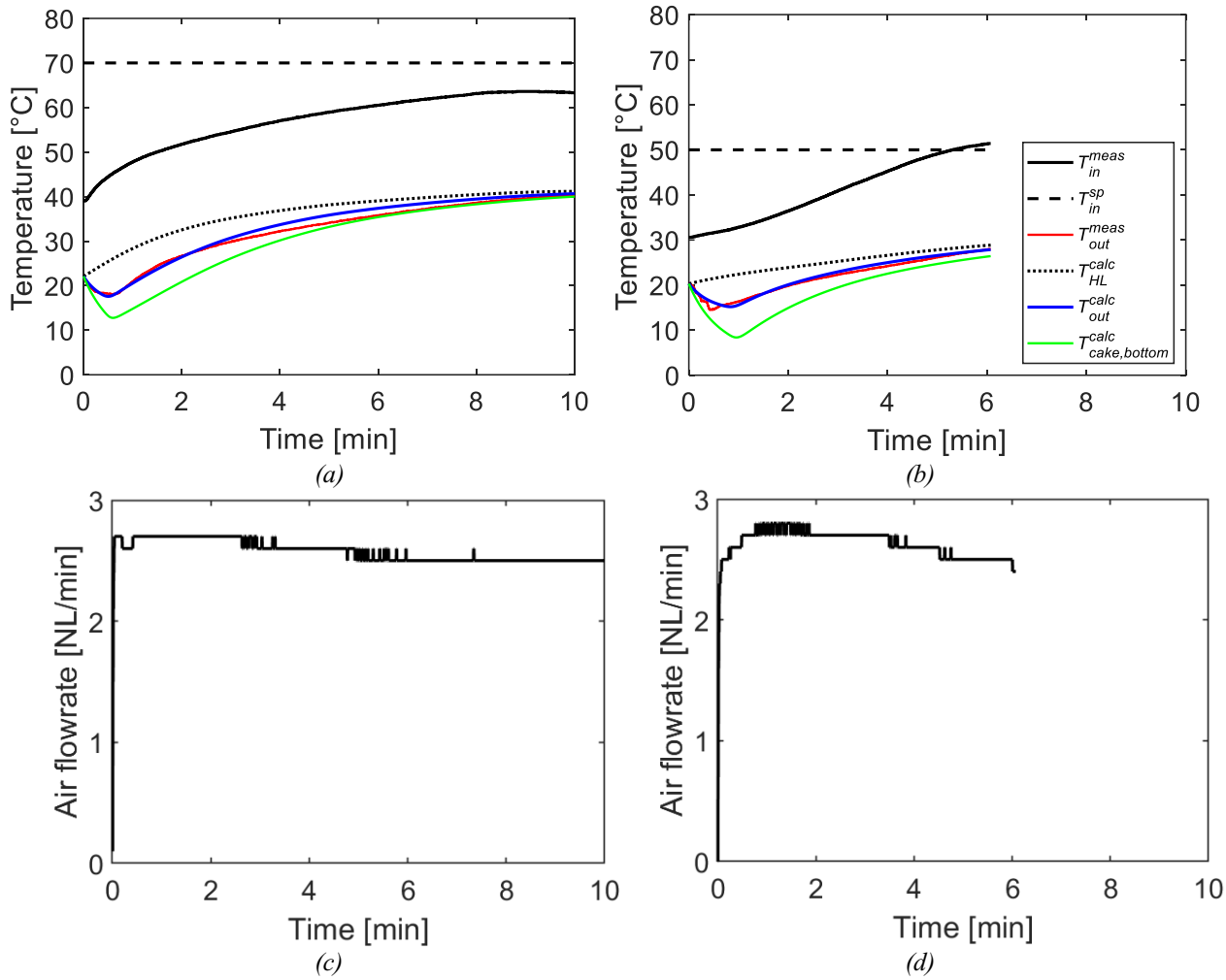


Figure 5.5. *Drying experiments. Measured drying air inlet (T_{in}^{meas}) and outlet (T_{out}^{meas})temperatures, drying air inlet temperature set-point (T_{sp}^{meas}) vs model predictions of drying air temperature at cake top T_{HL}^{calc}, of drying air outlet temperature T_{out}^{calc}, and of drying air temperature at cake bottom $T_{cake,bottom}^{calc}$ for (a) #D1, and (b) #D2. The measured air flowrate $\dot{V}_{in}^{meas}$ is also reported for (c) #D1 and (d) #D2.*

are not univocally identifiable from the available set of measurements. For instance, if the actual cake temperature drop due to drying were lower, the optimal set of estimated parameters would be different. In any case, the experiments confirm the assumption introduced in §5.3.2 that the drying gas and the cake are not in (local) thermal equilibrium. Actually, the heat of vaporization spent by the solvent vaporized during experiments #D1-2 is much larger than the sensible heat lost by the gas and the cake, if calculated from a heat balance based on T_{out}^{calc} under the assumption of local thermal equilibrium between gas and cake. Since the energy for the vaporization of the solvents during drying can only come from the gas and from the cake, the only possible explanation is that, during drying, the cake reaches lower temperatures than the gas, providing to the vaporizing solvents the energy necessary to close the heat balance. This conclusion is also supported by the compatibility of the difference between $T_{cake,bottom}^{calc}$ and T_{out}^{calc} with values found in the literature for co-current rotary drying (Chhabra and Basavaraj, 2019) and with experiments on a similar convective cake drying process (Li, 2014). Considering also that it is not possible to insert thermocouples in the cake in the available carousel setup, and that the purpose of the simulator is providing a realistic digital framework for control strategy testing, we retain the estimated set of parameters of Table 5.3 in the simulator.

All the estimated parameters pass the *t*-test (Table 5.3). The estimated ξ is compatible with values from the literature (Kunii and Suzuki, 1967). The estimated $c_{p,s}$ is slightly larger than values found in the literature (Harada et al., 2013); this looks reasonable, because this parameter is partly compensating for inaccurate modeling of the energy balance. S_0^{dry1} and S_0^{dry2} are estimated to be, respectively, 8.9 % and 13.2 %, values close to typical deliquoring equilibrium saturations (Ripperger et al., 2013). Based on these results, S_∞ is assigned a value of 8.5 % (the weak dependence of S_∞ on the pressure gradient is neglected).

5.4.5 Disturbances

The disturbances of Table 5.2 assume a new value at every carousel cycle, but do not vary within the same cycle. The only exceptions are the liquid phase of the slurry, always fixed to pure ethanol, and the drying gas inlet composition, always fixed to air's. In this study, for simulations in normal operating conditions of the process, we assign Gaussian distributions $N(\mu, \sigma^2)$ of mean $\mu = 1$ [-] and variance $\sigma^2 = 4E\text{-}4$ [-] to h_M^{dist}, h_T^{dist}, V_{slurry}^{dist}, and ϵ^{dist}. Assigned a mean slurry concentration $\mu_{c_{slurry}}$ desired for a certain simulation, c_{slurry} is calculated as $c_{slurry} = N(\mu_{c_{slurry}}, \mu_{c_{slurry}}^2 4E\text{-}4)$ kg/m^3.

A realistic routine simulating filter mesh fouling and cleaning-in-place is implemented in the simulator (Figure 5.6). $R_{m,1}$, $R_{m,2}$, $R_{m,3}$, and $R_{m,4}$ are first initialized, sampling each one from N(3E9, 1E18) [1/m]. The chosen distribution of resistance of clean meshes reproduces the variability estimated from the filtration experiments (Table 5.3). At the end of each cycle during which some material is processed in Station i (i.e., when the station is not empty), the corresponding $R_{m,i}$ increases

of 2E9 1/m. During the first cycle of carousel operation, only the first processing station works, as no material has reached the subsequent stations, yet. Just from the fourth cycle on, all stations V102-V105 will be fully operational. When the filter mesh resistance reaches a threshold value (set to 1.2E10 1/m), the mesh cleaning-in-place procedure is triggered. No additional material is loaded into the carousel for the following three cycles, which are required to complete the processing and discharge of the batches of slurry trapped into the carousel ports. Then, the filter mesh resistances are re-initialized, sampling again from N(3E9, 1E18) [1/m], and the feeding and fouling routines are repeated again. Note that in the version of *ContCarSim* used for generating the results presented in this Chapter, the idle time for mesh cleaning has been assumed to be null[1]. A null duration has also been assigned to the inter-cycle idle time for routines such as carousel rotation and piston ejection.

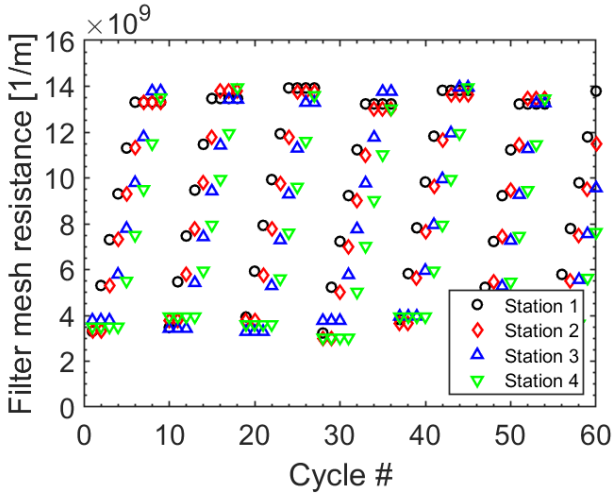


Figure 5.6. *Sample profile of filter mesh resistance disturbance implemented in the simulator.*

5.5 Quality-by-Design and Quality-by-Control challenges

The simulator proposed in this study can be used for testing different control strategies within the QbD and QbC frameworks.

The general objective of the process is delivering dry cakes meeting the target quality, namely a residual ethanol content (critical quality attribute) below 0.5%w. This is achieved by setting suitable values for the operating variables, for a given inlet slurry concentration (critical material attribute of the process). Following the QbD jargon, the critical process parameters are identified as V_{slurry}^{sp} and Δt_{cycle}^{sp}, since they significantly impact the residual ethanol in the discharged cake. $T_{in,g}^{sp}$ and ΔP^{sp}, instead, are identified as the control variables, which affect the product quality to a smaller extent, compared to the critical process parameters. Control routines can be implemented in the simulator for adjusting the desired values of the operating variables (set-point of the relevant controllers). The

[1]In the final release of *ContCarSim*, it will be possible to set non-null durations for the inter-cycle and mesh cleaning idle times.

simulator also features blocks specifically dedicated to the implementation of state and parameter estimation routines. Note that, although the simulator makes the values of multiple outputs available, only the simulated measurements can be used by the developed control and estimation routines, as in real life carousels.

The following specific challenges are envisioned for the simulator:

1. Open-loop operation. Determination of operating points delivering the target product quality and description of the design space of the unit;
2. State estimation and soft-sensing. Implementation of state estimators, soft sensors, and real time parameter estimation routines for monitoring key process variables, such as the ethanol content in the cake being dried and the resistance of the filter meshes;
3. Closed-loop control by manipulation of $T_{in,g}^{sp}$ and P_{compr}^{sp}. Implementation of (model-free and model-based) control routines for automatic adjustment of the control variables (i.e., $T_{in,g}^{sp}$ and ΔP^{sp}) to meet the target product quality in response to disturbances (in disturbance scenarios 0-2) and/or to changes in the inlet slurry concentration. Description of the design space of the unit with such closed-loop routines in place;
4. Closed-loop control by manipulation of V_{slurry}^{sp} and Δt_{cycle}^{sp}.. Implementation of (model-free and model-based) control routines for automatic adjustment of the critical process parameters (i.e., V_{slurry}^{sp} and Δt_{cycle}^{sp}), to meet the target product quality in response to disturbances (in disturbance scenarios 0-2) and/or to changes in the inlet slurry concentration;
5. Throughput maximization under a reference scenario. We define the specific objective of implementing closed-loop control routines acting on V_{slurry}^{sp} and Δt_{cycle}^{sp} to maximize the cumulative mass of the cakes (meeting the target quality) discharged by the carousel during a simulation of 1 hour duration, with c_{slurry}^{nom} equal to 250 kg/m^3, $T_{in,g}^{sp}$ fixed to 323 K, and P_{compr}^{sp} set to 10^5 Pa$_g$, and with the inter-cycle and mesh cleaning idle times set to zero. The maximum throughput that can be achieved in these conditions should be assessed under disturbance scenarios 0-2;
6. Throughput maximization under a general scenario. Implementing closed-loop control routines acting on $T_{in,g}^{sp}$, V_{slurry}^{sp}, ΔP^{sp}, and Δt_{cycle}^{sp}, to maximize the cumulative mass of the cakes (meeting the target quality) discharged by the carousel in a given timeframe, under disturbance scenarios 0-2, for different c_{slurry}^{nom} and inter-cycle and mesh cleaning idle times.

In addition to the listed tasks, the simulator can also be used for generating data for data analytics studies, or for benchmarking fault detection, identification and diagnosis methodologies.

Despite its advanced features, the simulator has some limitations. The assumption of perfect control of the operating variables may not be met in physical carousels, especially for the inlet drying air temperature. Moreover, all simulated measurements are generated without any delay or sensor dynamics. Another limitation involves filter mesh fouling. In practice, the actual fouling increase from cycle to cycle depends on the current operating conditions, especially the fed slurry volume and

concentration. Since not enough data were available for characterizing the fouling dependence on the operating conditions, the fouling schedule is kept constant in the simulator (§5.4.5).
Despite these limitations, *ContCarSim* is a realistic simulator that can be used for promoting the adoption of advanced control strategies in pharmaceutical manufacturing on the one hand, and to improve the operation of the novel carousel technology for continuous filtration-drying on the other hand.

5.6 A Quality-by-Control framework

We propose a three-layered (Layer 0, 1 and 2) control system for the carousel technology (Figure 5.7), based on the recently proposed QbC framework (Su et al., 2019b). The set-points of the control variables are fixed: ΔP^{sp} to 1 bar and $T_{in,g}^{sp}$ to 50°C. In the remainder of this chapter, the carousel simulator is used for implementing and testing the control strategy, and is referred to as "the carousel".

5.5.1 Layer 0

Layer 0 of the control system consists of the built-in controls of the carousel and of the ancillary equipment implemented in the simulator (§3.1). Conducting carousel operation with only Layer 0 in

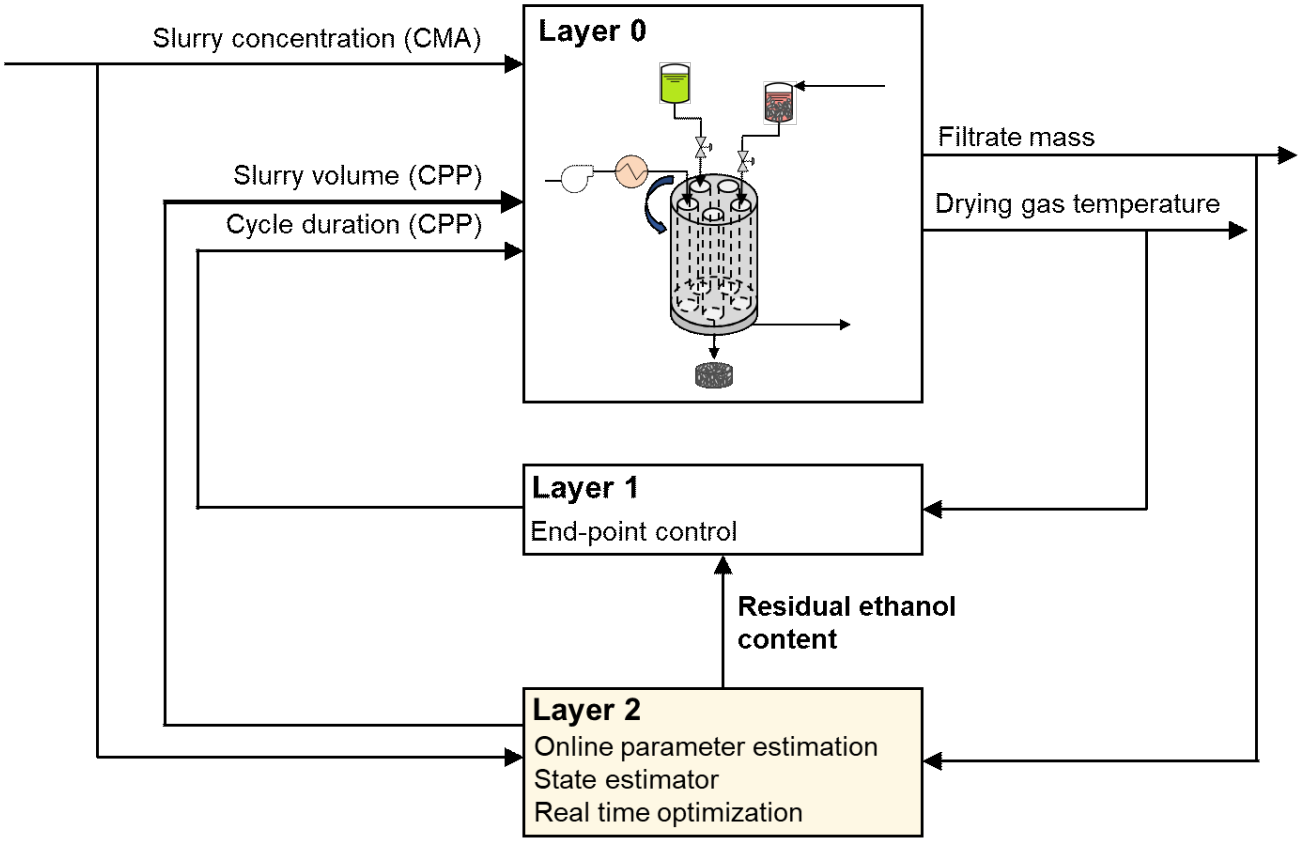


Figure 5.7. *Proposed QbC strategy for the carousel.*

place is equivalent to adopting a QbD approach, in which the CPPs are adjusted at open-loop within the DS. We determine the probabilistic DS of the process with only Layer 0 implemented, for benchmarking the three-layered control system proposed in this study with the traditional QbD open-loop approach. To this purpose, we build a three-dimensional grid whose axes are: V_{slurry}^{sp} (varying from 0.5 mL to 10 mL, step 0.5 mL), Δt_{cycle}^{sp} (varying from 5 s to 300 s, step 5 s), and $\mu_{c_{slurry}}$ (varying from 50 kg/m^3 to 250 kg/m^3, step 50 kg/m^3). For each point of the grid, we carry out a Monte Carlo simulation with 400 realizations. For each realization, the cycle-varying disturbances (c_{slurry}, h_M^{dist}, h_T^{dist}, V_{slurry}^{dist}, α^{dist}, ϵ^{dist} and filter mesh resistance) are sampled from the respective probability distributions[2] (§5.4.5). Then, the product CQA is calculated for the given set of grid point conditions and disturbances. The probability of attaining the target quality in a given grid point corresponds to the percentage of realizations satisfying the target CQA (i.e., residual ethanol content in discharged cake < 0.5 wt%). The DS (Figure 5.8) corresponds to the region of the grid where such probability is greater than 90%.

For simplicity, we assume that a fixed $\mu_{c_{slurry}}$ equal to 250 kg/m^3 comes from upstream. Figure 5.9a shows the section of the DS of Figure 5.8 corresponding to this inlet concentration. When a control strategy centered only on Layer 0 is used, the operating conditions (V_{slurry}^{sp} and Δt_{cycle}^{sp}) are assigned offline, and maintained constant during process operation, unless significant variations in c_{slurry} are

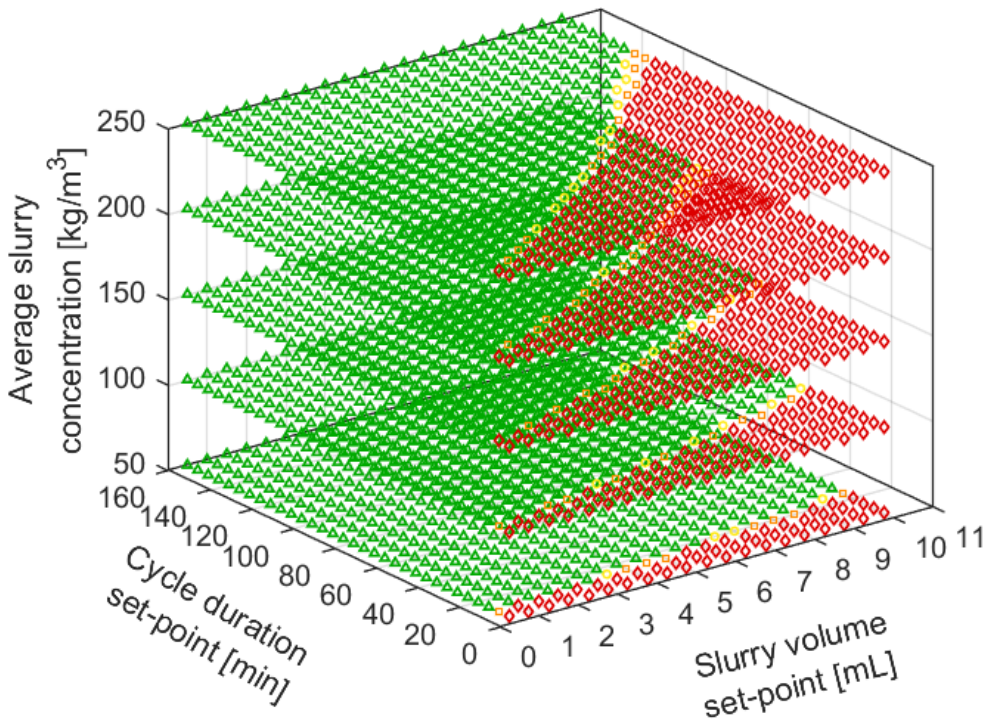


Figure 5.8. *Probabilistic DS with a control system consisting of only Layer 0. Green triangles: probability ≥ 90%, yellow circles: 80% ≤ probability < 90%, orange squares: 60% ≤ probability < 80% and red diamonds: probability < 60%.*

[2]All disturbances sampled from respective normal distribution, while the filter mesh resistance is sampled from U[3E9, 14E10] 1/m

registered. Within the DS, we select the combination of V_{slurry}^{sp} and Δt_{cycle}^{sp} that maximize the carousel throughput T at each cycle:

$$T = \frac{V_{slurry}\, c_{slurry}}{\Delta t_{cycle}} \quad . \tag{5.21}$$

The maximum T in the DS is achieved at the DS boundary. However, in order to minimize the risk of obtaining an out-of-specification product, we select an operating point (represented by a grey circle in Figure 5.9b) close to the maximum value of T, but slightly more inside the DS, namely $V_{slurry}^{sp} =$ 6 mL and $\Delta t_{cycle}^{sp} = 80$ s. Note that, if mesh cleaning did not occur instantaneously, as considered in the current implementation of the simulator (§5.4.5), the idle time for mesh cleaning would have to be accounted for in the denominator of Equation 5.21.

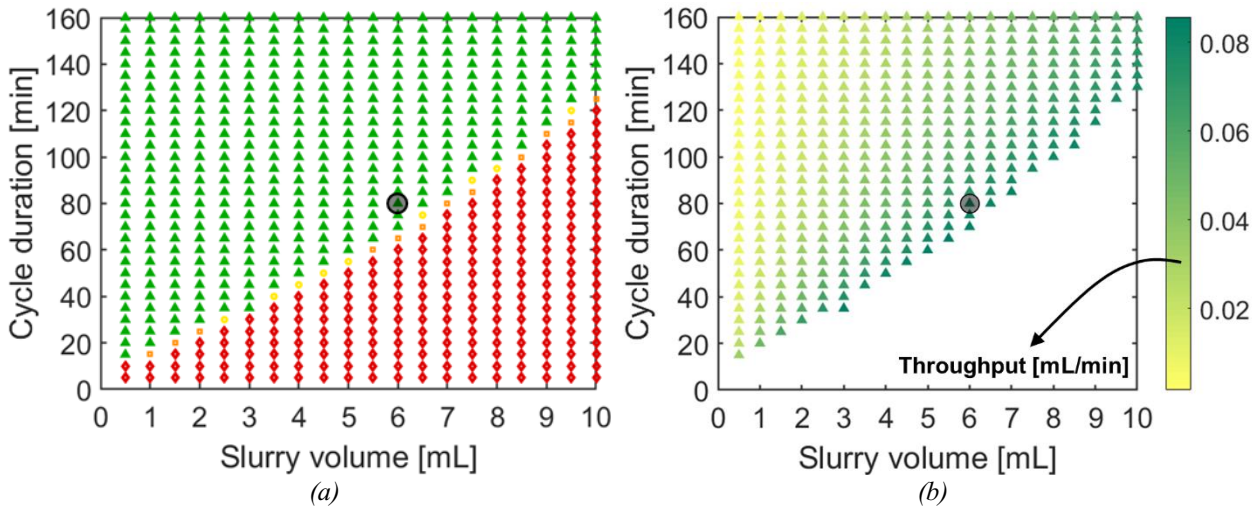


Figure 5.9. *(a)*: *probabilistic DS at a slurry inlet concentration* $\mu_{c_{slurry}} = 250$ *kg/m³ with a control system consisting of only Layer 0. (b): slurry throughput in the DS. The selected operating conditions for carousel open-loop operation are highlighted by a grey circle.*

5.5.2 Layer 1

In Layer 1 of the control system, an end-point controller is implemented to close the loop on the CQA. For the first three cycles after a cleaning-in-place procedure, cake drying is not carried out, and Δt_{cycle}^{sp} is fixed to 30 s. From the fourth cycle on, cake drying is carried out, and Δt_{cycle}^{sp} is automatically adjusted to terminate the cycle when $w_{EtOH,cake,4}$ reaches the target quality threshold. Since no real time measurements are available in the carousel for $w_{EtOH,cake,4}$, the cycle termination time is inferred from $y_{T_{out}}$. Figure 5.10 shows the relation between $w_{EtOH,cake,4}$ and $y_{T_{out}}$ for three batches of slurry, processed with filter meshes that present increasing resistances (3E9 1/m, 9E9 1/m, and 15E9 1/m). The selected values of mesh resistance correspond to the range of variability encountered during carousel operation, in between two cleaning-in-place procedures. All the other cycle-varying disturbances are kept to the mean value of the respective probability distributions. For all batches,

V_{slurry}^{sp} is set to 6 mL, the value selected in §5.5.1 for open-loop operation with Layer 0-only control. The target quality value is reached by $w_{EtOH,cake,4}$, in all the three batches, after that the air outlet temperature inversion has been registered, although for slightly different values of $y_{T_{out}}$ (Figure 5.10). We fix the cycle end-point of Layer 1 to the instant of time when a value of $y_{T_{out}}$ equal to 18.7°C is recorded after the temperature inversion. This is a conservative choice, as for clean filter meshes (e.g., resistance equal to 3E9 1/m), the cycle could be terminated earlier. Installation of a composition analyzer for the outlet drying air could further enhance the promptness of the detection of the end-point.

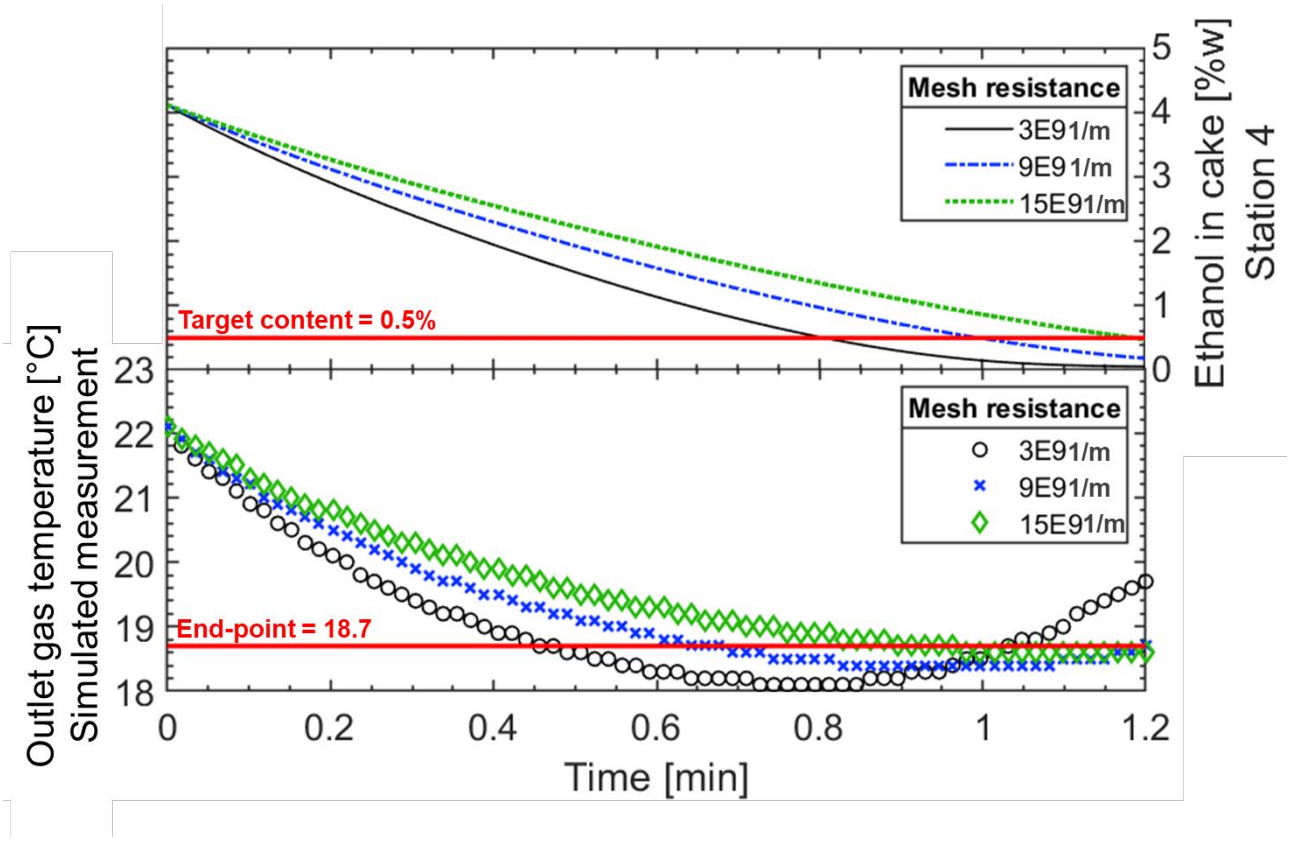


Figure 5.10. R*esidual ethanol in the cake* ($w_{EtOH,cake,4}$) *vs outlet gas temperature measurement during drying* ($y_{T_{out}}$) *for* $R_{m,4}$ *=3E9, 9E9 and 15E9 1/m.* $V_{slurry}^{sp} = 6$ *mL. Other operating conditions and settings of the simulator are described in §5.4. The temperature end-point for the Layer 1 end-point controller is also reported.*

5.5.3 Layer 2

In Layer 2 of the control system, advanced model-based techniques are implemented for process monitoring and control; namely, we use: *i)* online parameter estimation, *ii)* state estimation, and *iii)* RTO.

5.5.3.1 Online parameter estimation

At the end of every cycle, online parameter estimation is carried out for estimating the current

resistance of the mesh of the first station ($R_{m,1}^{estim}$), together with the porosity (ϵ^{estim}) and the specific resistance (α^{estim}) of the wet cake that has just formed in V102. Porosity is estimated as:

$$\epsilon^{estim} = 1 - \frac{y_{V_{slurry}}\, y_{c_{slurry}}}{\rho_s\, y_{H_{cake}}\, A} \quad . \tag{5.22}$$

MLE is used for estimating α^{estim} and $R_{m,1}^{estim}$ from the $y_{M_{filt}}$ profile, with the same procedure followed for offline parameter estimation with the filtration experiments (§5.4.2). Note that, although $y_{M_{filt}}$ is the sum of the weight of the filtrate collected from V102-V105, here we assume it to be equal to the filtrate from V105. This is a reasonable assumption, considering that, with the operating conditions used in this study, filtration always finishes in V102, and the amount of filtrate collected from the deliquoring in V103-V104 is several orders of magnitude smaller than the one from the actual filtration, occurring in V102.

Estimation of the resistance $R_{m,4}^{estim}$ of the filter mesh of Station 4, is also performed, factoring $y_{H_{cake}}$, $y_{\dot{V}_{dryer}}$, and α^{estim} into Equation 5.9. From $R_{m,1}^{estim}$, $R_{m,4}^{estim}$, $y_{c_{slurry}}$, $y_{H_{cake}}$, $y_{V_{slurry}}$, α^{estim} and ϵ^{estim}, the process is monitored to detect the occurrence of special cause variation from normal operating conditions. Moreover, the sequence of estimations of $R_{m,1}^{estim}$ and $R_{m,4}^{estim}$ for subsequent cycles is used for inferring the evolution of the filter mesh resistance fouling.

At the beginning of every cycle in which a batch is transferred in V105 for drying, the estimated saturation profile of the cake at the beginning of drying ($S_0^{estim}(z)$) is estimated by factoring α^{estim}, ϵ^{estim} and the relevant thermophysical and process parameters in the filtration and deliquoring models. Then, an extended Kalman filter (EKF) is run for estimating the ethanol content in the cake during drying ($w_{EtOH,cake,4}^{estim}$), as discussed in the next subsection.

5.5.3.2 State estimation

The EKF, based on the heat loss and drying models (§5.3.2), makes use of the estimated α^{estim}, ϵ^{estim} and S_0^{estim} and of the process measurements $y_{T_{in}}$, $y_{T_{out}}$, and $y_{\dot{V}_{dryer}}$ to estimate $w_{EtOH,cake,4}^{estim}$. The EKF also provides the standard deviation of the estimation error ($\sigma_{EtOH,cake,4}^{estim}$), based on the model and measurement error variance, and on the estimation uncertainty on the initial saturation profile $S_0^{estim}(z)$, as further detailed at the end of this subsection.

When Layer 2 is activated, $w_{EtOH,cake,4}^{estim}$ is used in the end-point controller of Layer 1, instead of $y_{T_{out}}$. Following the risk-based approach promoted by the QbD paradigm, the estimation uncertainty is accounted by triggering the carousel rotation only when (Figure 5.11):

$$w_{EOH,cake,4}^{estim} + 2\,\sigma_{EtOH,cake,4}^{estim} > 0.005 \quad . \tag{5.23}$$

Mathematical details on the implementation of the EKF follow, for the remainder of this subsection. Let us consider the nonlinear process model of Equation 5.24, represented by a set of ordinary differential equations:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t) + \mathbf{w}(t), \tag{5.24}$$

where $\mathbf{x}$ is the state vector, $\mathbf{u}$ is the input vector, $\mathbf{f}$ is a vector of nonlinear functions, and the process noise $\mathbf{w}$ is considered to follow an $\mathrm{N}(0, \mathbf{Q}(t))$ distribution, where $\mathbf{Q}$ is the model error variance. Let us define a measurement model $\mathbf{h}$, relating the vector $\mathbf{y}$ of measurements from a plant, available at finite sampling times t_k:

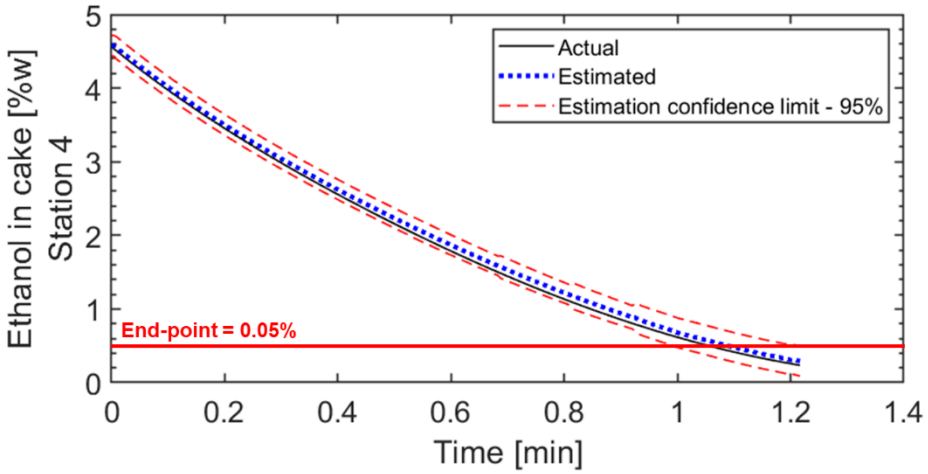


Figure 5.11. *Layer 2: the state estimator estimates the residual ethanol in the cake during the drying process in Station 4. Confidence limits for the estimations are also provided by the estate estimator. The cycle end is triggered through a carousel rotation when the upper confidence limit of the estimated residual ethanol in the cake reaches the target quality value of 0.05%w.*

$$\mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t), \mathbf{u}(t), t) + \mathbf{v}(t) \ , \tag{5.25}$$

where the measurement noise $\mathbf{v}$ is assumed to follow an $\mathrm{N}(0, \mathbf{R})$ distribution, where $\mathbf{R}$ is the measurement error variance. The discrete-time data EKF algorithm (Ray, 1981), given the models of Equations 5.24-25, provides $\hat{\mathbf{x}}$, estimation of the state vector, and $\mathbf{P}$, the estimation error covariance, through a series of subsequent prediction and updates steps. First, the EKF is initialized through the initial states estimation $\hat{\mathbf{x}}_\mathbf{0}$ and the initial estimation error covariance $\mathbf{P_0}$. A prediction step follows. For a general time interval in between two sampling times t_{k-1} and t_k, prediction steps are carried out, by integrating Equations 5.26-27 from, respectively, the updated estimations of states $\hat{\mathbf{x}}(t_{k-1}|t_{k-1})$ and estimation error covariance $\mathbf{P}(t_{k-1}|t_{k-1})$ at t_{k-1}, to yield the predictions of the states $\hat{\mathbf{x}}(t_k|t_{k-1})$ and of the estimation error covariance $\mathbf{P}(t_k|t_{k-1})$ at t_k:

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{f}(\hat{\mathbf{x}}(t), \mathbf{u}(t), t) \tag{5.26}$$

$$\dot{\mathbf{P}}(t) = \mathbf{FP} + \mathbf{PF}^{\mathrm{T}} + \mathbf{Q}(t) \ , \tag{5.27}$$

where $\mathbf{F}$ is the Jacobian matrix of the process model:

$$\mathbf{F} = \left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right)_{\hat{\mathbf{x}}(t),\mathbf{u}(t),t} \tag{5.28}$$

At each sampling time t_k, the estimations are updated through:

$$\hat{\mathbf{x}}(t_k|t_k) = \hat{\mathbf{x}}(t_k|t_{k-1}) + \mathbf{K}(t_k)\,[\mathbf{y}(t_k) - \mathbf{h}(\hat{\mathbf{x}}(t_k|t_{k-1}), \mathbf{u}(t_k), t_k)] \tag{5.29}$$

$$\mathbf{P}(t_k|t_k) = \mathbf{P}(t_k|t_{k-1}) - \mathbf{K}(t_k)\mathbf{H}^{\mathbf{T}}\mathbf{P}(t_k|t_{k-1})\,, \tag{5.30}$$

where the Kalman gain **K** and the Jacobian matrix **H**, both at t_k, are respectively calculated following:

$$\mathbf{K}(t_k) = \mathbf{P}(t_k|t_{k-1})\,\mathbf{H}^{\mathbf{T}}[\mathbf{H}\,\mathbf{P}(t_k|t_{k-1})\,\mathbf{H}^{\mathbf{T}} + \mathbf{R}]^{-1} \tag{5.31}$$

$$\mathbf{H} = \left(\frac{\partial \mathbf{h}}{\partial \mathbf{x}}\right)_{\hat{\mathbf{x}}(t_k|t_{k-1}),\mathbf{u}(t_k),t_k} \tag{5.32}$$

The implementation of Equations 5.24-5.32 for state estimation of cake drying in Layer 2 of the carousel control system is performed as follow. The same drying model used in the carousel simulator (§5.3.2) is used for state estimation within the EKF (Equation 5.9 is not used, as u_g is directly computed from $y_{\dot{V}_{dryer}}$, input of the EKF) . Hence, the model vector **f** is constituted by the space-discretized version of Equations 5.5-5.9. For computational burden reasons, the discretization grid has 10 points: $\mathbf{f} \in \mathbb{R}^{40}$, independently from the cake height. Note that in the simulator (§5.3.2), the equations are discretized with a finer grid (of spacing 0.3 mm), originating process-model mismatch. The following parameters are employed in the equations of the EKF model, instead of the corresponding ones in the carousel simulator, further increasing the process-model mismatch: *i)* ϵ^{estim} (instead of ϵ), *ii)* α^{estim} (instead of α), *iii)* $y_{H_{cake}}$ (instead of H_{cake}), *iv)* $R_{m,4}^{estim}$ (instead of $R_{m,4}$) and *v)* u_g is calculated from $y_{\dot{V}_{dryer}}$, affected by noise (instead than from Equation 5.9); also, *vi)* $h_{M,dist}$ and *vii)* $h_{T,dist}$ are set to 1 [-] (instead of using the actual local value of the disturbance, §5.4.5). All the other parameters are taken as in the simulator (§5.4). The corresponding states vector $\mathbf{x} \in \mathbb{R}^{40}$ is:

$$\mathbf{x} = [w_{EtOH,cake}(t,z_1)\;\; w_{EtOH,cake}(t,z_2)\;\; \dots\;\; w_{EtOH,cake}(t,z_{10})\;\; w_{EtOH,g}(t,z_1)\;\; w_{EtOH,g}(t,z_2)\;\; \dots\;\; w_{EtOH,g}(t,z_{10})\;\; T_{cake}(t,z_1)\;\; T_{cake}(t,z_2)\;\; \dots\;\; T_{cake}(t,z_{10})\;\; T_g(t,z_1)\;\; T_g(t,z_2)\;\; \dots\;\; T_g(t,z_{10})]\,, \tag{5.33}$$

where $z_1, z_2, \dots, z_{10}$ are the coordinates of the ten nodes of the discretization grid. The inputs from the process used by the EKF are:

$$\mathbf{u} = [y_{\dot{V}_{dryer}}\;\; y_{H_{cake}}\;\; y_{T_{in}}] \tag{5.34}$$

In the EKF, the heat loss model of Equations 5.18-5.21 (calibration parameters as in §5.4.3) is used for calculating T_{HL}^{calc} from $y_{T_{in}}$. T_{HL}^{calc} is then used as boundary condition: $T_g(t, z = 0) = T_{HL}^{calc}$. The output vector used within the EKF is $\mathbf{y} = y_{T_{out}}$, and the corresponding measurement model is $\mathbf{h} = T_g(z = z_{10})$.

In the initialization of the EKF, $\mathbf{x_0}$ is built with the structure of Equation 5.33. For every grid point z_i: $w_{EtOH,cake}(0, z_i)$ is calculated through a mass balance from $S_0^{estim}(z_i)$, $w_{EtOH,g}(0, z_i) = 0$, and $T_{cake}(t, z_i) = T_g(t, z_i) = 295$ K. Practically speaking, the only uncertain elements of $\mathbf{x_0}$ are $w_{EtOH,cake}(0, z_i)$, for each z_i node of the discretization grid. From considerations on the error propagation from $S_0^{estim}(z)$ and based on the results of simulations carried out for this purpose, the estimation variance of $w_{EtOH,cake}(0, z_i)$ is approximated to 1E-6. Hence, $\mathbf{P_0} \in \mathbb{R}^{40\,x\,40}$ is defined as a diagonal matrix, whose only non-null entries are the first 10 elements of the diagonal, set to 1E-6. Based on the measurement noise of TI102: $\mathbf{R} = 0.01$.

$\mathbf{Q}(t_k)$ is calculated as a time-variant matrix, updated at every t_k, through the algorithm proposed by Valappil and Georgakis (2000):

$$\mathbf{Q}(t_k) = \mathbf{J}_{p,nom}(t_k)\, \mathbf{C}_p\, \mathbf{J}_{p,nom}^{\mathbf{T}}(t_k)\,, \tag{5.35}$$

where $\mathbf{C}_p$ is the covariance matrix of the uncertain parameters $\mathbf{p}$, and $\mathbf{J}_{p,nom}(t_k)$ is the Jacobian of $\mathbf{f}$ with respect to $\mathbf{p}$:

$$\mathbf{J}_{p,nom}(t_k) = \left(\frac{\partial \mathbf{f}}{\partial \mathbf{p}}\right)_{\hat{\mathbf{x}}(t_k|t_{k-1}),\mathbf{u}(t_k),t_k} \tag{5.36}$$

Vector $\mathbf{p}$ contains the parameters that, among the set of mismatched parameters *i-vii*, affect the most the process-model mismatch in the EKF framework: $\mathbf{p} = [u_g \quad h_M \quad h_T \quad \epsilon]$. As suggested in the literature (Valappil and Georgakis, 2000) for systems affected only by parametric mismatch (neglecting the structural mismatch introduced by the approximation in the discretization grid), $\mathbf{C}_p$ is built as a diagonal matrix, presenting the variance of the parametric fluctuations as diagonal elements: $\mathbf{C}_p = diag([4E\text{-}6 \quad 1.6E\text{-}21 \quad 2.5E\text{-}5 \quad 4.9E\text{-}5])$. Note that Equations 5.35-5.36 inherently yield the local variance of the model error, and provide a reliable (Schneider and Georgakis, 2013) estimation of $\mathbf{Q}(t_k)$. Hence, the designed $\mathbf{Q}$ and $\mathbf{R}$ matrices allow to effectively propagate the uncertainty $\mathbf{P_0}$ in the initial state estimate $\mathbf{x_0}$ into the estimation error covariance $\mathbf{P}$ of the estimated states $\mathbf{x}$.

The EKF update step is called every 5 s during the process, even though the measurements are sampled every 0.25 s. The EKF provides to the end-point controller of Layer 1, at time t, $w_{EOH,cake,4}^{estim}(t)$, obtained as average of the ten $w_{EtOH,cake}(t, z_i)$ in the state vector. The estimation standard deviation $\sigma_{EtOH,cake,4}^{estim}$ is instead directly obtained from $\mathbf{P}$.

5.5.3.3 Real time optimization

While Δt_{cycle}^{sp} is the only manipulated variable adjusted by the controllers presented so far, the RTO routine of Layer 2, presented in this subsection, is meant to optimize V_{slurry}^{sp}, at every cycle. This is done by solving the following optimization problem, before the beginning of every cycle:

$$\max_{T^{sp},V_{slurry}^{sp}} \quad T^{sp} \tag{5.37}$$

subject to: $$w_{EOH,cake,final}^{calc}\left(T^{sp}, V_{slurry}^{sp}, \mathbf{d^{estim}}\right) < 0.005 \tag{5.38a}$$

$$0\ mL < V_{slurry}^{sp} < 10\ mL\ , \tag{5.38b}$$

where: $\mathbf{d}^{estim}$ is the estimated vector of the process disturbances, $w_{EOH,cake,final}^{calc}\left(T, V_{slurry}^{sp}, \mathbf{d}^{estim}\right)$ is the prediction of the mass fraction of ethanol in the final cake, and T^{sp} is the theoretical throughput:

$$T^{sp} = \frac{V_{slurry}^{sp}\ \mu_{c_{slurry}}}{\Delta t_{cycle}^{sp}} \quad . \tag{5.39}$$

Note that in the RTO Problem of Equations 5.37-5.39, T^{sp} is selected as one of the optimization variables in order to obtain a linear objective function (Equation 5.37), even though it is not an explicit input of the carousel simulator. The solution of the RTO yields the optimal V_{slurry}^{sp} to feed at the current cycle, and (indirectly through Equation 5.39) the optimal Δt_{cycle}^{sp} to use for processing the current batch. The actual Δt_{cycle}^{sp} used in the process will be manipulated by the Layer 1 end-point controller, to meet the target CQA in the cake being dried in V105. For the first three cycles after a cleaning operation, when no cake is being dried in V105, the optimal Δt_{cycle}^{sp} found at the end of the cleaning (namely, a state of clean meshes) is used.

Within the RTO Problem, mathematical models are needed for the computation of $w_{EOH,cake,final}^{calc}$ in Equation 5.38a. For this purpose, we use the filtration, deliquoring and drying models presented in §5.3.2, arranged as in the carousel simulator. For the calculation of $w_{EOH,cake,final}^{calc}$, all inputs of the simulator (Table 5.2) have to be provided (calibration parameters, thermophysical properties, set-points of the controlled variables, and disturbances). The actual calibration parameters and thermophysical properties implemented in the carousel simulator are used. Two of the set-points of the controlled variables are known (ΔP^{sp} and $T_{in,g}^{sp}$), while the other two are obtained from the optimization variables: V_{slurry}^{sp} explicitly, and Δt_{cycle}^{sp} implicitly, through Equation 5.39 (recall that, in this case study, $\mu_{c_{slurry}}$ is fixed to 250 kg/m^3). The actual values of the elements of the disturbances vector $\mathbf{d}$ are, instead, generally unknown, thus originating process-model mismatch. The elements of $\mathbf{d}$ are (Table 5.2):

$$\mathbf{d} = [c_{slurry}\ \ h_M^{dist}\ \ h_T^{dist}\ \ R_{m,1}\ \ R_{m,2}\ \ R_{m,3}\ \ R_{m,4}\ \ V_{slurry}^{dist}\ \ \mathbf{w}_{in,g}\ \ \mathbf{w}_{l,slurry}\ \ \epsilon^{dist}]\ , \tag{25}$$

and they are estimated in $\mathbf{d}^{estim}$ as follows. Since the RTO problem is solved before the beginning of every cycle, c_{slurry}, V_{slurry}^{dist} and ϵ^{dist} cannot be inferred from the online parameter estimation routines, and they are approximated by the mean value of their probability distributions (respectively: $\mu_{c_{slurry}}$, 1 [-] and 1 [-]). We also approximate both h_M^{dist} and h_T^{dist} to 1 [-], the mean value of their probability distributions. We fix $\mathbf{w}_{in,g}$ to air composition, and $\mathbf{w}_{l,slurry}$ to pure ethanol. The filter mesh resistances estimations in $\mathbf{d}^{estim}$ are intended as the estimation of the resistance that the slurry loaded in the current cycle will encounter when being processed in Station i. Considering the fouling schedule implemented in the simulator (Section 3.3), for a generic cycle, $R_{m,i}$ is approximated as: $R_{m,1}^{estim}$ + 2E9 [1/m], for i =1, 2, 3, 4. Instead, in the specific case of cycles initiated immediately after a cleaning-in-place routine, $R_{m,i}$ (for i =1, 2, 3, 4) is directly approximated to 3E9 1/m.

The RTO problem of Equations 5.37-5.39 is solved with the fmincon optimizer implemented in MATLAB. To enhance the robustness of the RTO solution, at every cycle we solve the optimization problem three times, varying the initial guess provided to the optimizer among a set of three initial points belonging to the DS. The (feasible) solution leading to the largest T^{sp} is implemented in the process. Both EKF and RTO computational times are compatible with real time implementation.

5.7 Control strategy response to disturbances and faults: Quality-by-Design vs Quality-by-Control

The simulator calibrated on the paracetamol/ethanol slurry system is used to generate data for 1 h of carousel operation, under the set of disturbances occurring in normal operating conditions (§5.4.5). The simulation is run five times, comparing the performance of different control strategies (§5.6). Figure 5.12 shows the residual ethanol content in the discharged cakes, the cycle duration and the fed slurry volume, obtained during the simulations under the following control strategies: *i)* Layer 0-only (denoted as *Layer 0*), *ii)* Layers 0-1 (denoted as *Layer 1*), *iii)* Layers 0-1 + an EKF at Layer 2 used as soft sensor for end-point control (denoted as *Layer 2 – only EKF*), *iv)* Layers 0-1 + a perfect state estimator at Layer 2 used as soft sensor for end-point control (denoted as *Layer 2 – perfect estimator*), and *v)* a complete Layers 0-1 implementation, with both EKF and RTO routines at Layer 2 (denoted as *Layer 2 – EKF + RTO*). Note that the perfect state estimator of the Layer 2 – perfect estimator control strategy is an ideal mathematical tool that could be implemented, for comparison purposes, just because the plant object of this study is virtual. In real world case study, no perfect state estimator can be implemented. The final throughput achieved with the different control strategies is reported in Table 5.4.

Focusing first on the fed slurry volume panel of Figure 5.12, it can be assessed that all the control strategies, except for Layer 2 – EKF + RTO, always use a constant value for V_{slurry}^{sp} (i.e., 6 mL; §5.6.1). The optimal V_{slurry}^{sp} calculated by the RTO system of the Layer 2 – EKF + RTO control strategy is, instead, always lower than 6 mL. Note that the three consecutive cycles during which the fed slurry is null (e.g., cycle #7-9), appearing periodically with all control strategies, correspond to the cycles of the cleaning-in-place routine (§5.4.5). When the RTO routine is turned on, the optimizer selects larger optimal V_{slurry}^{sp} for increasing filter mesh fouling. Considering for example the process cycles #10-15, V_{slurry}^{sp} increases when operating under the Layer 2 – EKF + RTO control strategy from cycle #10 (clean meshes conditions) up to cycle #15, the last one before mesh cleaning. Even though the increasing trend of the optimal V_{slurry}^{sp} in between two cleaning-in-place procedures is consistent during carousel operation, slightly different optimal V_{slurry}^{sp} are found for different cycles occurring after a given number of cycle after the cleaning-in-place (e.g., cycles #10, #19, and #28). This is because the optimal V_{slurry}^{sp} calculated by the RTO system depends on the whole set of estimated disturbances $\mathbf{d}^{estim}$, and not only on the local fouling conditions.

Considering the cycle duration panel of Figure 5.12b, all control strategies adjust Δt_{cycle}^{sp} during operation, except for the Layer 0 one. During the first three cycles after a cleaning-in-place procedure, Layer 1 and Layer 2 control strategies use a lower Δt_{cycle}^{sp} than for other cycles, as outlined in §5.6.2 and §5.6.3, since only filtration and deliquoring are carried out. The Layer 2 – EKF + RTO control strategy selects a growing Δt_{cycle}^{sp} with the passing of cycles after a cleaning-in-place procedure, consistently with the increasing V_{slurry}^{sp} profile found from the solutions of the RTO problem for increasing fouling mesh conditions. All the other Layer 1 and Layer 2 control strategies do not present clear Δt_{cycle}^{sp} patterns.

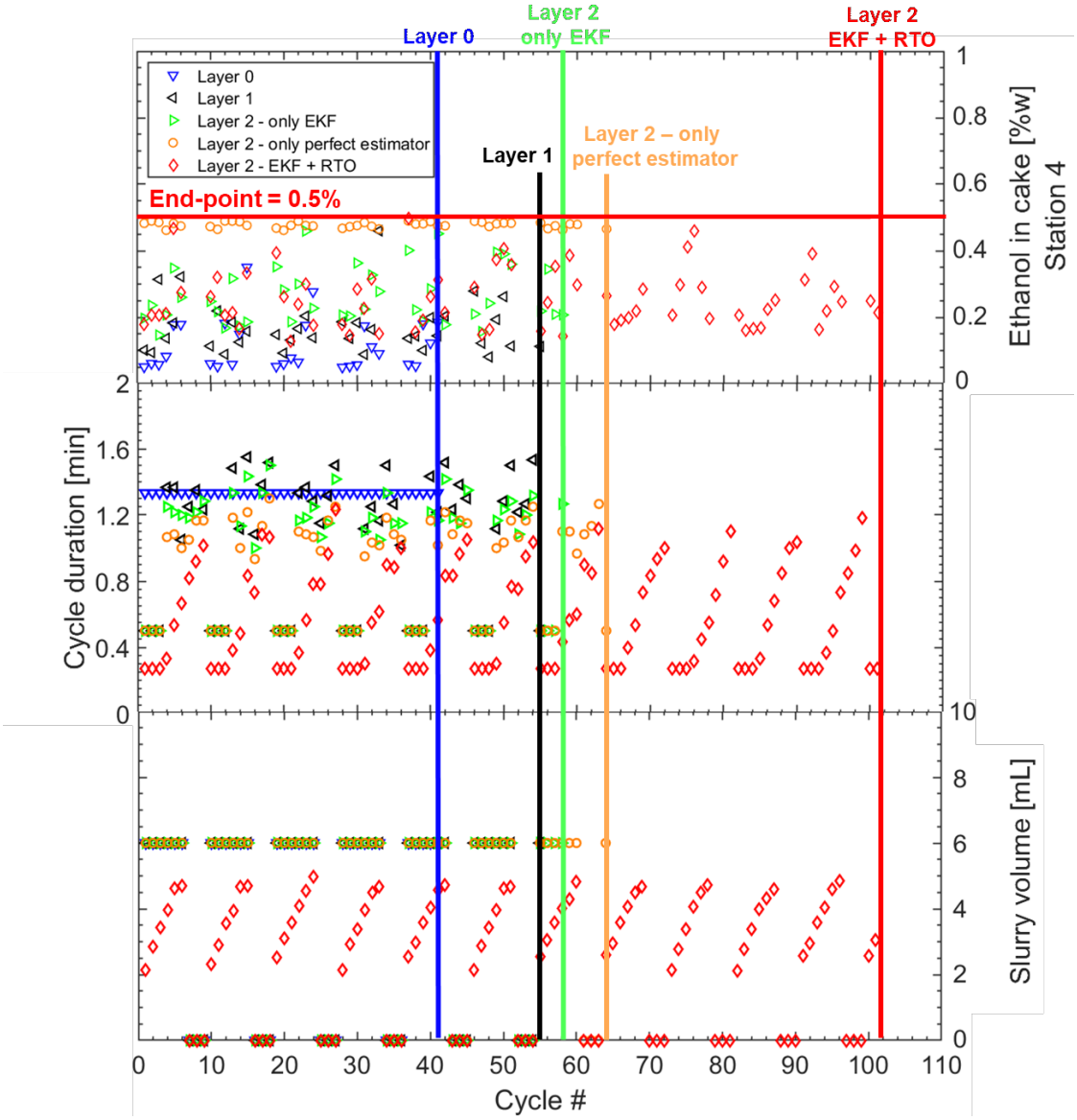


Figure 5.12. *Comparison of carousel performance under different control strategies: residual ethanol in discharged cakes, cycles duration and fed slurry volumes following 1 h of normal operating conditions operation.*

Table 5.4. *Comparison of carousel performance under different control strategies: slurry throughput achieved in 1 h of normal operating conditions operation.*

Control strategy	Throughput [mL]	% Layer 0 throughput
Layer 0	180.1	100
Layer 1	228.3	127
Layer 2		
only EKF	240.3	133
only perfect estimator	258.4	143
EKF + RTO	251.5	140

Figure 5.12 also shows that all the control strategies allow meeting the quality requirements on the residual ethanol content in all the discharged cakes. The residual ethanol content is the lowest when operating with the Layer 0 control strategy, as it is the most conservative approach, drying the cakes more than needed to compensate for the lack of feedback control routines. On the other hand, the Layer 2 – perfect estimator control strategy allows discharging cakes with the maximum acceptable residual ethanol content, avoiding to spend any time more than needed on drying. The other Layer 1 and Layer 2 control strategies behave intermediately between these two extremes. Since drying is the slowest processing step in the carousel, drying longer than needed implies being able to process a sub-optimal slurry throughput.

As confirmed from Table 5.4, more advanced control strategies allow processing a larger throughput. The Layer 0 control strategy leads to the poorest performance, as the operating conditions are fixed, instead of being adapted to optimize the process performance. Introduction of the Layer 1 end-point temperature controller allows processing, in the same time frame, 11 more cakes (Figure 5.12) of the same mass, corresponding to a throughput increase of about 27%. With the introduction of the EKF to directly monitor the residual ethanol in the cake, 3 more cakes (still of the same mass) are obtained (+33% throughput compared to Layer 0-only). If, instead of the EKF, an (ideal) perfect state estimator was available for monitoring the residual ethanol in the cake being dried, additional 8 cakes of the same mass could be processed in the carousel in the same time frame, corresponding to a throughput increase of about 43% with respect to Layer 0-only operation.

The throughput increase due to the implementation of (even simple temperature-based) end-point controllers is very large, due to the benefits of regulating Δt_{cycle}^{sp} based on the inferred drying duration, as opposite to keeping Δt_{cycle}^{sp} fixed, as in Layer 0. The more accurate the estimation, the larger the throughput. With the perfect state estimator (circles in Figure 5.12), the carousel rotation is triggered immediately after that the cake being dried reaches the target ethanol content. With the end-point controller based on the EKF, the estimation uncertainty is instead considered when triggering the cycle rotation (Equation 5.23), leading to drying the cakes more than needed. When resorting to the end-point controller on temperature, the cakes are dried even longer (Figure 5.12), as the end-point on temperature is more unreliable, and a conservative approach must be adopted (Figures 5.10). Another drawback of end-point temperature control is that the drying air outlet temperature at the end of drying depends non-linearly on the cake mass. Hence, for mid/large variations of c_{slurry} and

V_{slurry}, the end-point of Figure 5.10 is not valid anymore, and should be re-assessed. End-point control based on state estimation does not suffer from this limitation, hence it can be coupled with control routines adjusting V_{slurry}^{sp}, too.

This is the case of the Layer 2 – EKF + RTO control strategy, which, instead, manipulates both Δt_{cycle}^{sp} and V_{slurry}^{sp} (diamonds in Figure 5.12), aiming to maximize the slurry throughput. With this approach, as many as 101 cakes are obtained from the carousel operation in a 1 h time frame. Even though the cakes mass varies, depending on the local V_{slurry}^{sp} (Figure 5.12), the final throughput is 40% larger than with the Layer 0-only approach.

Note that the RTO implementation of Equations 5.37-39 optimizes the operating conditions for one batch at a time. More advanced RTO implementations can be conceived, for instance by optimizing the operating conditions for a series of consecutive batches, and factoring the triggering of the cleaning-in-place procedure inside the optimization problem. However, this is beyond the scope of this study, which aims at providing an overview of the benefits of active process control for the carousel operation.

5.8 Conclusions

We developed a comprehensive simulator for the carousel technology, and calibrated it with an experimental campaign on a process for isolating paracetamol from a paracetamol/ethanol slurry system. Then, we conceived a three-layered (Layers 0-1-2) control strategy for the unit, following the recent QbC paradigm. The control strategy includes at Layer 1 an end-point controller, for automatically triggering the end of a processing cycle and the consequent product discharge when the product CQAs are met in the final processing step of the unit. Layer 2 features an EKF, used as soft sensor for monitoring the product CQAs and to support the end-point controller of Layer 1, and an RTO routine, to optimize the amount of slurry to be fed to the unit at every processing cycle. Following a risk analysis-based approach promoted by the QbD initiative, the estimation uncertainty computed by the EKF is accounted for when using the EKF estimations for identifying the cycle end-point.

The conceived control strategy was implemented in the simulator, and benchmarked against other control strategies of growing complexity under a set of disturbances occurring under normal operation of the carousel, such as filter mesh fouling. The proposed control system demonstrated to be able to always guarantee the target product quality, and led to a significantly larger slurry throughput, compared to the traditional QbD approach of operating in a fixed point of the DS. The complete Layers 0-1-2 control strategy leads to achieving the largest slurry throughput. Nonetheless, even the inclusion of a simple Layer 1 end-point controller, using temperature measurements instead than EKF estimations for detecting the end of a cycle, significantly improved the process performance, compared to traditional open-loop operation.

Conclusions and future perspectives

This Dissertation makes several steps forward in the modernization of pharmaceutical development and manufacturing by promoting the use of mathematical methodologies for digitally-supported operation design, process monitoring and process control. The general objectives of reducing pharmaceutical development time and cost, and of increasing product quality attainment performances in pharmaceutical manufacturing through mathematical modelling are achieved by pursuing the following six specific objectives:

1. development and implementation of advanced process monitoring methodologies for pharmaceutical manufacturing;
2. model-based pharmaceutical operation design;
3. control-relevant modeling of pharmaceutical processes;
4. mathematical modeling for aiding the transition to continuous pharmaceutical manufacturing;
5. development and use of knowledge-driven models for supporting pharmaceutical development and manufacturing;
6. life cycle approach to mathematical modelling of pharmaceutical processes.

Table C.1 outlines the main achievements of this Dissertation, arranged chapter-wise (Table 1.6 in §1.7 provides additional details on the content of each chapter).

In Chapter 1, a thorough review on the historical evolution and on the state of implementation of Quality-by-Design is provided. A special focus is drawn on the emerging pharmaceutical technology and on state-of-the-art mathematical modeling for pharmaceutical development and manufacturing.

Chapters 2 and 3 are centered around the topic of **process monitoring.** Case studies applying to pharmaceutical manufacturing monitoring techniques commonly applied in other industries are presented, and a novel hybrid data-driven/knowledge-driven methodology for fault detection and diagnosis is introduced. The results demonstrate that advanced process monitoring techniques in pharmaceutical manufacturing can significantly improve the product quality attainment performances, and support process capability enhancement towards the six sigma standard. It is expected that the promising achievements reported in Chapters 2-3, partly resulting from industrial collaborations, will promote the adoption of these type of techniques for process monitoring by the pharmaceutical industry. Moreover, the novel hybrid approach to process monitoring presented in this Dissertation is expected to gain the interest of other industries beside the pharmaceutical sector, as it demonstrated superior performances compared to state-of-the-art standalone data-driven or knowledge-driven methodologies in fault detection and diagnosis.

Chapters 4 and 5 both involve model-based studies around a novel carousel for **continuous intensified filtration-washing-drying of crystallization slurries**. A **life cycle approach to mathematical modeling** is followed. A mathematical model for the unit is developed

Table C.1. Summary of the main achievements of this Dissertation, with indication of their relevant references, organized chapter-wise.

Chapter	Specific aim	Main activity	Main achievement	Reference
1	1-6	Review	• Review on state of implementation of Quality-by-Design, on emerging pharmaceutical technology and on state-of-the-art mathematical modeling for pharmaceutical development and manufacturing	Destro, F. and M. Barolo. A perspective on mathematical modeling for the digitalization of pharmaceutical development and manufacturing. *In preparation*
2	1, 4, 5	Process monitoring	• State estimator effectively reconciles noisy measurements from loss-in-weight feeders, and provides better powder composition estimations compared to traditional approaches based on statistical filtering • First proof of concept of state estimation for monitoring powder composition in direct compression lines, tested on data from pilot plant	Destro, F., S. García Muñoz, F. Bezzo, and M. Barolo (2021). Improving powder feeding monitoring in continuous solid dosage forms manufacturing through state estimation. *Int. J. Pharm.*, **605**, 120808.
3	1, 4, 5	Process monitoring	• Novel hybrid data-driven/knowledge-driven approach to process monitoring • Proposed hybrid approach outperforms traditional data-driven and knowledge-driven approaches in fault detection and diagnosis • Available first-principles knowledge on the system factored in the monitoring system, coherently with Quality-by-Design paradigm	Destro, F., P. Facco, S. García-Muñoz, F. Bezzo, and M. Barolo (2020). A hybrid framework for process monitoring: Enhancing data-driven methodologies with state and parameter estimation. *J. Process Control*, **92**, 333-351. Destro, F., A. J. Salmon, P. Facco, C. C. Pantelides, F. Bezzo, M. Barolo (2020). Monitoring a segmented fluid bed dryer by hybrid data-driven/knowledge-driven modeling. *IFAC-PapersOnLine*, **53**, 11638-11643.
4	2, 4, 5, 6	Modeling, Operation design	• Study on novel unit for continuous isolation of crystals from slurries through intensified filtration-washing-drying. The unit tackles bottleneck in end-to-end continuous pharmaceutical manufacturing • Mechanistic model of the unit developed and used for determination of probabilistic design space • Execution of filtration and drying experiments, respectively, on Nutsche filter and thermogravimetric analyzer, for model calibration and validation;	Destro, F., I. Hur, V. Wang, M. Abdi, X. Feng, E. Wood, S. Coleman, P. Firth, A. Barton, M. Barolo, and Z. K. Nagy (2021). Mathematical modeling and digital design of an intensified filtration-washing-drying unit for pharmaceutical continuous manufacturing. *Chem. Eng. Sci.* **244** (23), 116803. Destro, F., V. Wang, M. Abdi, X. Feng, E. Wood, S. Coleman, P. Firth, A. Barton, M. Barolo, and Z. K. Nagy. An intensified unit for continuous integrated filtration, washing and drying of drug substances: mathematical modelling and design space identification. *IFAC PapersOnLine*, **54**, 85-90. Laky, D. J., D. M. Casas-Orozco, F. Destro, M. Barolo, Z. K. Nagy (In press), Integrated synthesis, crystallization and drying of active pharmaceutical

				ingredients: a model-based framework for process optimization and control. Z. K. Nagy, R. Ramandrachan, P. Pardalos (Ed.), In: *Optimization of Pharmaceutical processes, models and methods*, Springer Nature, Basingstoke (UK). *In press*
5	3, 4, 5, 6	Modeling, Process control	• Real time carousel simulator for control-relevant applications developed from carousel model of Chapter 4 (*life cycle* approach to process modeling) • Filtration and drying experiments on a physical carousel carried out for model calibration and validation • Closed-loop control system for the unit conceived and tested on the simulator (first closed-loop control study on this type of unit). • Proposed control system includes model-based routines, (i.e., state estimation and real time optimization), based on developed mathematical model of the unit	Destro, F., Z. K. Nagy, and M. Barolo. Quality-by-Control of intensified continuous filtration-drying of active pharmaceutical ingredients. *In preparation* Destro, F., Z. K. Nagy, and M. Barolo. A simulator for control-relevant applications in intensified continuous filtration-drying of active pharmaceutical ingredients. *In preparation* Destro, F., V. Wang, M. Abdi, X. Feng, E. Wood, S. Coleman, P. Firth, A. Barton, M. Barolo, and Z. K. Nagy (2021), Quality-by-Control of continuous drug substance isolation: study on a novel unit for integrated filtration-drying, In: *Computer-Aided Chemical Engineering 50*, *Proc. of the 31st European Symposium on Computer Aided Process Engineering* (M. Türkay, R. Gani, Eds.), Elsevier, Amsterdam (The Netherlands), 1363-1369.

and used for offline design of the unit operations. Then, the model is further developed, and its use to support real time operation with soft-sensing and model-based process control is demonstrated. Chapters 4-5 prove that the use of mathematical modeling and active process control can significantly reduce the time and investment needed for process development, and improve the quality attainment performance in the manufacturing stage. Since the results demonstrate that the investment needed to implement this type of approaches in industrial lines is outweighed by the benefits, it is expected that the reported studies will promote the adoption of these techniques by practitioners. Moreover, the carousel unit object of the research work reported in Chapters 4-5 is a novel technology, one of the few available in the marketing for *continuous* filtration-drying. Since these operations are a current technological gap for the implementation of **end-to-end continuous processing**, the developed mathematical framework will also promote future studies on this unit and the transition to continuous processing. The developed carousel simulator (*ContCarSim*) will be made available to the public (www.github.com/francescodestro/ContCarSim) as a benchmark for Quality-by-Design and Quality-by-Control studies. *ContCarSim* is one of the first benchmarks for design- and control-relevant applications for pharmaceutical processes.

More in detail, in Chapter 2, a novel approach is presented for **monitoring the composition of powder mixtures in continuous tableting lines.** As shown in the chapter, the traditional approach of using statistical filters for calculating the powder mass flows (and, in turn, the powder composition) from the time series of mass measurements can lead to inaccuracies, even though this is done by default by loss-in-weight feeders software. A novel monitoring system rooted on a state estimator is proposed, which effectively reconciles noisy mass measurements coming from loss-in-weight feeders through a model-based optimization strategy. The API concentration in the final tablets estimated through the proposed monitoring approach present high consistency with HPLC and spectroscopic measurements collected downstream. With the proposed methodology the powder composition is estimated continuously, with a significant advantage over sampled measurements, such as those from HPLC. The results show that state estimation is a suitable way to obtain maximum synergy from the resources invested in the development of a mechanistic model for a process, and the resources dedicated to the implementation of PAT.

In Chapter 3, a **novel methodology for multivariate process monitoring based on hybrid modeling** is proposed. Real-time deterministic information about the process is first obtained from a state estimator in the form of estimated states, noise-deprived measurements, and possibly estimated parameters. The information is then passed to a data-driven model, where it is exploited, in conjunction with the available field measurements, for multivariate fault detection and diagnosis. When tested on batch and continuous case studies, the hybrid monitoring approach typically allows for earlier fault detection than with traditional data-driven and knowledge-driven approaches taken in isolation, even when the state estimator did not perform entirely satisfactorily. Moreover, the hybrid approach significantly facilitates fault diagnosis. The proposed hybrid monitoring framework is expected to gain the interest of the pharmaceutical community, as it perfectly meets the request of

the Quality-by-Design initiative to factor the available knowledge on the physics of the process into the control strategy.

Chapter 4 presents a mathematical model for the novel continuous intensified carousel for the isolation of crystals from slurries. Independent filtration, deliquoring, washing and drying modeling components are developed and combined to mimic carousel operation. A cake properties module estimates the crystals cake physical properties from the slurry crystal size and shape distribution. The carousel model tracks dynamic profiles of cake composition throughout all carousel processing, overcoming the limitations of short-cut methods traditionally used for designing this type of process. After successfully calibrating the model with filtration (on a Nutsche filter) and drying experiments (on a thermogravimetric analyzer), it is used for describing the **probabilistic design space** for a process for isolating paracetamol from a slurry system. The maximum throughput of the slurry system under investigation that can be processed in the carousel with acceptable probability of meeting the product quality specifications is also determined.

In Chapter 5, the carousel model introduced in Chapter 4 is further developed for **process control-relevant studies** on the unit. An advanced real time simulator of the carousel is implemented, with digital routines simulating sensors and actuators present in physical units for the purpose of control strategy testing. A pilot scale carousel is used for carrying out filtration and drying experiments on a paracetamol/ethanol slurry system, to calibrate and validate the simulator. Following the recent Quality-by-Control paradigm, a closed-loop quality control strategy is proposed for this type of units for the first time. The proposed control strategy also features model-based control and monitoring elements, such as state estimation and real time optimization, which exploit the developed mechanistic model of the carousel, following a life cycle approach to pharmaceutical process modeling. When tested on the simulator, under a set of disturbances known to affect the unit operation (e.g., filter mesh fouling), the proposed closed-loop quality control system proves improved control performance with respect to the traditional QbD approach, consisting on operating at open-loop within the design space.

Different **research areas for future investigations** stem from the results presented and discussed in this Dissertation:

- Regarding the approach presented in Chapter 2 for powder composition monitoring in continuous direct compression lines, future studies can involve tests of the methodology on datasets from manufacturing lines, and, eventually, the implementation of the state estimator on a physical line;
- The results on state estimation of Chapter 2 represent the first step in the journey of adopting model-based closed-loop quality control and advanced process monitoring in continuous direct compression lines. The use of the state estimator within a model predictive control framework, or for hybrid process monitoring with the approach presented in Chapter 3, is also envisioned;

- Following the promising results obtained with the hybrid monitoring approach of Chapter 3, future studies can involve tests of the approach on additional processes, including on physical systems (e.g., direct compression lines, such as exploiting the state estimation framework of Chapter 2);
- Still on the hybrid monitoring side, an important feature to be added to the framework is a strategy for automatic selection of the variables to be actively included within the multivariate monitoring framework. Actually, for large systems, it is impractical to include all the states and parameters obtained from a state estimator in the data matrix for hybrid process monitoring, as proposed in Chapter 3. Routines for maintenance of the hybrid monitoring framework should also be investigated;
- Regarding the work on the continuous carousel for integrated filtration-drying (Chapters 4 and 5), significant research work can originate from the obtained results. First, the developed simulation studies on operation design and process control for the paracetamol/ethanol slurry system can be validated on a physical unit. Other control strategies for the unit can also be designed, tested on the simulator, and then benchmarked against the control strategy introduced in Chapter 5. In particular, space of improvement can be seen through the implementation of advanced optimization frameworks at Layer 2. Case studies on other slurry systems can also be conducted. An interesting application regards the investigation of processes where washing is a critical step (e.g., when impurities adsorb the solid phase);
- The integration of a crystallizer in the simulator of Chapter 5 represents another major step forward to be achieved in future research work. Integrated design of crystallization-filtration-drying processes, operation and control system is acknowledged as a research topic of pivotal importance in pharmaceutical development and manufacturing. Even a relatively simple crystallizer model, factored into the carousel simulator, would allow evaluating the mutual interaction between crystallization, filtration and drying design and operating parameters, opening up many possibilities for optimization and control studies.

Appendix A

Penicillin manufacturing: detailed and simplified models

The fed-batch process for the manufacturing of penicillin by biomass fermentation is simulated using the detailed model by Birol et al. (2002):

$$\frac{dX}{dt} = \mu_x X - \frac{X}{V}\frac{dV}{dt} \tag{A.1}$$

$$\frac{dP}{dt} = \mu_P X - K\,P - \frac{P}{V}\frac{dV}{dt} \tag{A.2}$$

$$\frac{dS}{dt} = -\frac{\mu_x}{Y_{X/S}} X - \frac{\mu_P}{Y_{P/S}} X - m_x X + F\frac{s_F}{V} - \frac{S}{V}\frac{dV}{dt} \tag{A.3}$$

$$\frac{dC_{O_2}}{dt} = \frac{\mu_x}{Y_{X/O}} X - \frac{\mu_P}{Y_{P/O}} X - m_o X + K_{la}\left(c^*_{O_2} - c_{O_2}\right) - \frac{C_{O_2}}{V}\frac{dV}{dt} \quad , \tag{A.4}$$

with:

$$K_{la} = \alpha_1\sqrt{f_g}\left(\frac{P_w}{V}\right)^{\alpha_2} . \tag{A.5}$$

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \frac{U}{s_F} + F_{a/b} - V\lambda(e^{\frac{5(T-T_0)}{T_v-T_0}} - 1) \tag{A.6}$$

$$\frac{dc_{CO_2}}{dt} = \alpha\frac{dX}{dt} + \beta X + \gamma \tag{A.7}$$

$$\mu_P = \mu_{P,max}\ \frac{S}{k_p + S(1+S/K_I)}\frac{C^p_{O_2}}{k_{op} X + C^p_{O_2}} \tag{A.8}$$

$$\mu_X = \left[\frac{\mu_{x,max}}{1+\frac{K_1}{[H^+]}+\frac{[H^+]}{K_2}}\right]\frac{S}{k_x X + S}\frac{C_{O_2}}{k_{ox} X + C_{O_2}}\left\{\left[k_g \exp\left(-\frac{E_g}{RT}\right)\right] - \left[k_d \exp\left(-\frac{E_d}{RT}\right)\right]\right\} \tag{A.9}$$

$$\frac{\mathrm{d}[H^+]}{\mathrm{d}t} = \gamma\left(\mu X - \frac{FX}{V}\right) + \frac{\left[\frac{-\delta+\sqrt{(\delta^2+4\cdot 10^{-14})}}{2} - [H^+]\right]}{\Delta t}, \tag{A.10}$$

with:

$$\delta = \left[\frac{10^{-14}}{[\mathrm{H}^+]} - [\mathrm{H}^+]\right] V - \frac{C_{a/b}(F_a+F_b)\Delta t}{V+(F_a+F_b)\Delta t}. \quad (A.11)$$

$$\frac{\mathrm{d}Q_r}{\mathrm{d}t} = r_{q1}\frac{\mathrm{d}X}{\mathrm{d}t}V + r_{q2}XV. \quad (A.12)$$

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \frac{F}{s_F}\left(T_f - T\right) + \frac{1}{V\rho c_p}\left[Q_r - \frac{aF_c^{b+1}(T-T_c)}{F_c+\left(\frac{aF_c^b}{2\rho c_{pc}}\right)}\right]. \quad (A.13)$$

We report in Table A.1 the symbols used for the states and outputs, and in Tables A.2 and A.3 those used for the parameters and the inputs, respectively.

To allow for more variability under NOC, thus making the monitoring problem more challenging, we detune the pH and temperature loops with respect to the original tuning in (Birol et al., 2002). We also include process noise as fluctuations in the inputs, as reported in Table A.3.

Table A.1. *Case study 2: states of the detailed model, with indication on whether they are measured, and whether or not they are included also in the simplified model.*

State variable	Symbol	Initial value	Units	Measured?	In the FPM?
Biomass concentration	X	0.1	g_X L^{-1}	No	Yes
Penicillin concentration	P	0	g_P L^{-1}	No	Yes
Substrate concentration	S	15	g_S L^{-1}	No	Yes
Dissolved oxygen concentration	c_{O_2}	1.16	g_{O_2} L^{-1}	Yes	Yes
Volume	V	100	L	Yes	Yes
CO_2 concentration	c_{CO_2}	0.5	mmol$_{CO_2}$ L^{-1}	Yes	Yes
Hydrogen ion concentration	$[\mathrm{H}^+]$	$10^{-5.1}$	mol L^{-1}	Yes	No
Temperature	T	297	K	Yes	No
Heat released	Q_r	0	cal	No	No

The faulty batches discussed in Section 4.2 are initialized as the NOC ones, and then the following changes are considered in the detailed model:

- Fault #1: aeration rate decreases by 0.01 L h^{-1} from 150 h to 300 h;
- Fault #2: substrate feed concentration decreases by 0.50 g_S L^{-1} h^{-1} from 150 h to 300 h;
- Fault #3: $\mu_{X,\max}$ decreases by 3.2E-4 h^{-2} from 15 h to 140 h;
- Fault #4: T_c = 298 K from the beginning of the batch.

The simplified model by Bajpai and Reuss (Bajpai and Reuß, 1980), complemented with the CO_2 balance and a simplified version of the volume loss equation from Birol et al. (Birol et al., 2002), are used as the FPM. Namely, the FPM is composed by Equations A.1-4, Equation A.7-8 and the following equations as the volume balance and the biomass growth kinetics:

$$\frac{dV}{dt} = F - \lambda V \quad (A.14)$$

$$\mu_X = \mu_{X,\max}\frac{S}{k_x X+S}\frac{C_{O_2}}{k_{ox} X+ C_{O_2}} \quad (A.15)$$

States, parameters and inputs are set as in Tables A.1-3. Note that the FPM retains only 6 of the 9 states of the process. The main sources of parametric and structural mismatch are as follows:

- protons, heat and energy balances (Equations A.10-13) are not included in the FPM;

Table A.2. *Case study 2: parameters the detailed model, with indication on whether or not they are included also in the simplified model.*

Parameter	Symbol	Units	Value	In the FPM?
Penicillin hydrolysis rate constant	K	h^{-1}	0.04	Yes
Yield of biomass on substrate	$Y_{X/S}$	$g_X\ g_S^{-1}$	0.45	Yes
Yield of penicillin on substrate	$Y_{P/S}$	$g_P\ g_S^{-1}$	0.90	Yes
Maintenance coefficient on substrate	m_x	h^{-1}	0.014	Yes
Feed substrate concentration	s_F	$g_S\ L^{-1}$	600	Yes
Yield of biomass on oxygen	$Y_{X/O}$	$g_X\ g_{O_2}^{-1}$	0.04	Yes
Yield of product on oxygen	$Y_{P/O}$	$g_P\ g_{O_2}^{-1}$	0.20	Yes
Maintenance requirement of oxygen	m_o	h^{-1}	0.467	Yes
Solubility of oxygen in broth	$c_{O_2}^*$	$g_{O_2} L^{-1}$	1.16	Yes
Volume loss parameter	λ	h^{-1}	2.5E-4	Yes
Constant relating CO_2 to growth	α	$mmol_{CO_2}\ g_X^{-1}\ h^{-1}$	0.143	Yes
Constant relating CO_2 to maintenance energy	β	$mmol_{CO_2}\ g_X^{-1}\ h^{-1}$	4E7	Yes
Constant relating CO_2 to penicillin production	γ	$mmol_{CO_2}\ L^{-1}\ h^{-1}$	10E-4	Yes
Maximum specific biomass growth rate	$\mu_{X,\max}$	h^{-1}	0.092	Yes
Contois saturation constant	k_x	$g_X\ L^{-1}$	0.15	Yes
Oxygen limitations constants (no limitation)	K_{ox}, K_{op}	g_X^{-1} L	0	Yes
Oxygen limitations constants (with limitation)	K_{ox}, K_{op}	g_X^{-1} L	2E-2, 5E-4	Yes
Maximum specific rate of product formation	$\mu_{P,\max}$	h^{-1}	0.005	Yes
Inhibition constant	k_p	$g_S\ L^{-1}$	0.0002	Yes
Inhibition constant for product formation	K_I	$g_S\ L^{-1}$	0.10	Yes
Constant	P	-	3	Yes
Feed temperature of substrate	T_f	K	298	No
Constant for μ	K_1	[mol/L]	1E-10	No
Constant for μ #2	K_2	[mol/L]	7E-5	No
Arrhenius constant for growth	k_g	-	7E3	No
Activation energy for growth	E_g	cal mol^{-1}	5100	No
Arrhenius constant for cell death	k_d	-	1E33	No
Activation energy for cell death	E_d	cal mol^{-1}	50000	No
Density × heat capacity of the medium	ρc_p	cal $°C^{-1}\ L^{-1}$	1/1500	No
Density × heat capacity of the cooling liquid	ρc_{pc}	cal $°C^{-1}\ L^{-1}$	1/2000	No
Yield of heat generation	r_{q1}	cal g_X^{-1}	60	No
Constant in heat generation	r_{q2}	cal $g_X^{-1}\ h^{-1}$	1.6783×10^{-4}	No
Heat transfer coefficient	a	cal $h^{-1}\ °C^{-1}$	1000	No
Constant	b	-	0.60	No
Constant for K_{la}	α_1	-	70	No
Constant for K_{la}	α_2	-	0.4	No
Proportionality constant	γ	mol $[H^+]g_X^{-1}$	1E-5	No
Cooling water temperature	T_c	K	290	No

Table A.3. *Case study 2: input variables of the detailed and simplified models, with indication on whether or not they are measured. In the detailed model, fluctuations are added around the nominal values as smoothed pseudo-random binary signals with the indicated maximum/minimum amplitudes. In the simplified model, constant (nominal) values are used.*

Input variable	Symbol	Units	Nominal value	Max/min amplitude	Measured?	In the FPM?
Feed substrate	F	$L\ h^{-1}$	0.045*	8E-4	Yes	Yes
Aeration rate	f_g	$L\ h^{-1}$	8	0.08	No	No
Agitator power	P_w	W	30	0.45	No	No

*F = 0 $L\ h^{-1}$ in the first 50 h of the batch

- Equations A.14-16 do not consider the effects of pH and temperature on the kinetic parameters and on the volume loss by evaporation, which are significant in the process and are accounted for in the detailed model;

- in the process, the mass transfer coefficient K_{la} depends on two inputs (f_g and P_w, Equation A.5), which are subject to small fluctuations. The FPM neglects this dependency, assuming K_{la} = 120.3 h^{-1}.

References

Abboud, L. and Hensley, S. (2003). New prescription for drug makers: update the plants. *Wall Str. J. 09/23/2003*. Available at: https://www.wsj.com/articles/SB10625358403931000. Last accessed: 09/29/2021.

Accumulus Sinergy (2021). Ten leading biopharma companies announce formation of accumulus synergy to develop global data sharing platform. Available at: https://www.accumulus.org/press-release/ten-leading-biopharma-companies-announce-formation-of-accumulus-synergy-inc-to-develop-global-data-sharing-platform/. Last accessed: 09/29/2021.

Acevedo, D., Peña, R., Yang, Y., Barton, A., Firth, P., and Nagy, Z.K. (2016). Evaluation of mixed suspension mixed product removal crystallization processes coupled with a continuous filtration system. *Chem. Eng. Process. Process Intensif.* **108**, 212–219.

Addison, P.S. (2017). *The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance*. CRC press, Boca Raton (FL; USA).

Al, R., Behera, C.R., Zubov, A., Gernaey, K. V, and Sin, G. (2019). Meta-modeling based efficient global sensitivity analysis for wastewater treatment plants–An application to the BSM2 model. *Comput. Chem. Eng.* **127**, 233–246.

Am Ende, M.T. (2019). *Chemical Engineering in the Pharmaceutical Industry: Drug Product Design, Development, and Modeling*. John Wiley & Sons, Inc., New York (NY; USA).

Armstrong, C.T., Pritchard, C.Q., Cook, D.W., Ibrahim, M., Desai, B.K., Whitham, P.J., Marquardt, B.J., Chen, Y., Zoueu, J.T., and Bortner, M.J. (2019). Continuous flow synthesis of a pharmaceutical intermediate: a computational fluid dynamics approach. *React. Chem. Eng.* **4**, 634–642.

Babnik, S., Erkalvec Zajec, V., Oblak, B., Likozar, B., and Pohar, A. (2020). A review of computational fluid dynamics (CFD) simulations of mixing in the pharmaceutical industry. *Biomed. J. Sci. Tech. Res.* **3**, 20732–20736.

Bajpai, R.K. and Reuß, M. (1980). A mechanistic model for penicillin production. *J. Chem. Technol. Biotechnol.* **30**, 332–344.

Baklouti, I., Mansouri, M., Hamida, A. Ben, Nounou, H., and Nounou, M. (2018). Monitoring of wastewater treatment plants using improved univariate statistical technique. *Process Saf. Environ. Prot.* **116**, 287–300.

Baklouti, I., Mansouri, M., Hamida, A. Ben, Nounou, H.N., and Nounou, M. (2019). Enhanced operation of wastewater treatment plant using state estimation-based fault detection strategies. *Int. J. Control* **94**(2), 300–311.

Ballentine, C. (1981). Taste of raspberries, taste of death: the 1937 elixir sulfanilamide incident. *FDA Consum. Mag.* **15**.

Bamberger, J.A. and Greenwood, M.S. (2004a). Measuring fluid and slurry density and solids concentration non-invasively. *Ultrasonics* **42**, 563–567.

Bamberger, J.A. and Greenwood, M.S. (2004b). Using ultrasonic attenuation to monitor slurry mixing in real time. *Ultrasonics* **42**, 145–148.

Bana, P., Örkényi, R., Lövei, K., Lakó, Á., Túrós, G.I., Éles, J., Faigl, F., and Greiner, I. (2017). The route from problem to solution in multistep continuous flow synthesis of pharmaceutical compounds. *Bioorg. Med. Chem.* **25**, 6180–6189.

Bano, G., Facco, P., Bezzo, F., and Barolo, M. (2018a). Probabilistic design space determination in pharmaceutical product development: A Bayesian/latent variable approach. *AIChE J.* **64**, 2438–2449.

Bano, G., Facco, P., Ierapetritou, M., Bezzo, F., and Barolo, M. (2019). Design space maintenance

by online model adaptation in pharmaceutical manufacturing. *Comput. Chem. Eng.* **127**, 254–271.

Bano, G., Facco, P., Meneghetti, N., Bezzo, F., and Barolo, M. (2017). Uncertainty back-propagation in PLS model inversion for design space determination in pharmaceutical product development. *Comput. Chem. Eng.* **101**, 110–124.

Bano, G., Wang, Z., Facco, P., Bezzo, F., Barolo, M., and Ierapetritou, M. (2018b). A novel and systematic approach to identify the design space of pharmaceutical processes. *Comput. Chem. Eng.* **115**, 309–322.

Bard, Y. (1974). *Nonlinear parameter estimation*. Academic Press, Cambridge (MA; USA).

Bascone, D., Galvanin, F., Shah, N., and Garcia-Munoz, S. (2020). A hybrid mechanistic-empirical approach to the modelling of twin screw feeders for continuous tablet manufacturing. *Ind. Eng. Chem. Res.* **59**(14), 6650-6661.

Baumann, M. and Baxendale, I.R. (2015). The synthesis of active pharmaceutical ingredients (APIs) using continuous flow chemistry. *Beilstein J. Org. Chem.* **11**, 1194–1219.

Belmir, H., Abourriche, A., Bennamara, A., Saffaj, T., and Ihssane, B. (2021). Using design space and response surface methodology for developing a liquid chromatography method for simultaneous determination of five statins in pharmaceutical form. *Acta Chromatogr.* **33**, 345–353.

Benyahia, B., Lakerveld, R., and Barton, P.I. (2012). A plant-wide dynamic model of a continuous pharmaceutical process. *Ind. Eng. Chem. Res.* **51**, 15393–15412.

Bhosekar, A. and Ierapetritou, M. (2018). Advances in surrogate based modeling, feasibility analysis, and optimization: a review. *Comput. Chem. Eng.* **108**, 250–267.

Biegler, L.T. (2010). *Nonlinear programming: concepts, algorithms, and applications to chemical processes*. Society for Industrial and Applied Mathematics, Philadelphia (PA; USA).

Biegler, L.T. (2007). An overview of simultaneous strategies for dynamic optimization. *Chem. Eng. Process. Process Intensif.* **46**, 1043–1053.

Biegler, L.T., Yang, X., and Fischer, G.A.G. (2015). Advances in sensitivity-based nonlinear model predictive control and dynamic real-time optimization. *J. Process Control* **30**, 104–116.

Bird, R.B., Lightfoot, E.N., and Stewart, W.E. (2006). *Transport Phenomena*. John Wiley & Sons, Inc., New York (NY; USA).

Birol, G., Ündey, C., and Çinar, A. (2002). A modular simulation package for fed-batch fermentation: Penicillin production. *Comput. Chem. Eng.* **26**, 1553–1565.

Blackshields, C.A. and Crean, A.M. (2018). Continuous powder feeding for pharmaceutical solid dosage form manufacture: a short review. *Pharm. Dev. Technol.* **23**, 554–560.

Blanke, M., Kinnaert, M., Lunze, J., Staroswiecki, M., and Schröder, J. (2006). D*iagnosis and fault-tolerant control*. Springer, Berlin (Germany).

Boni, A.A. (2016). emerging business models and strategies to accelerate innovation in the biopharmaceutical industry. *J. Commer. Biotechnol.* **22**.

Bonvin, D., Georgakis, C., Pantelides, C.C., Barolo, M., Grover, M.A., Rodrigues, D., Schneider, R., and Dochain, D. (2016). Linking models and experiments. *Ind. Eng. Chem. Res.* **55**, 6891–6903.

Bosca, S. and Fissore, D. (2011). Design and validation of an innovative soft-sensor for pharmaceuticals freeze-drying monitoring. *Chem. Eng. Sci.* **66**, 5127–5136.

Bostijn, N., Dhondt, W., Vervaet, C., and De Beer, T. (2019). PAT-based batch statistical process control of a manufacturing process for a pharmaceutical ointment. *Eur. J. Pharm. Sci.* **136**.

Boukouvala, F., Muzzio, F.J., and Ierapetritou, M.G. (2017). Methods and tools for design space identification in pharmaceutical development. In: *Comprehensive quality by design for pharmaceutical product development and manufacture* (G. V. Reklaitis, C. Seymour, S. García-Munoz, Eds.). John Wiley & Sons, Inc., New York (NY; USA), 95–123.

Boukouvala, F., Muzzio, F.J., and Ierapetritou, M.G. (2010). Design space of pharmaceutical processes using data-driven-based methods. *J. Pharm. Innov.* **5**, 119–137.

Bourcier, D., Féraud, J.P., Colson, D., Mandrick, K., Ode, D., Brackx, E., and Puel, F. (2016). Influence of particle size and shape properties on cake resistance and compressibility during pressure filtration. *Chem. Eng. Sci.* **144**, 176–187.

Box, G.E.P. and Wilson, K.B. (1992). On the experimental attainment of optimum conditions, In: *Breakthroughs in statistics* (S. Kotz and N. L. Johnson, Eds.). Springer, Berlin (Germany), 270–310.

Burcham, C.L., Florence, A.J., and Johnson, M.D. (2018). Continuous manufacturing in pharmaceutical process development and manufacturing. *Annu. Rev. Chem. Biomol. Eng.* **9**, 253–281.

Burgschweiger, J. and Tsotsas, E. (2002). Experimental investigation and modelling of continuous fluidized bed drying under steady-state and dynamic conditions. *Chem. Eng. Sci.* **57**, 5021–5038.

Caccavale, F., Pierri, F., Iamarino, M., and Tufano, V. (2009). An integrated approach to fault diagnosis for a class of chemical batch processes. *J. Process Control* **19**, 827–841.

Camacho, J., Picó, J., and Ferrer, A. (2008). Bilinear modelling of batch processes. Part I: Theoretical discussion. *J. Chemom.* **22**, 299–308.

Cartwright, J.J., Robertson, J., D'Haene, D., Burke, M.D., and Hennenkamp, J.R. (2013). Twin screw wet granulation: Loss in weight feeding of a poorly flowing active pharmaceutical ingredient. *Powder Technol.* **238**, 116–121.

Casas-Orozco, D., Laky, D., Wang, V., Abdi, M., Feng, X., Wood, E., Laird, C., Reklaitis, G. V, and Nagy, Z.K. (2021). PharmaPy: an object-oriented tool for the development of hybrid pharmaceutical flowsheets. *Comput. Chem. Eng.* 153, 107408.

Castagnoli, C., Yahyah, M., Cimarosti, Z., and Peterson, J.J. (2010). Application of quality by design principles for the definition of a robust crystallization process for casopitant mesylate. *Org. Process Res. Dev.* **14**, 1407–1419.

Celikovic, S., Kirchengast, M., Rehrl, J., Kruisz, J., Sacher, S., Khinast, J., and Horn, M. (2020). Model predictive control for continuous pharmaceutical feeding blending units. *Chem. Eng. Res. Des.* **154**, 101–114.

Centers, E.R. (2015). FDA approves tablet production on continuous manufacturing line. Available at: https://www.pharmtech.com/view/fda-approves-tablet-production-janssen-continuous-manufacturing-line. Last accessed: 06/01/2021.

Cervera-Padrell, A E, Nielsen, J.P., Jønch Pedersen, M., Müller Christensen, K., Mortensen, A.R., Skovby, T., Dam-Johansen, K., Kiil, S., and Gernaey, K. V (2012). Monitoring and control of a continuous Grignard reaction for the synthesis of an active pharmaceutical ingredient intermediate using inline NIR spectroscopy. *Org. Process Res. Dev.* **16**, 901–914.

Cervera-Padrell, Albert E., Skovby, T., Kiil, S., Gani, R., and Gernaey, K. V. (2012). Active pharmaceutical ingredient (API) production involving continuous processes - A process system engineering (PSE)-assisted design framework. *Eur. J. Pharm. Biopharm.* **82**, 437–456.

Chatterjee, S., Moore, C.M. V., and Nasr, M.M. (2017). An overview of the role of mathematical models in implementation of quality by design paradigm for drug development and manufacture. *In: Comprehensive quality by design for pharmaceutical product development and manufacture* (G. V. Reklaitis, C. Seymour, S. García-Munoz, Eds.). John Wiley & Sons, Inc., New York (NY; USA), 9–24.

Chatzizacharia, K.A. and Hatziavramidis, D.T. (2014). Design space approach for pharmaceutical tablet development. *Ind. Eng. Chem. Res.* **53**, 12003–12009.

Che Mid, E. and Dua, V. (2017). Model-based parameter estimation for fault detection using multiparametric programming. *Ind. Eng. Chem. Res.* **56**, 8000–8015.

Chen, Y., Yang, O., Sampat, C., Bhalode, P., Ramachandran, R., and Ierapetritou, M. (2020). Digital twins in pharmaceutical and biopharmaceutical manufacturing: A literature review. *Processes* **8**, 1–33.

Cheng, Y.S., Lam, K.W., Ng, K.M., and Wibowo, C. (2010). Workflow for managing impurities in an integrated crystallization process. *AIChE J.* **56**, 633–649.

Chhabra, R. and Basavaraj, M.G. (2019). Drying, in: *Coulson and Richardson's chemical engineering:*

Volume 2: Particulate systems and particle technology. Elsevier, Amsterdam (The Netherlands), 901–968.

Claßen, J., Aupert, F., Reardon, K.F., Solle, D., and Scheper, T. (2017). Spectroscopic sensors for in-line bioprocess monitoring in research and pharmaceutical industrial application. *Anal. Bioanal. Chem.* **409**, 651–666.

Collins, P.C. (2018). Chemical engineering and the culmination of quality by design in pharmaceuticals. *AIChE J.* **64**, 1502–1510.

Colucci, D., Prats-Montalbán, J.M., Ferrer, A., and Fissore, D. (2021). On-line product quality and process failure monitoring in freeze-drying of pharmaceutical products. *Dry. Technol.* **39**, 134–147.

Continuus Pharmaceuticals (2021). CONTINUUS Pharmaceuticals Secures $69.3 Million Government Contract to Manufacture Critical Medicines in the U.S. Available at: https://www.continuuspharma.com/continuus-pharmaceuticals-secures-69-3-million-government-contract-to-manufacture-critical-medicines-in-the-u-s/. Last accessed: 09/29/2021.

Cook, J., Cruañes, M.T., Gupta, M., Riley, S., and Crison, J. (2014). Quality-by-design: Are we there yet? *AAPS PharmSciTech* **15**, 140–148.

Cornish, A.R.H. (1965). Note on minimum possible rate of heat transfer from a sphere when other spheres are adjacent to it. *Trans. Inst. Chem. Eng.* **43**, T332–T333.

Davis, J., Edgar, T., Graybill, R., Korambath, P., Schott, B., Swink, D., Wang, J., and Wetzel, J. (2015). Smart manufacturing. *Annu. Rev. Chem. Biomol. Eng.* **6**, 141–160.

De-Luca, R., Bano, G., Tomba, E., Bezzo, F., and Barolo, M. (2020). Accelerating the development and transfer of freeze-drying operations for the manufacturing of biopharmaceuticals by model-based design of experiments. *Ind. Eng. Chem. Res.* **59**, 20071–20085.

De Beer, T., Burggraeve, A., Fonteyne, M., Saerens, L., Remon, J.P., and Vervaet, C. (2011). Near infrared and Raman spectroscopy for the in-process monitoring of pharmaceutical production processes. *Int. J. Pharm.* **417**, 32–47.

De Beer, T.R.M., Bodson, C., Dejaegher, B., Walczak, B., Vercruysse, P., Burggraeve, A., Lemos, A., Delattre, L., Heyden, Y. Vander, Remon, J.P., Vervaet, C., and Baeyens, W.R.G. (2008). Raman spectroscopy as a process analytical technology (PAT) tool for the in-line monitoring and understanding of a powder blending process. *J. Pharm. Biomed. Anal.* **48**, 772–779.

Delgado-Aguiñaga, J.A., Besançon, G., Begovich, O., and Carvajal, J.E. (2016). Multi-leak diagnosis in pipelines based on Extended Kalman Filter. *Control Eng. Pract.* **49**, 139–148.

Deloitte (2021). Seeds of change Measuring the return from pharmaceutical innovation 2020. Available at: https://www2.deloitte.com/content/dam/Deloitte/us/Documents/deloitte-uk-measuring-the-return-from-pharmaceutical-innovation-2021.pdf. Last accessed: 09/29/2021.

Deloitte (2020a). 2020 Global life sciences outlook. Available at: https://www2.deloitte.com/br/en/pages/life-sciences-and-healthcare/articles/global-life-sciences-sector-outlook.html. Last accessed: 09/29/2021.

Deloitte (2020b). Bringing new therapies to patients. Transforming clinical development. Available at: https://www2.deloitte.com/us/en/insights/industry/life-sciences/future-of-drug-discovery.html. Last accessed: 09/29/2021. Deloitte (2019a). Ten years on - Measuring the return from pharmaceutical innovation 2019. *Deloitte Cent. Heal. Solut.* 1–52.

Deloitte (2019a). Ten years on - Measuring the return from pharmaceutical innovation 2019. Available at: https://www2.deloitte.com/ch/en/pages/life-sciences-and-healthcare/articles/measuring-the-return-from-pharmaceutical-innovation-2019.html. Last accessed: 09/29/2021.

Deloitte (2019b). Intelligent drug discovery powered by AI - A report from the Deloitte Centre for Health Solutions. Available at: https://www2.deloitte.com/us/en/insights/industry/life-sciences/artificial-intelligence-biopharma-intelligent-drug-discovery.html. Last accessed: 09/29/2021.

Deshpande, A.P., Patwardhan, S.C., and Narasimhan, S.S. (2009). Intelligent state estimation for fault tolerant nonlinear predictive control. *J. Process Control* **19**, 187–204.

Domokos, A., Nagy, B., Gyürkés, M., Farkas, A., Tacsi, K., Pataki, H., Liu, Y.C., Balogh, A., Firth,

P., Szilágyi, B., Marosi, G., Nagy, Zoltán K., and Nagy, Zsombor Kristóf (2020). End-to-end continuous manufacturing of conventional compressed tablets: From flow synthesis to tableting through integrated crystallization and filtration. *Int. J. Pharm.* **581**, 119297.

Drăgoi, E.N., Curteanu, S., and Fissore, D. (2013). On the use of artificial neural networks to monitor a pharmaceutical freeze-drying process. *Dry. Technol.* **31**, 72–81.

Dumarey, M., Hermanto, M., Airiau, C., Shapland, P., Robinson, H., Hamilton, P., and Berry, M. (2019). Advances in Continuous Active Pharmaceutical Ingredient (API) Manufacturing: Real-t ime Monitoring Using Multivariate Tools. *J. Pharm. Innov.* **14**, 359–372.

ECDC (2021). Rollout of COVID-19 vaccines in the EU/EEA : challenges and good practice. Available at: https://www.ecdc.europa.eu/en/publications-data/rollout-covid-19-vaccines-eueea-challenges-and-good-practice. Last accessed: 09/29/2021.

EFPIA (2021). The pharmaceutical industry: a key asset to scientific and medical progress the pharmaceutical industry. Available at: https://www.efpia.eu/publications/downloads/efpia/2020-the-pharmaceutical-industry-in-figures/. Last accessed: 09/29/2021.

EMA (2018). The EU Innovation Network (EU-IN). Available at: https://www.ema.europa.eu/en/documents/presentation/presentation-session-1-what-eu-innovation-network_en.pdf. Last accessed: 09/29/2021.

Engisch, W.E. and Muzzio, F.J. (2012). Method for characterization of loss-in-weight feeder equipment. *Powder Technol.* **228**, 395–403.

EvaluatePharma (2020a). World Preview 2020, Outlook to 2026, EvaluatePharma. Available at: https://fondazionecerm.it/wp-content/uploads/2020/07/EvaluatePharma-World-Preview-2020_0.pdf. Last accessed: 09/29/2021.

EvaluatePharma (2020b). Evaluate Vantage 2021 Preview. Available at: https://www.evaluate.com/thought-leadership/vantage/evaluate-vantage-2021-preview. Last accessed: 09/29/2021.

Facco, P., Dal Pastro, F., Meneghetti, N., Bezzo, F., and Barolo, M. (2015). Bracketing the design space within the knowledge space in pharmaceutical product development. *Ind. Eng. Chem. Res.* **54**, 5128–5138.

Facco, P., Santomaso, A.C., and Barolo, M. (2017). Artificial vision system for particle size characterization from bulk materials. *Chem. Eng. Sci.* **164**, 246–257.

FDA (2020). Center for Drug Evaluation and Research - Office of Pharmaceutical Quality - 2020 Annual Report. Assuring quality medicines are available to the American public. Available at: https://www.fdanews.com/ext/resources/files/2021/02-11-21-OPQAnnualReport2020.pdf?1613087522. Last accessed: 09/29/2021.

FDA (2019). Quality Considerations for Continuous Manufacturing. Guidance for Industry (Draft guidance). Available at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/quality-considerations-continuous-manufacturing. Last accessed: 09/29/2021.

FDA (2017). Advancement of Emerging Technology Applications for Pharmaceutical Innovation and Modernization. Guidance for Industry. Available at: https://www.fda.gov/files/drugs/published/Advancement-of-Emerging-Technology-Applications-for-Pharmaceutical-Innovation-and-Modernization-Guidance-for-Industry.pdf. Last accessed: 09/29/2021.

FDA (2013a). Strategic Plan for Preventing and Mitigating Drug Shortages. Available at: https://www.fda.gov/files/drugs/published/Strategic-Plan-for-Preventing-and-Mitigating-Drug-Shortages.pdf. Last accessed: 09/29/2021.

FDA (2013b). Approval Package for Gazyva (obinutuzumab). Application Number: 125486Orig1s000. Available at: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2013/125486Orig1s000Approv.pdf. Last accessed: 09/29/2021.

FDA (2006). Approval Letter for Januvia (sitagliptin phosphate). Application Number: 021995. Available at: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2006/021995s000TOC.cfm. Last accessed: 09/29/2021.

FDA (2004a). Pharmaceutical current good manufacturing practices (cGMPs) for the 21st century—a risk based approach. Available at: https://www.fda.gov/media/77391/download. Last accessed: 09/29/2021.

FDA (2004b). PAT — A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance. Available at: https://www.fda.gov/media/71012/download. Last accessed: 09/29/2021.

Fisher, A.C., Kamga, M.H., Agarabi, C., Brorson, K., Lee, S.L., and Yoon, S. (2019). The current scientific and regulatory landscape in advancing integrated continuous biopharmaceutical Manufacturing. *Trends Biotechnol.* **37**, 253–267.

Fisher, A.C., Lee, S.L., Harris, D.P., Buhse, L., Kozlowski, S., Yu, L., Kopcha, M., and Woodcock, J. (2016). Advancing pharmaceutical quality: An overview of science and research in the U.S. FDA's Office of Pharmaceutical Quality. *Int. J. Pharm.* **515**, 390–402.

Floudas, C.A., Gümüş, Z.H., and Ierapetritou, M.G. (2001). Global optimization in design under uncertainty: feasibility test and flexibility index problems. *Ind. Eng. Chem. Res.* **40**, 4267–4282.

Franceschini, G. and Macchietto, S. (2008). Model-based design of experiments for parameter precision: State of the art. *Chem. Eng. Sci.* **63**, 4846–4872.

Gabor, J.D., Botterill, J.S.M., and Rohsenow, W.M. (1985). Heat transfer in fluidized and packed beds. In: *Handbook of Heat Transfer Applications* (W. M. Rohsenow, J. P. Hartnett, and Y. I. Cho, Eds.), McGraw-Hill, New York (NY; USA), 1–6.

Gagnon, F., Desbiens, A., Poulin, É., Lapointe-Garant, P.-P., and Simard, J.-S. (2017). Nonlinear model predictive control of a batch fluidized bed dryer for pharmaceutical particles. *Control Eng. Pract.* **64**, 88–101.

Galata, D.L., Mészáros, L.A., Kállai-Szabó, N., Szabó, E., Pataki, H., Marosi, G., and Nagy, Z.K. (2021). Applications of machine vision in pharmaceutical technology: a review. *Eur. J. Pharm. Sci.* **159**, 105717.

Gao, Z., Cecati, C., and Ding, S.X. (2015). A survey of fault diagnosis and fault-tolerant techniques-part I: Fault diagnosis with model-based and signal-based approaches. *IEEE Trans. Ind. Electron.* **62**, 3757–3767.

García-Muñoz, S. (2009). Establishing multivariate specifications for incoming materials using data from multiple scales. *Chemom. Intell. Lab. Syst.* **98**, 51–57.

García-Muñoz, S., Butterbaugh, A., Leavesley, I., Manley, L.F., Slade, D., and Bermingham, S. (2018). A flowsheet model for the development of a continuous process for pharmaceutical tablets: An industrial perspective. *AIChE J.* **64**, 511–525.

García-Muñoz, S. and Carmody, A. (2010). Multivariate wavelet texture analysis for pharmaceutical solid product characterization. *Int. J. Pharm.* **398**, 97–106.

García-Muñoz, S., Dolph, S., and Ward II, H.W. (2010). Handling uncertainty in the establishment of a design space for the manufacture of a pharmaceutical product. *Comput. Chem. Eng.* **34**, 1098–1107.

García-Muñoz, S. and Gierer, D.S. (2010). Coating uniformity assessment for colored immediate release tablets using multivariate image analysis. *Int. J. Pharm.* **395**, 104–113.

García-Muñoz, S., Luciani, C. V., Vaidyaraman, S., and Seibert, K.D. (2015). Definition of Design Spaces Using Mechanistic Models and Geometric Projections of Probability Maps. *Org. Process Res. Dev.* **19**, 1012–1023.

Garcia-Munoz, S. and Settell, D. (2009). Application of multivariate latent variable modeling to pilot-scale spray drying monitoring and fault detection: monitoring with fundamental knowledge. *Comput. Chem. Eng.* **33**, 2106–2110.

Geladi, P., Isaksson, H., Lindqvist, L., Wold, S., and Esbensen, K. (1989). Principal component analysis of multivariate images. *Chemom. Intell. Lab. Syst.* **5**, 209–220.

Geladi, P. and Kowalski, B.R. (1986). Partial Least-Squares Regression: a Tutorial. *Anal. Chim. Acta* **186**.

Gernaey, K. V and Gani, R. (2010). A model-based systems approach to pharmaceutical product-process design and analysis. *Chem. Eng. Sci.* **65**, 5757–5769.

Gertler, J. (1998). *Fault detection and diagnosis in engineering systems*. CRC press, Boca Raton (FL; USA).

Ghosh, K., Ng, Y.S., and Srinivasan, R. (2011). Evaluation of decision fusion strategies for effective collaboration among heterogeneous fault diagnostic methods. *Comput. Chem. Eng.* **35**, 342–355.

Giridhar, A. and Reklaitis, G. V (2020). Real-time optimization: how to change setpoints in pharmaceutical manufacturing, in: *Continuous Pharmaceutical Processing*. Springer, Berlin (Germany), 429–440.

Glicksman, L.R. and Joos, F.M. (1980). Heat and mass transfer in fixed beds at low Reynolds numbers. *J. Heat Transfer*, **102**(4), 736-741.

GlobalData Healthcare (2020). Continuous manufacturing: could 2021 be year of the first biologic approval?. Available at: https://www.pharmaceutical-technology.com/comment/continuous-manufacturing-of-biologics-2020/#:~:text=The%20International%20Council%20for%20Harmonization,1)%2C%20all%20small%20molecules. Last accessed: 09/29/2021.

Goyal, V. and Ierapetritou, M.G. (2002). Determination of operability limits using simplicial approximation. *AIChE J.* **48**, 2902–2909.

Grangeia, H.B., Silva, C., Simões, S.P., and Reis, M.S. (2020). Quality by design in pharmaceutical manufacturing: A systematic review of current status, challenges and future perspectives. *Eur. J. Pharm. Biopharm.* **147**, 19–37.

Green, D.W. and Southard, M.Z. (2019). *Perry's chemical engineers' handbook*. McGraw-Hill, New York (NY; USA).

Grieves, M. (2014). Digital twin: manufacturing excellence through virtual factory replication. Available at: https://theengineer.markallengroup.com/production/content/uploads/2014/12/Digital_Twin_White_Paper_Dr_Grieves.pdf. Last accessed: 09/29/2021.

Grossmann, I.E., Calfa, B.A., and Garcia-Herreros, P. (2014). Evolution of concepts and models for quantifying resiliency and flexibility of chemical processes. *Comput. Chem. Eng.* **70**, 22–34.

Grossmann, I.E. and Daichendt, M.M. (1996). New trends in optimization-based approaches to process synthesis. *Comput. Chem. Eng.* **20**, 665–683.

Grossmann, I.E. and Morari, M. (1983). Operability, resiliency, and flexibility: Process design objectives for a changing world.

Gupta, A., Giridhar, A., Venkatasubramanian, V., and Reklaitis, G. V (2013). Intelligent alarm management applied to continuous pharmaceutical tablet manufacturing: an integrated approach. *Ind. Eng. Chem. Res.* **52**, 12357–12368.

Gursch, J., Hohl, R., Dujmovic, D., Brozio, J., Krumme, M., Rasenack, N., and Khinast, J. (2016). Dynamic cross-flow filtration: Enhanced continuous small-scale solid-liquid separation. *Drug Dev. Ind. Pharm.* **42**, 977–984.

Gyürkés, M., Madarász, L., Köte, Á., Domokos, A., Mészáros, D., Beke, Á.K., Nagy, B., Marosi, G., Pataki, H., Nagy, Z.K., and Farkas, A. (2020). Process design of continuous powder blending using residence time distribution and feeding models. *Pharmaceutics* **12**, 1–20.

Haas, N.T., Ierapetritou, M., and Singh, R. (2017). Advanced model predictive feedforward/feedback control of a tablet press. *J. Pharm. Innov.* **12**, 110–123.

Halemane, K.P. and Grossmann, I.E. (1983). Optimal process design under uncertainty. *AIChE J.* **29**, 425–433.

Hamburg, M. (2012). 50 years after thalidomide: why regulation matters. Available at: http://www.gmptrainingsystems.com/files/u2/pdf/Hamburg_Why_Regulation_Matters.pdf. Last accessed: 09/29/2021.

Hanson, J. (2018). Control of a system of loss-in-weight feeders for drug product continuous manufacturing. *Powder Technol.* **331**, 236–243.

Harada, T., Kawakami, K., Yoshihashi, Y., Yonemochi, E., Terada, K., and Moriyamab, H. (2013).

Practical approach for measuring heat capacity of pharmaceutical crystals/glasses by modulated-temperature differential scanning calorimetry. *Chem. Pharm. Bull.* **61**, 315–319.

Harinath, E., Foguth, L.C., and Braatz, R.D. (2016). Maximization of ellipsoidal design space for continuous-time systems: A robust optimal control approach. In: *Proceedings of the 2016 American Control Conference (ACC)*. 3850–3855.

Hart, W.E., Laird, C.D., Watson, J.-P., Woodruff, D.L., Hackebeil, G.A., Nicholson, B.L., and Siirola, J.D. (2017). *Pyomo-optimization modeling in Python*. Springer, Berlin (Germany).

Haseltine, E.L. and Rawlings, J.B. (2005). Critical evaluation of Extended Kalman Filtering and Moving-Horizon Estimation 2451–2460.

Haynes, W.M. (2014). *CRC handbook of chemistry and physics*. CRC press, Boca Raton (FL; USA).

He, Q.P. and Wang, J. (2018). Statistical process monitoring as a big data analytics tool for smart manufacturing. *J. Process Control* **67**, 35–43.

Huang, J., Kaul, G., Cai, C., Chatlapalli, R., Hernandez-Abad, P., Ghosh, K., and Nagi, A. (2009). Quality by design case study: an integrated multivariate approach to drug product and process development. *Int. J. Pharm.* **382**, 23–32.

Huang, Y., Sheriff, M.Z., Bachawala, S., Gonzalez, M., and Nagy, Z.K. (2021). Evaluation of a combined MHE-NMPC approach to handle plant-model mismatch in a rotary tablet press. *Processes* **9**(9), 1612.

Huggins, S., Cosbie, A., and Gaertner, J. (2019). Filtration case studies. In: C*hemical engineering in the pharmaceutical industry: active pharmaceutical ingredients* (D. J. am Ende and M. T. am Ende, Eds.), John Wiley & Sons, Inc., New York (NY; USA), 833–845.

IIBM Business Consulting Services (2005). Transforming industrialization. A new paradigm for pharmaceutical development. Available at: http://www-935.ibm.com/services/uk/igs/pdf/ge510-3997-transforming-industrialization.pdf. Last accessed: 06/05/2021.

ICH (2021). Q13 - Continuous manufacturing of drug substances and drug products. Available at: https://database.ich.org/sites/default/files/ICH_Q13_Step2_DraftGuideline_%202021_0727.pdf. Last accessed: 09/29/2021.

ICH (2019). Q12 - Technical and regulatory considerations for pharmaceutical product lifecycle management. Available at: https://www.ich.org/page/quality-guidelines. Last accessed: 09/29/2021.

ICH (2017). Q11 - Questions and answers. Available at: https://www.ich.org/page/quality-guidelines. Last accessed: 09/29/2021.

ICH (2016). Q3C(R6) - Impurities: guideline for residual solvents. Available at: https://www.ich.org/page/quality-guidelines. Last accessed: 09/29/2021.

ICH (2012a). Q11 - Development and manufacture of drug substances–chemical and biotechnological/biological entities. Available at: https://www.ich.org/page/quality-guidelines. Last accessed: 09/29/2021.

ICH (2012b). Q11 - Development and manufacture of drug substances–chemical and biotechnological/biological entities. Available at: https://www.ich.org/page/quality-guidelines. Last accessed: 09/29/2021.

ICH (2011). Q8, Q9, Q10 - Points to consider. Available at: https://www.ich.org/page/quality-guidelines. Last accessed: 09/29/2021.

ICH (2010). Q8, Q9, Q10 - Questions and Answers. Available at: https://www.ich.org/page/quality-guidelines. Last accessed: 09/29/2021.

ICH (2009). Q8(R2) - Pharmaceutical development. *Q8(R2)*. Available at: https://www.ich.org/page/quality-guidelines. Last accessed: 09/29/2021.

ICH (2008). Q10 - Pharmaceutical quality system. Available at: https://www.ich.org/page/quality-guidelines. Last accessed: 09/29/2021.

ICH (2005). Q9 - Quality Risk Management. Available at: https://www.ich.org/page/quality-guidelines. Last accessed: 09/29/2021.

ICH (2000). Q7 - Good Manufacturing Practice Guide for Active Pharmaceutical Ingredients. Available at: https://www.ich.org/page/quality-guidelines. Last accessed: 09/29/2021.

Ierapetritou, M., Muzzio, F., and Reklaitis, G. (2016). Perspectives on the continuous manufacturing

of powder-based pharmaceutical processes. *AIChE J.* **62**, 1846–1862.

IFPMA (2021). The pharmaceutical industry and global health. Facts and figures., Annual report from the Internation Federation of Pharmaceutical Manufacturers & Associations. Available at: https://www.ifpma.org/wp-content/uploads/2021/04/IFPMA-Facts-And-Figures-2021.pdf. Last accessed: 09/29/2021.

IQVIA Institute (2020). Global Medicine Spending and Usage Trends 27. Available at: https://www.iqvia.com/insights/the-iqvia-institute/reports/global-medicine-spending-and-usage-trends. Last accessed: 09/29/2021.

Isaksson, A.J., Harjunkoski, I., and Sand, G. (2018). The impact of digitalization on the future of control and operations. *Comput. Chem. Eng.* **114**, 122–129.

ISPE (2017). Drug shortages. A report from Pew Charitable Trusts and the International Society for Pharmaceutical Engineering. Available at: https://www.pewtrusts.org/en/research-and-analysis/reports/2017/01/drug-shortages. Last accessed: 09/29/2021.

Jackson, J.E. (1991). *A user's guide to principal components*. John Wiley & Sons, Inc., New York (NY; USA).

Jelsch, M., Roggo, Y., Kleinebudde, P., and Krumme, M. (2021). Model predictive control in pharmaceutical continuous manufacturing: a review from a user's perspective. *Eur. J. Pharm. Biopharm.,* **159**, 137-142.

Jia, R., Mao, Z., Chang, Y., and Zhao, L. (2011). Soft-sensor for copper extraction process in cobalt hydrometallurgy based on adaptive hybrid model. *Chem. Eng. Res. Des.* **89**, 722–728.

Jiang, Q., Yan, X., and Huang, B. (2019). Review and perspectives of data-driven distributed monitoring for industrial plant-wide processes. *Ind. Eng. Chem. Res.* **58**, 12899–12912.

Johnson & Johnson (2021). Johnson & Johnson Statement on U.S. COVID-19 Vaccine Manufacturing. Available at: https://www.jnj.com/johnson-johnson-statement-on-u-s-covid-19-vaccine-manufacturing. Last accessed: 09/29/2021.

Juran, J.M. (1992). *Juran on quality by design: the new steps for planning quality into goods and services*. Simon and Schuster, New York (NY; USA).

Kagermann, H., Lukas, W.-D., and Wahlster, W. (2011). Industrie 4.0: Mit dem Internet der Dinge auf dem Weg zur 4. industriellen Revolution. *VDI nachrichten* **13**, 2–3.

Kajiwara, E., Kamizato, H., and Shikano, M. (2020). Impact of quality by design development on the review period of new drug approval and product quality in Japan. *Ther. Innov. Regul. Sci.* **54**, 1192–1198.

Kamyar, R., Lauri Pla, D., Husain, A., Cogoni, G., and Wang, Z. (2021). Soft sensor for real-time estimation of tablet potency in continuous direct compression manufacturing operation. *Int. J. Pharm.* **602**, 120624.

Khafagy, E.-S., Fayed, M.H., Alrabahi, S.H., Gad, S., Alshahrani, S.M., and Aldawsari, M. (2020). Defining design space for optimization of escitalopram ultra-fast melting tablet using suspension spray-coating technique: In-vitro and in-vivo evaluation. *J. Drug Deliv. Sci. Technol.* **57**, 101631.

Kim, E.J., Kim, J.H., Kim, M.S., Jeong, S.H., and Choi, D.H. (2021). Process analytical technology tools for monitoring pharmaceutical unit operations: A control strategy for continuous process verification. *Pharmaceutics* **13**.

Kirchengast, M., Celikovic, S., Rehrl, J., Sacher, S., Kruisz, J., Khinast, J., and Horn, M. (2019). Ensuring tablet quality via model-based control of a continuous direct compaction process. *Int. J. Pharm.* **567**, 118457.

Kirdar, A.O., Conner, J.S., Baclaski, J., and Rathore, A.S. (2007). Application of multivariate analysis toward biotech processes: case study of a cell-culture unit operation. *Biotechnol. Prog.* **23**, 61–67.

Kishida, M. and Braatz, R.D. (2012). A model-based approach for the construction of design spaces in quality-by-design, In: *Proceedings of the 2012 American Control Conference (ACC),* 1513–1518.

Kona, R., Qu, H., Mattes, R., Jancsik, B., Fahmy, R.M., and Hoag, S.W. (2013). Application of in-

line near infrared spectroscopy and multivariate batch modeling for process monitoring in fluid bed granulation. *Int. J. Pharm.* **452**, 63–72.

Kotidis, P., Demis, P., Goey, C.H., Correa, E., McIntosh, C., Trepekli, S., Shah, N., Klymenko, O. V, and Kontoravdi, C. (2019). Constrained global sensitivity analysis for bioprocess design space identification. *Comput. Chem. Eng.* **125**, 558–568.

Kourti, T. (2019). Pharmaceutical manufacturing: the role of multivariate analysis in design space, control strategy, process understanding, troubleshooting, and optimization. In: C*hemical engineering in the pharmaceutical industry: active pharmaceutical ingredients* (D. J. am Ende and M. T. am Ende, Eds.), 601–629.

Kourti, T. and Davis, B. (2012). The business benefits of quality by design. *Pharm Eng* **32**, 1–10.

Kourti, T., Lepore, J., Liesum, L., Nasr, M., Chatterjee, S., Moore, C.M.V., and Korakianiti, E. (2015). Scientific and regulatory considerations for implementing mathematical models in the quality by design (QbD) framework. *Pharm. Eng.* **35**, 80–88.

Kresta, J. V, Macgregor, J.F., and Marlin, T.E. (1991). Multivariate statistical monitoring of process operating performance. *Can. J. Chem. Eng.* **69**, 35–47.

Kritzinger, W., Karner, M., Traar, G., Henjes, J., and Sihn, W. (2018). Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine* **51**, 1016–1022.

Kruisz, J., Rehrl, J., Faulhammer, E., Witschnigg, A., and Khinast, J.G. (2018). Material tracking in a continuous direct capsule-filling process via residence time distribution measurements. *Int. J. Pharm.* **550**, 347–358.

Ku, W., Storer, R.H., and Georgakis, C. (1995). Disturbance detection and isolation by dynamic principal component analysis. *Chemom. Intell. Lab. Syst.* **30**, 179–196.

Ku, W., Storer, R.H., and Georgakis, C. (1994). Uses of state estimation for statistical process control. *Comput. Chem. Eng.* **18**, S571–S575.

Kucherenko, S., Giamalakis, D., Shah, N., and García-Muñoz, S. (2020). Computationally efficient identification of probabilistic design spaces through application of metamodeling and adaptive sampling. *Comput. Chem. Eng.* **132**, 106608.

Kunii, D. and Suzuki, M. (1967). Particle-to-fluid heat and mass transfer in packed beds of fine particles. *Int. J. Heat Mass Transf.* **10**, 845–852.

Kusumo, K.P., Gomoescu, L., Paulen, R., Garciá Munõz, S., Pantelides, C.C., Shah, N., and Chachuat, B. (2020). Bayesian Approach to Probabilistic Design Space Characterization: A Nested Sampling Strategy. *Ind. Eng. Chem. Res.* **59**, 2396–2408.

Lakerveld, R., Benyahia, B., Braatz, R.D., and Barton, P.I. (2013). Model-based design of a plant-wide control strategy for a continuous pharmaceutical plant. *AIChE J.* **59**, 3671–3685.

Laky, D., Xu, S., Rodriguez, J.S., Vaidyaraman, S., Garciá Munõz, S., and Laird, C. (2019). An optimization-based framework to define the probabilistic design space of pharmaceutical processes with model uncertainty. *Processes* **7**(2), 96.

Lapidus, L. and Amundson, N.R. (1952). Mathematics of adsorption in beds. VI. The effect of longitudinal diffusion in ion exchange and chromatographic columns. *J. Phys. Chem.* **56**, 984–988.

Lebrun, P., Krier, F., Mantanus, J., Grohganz, H., Yang, M., Rozet, E., Boulanger, B., Evrard, B., Rantanen, J., and Hubert, P. (2012). Design space approach in the optimization of the spray-drying process. *Eur. J. Pharm. Biopharm.* **80**, 226–234.

Lee, B.W., Peterson, J.J., Yin, K., Stockdale, G.S., Liu, Y.C., and O'Brien, A. (2020). System model development and computer experiments for continuous API manufacturing. *Chem. Eng. Res. Des.* **156**, 495–506.

Lee, S.L., O'Connor, T.F., Yang, X., Cruz, C.N., Chatterjee, S., Madurawe, R.D., Moore, C.M.V., Yu, L.X., and Woodcock, J. (2015). Modernizing pharmaceutical manufacturing: from batch to continuous production. *J. Pharm. Innov.* **10**, 191–199.

Li, W. (2014). Drying of pharmaceutical powders using an agitated filter dryer. *Ph.D. Thesis*, University

of Leeds, Leeds (UK).
Li, X., Yang, X., Yang, Y., Bennett, I., Collop, A., and Mba, D. (2019). Canonical variate residuals-based contribution map for slowly evolving faults. *J. Process Control* **76**, 87–97.
Ling, C. and Kravaris, C. (2016). State Observer Design for Monitoring the Degree of Polymerization in a Series of Melt Polycondensation Reactors. *Processes* **4**, 4.
Liotta, V., Georgakis, C., and El-Aasser, M.S. (1997). Real-time estimation and control of particle size in semi-batch emulsion polymerization, in: Proceedings of the 1997 American Control Conference (Cat. No. 97CH36041). IEEE, pp. 1172–1176.
Lipsky, M.S. and Sharp, L.K. (2001). From idea to market: the drug approval process. *J. Am. Board Fam. Pract.* **14**, 362–367.
Liu, H., Galbraith, S.C., Ricart, B., Stanton, C., Smith-Goettler, B., Verdi, L., O'Connor, T., Lee, S., and Yoon, S. (2017). Optimization of critical quality attributes in continuous twin-screw wet granulation via design space validated with pilot scale experimental data. *Int. J. Pharm.* **525**, 249–263.
Liu, H., Zhang, Z., and Linhardt, R.J. (2009). Lessons learned from the contamination of heparin. *Nat. Prod. Rep.* **26**, 313–321.
Liu, J., Su, Q., Moreno, M., Laird, C., Nagy, Z., and Reklaitis, G. (2018). Robust state estimation of feeding–blending systems in continuous pharmaceutical manufacturing. *Chem. Eng. Res. Des.* **134**, 140–153.
Liu, S., Papageorgiou, L.G., and Shah, N. (2020). Optimal design of low-cost supply chain networks on the benefits of new product formulations. *Comput. Ind. Eng.* **139**, 106189.
Liu, Y.C., Domokos, A., Coleman, S., Firth, P., and Nagy, Z.K. (2019). Development of continuous filtration in a novel continuous filtration carousel integrated with continuous crystallization. *Org. Process Res. Dev.* **23**, 2655–2665.
Liu, Z., Bruwer, M.-J., MacGregor, J.F., Rathore, S.S.S., Reed, D.E., and Champagne, M.J. (2011a). Scale-up of a pharmaceutical roller compaction process using a joint-Y partial least squares model. *Ind. Eng. Chem. Res.* **50**, 10696–10706.
Liu, Z., Bruwer, M.-J., MacGregor, J.F., Rathore, S.S.S., Reed, D.E., and Champagne, M.J. (2011b). Modeling and optimization of a tablet manufacturing line. *J. Pharm. Innov.* **6**, 170–180.
López-Negrete, R. and Biegler, L.T. (2012). A moving horizon estimator for processes with multi-rate measurements: a nonlinear programming sensitivity approach. *J. Process Control* **22**, 677–688.
Luyben, W.L. and Luyben, M.L. (1997). *Essentials of process control*. McGraw-Hill, New York (NY; USA).
MacGregor, J.F. and Bruwer, M.J. (2008). A framework for the development of design and control spaces. *J. Pharm. Innov.* **3**, 15–22.
Mascia, S., Heider, P.L., Zhang, H., Lakerveld, R., Benyahia, B., Barton, P.I., Braatz, R.D., Cooney, C.L., Evans, J.M.B., Jamison, T.F., Jensen, K.F., Myerson, A.S., and Trout, B.L. (2013). End-to-end continuous manufacturing of pharmaceuticals: Integrated synthesis, purification, and final dosage formation. *Angew. Chemie - Int. Ed.* **52**, 12359–12363.
Mathe, R., Casian, T., and Tomuţă, I. (2020). Multivariate feed forward process control and optimization of an industrial, granulation based tablet manufacturing line using historical data. *Int. J. Pharm.* **591**, 119988.
McKinsey & Company (2021). Smart quality: Reimagining the way quality works. Available at: https://www.mckinsey.com/industries/life-sciences/our-insights/smart-quality-reimagining-the-way-quality-works. Last accessed: 09/29/2021.
McWilliams, J.C., Allian, A.D., Opalka, S.M., May, S.A., Journet, M., and Braden, T.M. (2018). The evolving state of continuous processing in pharmaceutical API manufacturing: a survey of pharmaceutical companies and contract manufacturing organizations. *Org. Process Res. Dev.* **22**, 1143–1166.
Mesbah, A., Paulson, J.A., Lakerveld, R., and Braatz, R.D. (2017). Model Predictive Control of an

Integrated Continuous Pharmaceutical Manufacturing Pilot Plant. *Org. Process Res. Dev.* **21**, 844–854.

Miller, P., Swanson, R.E., and Heckler, C.E. (1998). Contribution plots: a missing link in multivariate quality control. *Appl. Math. Comput. Sci.* **8**, 775–792.

Milmo, S. (2014). Quality by Design—Bridging the Gap between Concept and Implementation. *Pharm. Technol.* **38**, 18–20.

Minatovicz, B., Bogner, R., and Chaudhuri, B. (2021). Use of a Design of Experiments (DoE) Approach to Optimize Large-Scale Freeze-Thaw Process of Biologics. *AAPS PharmSciTech* **22**, 153.

Mohd, J., Hoang, N.H., Hussain, M.A., and Dochain, D. (2015). Review and classification of recent observers applied in chemical process systems. *Comput. Chem. Eng.* **76**, 27–41.

Monroy, I., Villez, K., Graells, M., and Venkatasubramanian, V. (2012). Fault diagnosis of a benchmark fermentation process: a comparative study of feature extraction and classification techniques. *Bioprocess Biosyst. Eng.* **35**, 689–704.

Montes, F.C.C., Gernaey, K., and Sin, G. (2018). Dynamic plantwide modeling, uncertainty, and sensitivity analysis of a pharmaceutical upstream synthesis: ibuprofen case study. *Ind. Eng. Chem. Res.* **57**, 10026–10037.

Moreno, M., Liu, J., Su, Q., Leach, C., Giridhar, A., O'Connor, T., Yazdanpanah, N., Nagy, Z.K., and Reklaitis, G. V. (2018). Steady-State Data Reconciliation Framework for a Direct Continuous Tableting Line. *J. Pharm. Innov.* **14**(3), 221-238.

Muskat, M. and Meres, M.W. (1936). The flow of heterogeneous fluids through porous media. *Physics* **7**, 346–363.

Myerson, A.S., Krumme, M., Nasr, M., Thomas, H., and Braatz, R.D. (2015). Control systems engineering in continuous pharmaceutical manufacturing may 20-21, 2014 continuous manufacturing symposium. *J. Pharm. Sci.* **104**, 832–839.

Nagy, B., Farkas, A., Gyürkés, M., Komaromy-Hiller, S., Démuth, B., Szabó, B., Nusser, D., Borbás, E., Marosi, G., and Nagy, Z.K. (2017). In-line Raman spectroscopic monitoring and feedback control of a continuous twin-screw pharmaceutical powder blending and tableting process. *Int. J. Pharm.* **530**, 21–29.

Nagy, B., Szilágyi, B., Domokos, A., Tacsi, K., Pataki, H., Marosi, G., Nagy, Zsombor Kristóf, and Nagy, Zoltan K. (2021). Modeling of pharmaceutical filtration and continuous integrated crystallization-filtration processes. *Chem. Eng. J.* **413**.

Nagy, Z.K. and Braatz, R.D. (2012). Advances and new directions in crystallization control. *Annu. Rev. Chem. Biomol. Eng.* **3**, 55–75.

Nagy, Z.K., Fevotte, G., Kramer, H., and Simon, L.L. (2013). Recent advances in the monitoring, modelling and control of crystallization systems. *Chem. Eng. Res. Des.* **91**, 1903–1922.

Natof, T. and Pellegrini, M. V (2021). Food & Drug Administration Recalls. Available at: https://www.ncbi.nlm.nih.gov/books/NBK570589/. Last accessed: 09/29/2021. Nicholson, B., López-Negrete, R., and Biegler, L.T. (2014). On-line state estimation of nonlinear dynamic systems with gross errors. *Comput. Chem. Eng.* **70**, 149–159.

Nicholson, B., Siirola, J.D., Watson, J.-P., Zavala, V.M., and Biegler, L.T. (2018). pyomo. dae: A modeling and automatic discretization framework for optimization with differential and algebraic equations. *Math. Program. Comput.* **10**, 187–223.

Nikolakopoulou, A., von Andrian, M., and Braatz, R.D. (2020). Fast model predictive control of startup of a compact modular reconfigurable system for continuous-flow pharmaceutical manufacturing. In: *Proceedings of the 2020 American Control Conference (ACC)*, 2778–2783.

Nomikos, P. and Macgregor, J.F. (1995). Multivariate Processes SPC Charts for Monitoring. *Technometrics* **37**, 41–59.

O'Connor, T.F., Yu, L.X., and Lee, S.L. (2016). Emerging technology: A key enabler for modernizing pharmaceutical manufacturing and advancing product quality. *Int. J. Pharm.* **509**,

492–498.
Ochoa, M.P., García-Muñoz, S., Stamatis, S., and Grossmann, I.E. (2021). Novel flexibility index formulations for the selection of the operating range within a design space. *Comput. Chem. Eng.* **149**, 107284.
Öner, M., Stocks, S.M., and Sin, G. (2020). Comprehensive sensitivity analysis and process risk assessment of large scale pharmaceutical crystallization processes. *Comput. Chem. Eng.* **135**.
Orehek, J., Teslić, D., and Likozar, B. (2020). Continuous crystallization processes in pharmaceutical manufacturing: A Review. *Org. Process Res. Dev.* **25**, 16–42.
Ottoboni, S., Price, C.J., Steven, C., Meehan, E., Barton, A., Firth, P., Mitchell, A., and Tahir, F. (2019). Development of a Novel Continuous Filtration Unit for Pharmaceutical Process Development and Manufacturing. *J. Pharm. Sci.* **108**, 372–381.
Ottoboni, S., Shahid, M., Steven, C., Coleman, S., Meehan, E., Barton, A., Firth, P., Sutherland, R., and Price, C.J. (2020a). Developing a batch isolation procedure and running it in an automated semicontinuous unit: AWL CFD25 Case Study. *Org. Process Res. Dev.* **24**, 520–539.
Ottoboni, S., Simurda, M., Wilson, S., Irvine, A., Ramsay, F., and Price, C.J. (2020b). Understanding effect of filtration and washing on dried product: Paracetamol case study. *Powder Technol.* **366**, 305–323.
Pantelides, C.C. and Renfro, J.G. (2013). The online use of first-principles models in process operations: Review, current status and future needs. *Comput. Chem. Eng.* **51**, 136–148.
Papadakis, E., Woodley, J.M., and Gani, R. (2018). Perspective on PSE in pharmaceutical process development and innovation. In: *Computer Aided Chemical Engineering* **41**, Elsevier, Amsterdam (The Netherlands), 597-656.
Papathanasiou, M.M., Burnak, B., Katz, J., Shah, N., and Pistikopoulos, E.N. (2019). Assisting continuous biomanufacturing through advanced control in downstream purification. *Comput. Chem. Eng.* **125**, 232–248.
Papathanasiou, M.M., Steinebach, F., Morbidelli, M., Mantalaris, A., and Pistikopoulos, E.N. (2017). Intelligent, model-based control towards the intensification of downstream processes. *Comput. Chem. Eng.* **105**, 173–184.
Paulson, J.A., Streif, S., Findeisen, R., Braatz, R.D., and Mesbah, A. (2018). Fast stochastic model predictive control of end-to-end continuous pharmaceutical manufacturing. In: *Computer Aided Chemical Engineering* **41**, Elsevier, Amsterdam (The Netherlands), 353–378.
Pérez, G., Garmendia, M., Reynaud, J.F., Crego, J., and Viscarret, U. (2015). Enhanced closed loop State of Charge estimator for lithium-ion batteries based on Extended Kalman Filter. *Appl. Energy* **155**, 834–845.
Peterson, J.J. (2008). A Bayesian approach to the ICH Q8 definition of design space. *J. Biopharm. Stat.* **18**, 959–975.
Peterson, J.J. (2004). A posterior predictive approach to multiple response surface optimization. *J. Qual. Technol.* **36**, 139–153.
Pirnay, H., López-Negrete, R., and Biegler, L.T. (2012). Optimal sensitivity based on IPOPT. *Math. Program. Comput.* **4**, 307–331.
Pla, D.L., Kamyar, R., Hashemian, N., Mehdizadeh, H., and Moshgbar, M. (2018). Moisture soft sensor for batch fluid bed dryers: A practical approach. *Powder Technol.* **326**, 69–77.
Plumb, K. (2005). Continuous processing in the pharmaceutical industry: Changing the mind set. *Chem. Eng. Res. Des.* **83**, 730–738.
Qin, S.J. (2012). Survey on data-driven industrial process monitoring and diagnosis. *Annu. Rev. Control* **36**, 220–234.
Quatrini, E., Li, X., Mba, D., and Costantino, F. (2020). Fault Diagnosis of a Granulator Operating under Time-Varying Conditions Using Canonical Variate Analysis. *Energies* **13**, 4427.
Rantanen, J. and Khinast, J. (2015). The Future of Pharmaceutical Manufacturing Sciences. *J. Pharm. Sci.* **104**, 3612–3638.
Rao, C. V (2000). Moving horizon strategies for the constrained monitoring and control of nonlinear

discrete-time systems. *PhD Thesis*, The University of Wisconsin - Madison, Madison (WI; USA).
Rao, C. V and Rawlings, J.B. (2002). Constrained Process Monitoring : Moving-Horizon Approach **48**.
Rathore, A.S., Singh, S.K., Kumar, J., and Kapoor, G. (2018). Implementation of QbD for manufacturing of biologics - Has it met the expectations?. In: *Biopharmaceutical processing: development, design, and implementation of manufacturing processes* (Jagschies G., Lindskog E., Łącki, K., and Galliher, P., Eds.), Elsevier, Amsterdam (The Netherlands), 1051-1073.
Rawlings, J.B., Mayne, D.Q., and Diehl, M. (2017). Model predictive control: theory, computation, and design. Nob Hill Publishing Madison, WI.
Rawlings, J.B., Miller, S.M., and Witkowski, W.R. (1993). Model identification and control of solution crystallization processes: A Review. *Ind. Eng. Chem. Res.* **32**, 1275–1296.
Ray, W.H. (1981). *Advanced process control*. McGraw-Hill, New York (NY; USA).
Rehrl, J., Karttunen, A.P., Nicolaï, N., Hörmann, T., Horn, M., Korhonen, O., Nopens, I., De Beer, T., and Khinast, J.G. (2018). Control of three different continuous pharmaceutical manufacturing processes: use of soft sensors. *Int. J. Pharm.* **543**, 60–72.
Rehrl, J., Kruisz, J., Sacher, S., Khinast, J., and Horn, M. (2016). Optimized continuous pharmaceutical manufacturing via model-predictive control. *Int. J. Pharm.* **510**, 100–115.
Reinhardt, I.C., Oliveira, J.C., and Ring, D.T. (2020). Current perspectives on the development of Industry 4.0 in the pharmaceutical sector. *J. Ind. Inf. Integr.* **18**, 100131.
Reis, M.S., Gins, G., and Rato, T.J. (2019). Incorporation of process-specific structure in statistical process monitoring : A review. *J. Qual. Technol.* 1–15.
Reklaitis, G. V, Seymour, C., and García-Munoz, S. (2017). *Comprehensive Quality by Design for Pharmaceutical Product Development and Manufacture*. John Wiley & Sons, Inc., New York (NY; USA).
Ricker, N.L. and Lee, J.H. (1995). Nonlinear modeling and state estimation for the Tennessee Eastman challenge process. *Comput. Chem. Eng.* **19**, 983–1005.
Ripperger, S., Gösele, W., Alt, C., and Loewe, T. (2000). Filtration, 1. fundamentals. *Ullmann's Encycl. Ind. Chem.* 1–38.
Rogers, A. and Ierapetritou, M. (2015a). Challenges and opportunities in modeling pharmaceutical manufacturing processes. *Comput. Chem. Eng.* **81**, 32–39.
Rogers, A. and Ierapetritou, M. (2015b). Feasibility and flexibility analysis of black-box processes part 2: Surrogate-based flexibility analysis. *Chem. Eng. Sci.* **137**, 1005–1013.
Rossi, R.C., Dias, C.L., Donato, E.M., Martins, L.A., Bergold, A.M., and Fröehlich, P.E. (2007). Development and validation of dissolution test for ritonavir soft gelatin capsules based on in vivo data. *Int. J. Pharm.* **338**, 119–124.
Rusinov, L.A., Vorobiev, N. V., and Kurkina, V. V. (2013). Fault diagnosis in chemical processes and equipment with feedbacks. *Chemom. Intell. Lab. Syst.* **126**, 123–128.
Sagmeister, P., Wechselberger, P., Jazini, M., Meitz, A., Langemann, T., and Herwig, C. (2013). Soft sensor assisted dynamic bioprocess control: Efficient tools for bioprocess development. *Chem. Eng. Sci.* **96**, 190–198.
Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S. (2008). *Global sensitivity analysis: the primer.* John Wiley & Sons, Inc., New York (NY; USA).
Sansana, J., Joswiak, M.N., Castillo, I., Wang, Z., Rendall, R., Chiang, L.H., and Reis, M.S. (2021). Recent trends on hybrid modeling for Industry 4.0. *Comput. Chem. Eng.* 107365.
Sarkar, A., Shoemaker, B., Doshi, P., am Ende, M.T., Jajcevic, D., Böhling, P., Toson, P., Zadravec, M., and Khinast, J.G. (2019). Multiscale modeling of a pharmaceutical fluid bed coating process using CFD/DEM and population balance models to predict coating uniformity. *Chem. Eng. Pharm. Ind. Drug Prod. Des. Dev. Model.* 419–450.
Sarkis, M., Bernardi, A., Shah, N., and Papathanasiou, M.M. (2021). Emerging challenges and opportunities in pharmaceutical manufacturing and distribution. *Processes* **9**, 457.

Sarraguça, M.C., Ribeiro, P.R.S., Santos, A.O., Silva, M.C.D., and Lopes, J.A. (2014). A PAT approach for the on-line monitoring of pharmaceutical co-crystals formation with near infrared spectroscopy. *Int. J. Pharm.* **471**, 478–484.

Scali, C., Morretta, M., and Semino, D. (1997). Control of the quality of polymer products in continuous reactors: Comparison of performance of state estimators with and without updating of parameters. *J. Process Control* **7**, 357–369.

Schaber, S.D., Gerogiorgis, D.I., Ramachandran, R., Evans, J.M.B., Barton, P.I., and Trout, B.L. (2011). Economic analysis of integrated continuous and batch pharmaceutical manufacturing: a case study. *Ind. Eng. Chem. Res.* **50**, 10083–10092.

Schneider, R. and Georgakis, C. (2013). How to NOT make the extended kalman filter fail. *Ind. Eng. Chem. Res.* **52**, 3354–3362.

Seborg, D.E., Edgar, T.F., Mellichamp, D.A., and Doyle, F.J. (2017). *Process Dynamics and Control*, John Wiley & Sons, Inc., New York (NY; USA).

Sen, M., Chaudhury, A., Singh, R., John, J., and Ramachandran, R. (2013a). Multi-scale flowsheet simulation of an integrated continuous purification-downstream pharmaceutical manufacturing process. *Int. J. Pharm.* **445**, 29–38.

Sen, M., Rogers, A., Singh, R., Chaudhury, A., John, J., Ierapetritou, M.G., and Ramachandran, R. (2013b). Flowsheet optimization of an integrated continuous purification-processing pharmaceutical manufacturing operation. *Chem. Eng. Sci.* **102**, 56–66.

Sen, M., Singh, R., and Ramachandran, R. (2014). A Hybrid MPC-PID control system design for the continuous purification and processing of active pharmaceutical ingredients. *Processes* **2**, 392–418.

Sencar, J., Hammerschmidt, N., and Jungbauer, A. (2020). Modeling the residence time distribution of integrated continuous bioprocesses. *Biotechnol. J.* **15**, 2000008.

Severson, K.A., VanAntwerp, J.G., Natarajan, V., Antoniou, C., Thömmes, J., and Braatz, R.D. (2018). A systematic approach to process data analytics in pharmaceutical manufacturing: The data analytics triangle and its application to the manufacturing of a monoclonal antibody. In: *Multivariate analysis in the pharmaceutical industry.* Elsevier, Amsterdam (The Netherlands), 295–312.

Shahmohammadi, A. and McAuley, K.B. (2020). Using prior parameter knowledge in model-based design of experiments for pharmaceutical production. *AIChE J.* **66**, e17021.

Shi, Z., Hermiller, J., and Muñoz, S.G. (2019). Estimation of mass-based composition in powder mixtures using Extended Iterative Optimization Technology (EIOT). *AIChE J.* **65**, 87–98.

Simon, D. (2006). *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons, Inc., New York (NY; USA).

Simon, L.L., Pataki, H., Marosi, G., Meemken, F., Hungerbühler, K., Baiker, A., Tummala, S., Glennon, B., Kuentz, M., and Steele, G. (2015). Assessment of recent process analytical technology (PAT) trends: a multiauthor review. *Org. Process Res. Dev.* **19**, 3–62.

Singh, R., Muzzio, F.J., Ierapetritou, M., and Ramachandran, R. (2015). A combined feed-forward/feed-back control system for a QbD-based continuous tablet manufacturing process. *Processes* **3**, 339–356.

Singh, R., Sahay, A., Karry, K.M., Muzzio, F., Ierapetritou, M., and Ramachandran, R. (2014a). 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.* **473**, 38–54.

Singh, R., Sahay, A., Muzzio, F., Ierapetritou, M., and Ramachandran, R. (2014b). A systematic framework for onsite design and implementation of a control system in a continuous tablet manufacturing process. *Comput. Chem. Eng.* **66**, 186–200.

Singh, R., Velazquez, C., Sahay, A., Karry, K.M., Muzzio, F.J., Ierapetritou, M.G., and Ramachandran, R. (2016). Advanced control of continuous pharmaceutical tablet manufacturing processes. In: *Process simulation and data modeling in solid oral drug development and manufacture*. Springer, Berling (Germany), 191–224.

Sinner, P., Stiegler, M., Herwig, C., and Kager, J. (2021). Noninvasive online monitoring of Corynebacterium glutamicum fed-batch bioprocesses subject to spent sulfite liquor raw material uncertainty. *Bioresour. Technol.* **321**, 124395.

Smiatek, J., Jung, A., and Bluhmki, E. (2020). Towards a digital bioprocess replica: computational approaches in biopharmaceutical development and manufacturing. *Trends Biotechnol.* **38**, 1141–1153.

Sobol, I.M. (1993). Sensitivity estimates for nonlinear mathematical models. *Math. Model. Comput. Exp* **1**, 407–414.

Souza, F.A.A., Araújo, R., and Mendes, J. (2016). Review of soft sensor methods for regression applications. *Chemom. Intell. Lab. Syst.* **152**, 69–79.

Steinwandter, V., Borchert, D., and Herwig, C. (2019). Data science tools and applications on the way to Pharma 4.0. *Drug Discov. Today* **24**, 1795–1805.

Su, Q., Bommireddy, Y., Shah, Y., Ganesh, S., Moreno, M., Liu, J., Gonzalez, M., Yazdanpanah, N., O'Connor, T., Reklaitis, G. V., and Nagy, Z.K. (2019a). Data reconciliation in the Quality-by-Design (QbD) implementation of pharmaceutical continuous tablet manufacturing. *Int. J. Pharm.* **563**, 259–272.

Su, Q., Ganesh, S., Moreno, M., Bommireddy, Y., Gonzalez, M., Reklaitis, G. V., and Nagy, Z.K. (2019b). A perspective on Quality-by-Control (QbC) in pharmaceutical continuous manufacturing. *Comput. Chem. Eng.* **125**, 216–231.

Su, Q., Moreno, M., Giridhar, A., Reklaitis, G. V., and Nagy, Z.K. (2017). A Systematic Framework for Process Control Design and Risk Analysis in Continuous Pharmaceutical Solid-Dosage Manufacturing. *J. Pharm. Innov.* **12**, 327–346.

Su, Q., Reklaitis, G. V, and Nagy, Z.K. (2020). Continuous Feeding-Blending in Pharmaceutical Continuous Manufacturing, in: Continuous Pharmaceutical Processing. Springer, pp. 193–226.

Suresh, P. and Basu, P.K. (2008). Improving pharmaceutical product development and manufacturing: Impact on cost of drug development and cost of goods sold of pharmaceuticals. *J. Pharm. Innov.* **3**, 175–187.

Suresh, P., Hsu, S., Akkisetty, P., Reklaitis, G. V, and Venkatasubramanian, V. (2010). OntoMODEL: Ontological Mathematical Modeling Knowledge Management in Pharmaceutical Product Development, 1: Conceptual Framework. *Ind. Eng. Chem. Res.* **49**, 7758–7767.

Swaney, R.E. and Grossmann, I.E. (1985). An index for operational flexibility in chemical process design. Part I: Formulation and theory. *AIChE J.* **31**, 621–630.

Tabora, J.E., Lora Gonzalez, F., and Tom, J.W. (2019). Bayesian probabilistic modeling in pharmaceutical process development. *AIChE J.* **65**, e16744.

Tahir, F., Islam, M.T., Mack, J., Robertson, J., and Lovett, D. (2019). Process monitoring and fault detection on a hot-melt extrusion process using in-line Raman spectroscopy and a hybrid soft sensor. *Comput. Chem. Eng.* **125**, 400–414.

Tarleton, S. and Wakeman, R. (2006). Solid/liquid separation: equipment selection and process design. Elsevier.

ter Horst, J.P., Turimella, S.L., Metsers, F., and Zwiers, A. (2021). Implementation of Quality by Design (QbD) Principles in Regulatory Dossiers of Medicinal Products in the European Union (EU) Between 2014 and 2019. *Ther. Innov. Regul. Sci.* **55**, 583–590.

Tidriri, K., Chatti, N., Verron, S., and Tiplica, T. (2016). Bridging data-driven and model-based approaches for process fault diagnosis and health monitoring: A review of researches and future challenges. *Annu. Rev. Control* **42**, 63–81.

Tomba, E., Barolo, M., and García-Muñoz, S. (2012). General framework for latent variable model inversion for the design and manufacturing of new products. *Ind. Eng. Chem. Res.* **51**, 12886–12900.

Tomba, E., Facco, P., Bezzo, F., and Barolo, M. (2013). Latent variable modeling to assist the implementation of Quality-by-Design paradigms in pharmaceutical development and

manufacturing: A review. *Int. J. Pharm.* **457**, 283–297.
Torjesen, I. (2021). Covid-19 vaccine shortages: what is the cause and what are the implications? *BMJ Br. Med. J.* **372**.
Treybal, R.E. (1980). Mass transfer operations. *New York* **466**.
Troup, G.M. and Georgakis, C. (2013). Process systems engineering tools in the pharmaceutical industry. *Comput. Chem. Eng.* **51**, 157–171.
Udugama, I.A., Lopez, P.C., Gargalo, C.L., Li, X., Bayer, C., and Gernaey, K. V. (2021). Digital Twin in biomanufacturing: challenges and opportunities towards its implementation. *Syst. Microbiol. Biomanufacturing* **1**, 257–274.
Ündey, C., Ertunç, S., and Çinar, A. (2003). Online batch/fed-batch process performance monitoring, quality prediction, and variable-contribution analysis for diagnosis. *Ind. Eng. Chem. Res.* **42**, 4645–4658.
Valappil, J. and Georgakis, C. (2000). Systematic estimation of state noise statistics for extended Kalman filters. *AIChE J.* **46**, 292–308.
Van Leer, B. (1974). Towards the ultimate conservative difference scheme. II. Monotonicity and conservation combined in a second-order scheme. *J. Comput. Phys.* **14**, 361–370.
Varshney, D., Bhushan, M., and Patwardhan, S.C. (2019). State and parameter estimation using extended Kitanidis Kalman filter. *J. Process Control* **76**, 98–111.
Velayudhan, A. (2014). Overview of integrated models for bioprocess engineering. *Curr. Opin. Chem. Eng.* **6**, 83–89.
Venkatasubramanian, V., Rengaswamy, R., Yin, K., and Kavuri, S.N. (2003a). A review of process fault detection and diagnosis Part I: Quantitative model-based methods. *Comput. Chem. Eng.* **27**, 293–311.
Venkatasubramanian, V., Rengaswamy, R., Yin, K., and Kavuri, S.N. (2003b). A review of fault detection and diagnosis. Part III: Process history based methods. *Comput. Chem. Eng.* **27**, 327–346.
Vishwasrao, S.S. and Singh, S. (2016). Current Perspective on Opportunities and Adoption Challenges of QbD Implementation in Pharmaceutical Product Development. *Inven. Rapid Pharm. Process Dev.* **2016**.
von Stosch, M., Oliveira, R., Peres, J., and Feyo de Azevedo, S. (2013). Hybrid semi-parametric modeling in process systems engineering: Past, present and future. *Comput. Chem. Eng.* **60**, 86–101.
Wächter, A. and Biegler, L.T. (2006). On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Program.* **106**, 25–57.
Wakeman, R. and Tarleton, S. (2005). *Solid/liquid separation: principles of industrial filtration*. Elsevier, Amsterdam (The Netherlands).
Wakeman, R.J. (1979). Low-pressure dewatering kinetics of incompressible filter cakes, I. Variable total pressure loss or low-capacity systems. *Int. J. Miner. Process.* **5**, 379–393.
Wakeman, R.J. (1976). Vacuum dewatering and residual saturation of incompressible filter cakes. *Int. J. Miner. Process.* **3**, 193–206.
Wakeman, R.J. and Attwood, G.J. (1990). Simulations of dispersion phenomena in filter cake washing. *Chem. Eng. Res. Des.* **68**, 161–171.
Wang, J. and Lakerveld, R. (2017). Continuous membrane-assisted crystallization to increase the attainable product quality of pharmaceuticals and design space for operation. *Ind. Eng. Chem. Res.* **56**, 5705–5714.
Wang, Z., Escotet-Espinoza, M.S., and Ierapetritou, M. (2017). Process analysis and optimization of continuous pharmaceutical manufacturing using flowsheet models. *Comput. Chem. Eng.* **107**, 77–91.
Westerhuis, J.A., Gurden, S.P., and Smilde, A.K. (2000). Generalized contribution plots in multivariate statistical process monitoring. *Chemom. Intell. Lab. Syst.* **51**, 95–114.
Wibowo, C., Chang, W.C., and Ng, K.M. (2001). Design of integrated crystallization systems. *AIChE*

J. **47**, 2474–2492.
WifOR (2020). The global economic impact of the pharmaceutical industry. Available at: https://www.wifor.com/uploads/2021/06/WifOR_Global_Economic_Footprint_Study_September_2020.pdf. Last accessed: 09/29/2021.
Wise, B.M. and Gallagher, N.B. (1996). The process chemometrics approach to process monitoring and fault detection. *J. Process Control* **6**, 329–348.
Wong, W.C., Chee, E., Li, J., and Wang, X. (2018). Recurrent neural network-based model predictive control for continuous pharmaceutical manufacturing. *Mathematics* **6**, 242.
Wood, B., Girard, K.P., Polster, C.S., and Croker, D.M. (2019a). Progress to date in the design and operation of continuous crystallization processes for pharmaceutical applications. *Org. Process Res. Dev.* **23**, 122–144.
Wood, B., Girard, K.P., Polster, C.S., and Croker, D.M. (2019b). Progress to Date in the Design and Operation of Continuous Crystallization Processes for Pharmaceutical Applications. *Org. Process Res. Dev.* **23**, 122–144.
Woodcock, J. (2004). The concept of pharmaceutical quality. *Am. Pharm. Rev.* **7**, 10–15.
Yan, B., Li, Y., Guo, Z., and Qu, H. (2014). Quality by design for herbal drugs: a feedforward control strategy and an approach to define the acceptable ranges of critical quality attributes. *Phytochem. Anal.* **25**, 59–65.
Yang, Y., Song, L., Gao, T., and Nagy, Z.K. (2015). Integrated upstream and downstream application of wet milling with continuous mixed suspension mixed product removal crystallization. *Cryst. Growth Des.* **15**, 5879–5885.
Yaws, C.L. (2009). *Diffusion Coefficient in Water - Organic Compounds, Transport Properties of Chemicals and Hydrocarbons*. William Andrew Inc., Norwich (NY; USA).
Yaws, C.L. and Le Xuan Dang, L. (2009). Viscosity of Liquid - Organic Compounds, Transport Properties of Chemicals and Hydrocarbons. William Andrew Inc.
Yu, A.B., Zou, R.P., and Standish, N. (1996). Modifying the linear packing model for predicting the porosity of nonspherical particle mixtures. *Ind. Eng. Chem. Res.* **35**, 3730–3741.
Yu, L.X. (2008). Pharmaceutical quality by design: Product and process development, understanding, and control. *Pharm. Res.* **25**, 781–791.
Yu, L.X., Amidon, G., Khan, M.A., Hoag, S.W., Polli, J., Raju, G.K., and Woodcock, J. (2014). Understanding pharmaceutical quality by design. *AAPS J.* **16**, 771–783.
Yu, L.X. and Kopcha, M. (2017). The future of pharmaceutical quality and the path to get there. *Int. J. Pharm.* **528**, 354–359.
Yu, L.X., Raw, A., Wu, L., Capacci-Daniel, C., Zhang, Y., and Rosencrance, S. (2019). FDA's new pharmaceutical quality initiative: Knowledge-aided assessment & structured applications. *Int. J. Pharm. X* **1**, 0–3.
Yu, L.X. and Woodcock, J. (2015). FDA pharmaceutical quality oversight. *Int. J. Pharm.* **491**, 2–7.
Zahel, T., Hauer, S., Mueller, E.M., Murphy, P., Abad, S., Vasilieva, E., Maurer, D., Brocard, C., Reinisch, D., and Sagmeister, P. (2017). Integrated process modeling—a process validation life cycle companion. *Bioengineering* **4**, 86.
Zavala, V.M., Laird, C.D., and Biegler, L.T. (2008). A fast moving horizon estimation algorithm based on nonlinear programming sensitivity. *J. Process Control* **18**, 876–884.
Zendehboudi, S., Rezaei, N., and Lohi, A. (2018). Applications of hybrid models in chemical, petroleum, and energy systems: A systematic review. *Appl. Energy* **228**, 2539–2566.
Zhao, Y., Drennen, J.K., Mohan, S., Wu, S., and Anderson, C.A. (2019). Feedforward and feedback control of a pharmaceutical coating process. *AAPS PharmSciTech* **20**, 1–12.
Zou, R.P. and Yu, A.B. (1996). Evaluation of the packing characteristics of mono-sized non-spherical particles. *Powder Technol.* **88**, 71–79.

Acknowledgements

In the last three years, I have met many people that made my PhD the invaluable journey that I feel so lucky to have lived.

My deepest gratitude goes to my supervisor, Prof. Massimiliano Barolo, who guided me in my first steps as a scientist with passion and dedication. Thank you, Max, for teaching (and showing) me that the key to research is in paying attention to details, and for always endorsing me in every challenge. I owe you a lot. Thanks to Prof. Fabrizio Bezzo and Prof. Pierantonio Facco, for their advice and mentoring during my PhD. Thank you for always being available for discussing, and for guiding me in my scientific and personal choices. Warm thanks go to past and present members of the CAPE-Lab, for making the group a friendly environment where research thrives: Gabriele, Riccardo, Federico, Christopher, Gianmarco, Alberto, Francesco, Elia, Francesca, and Andrea. Thanks to Ivo, for having been around during my undergraduate and postgraduate student career for eight (!) years now. You have been my, and everyone's, salvation multiple times, in science, and in facing everyday debacles.

Sincere thanks go to Prof. Zoltan Nagy for mentoring me during my stay at Purdue University and in my academic journey, and for teaching me what passion for research is. Thank you for having put the seeds of many research ideas in me, and for always being available for precious conversations. During my stay at Purdue, I met many people that made my American experience special, despite the almost total overlapping with the COVID-19 pandemics first wave. Special thanks go to Inyoung, Varun, Daniel, Dan, Yan-Shu, Jaron, Onni, Ayse, Wei-Lee, and all the members of Prof. Nagy's group. I will never be able to thank enough Francesco for the scientific support in my academic journey since we met at Purdue, and for the incredible personal help that, together with Jenny, gave me during the pandemics, and afterwards.

I am also deeply grateful to Prof. Costas Pantelides for having me as a Visiting PhD Student at Process Systems Enterprise, and for giving a huge contribution to my scientific growth in limited time. Sincere thanks go also to Andrew for his passionate and kind supervision, and to all the people at PSE for making the office such a pleasant place to work at.

Heartfelt thanks go to Dr. Sal García Muñoz for the precious inspiration and support he gave me during all my PhD. Thank you for making yourself available to becoming a reference person to me, through fruitful conversations and mentoring.

Special thanks to my roommate, Lorenzo, and to the old and new friends that have been close to me since my comeback to Padova: you all helped me to achieve the end of this journey surrounded by love and care.

Infine, grazie a tutta la mia famiglia per il supporto che ho ricevuto nella mia vita. Grazie mamma e papà.