Faculty of Pharmacy
of the
University of Porto

Faculty of Pharmaceutical
Sciences
Ghent University

Process analytical technology for batch and continuous pharmaceutical processes' supervision

Ana Filipa Tavares da Silva

Thesis of the 3rd Cycle of Studies for obtaining the PhD degree in Pharmaceutical Sciences in the specialty of Analytical Chemistry submitted to the Faculty of Pharmacy of the University of Porto and to the Faculty of Pharmaceutical Sciences – Ghent University

Supervisors:

Prof. Dr. Thomas De Beer (Ghent University)

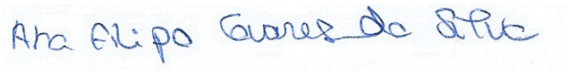
Prof. Dr. José Luís Fontes da Costa Lima (University of Porto)

Co-supervisors:

Prof. Dr. João Pedro Martins de Almeida Lopes (University of Lisbon)

Dr. Mafalda Sofia Coelho da Cruz Sarraguça (LAQV/REQUIMTE)

February 2017



Aos meus pais
À minha irmã

To my parents
To my sister

"This is how you do it: you sit down at the keyboard and you put one word after another until its done. It's that easy, and that hard."
- Neil Gaiman

Acknowledgements

First, I would like to express my sincere gratitude to my supervisors and co-supervisors. I would like to thank Prof. Thomas De Beer for the opportunity to start this joint PhD, for the support when I first arrived to Belgium, it was truly life-changing. Thank you as well for your supervision and for the possibility to learn from you. I am very thankful to Prof. João Pedro Martins de Almeida Lopes for his guidance, support and for sharing his expertise. I would like to deeply thank Dr. Mafalda Sarraguça without whom this work would not have been possible. Thank you for the support, supervision and encouragement. Thank you for the patience and availability. Thank you, Prof. José Luis, Fontes da Costa Lima for the support over the years. I thank them all for the opportunity they gave me to carry out the research for this thesis in their labs and learn from their vast scientific knowledge.

I would like to thank all members of the Laboratory of Pharmaceutical Process Analytical Technology and Pharmaceutical Technology at the University of Gent. A big thank you to Anneleen Burggraeve, Lien Saerens, Margot Fonteyne, Laurent Hansen, Elisabeth Peeters and Tinne Monteyne, for welcoming me when I first arrived to a complete new country I knew nothing about. Thank you for those who came later and I had the opportunity to work with, for which I am grateful (Pieter-Jan Van Bockstal, Niels Nicolaï, Séverine Mortier and Laurens De Meyer). I am very thankful to Valerie Vanhoorne, Nils Bostijn, Maxim Verstraeten and Jurgen Vercruysse for the support you gave for the work in this thesis. Last, but definitely not the least, thank you Fien De Leersnyder for your support and availability.

Thank you, Prof. João Luís Machado dos Santos and Dr. Karine Lopes Marques for introducing me to scientific research and for the support they gave me before I even knew what I wanted to do. Had we not crossed paths and I would probably not have followed this route.

I would also like to thank my colleagues at Janssen Pharmaceutica NV, for their warmth, understanding and patience. It was essential to hear your words of encouragement and support, especially in those more stressful days where no time ever seemed to be enough to get things done.

Agradeço aos meus amigos, em especial aos Ursos, pela paciência, apoio e amizade apesar da ausência e da distância que nos separa. Um agradecimento especial à Margarida Teixeira, à Sara Freixo pela paciência extra nos momentos mais stressantes.

Um obrigada à minha família. Especialmente aos meus pais e irmã por me darem a coragem para seguir este caminho. Obrigado pela paciência e apoio nestes anos. Longe ou perto os obstáculos vieram e foram-se ultrapassando e que assim seja sempre.

I would like to thank everyone else who has contributed somehow to the realization of this thesis

Publications

The following publications in international peer-reviewed journals are published/prepared under the scope of this PhD thesis:

A.F.T. Silva, A. Burggraeve, Q. Denon, P. Van der Meeren, N. Sandler, T. Van Den Kerkhof, M. Hellings, C. Vervaet, J.P. Remon, J.A. Lopes, T. De Beer, Particle sizing measurements in pharmaceutical applications: Comparison of in-process methods versus off-line methods, European Journal of Pharmaceutics and Biopharaceutics, 85 (2013) 1006-1018 (http://dx.doi.org/10.1016/j.ejpb.2013.03.032).

A.F.T. Silva, M.C. Sarraguça, P.R. Ribeiro, A.O. Santos, T. de Beer, J.A. Lopes, Statistical process control of cocrystallization processes: a comparison between OPLS and PLS, Internacional Journal of Pharmaceutics (Accepted manuscript, http://dx.doi.org/10.1016/j.ijpharm.2017.01.052).

A.F.T. Silva, J. Vercruysse, M.C. Sarraguça, J.L.F. Costa Lima, C. Vervaet, J.P. Remon, J.A. Lopes, T. De Beer, Understanding and monitoring a continuous pharmaceutical twin-screw granulation and drying process using multivariate data analysis – part I, (Submission in process).

A.F.T. Silva, J. Vercruysse, M.C. Sarraguça, J.L.F. Costa Lima, C. Vervaet, J.P. Remon, J.A. Lopes, T. De Beer, Understanding and monitoring a continuous pharmaceutical twin-screw granulation and drying process using multivariate data analysis – part II, (Submission in process).

A.F.T. Silva, M.C. Sarraguça, M. Fonteyne, J. Vercruysse, F. De Leersnyder, V. Vanhoorne, N. Bostijn, M. Verstraeten, C. Vervaet, J.P. Remon, T. De Beer, J.A. Lopes, Multivariate statistical process control approach to monitor a continuous pharmaceutical twin-screw granulation and drying process, (Submission in process).

Summary

Process analytical technology (PAT) methodologies are being increasingly adopted by drug products manufacturing companies for products and processes' monitoring and control. PAT aims at improving and updating outdated pharmaceutical manufacturing practices where quality control is based more on compliance with product specifications rather than on process understanding. As the vast majority of drug products today are manufactured relying on a sequence of batch or semi-continuous unit operations, PAT tools have been implemented by focusing on specific production process stages. As continuous manufacturing of drug products becomes a reality (increasing investment of industry in this manufacturing paradigm), many established approaches for process monitoring, tracing and control must be reevaluated to adjust to the new requirements of this manufacturing mode. PAT methods are core in this context, specifically in what concerns product quality assessment. Consequently, quality assessment in continuous manufacturing forces PAT tools implementation to reach another level. These new implementation requirements must be effectively identified and solved in order to reduce risks associated with inefficient assessment of products quality. This thesis intended to provide a deep insight on the implementation of PAT tools in the context of continuous manufacturing of solid dosage forms, paving the way to a more effective migration from the traditional batch-wise production paradigm.

Chapter 1 defines the general objectives that this thesis aimed at accomplishing while Chapter 2 serves the purpose of introducing the concepts and tools utilized to reach those goals.

The need for real-time quality assessment of processing materials during continuous manufacturing while complying with the regulatory agencies requirements is especially challenging. For instance, particle size distribution is a critical quality attribute (CQA) in many unit operations such as blending, granulation, spheronization or crystallization. Currently available tools for in-process particle size distribution monitoring utilize dissimilar working principles complicating the task of validation against well-established techniques such as laser diffraction and sieving. This problem was investigated by comparing particle size of several batches of Cellets™ and granules, determined with in-process techniques versus size obtained by laser diffraction and sieving and explaining the observed differences in results (Chapter 3).

In Chapter 4 and Chapter 5, PAT tools were investigated in the context of processes monitoring and supervision. In these chapters, PAT tools implemented in

specific sections of different manufacturing processes were used to generate process localized fingerprints that combined with multivariate data analysis (MVDA, e.g., principal component analysis (PCA), partial least squares (PLS) and some extensions of these methods) were utilized to ensure that steady state conditions were kept. Multivariate statistical process control concepts are applied. Chapter 4 provides a comparison between PLS and Orthogonal-PLS (OPLS) for batch statistical process control (BSPC) of a cocrystallization process considering near-infrared spectra acquired in-situ and in real-time. A systematic analysis of the operation of the granulation and drying stages of the continuous manufacturing line is described in Chapter 5 (ConsiGma-25™). Univariate data acquired from sensors spread in multiple locations of the line was analyzed with MVDA methods (PLS, PCA and OPLS). Multivariate statistical process control (MSPC) concepts were applied to guarantee that the process was kept in a state of control. Two different wet granulation processes were examined. In the first two studies, a formulation containing two Active Pharmaceutical Ingredients (API) was granulated with distilled water, and in the third study, a second placebo formulation consisting of milled lactose monohydrate was granulated using a 20% (w/w) PVP solution as a binder.

Differences in particle size results of in-process versus offline techniques were found, examined and explained. A thorough overview of the different equipments is provided which can be used as a tool for selection of the best instrument for a specific application. Both PLS and OPLS-based BSPC approaches were successful in terms of detecting process imposed disturbances in the studied cocrystallization process. The MVDA methods generated important process knowledge about the operation of the ConsiGma™-25 line. The developed MSPC approaches are a step toward a monitoring approach that links measured variables, process parameters, together with raw material properties, in a manner to predict, and ultimately control the critical product quality attributes.

Chapter 6 frames this dissertation in a board international context, describes its relevance, and sheds light on the future perspectives as a continuation of the developed work. Chapter 7 presents the general conclusion of this work.

The developed tools consolidate PAT in terms of the added-value to pharmaceutical manufacturing, especially when industry is gradually migrating towards continuous manufacturing. Implementation of this technology directly in the process stream will allow to obtain fundamental process knowledge, and open up the possibility to establish better control strategy, better process performance and, ultimately, improving product quality.

Keywords: Continuous manufacturing; Process analytical technology; Near-infrared spectroscopy; Multivariate data analysis; Multivariate statistical process control.

Resumo

As metodologias de Tecnologia Analítica de Processos (PAT) são cada vez mais utilizadas pela industria farmacêutica na monitorização e controlo de produtos e processos. As metodologias PAT visam melhorar e atualizar as práticas obsoletas de fabrico de produtos farmacêuticos em que, o controle de qualidade se baseia mais na conformidade com as especificações do produto final, em detrimento da compreensão do processo em si. Como a grande maioria dos produtos farmacêuticos são atualmente fabricados com base numa sequência de operações unitárias em lote, ou semi-contínuas, as ferramentas PAT têm sido implementadas focando etapas específicas do processo de produção. À medida que a fabricação contínua de medicamentos se torna uma realidade, (resultado de um investimento crescente da indústria farmacêutica neste paradigma de fabrico), muitas das abordagens de monitorização de processos, rastreamento e controle devem ser reavaliadas de modo a se ajustarem às exigências deste novo modo de produção. A necessidade de avaliar a qualidade de materiais e produtos durante o fabrico contínuo promove a implementação de ferramentas PAT. Os requisitos de implementação destas metodologias têm de ser eficazmente identificados e resolvidos, a fim de reduzir os riscos associados a uma avaliação ineficaz da qualidade dos produtos. Esta tese pretende fornecer uma visão detalhada sobre a implementação de ferramentas PAT no contexto da fabricação contínua de formas sólidas orais, abrindo o caminho para uma migração mais eficaz do tradicional paradigma de produção em lote para a produção contínua.

No Capítulo 1 são definidos os objetivos gerais que esta tese se propôs a realizar, enquanto o Capítulo 2 serve o propósito de introduzir os conceitos e ferramentas utilizados para alcançar esses objetivos.

É essencial avaliar em tempo real a qualidade dos materiais a serem processados durante a fabricação contínua, mas fazê-lo cumprindo os requisitos das agências reguladoras, é uma tarefa exigente. A distribuição de tamanho de partícula é um exemplo de um atributo crítico de qualidade (CQA) em muitas operações unitárias tais como mistura, granulação, esferonização, cristalização, entre outras. As ferramentas atualmente disponíveis para a monitorização da distribuição de tamanho de partícula durante o processo utilizam princípios de medição distintos, o que dificulta a tarefa de validá-los por comparação com técnicas bem estabelecidas, tais como a difração a laser e análise granulométrica utilizando tamises. Nesta dissertação, esta questão foi abordada experimentalmente, determinando o tamanho de partícula de diferentes lotes de Cellets™

e grânulos utilizando simultaneamente, as novas tecnologias destinadas a monitorização durante o processo, e técnicas tradicionais de difração a laser e análise granulométrica. Os resultados obtidos foram comparados e as diferenças observadas explicadas (Capítulo 3).

No Capítulo 4 e no Capítulo 5, as ferramentas PAT foram investigadas no contexto da monitorização e supervisão de processos. Nestes capítulos, ferramentas PAT implementadas em segmentos específicos de diferentes processos de fabricação foram utilizadas para gerar impressões digitais que, quando combinadas com análise de dados multivariada (MVDA, e.g. análise de componentes principais (PCA), mínimos quadrados parciais (PLS) e algumas extensões destes métodos) foram empregues para assegurar que o processo permanecia num estado estacionário. Para este fim foram utilizados conceitos de controle estatístico multivariado de processos (MSPC). No Capítulo 4 é feita uma comparação entre PLS e PLS ortogonal (OPLS) no controle estatístico multivariado em lote (BSPC) de um processo de co-cristalização monitorizado com espectroscopia de infravermelho próximo *in situ,* e em tempo real. No Capítulo 5 é feita uma análise sistemática do funcionamento dos estágios de granulação e secagem da linha de produção contínua (ConsiGma-25 ™). Dados univariados adquiridos a partir de sensores espalhados em múltiplos locais da linha foram analisados com métodos MVDA (PLS, PCA e OPLS). Foram aplicados conceitos de controle estatístico multivariado de processos para garantir que o processo permanecia em controlo. Dois diferentes processos de granulação a húmido foram estudados. Nos dois primeiros estudos, granulou-se uma formulação contendo duas substâncias ativas com água destilada e, no terceiro estudo, granulou-se uma segunda formulação placebo consistindo em monohidrato de lactose com uma solução de PVP a 20% (m/m) como aglutinante.

Foram encontradas, examinadas e explicadas as diferenças entre os resultados de tamanho de partícula obtidos pelas técnicas em processo versus as metodologias tradicionais off-line. É fornecida uma visão geral dos diferentes equipamentos que pode ser usada como uma ferramenta para seleção do melhor instrumento a ser usado numa aplicação específica. Ambos as abordagens baseadas em PLS e OPLS foram capazes de detetar os distúrbios impostos no processo de co-cristalização estudado. Os métodos MVDA forneceram informação relevante relativamente ao funcionamento da linha de produção contínua ConsiGma ™ -25. As abordagens MSPC apresentadas são um passo na direção de uma abordagem de monitorização mais abrangente capaz de conectar variáveis e parâmetros de processo, juntamente com as propriedades da matéria-prima, de forma a prever e controlar os atributos críticos da qualidade do produto final.

O Capítulo 6 enquadra esta dissertação no contexto internacional e descreve a sua relevância, apresentado também perspetivas futuras para a continuação do trabalho desenvolvido. O Capítulo 7 apresenta as conclusões gerais deste trabalho.

As ferramentas desenvolvidas reforçam o valor das metodologias PAT aplicadas ao fabrico de produtos farmacêuticos numa altura especialmente relevante, em que a indústria está no processo de migração gradual para um modo de fabrico contínuo. A implementação desta tecnologia diretamente no fluxo dos processos permitirá obter conhecimento essencial acerca do processo em si e abre a possibilidade de se estabelecerem estratégias de controlo mais eficazes, contribuindo para um melhor desempenho do processo, resultando num produto final de qualidade superior.

Palavras-chave: Produção contínua; Tecnologia analítica de processos; Espectroscopia de infravermelho próximo; Análise de dados multivariada; Controle estatístico multivariado de processos.

Samenvatting

Proces Analytische Technologie (PAT) wordt steeds meer toegepast door geneesmiddel producerende bedrijven voor het monitoren en de controle van de producten en het productieproces. PAT heeft als doel het verbeteren en updaten van verouderde farmaceutische manufacturing toepassingen waarbij kwalitet eerder gegarandeerd wordt op basis van controle op het eindproduct dan op basis van proceskennis. Aangezien een grote meerderheid van de geneesmiddelen vandaag de dag aangemaakt worden via een opeenvolging van verschillende batch of semi-continue processen, werden PAT toepassingen tot nu toe vooral ingebouwd in specifieke processtappen. Nu continue productie van geneesmiddelen werkelijkheid wordt (meer investeringen door de industrie), moeten veel gevestigde methodes voor procesbewaking, foutopsporing en controle opnieuw worden geëvalueerd en worden aangepast. PAT toepassingen zijn van groot belang in deze context, zeker wanneer het over de controle van productkwaliteit gaat. Dit zorgt er voor dat een verschillende aanpak voor PAT implementatie nodig is met betrekking tot kwaliteitscontrole in continue processen dan bij batch processen. Deze nieuwe implementatiemethodes moeten ontwikkeld en geëvalueerd worden om risico's geassocieerd met ondoelmatige controle van productkwaliteit te vermijden. Deze thesis is bedoeld om een dieper inzicht te verkrijgen in de implementatie van PAT toepassingen in de context van continue productieprocessen van vaste doseringsvormen, het pad effenend voor een meer effectieve migratie van het traditionele batchgewijze productiepatroon naar continue productiewijzen

Hoofdstuk 1 definieert de algemene doelstelling van deze thesis terwijl Hoofdstuk 2 een introductie geeft tot de concepten en toepassingen gebruikt om deze doelstelling te bereiken.

De nood aan real-time kwaliteitsbeoordeling van de verwerkte materialen tijdens de continue productie, de eisen van de regelgevende instantie in acht genomen, is bijzonder uitdagend. Zo zal deeltjesgroottedistributie een kritische kwaliteitseigenschap (CQA) zijn in veel eenheidsoperaties zoals menging, granulatie, sferonisatie of kristallisatie. De huidig beschikbare toepassingen voor inline deeltjesgrootte monitoring maken gebruik van verschillende meet- en rekenprincipes, wat de validatie ten opzichte van gevestigde offline technieken zoals zeven en laserdiffractie moeilijk maakt. Dit probleem werd onderzocht door de deeltjesgrootte van verschillende batchen Cellets™ en granules, verkregen door inline technieken te vergelijken met de deeltjesgrootte verkregen met behulp van offline technieken (Hoofdstuk 3).

In Hoofdstuk 4 en Hoofdstuk 5, werden PAT toepassingen onderzocht in de context van procescontrole en –toezicht. In deze hoofdstukken werden PAT toepassingen geïmplementeerd in welbepaalde secties van verschillende productieprocessen, hierdoor werden proces-gelokaliseerde vingerafdrukken gegenereerd die, gecombineerd met multivariate data-analyse (MVDA zoals principale componenten analyse (PCA), partiële kleinste kwadraten regressie (PLS) en enkele uitbreidingen op deze methodes) gebruikt werden om te verzekeren dat steady-state condities aangehouden werden. Multivariate statistische procescontrole begrippen werden toegepast. Hoofdstuk 4 geeft een vergelijking tussen PLS en Orthogonale-PLS (OPLS) voor batch statistische procescontrole (BSPC) van een co-kristallisatieproces, vanuit nabije infrarood spectra die inline werden verkregen tijdens het proces. Een systematische analyse van de granulatiestap en de droogstappen in een continue productielijn (ConsiGma™-25) werd beschreven in Hoofdstuk 5. Univariate data verkregen uit sensoren verspreid over verschillende locaties in de productielijn werden geanalyseerd door middel van MVDA methodes (PCA, PLS en OPLS). Het multivariate statistische procescontrole (MSPC) concept werd toegepast om er voor te zorgen dat het proces onder controle bleef. Twee verschillende granulatieprocessen werden onderzocht. In de eerste twee studies werd een formulatie bestaande uit twee actieve farmaceutische ingrediënten (API) gegranuleerd met gedestilleerd water en in een derde studie werd een placeboformulatie bestaande uit gemalen lactose monohydraat gegranuleerd met een PVP oplossing (20%(w/w)) als granulatievloeistof.

Er werden verschillen gezien tussen de deeltjesgrootte gemeten met inline en offline methodes. Deze verschillen werden onderzocht en verklaard. Een volledig en grondig overzicht van de gebruikte toepassingen werd voorzien, welke gebruikt kan worden als hulpmiddel om te selecteren welke toepassing het beste gebruikt wordt in een bepaalde situatie. Zowel PLS en PLS-gebaseerde BSPC benaderingen konden succesvol gebruikt worden om verstoringen, opgelegd aan het systeem, te detecteren in een bestudeerd co kristallisatie proces. De MVDA methodes genereerden belangrijke proceskennis over de ConsiGma™-25 lijn. Met de ontwikkelde MSPC aanpak staan we een stap dichter tot een controlemethode die gemeten variabelen, proces parameters en eigenschappen en de ruwe materialen linkt op een manier dat de CQA's voorspeld en uiteindelijk ook gecontroleerd kunnen worden.

Hoofdstuk 6 plaatst dit proefschrift in een bredere context, beschrijft de relevantie en werpt licht op de toekomstplannen in de vorm van een vervolg op het bestaande werk. Hoofdstuk 7 geeft de algemene conclusie van dit werk weer.

De ontwikkelde toepassingen maken PAT tot een toegevoegde en onmisbare waarde voor farmaceutische productie, zeker nu de industrie geleidelijk meer overstapt naar continue processen. Wanneer deze technologie direct in de processtroom wordt geïmplementeerd, zal het mogelijk zijn fundamentele proceskennis te verschaffen en maakt het het ontwikkelen van een betere controlestrategie en betere proces performance mogelijk, uiteindelijk leidend tot een betere productkwaliteit.

Kernwoorden; Continue productie; Proces analytische technologie; Nabije infrarood spectroscopie; Multivariate data-analyse; Multivariate statische proces controle.

Contents

List of Figures

List of Tables

List of Abbreviations

(alphabetic order)

ANN	Artificial Neural Networks
API	Active Pharmaceutical Ingredient
BSPC	Batch Statistical Process Control
C-SOC	Engineering Research Centre for Structured Organic Composites
C-SOPS	Center of Structured Organic Particulate Systems
cGMPs	Pharmaceutical Current Good Manufacturing Practices
CL	Chord Length
CLD	Chord Length Distribution
CPAC	Center for Process Analytical Chemistry
CPACT	Centre for Process Analysis and Control Technology
CPP	Critical Process Parameter
CPPR	Centre for Pharmaceutical Processing Research at Purdue University
CQA	Critical Quality Attribute
CUSUM	Multivariate Cumulative Sum
DSC	Differential Scanning Calorimetry
DoE	Design of Experiments
DCPT	Duquesne University Centre for Pharmaceutical Technology
DModX	Distance to the Model in the X-space
EMA	European Medicines Agency
EWMA	Multivariate Exponentially Weighted Moving Average
FBRM	Focused Beam Reflectance Measurements
FDA	Food and Drug Administration
FIFO	First-in-first out
GMP	Good Manufacturing Practices
HTZ	Hydrochlorothiazide
ICA	Independent Component Analysis
ISA	International Society of Automation
ICH	International Conference on Harmonization
KDE	Kernel Density Estimation
KICA	Independent Component Analysis with Kernel Density Estimation
KPCA	Principal Component Analysis with Kernel Density Estimation

K-OPLS	Kernel-based Orthogonal Partial Least Squares
LV	Latent Variable
MIT	Massachusetts Institute of Technology
MSPC	Multivariate Statistical Process Control
MVDA	Multivariate Data Analysis
NIR	Near Infrared Spectroscopy
NIPTE	National Institute for Pharmaceutical Technology and Education
NOC	Normal Operating Conditions
n-PLS	N-way PLS
PAT	Process Analysis Technology
PC	Principal Component
oPC	Orthogonal Principal Component
OPLS	Orthogonal Partial Least Squares
OPLS-DA	Orthogonal Partial Least Squares Discriminant Analysis
PABA	*p*-aminobenzoic acid
PCA	Principal Component Analysis
PCR	Principal Component Regression
PID	Proportional-Integral-Derivative
PLS	Partial Least Squares
pPC	Predictive Principal Component
PS	Particle Size
PSI	Photometric Stereo Imaging
PSD	Particle Size Distribution
QbD	Quality-by-Design
RCPE	Research Centre Pharmaceutical Engineering
RMSE	Root-Mean Square Error
RMSECV	Root-Mean Square Error of Cross-Validation
RMSEP	Root-Mean Square Error of Prediction
rpm	Rotations per minute
RTRT	Real-Time Release Testing
SFV	Spatial Filtering Velocimetry
SNV	Standard Normal Variate
SPC	Statistical Process Control
SPE	Squared Prediction Error
SVM	Support Vector Machines
XRPD	X-Ray Powder Diffraction

List of Symbols

π	Pi
ρ	Particle density
a	Principal Component or Latent Variable
A	Number of Principal Components or Latent Variables
$area$	Peak area
AAR	Average aspect ratio
c	Calibration constant
d	Particle diameter
D_{max}	Maximum diameter
D_{min}	Minimum diameter
$DModX_{abs}$	Absolute distance to the model in the X-space
E	X-residuals matrix
F	Y-residuals matrix
j	Sample/observation
k	X-variable
K	Number of X-variables
M	Number of Y-variables
$Mass$	Calculated particle mass
N	Number of sample/Observations
P	X-loadings
Q	Y- weights matrix
Q_j	Q-statistic for sample/observation *j*
Q^2	Variance predicted by the model according to cross-validation)
R^2/R^2X	Variance in the X-data captured by the model
R^2Y	Variance in the Y-data captured by the model
s_0	Relative standard deviation of the model in the X-space
s_a^2	Estimate of variance of the corresponding latent variable *a*
T	X-scores
$t_{j,a}^2$	Score value for sample/observation *j* for latent variable *a*
T_j^2	Hotelling's T^2 statistic for sample/observation *j*
U	Y-scores
V	Volume-based particle size

X	X-matrix
$\hat{x}_{j,k}$	Data estimation using the model with *A* components for sample *j* and variable *k*
Y	Y-matrix

Chapter 1

Objectives

This thesis aims at demonstrating the usefulness of the use of Process Analytical Technology (PAT) concepts to supervise batch and continuous pharmaceutical processes in order to achieve an in-depth level of process understanding. The ultimate goal of PAT is to provide a better process understanding, monitoring the process and, ultimately, allowing to steer it. A PAT environment includes real-time monitoring of the process including at-line, on-line and in-line measurements. Even with the most elementary application of PAT, one can gain a deeper knowledge, enabling the possibility of applying changes during manufacture to keep the process running within the desired operating space. Better control is the prime objective, as it will improve product quality, resulting in less waste, increase safety of operations, and thus increase profitability. It is a step forward in the direction of implementation innovative tools for process supervision in the pharmaceutical industry.

Particle size in one of the most important quality attributes in the manufacturing of solid dosage forms. To be able to apply PAT concepts to monitor this quality attribute it is necessary to have techniques capable of measuring particle size in real-time during processing. One of the objectives of this thesis is studying several in-process techniques and compare the results to well-established off-line methods to determine particle size.

A central topic throughout the work developed for this thesis is applying multivariate statistical process control (MSPC) strategies to monitor batch and continuous manufacturing processes. Data collected from in-process analyzers such as, near infrared spectroscopy, as well as, data from univariate sensors (e.g., temperatures, pressures, etc.) is analyzed with multivariate data analysis methods, for instance, principal component analysis (PCA), partial least squares (PLS).

The development of these technologies provides scientific-based solutions with applications in the pharmaceutical industry, resulting in the substitution of off-line reference methods and attaining an improved process knowledge. The work developed in this thesis aims at supporting the transition of the manufacturing paradigm from batch to continuous production. A shift that is highly beneficial to the pharmaceutical industry from the economic and the product quality point of view.

Chapter 2

Introduction

2.1 Pharmaceutical industry: need for innovation

The pharmaceutical industry is continuously confronted with the challenge to improve product quality while lowering production costs. As in any other industry, the fundamental aim is to produce an efficient, and safe product, satisfactory to the customer (i.e. within prescribed quality specifications), and at a minimum cost.

Pharmaceutical industry lived profitable years in the 1980s and 1990s. The many breakthroughs in medicines, high product margins and favorable and growing global economy were the main contributing factors [1]. However, in the last decades, the overall economic downturn, the rising cost of healthcare and costs associated with the development and sales of pharmaceuticals and the competition with generic medicines, generated a marked profitability gap between the increasing costs and decreasing market prices [2].

A successful and continuous introduction of new drugs is essential to sustain competitive advantage for the firms. However, new drugs are complex to create. This a laborious, long-lasting, and expensive process with very uncertain outcomes. One in every 5,000 to 10,000 potential compounds investigated by the US-based pharmaceutical companies is granted Food and Drug Administration (FDA) approval. The odds of a promising drug to make it through the sequential stages of the drug development process are about one in five. To FDA approval, and the rights to market a drug takes in average about 15 years, most of which dedicated to clinical trials. At the time of obtaining FDA approval, the clock on market exclusivity starts ticking [1], and currently, patent protection times are decreasing, due to longer development times [3]. The pharmaceutical industry

needs to find alternatives to strengthen its position on other than drug discovery through cost control and, on the other hand, through innovation, brand promotion and services [2]. Pharmaceutical companies must create new ways and bring new ideas. In addition to profits, innovation brings added value by contributing significantly to the economy, and society, by creating highly-skilled jobs and increasing scientific knowledge. Innovation in medicines brings also significant value in terms of the population's health, increasing survival rates and bettering quality of life which reduces the of costs related with diseases. Powerful, diversified, influential and globalized the pharmaceutical industry plays a crucial role in the economy [4].

2.2 Pharmaceutical Manufacturing: from Batch to Continuous Production

The pharmaceutical industry is showing increasing interest in switching from traditional batch manufacturing to continuous production [5]. This change is mainly driven by two highly compelling reasons, fortified by the current legislative and economical conjecture: improvement of guaranteed product quality and reduction of production costs [5].

Today, pharmaceutical production plants of active pharmaceutical ingredients (API's) and drug products are still mostly batch manufacturing plants [6]. In batch manufacturing, all materials are charged before the start of the process, and discharged at the end (Figure 1). On the other hand, in continuous manufacturing materials follow the "first-in-first-out" (FIFO) principle, being continuously charged and discharged from the process. FIFO is a concept of lean manufacturing [7] an idea developed by Toyota which aims at simultaneously reducing waste and boosting productivity. A manufacturing line can work in more than one mode, multimode processes also exist [8].

In a traditional batch pharmaceutical manufacturing process, raw materials, intermediates and final products are usually tested off-line and stored before they are sent to the next step. In continuous manufacturing, raw materials are fed continuously to the process and intermediate products are sent continuously to the next step for further processing. Therefore, continuous manufacturing, requires a higher level of design and in-process control to ensure a product of quality. A fully integrated continuous installation includes continuous unit operations, connected in an integrated manufacturing process. In a fully integrated line, PAT concepts are utilized to provide real-time data for process monitoring and control, and engineering process control systems are implemented to mitigate the impact of raw material and process variability on the finished products' quality

[9]. This comes in line with Quality-by-Design (QbD) paradigm for pharmaceutical processes development framed by the International Conference on Harmonization (ICH) Q8 guideline [10-12].

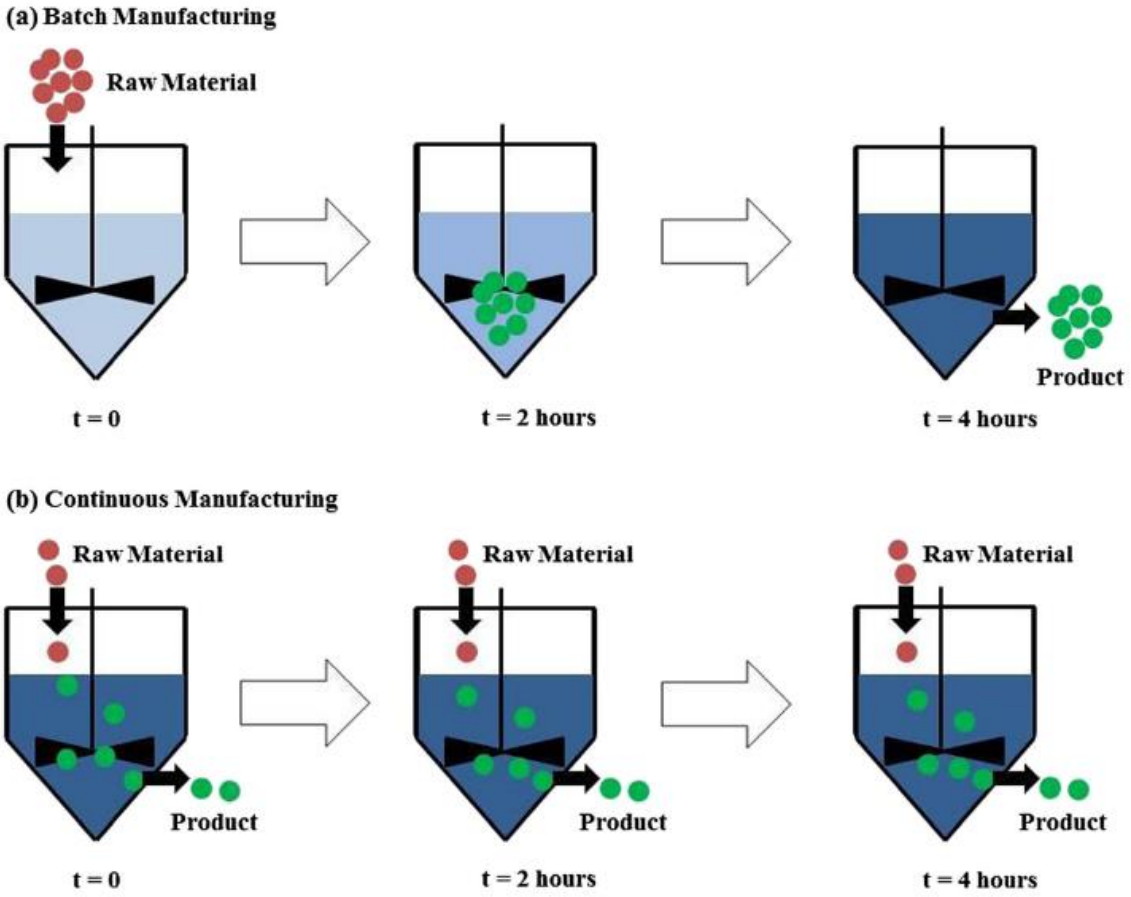


Figure 1 – Batch manufacturing: the material(s) is charged before the start of processing and the product is discharged at the end of processing. (b) Continuous manufacturing: material(s) and the product are simultaneously charged and discharged from the process, respectively. Adapted from [13].

Batch manufacturing offers the advantage that a batch can be accepted or rejected simplifying quality control. However, this fact should not pose a barrier to implementing continuous manufacturing. According to the FDA, the definition of a batch is not dependent on the manufacturing mode (batch or continuous), and is instead demarcated by the quantity of manufactured drug. A batch is 'a specific quantity of a drug or other material that is intended to have uniform character and quality, within specified acceptance limits, and is produced according to a single manufacturing order during the same cycle of manufacture' [14]. From a regulatory point of view, this definition does not pose at all an obstacle to adoption of continuous manufacturing, as a batch in a

continuous setting can simply be defined by a fixed quantity of product, or the amount produced in a predefined time interval [15].

Continuous manufacturing is being more commonly used in the pharmaceutical industry, and therefore, more attention is being paid to it. Several reasons can be attributed to this, among them are the possibility of product quality improvement, an enhanced cost-efficiency and reduced environmental impact.

Batch manufacturing plants operate intermittently and often, some or all process units are shut down and started up. Oppositely, continuous production plants are designed to work for twenty-four hours a day, seven days a week through the year. Down time is required for maintenance and, for some types of processes, catalyst regeneration. In a continuous process, materials are continuously processed and transferred thorough the system which, after a startup period, works at a steady state condition. How to operate start-up and shutdown as fast as possible with minimal waste is essential [16] and a smart sequencing of unit operations during both, start-up and shutdown, was suggested to decrease losses. PAT systems can gather real-time data for process and product monitoring, and ultimately, process steering an product quality control. When implemented in a continuous setting, PAT reduces or even eliminates the need of storage times of intermediates and final product. The alternative, which is an off-line quality control strategy, would entirely eliminate the advantage of producing continuously. PAT makes continuous processing more cost-efficient than batch processing, and leads to faster response times, and a faster final product release. A faster time-to-market is especially important for breakthrough therapies, for the treatment of life-threatening conditions, in preliminary clinical trials.

In continuous manufacturing, it is also potentially easier to scale up/down. The flexibility of a continuous process relies on the fact that, to obtain higher product throughput, there is only the need to increase the process time thus avoiding costly and time-expensive scale-up studies. This flexibility and robustness allows the same equipment to be used for both, small-scale batches for development and clinical trials, and for large-scale production. Being able to produce more or less product, in such straightforward manner, also allows to easily adapt to market demands, avoiding both drug shortages and overstock.

In continuous production, steps occur sequentially without the need of isolated or dedicated modules, greatly reducing plant footprint, and resulting in considerable saving in invested capital and operating expenses. A complete raw material to final product continuous manufacturing process can be installed in a single facility, rendering transport between several installations for further processing unnecessary. Labor costs due to

storage and handling of intermediates, stockpiling and handling of materials are hence also reduced.

Regardless of the numerous advantages of continuous manufacturing, there are still some challenges related to its employment. For many years, the pharmaceutical industry has been reluctant to adopt continuous manufacturing due to the strict regulatory environment. Once a process was approved, changes were avoided, hindering the introduction of new technologies. Companies with an established batch asset-base, wherein a huge capital was invested during the 80s and 90s, were also not willing to invest time and money in this new manufacturing concept without sufficient proof. Moreover, continuous technologies were not widely available, unlike batch equipment, and no one wanted to be the first to deal with the technology and regulatory hurdles that this change would bring [6, 17-19]. The required initial investment in the construction of facilities, new generation of equipment, sensors and automation, and generation of process knowledge were, and still are, unappealing to companies. It is also an important challenge that accurate process operation models of the steps in a continuous process need to be developed, together with powder characterization and handling, especially for low-dose production, with several solutions already proposed to this end [16]. Engineers, scientists and regulators need to adopt a new mindset and upskill in statistics to better analyze and understand process data [16].

Despite the challenges, they are resolvable, and clearly outweighed by the advantages offered by manufacturing continuously (Table 1). This was recognized by several big pharmaceutical companies. Johnson and Johnson aims to manufacture 70% of its highest-volume products continuously within 7 years from now [20]. Eli Lilly will invest €35 million in a purpose-built continuous API production facility in Ireland, for products in the company's late-stage pipeline [21]. GlaxoSmithKline has a continuous processing plant underway at its site in Jurong, Singapore, to make API's for its respiratory drugs. AstraZeneca has also recently invested in a continuous wet granulation unit [22]. The FDA recently approved two continuous manufacturing processes [23, 24]. The first approval ever of a production method change from "batch" to continuous manufacturing was conceded to Janssen: Pharmaceutical Companies of Johnson & Johnson. This approval is for manufacturing medication for the treatment of HIV-1 infection, Prezista™ (Darunavir), a direct compression product, at Janssen's Gurabo, Puerto Rico, plant. The second FDA approval was to Vertex, for producing their breakthrough cystic fibrosis therapy Orkambi™, a continuous wet granulation product. Approval was received by FDA in July 2015.

Table 1 – Advantages and challenges to the implementation of continuous manufacturing.

Advantages	Challenges
Product quality	**Product quality**
• PAT implementation: constant raw material and process monitoring and an improved control	• Simpler quality control (accept or reject an entire batch)
• Reduced variability: better reproducibility	**Business and organizational**
• Improved robustness	• Previous large investment in batch processes in the 80's and 90's is a hurdle for introduction of CM
• Compliant with Quality-by-Design	• Shift the mindset of process and formulation scientists, quality units within companies and regulators
• Faster release of breakthrough products	**Economic**
Cost-effectiveness	• Initial investment in facilities
• Improved flexibility	• Acquisition of equipment, sensors, automation and systems' integration and control
• Reduced production times	• Investment on the creation of process knowledge (process phases, raw materials, quality and environmental conditions)
• Elimination of intermediates and product storage times	
• Reduced intermediates and product handling	
• Shorter downtimes	
• Faster response times	
• Easier to scale-up	
• Reduction of time for product development	
• Reduced time-to-market (competitive advantage)	
• Adaptability to market demands	
• Capital savings on installation and operating expenses	
Environmental impact	
• Reduced footprint	
• Reduced energy consumption	

2.3 Process Analytical Technology (PAT)

PAT is defined as "a system for designing, analyzing, and controlling manufacturing through timely measurements of critical quality and performance attributes of raw and in-process materials and processes, with the goal of ensuring final product quality" [9]. The goal of PAT is to acquire a fundamental understanding of the manufacturing process and the manufactured product.

In the pharmaceutical industry, for many decades, quality and quality management activities were focused on compliance rather than on process understanding. As a result, business practices aim at minimizing regulatory risks. Additionally, the pharmaceutical industry is highly regulated and in the past, strict regulatory practices obliged processes to

operate under pre-defined and fixed operating conditions, disregard of variability of raw materials, and processing conditions. In some cases, this ended with out-of-specification products that had to be discarded causing major losses. Such stringent regulations, with little room for change, hindered innovations in this industry's manufacturing models and quality assurance [25].

The introduction of FDA's Pharmaceutical Quality for the 21st Century shifted that mindset by promoting a more efficient, agile, flexible pharmaceutical sector, that consistently and reliably produces high-quality drugs without an extensive regulatory burden. In August 2002, the FDA announced the "Pharmaceutical current Good Manufacturing Practices (cGMP's) for the 21st Century—a risk-based approach" [26]. Published later in 2004 these document aims at enhancing and modernizing pharmaceutical processes and product quality for veterinary and human drugs and selected human biological products such as vaccines.

The FDA Center for Drug Evaluation and Research (CDER) firstly proposed the creation of the PAT initiative with the goal of achieving significant health and economic benefits by applying modern process control and tests in pharmaceutical manufacturing. Shortly after, in 2001 the PAT subcommittee was formed, consisting in the four working groups with representatives from the FDA, experts from the industry and academic representatives: 1) PAT applications working group; 2) PAT products and the process development working group; 3) PAT process and analytical validation working group and 4) PAT chemometric working group. The comments and recommendations suggested by this subcommittee lead the FDA to release its PAT guidance. The introduction by the FDA of the PAT initiative [27] initiated a change in the pharmaceutical industry, toward a risk- and science-based approach for pharmaceutical processing. Real-time process monitoring and control, continuous improvement of processes and quick product technology transfer are now the focus and concepts like QbD, Design Space, Control Strategy, PAT, Process Signature became and remain hot topic for the pharmaceutical industry [28].

Following the PAT initiative, the ICH issued a guidance in which the concept of QbD was introduced. The QbD approach states that "quality should not be tested into products, it should be built-in". The introduction of the QbD concept opened the opportunity to design and validate processes, to run not only at one set of fixed processing conditions, but instead in a range of processing conditions known as the "design space". The delineation of a design space offered the possibility of responding to process disturbances, such as the variability in the raw materials, by establishing appropriate control loops, maintaining the process operation under optimal conditions with

a satisfactory product quality, without the need to revalidate the process when certain process parameters are manipulated. This greatly relieved the burden by discarding the necessity of regulatory post-approval changes.

Quality is built into pharmaceutical products through a comprehensive understanding of:

- The intended therapeutic objectives; patient population; route of administration; and pharmacological, toxicological, and pharmacokinetic characteristics of a drug.
- The chemical, physical, and biopharmaceutical characteristics of a drug.
- Design of a product and selection of product components and packaging based on drug attributes listed above.

  A process is generally considered well understood when:
- All critical sources of variability are identified and explained;
- variability is managed by the process and;
- product-quality attributes can be accurately and reliably predicted over the design space established for materials used, process conditions, manufacturing, environmental, and other conditions [29].

PAT is ideal to new products and processes development and R&D technology transfer, especially the ones in Phase I, II, and III of the clinical trials, since validation is simpler as the product and process are completely understood [27, 29].

The PAT initiative been supported by the European Medicines Agency (EMA) [30], that set up a PAT team in November 2003 to support the activities in the European Union, by the Japanese Ministry of Health, Labor, and Welfare, and by the ICH [10, 31].

From an implementation perspective, PAT can be visualized as the three-step process illustrated in Figure 2 [31]:

- **Design phase -** starts early in process development when the given unit operation is being designed and then later optimized and characterized. The critical quality attributes (CQA) that are being affected by the process step are identified along with the critical process parameters (CPP) that have been defined to affect the CQA. This type of knowledge is the essence of PAT and critical for the subsequent two phases.
- **Analyze phase -** in this phase a suitable analyzer is identified monitoring of CQA and CPP. PAT application can be implemented at-line (sample is removed, isolated from the process stream and analyzed close to it), on-line (sample removed for analysis from process stream and returned again), in-line (sample analyzed in place), and off-line (sample removed and analyzed away from process stream, normally later) [8, 18, 19]. For a PAT application, it is necessary for the

analytical results to be available in the time-frame necessary to allow real-time decision making.

- **Control phase** – this phase involves designing a control scheme based on process understanding. The data from the analyzer can be utilized for making real-time process decisions, allowing consistent process performance and product quality.

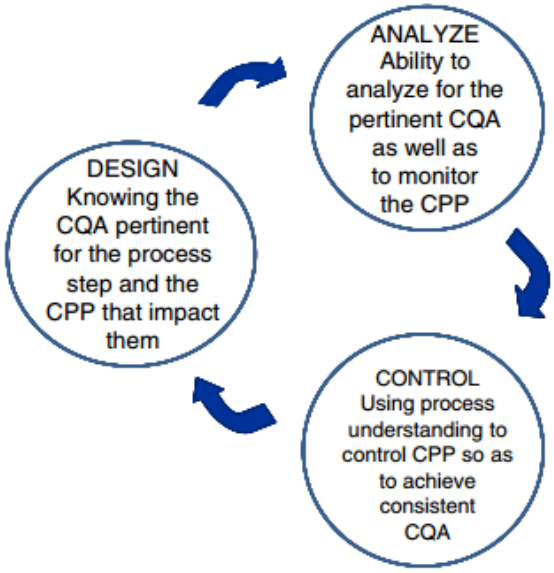


Figure 2 – Steps for PAT implementation, and the objective(s) of each step. Adapted from [31].

PAT tools can be categorized into the following categories [31]:

- Tools for design, data acquisition and analysis;
- Process analyzers;
- Process control tools and
- Continuous improvement and knowledge management tools.

Process analyzers are essential PAT tools enabling real-time product monitoring and they have advanced from simple univariate measurements such as pH and temperature to more complex information such as particle size and even multivariate information such as that extracted from spectroscopic techniques (e.g. Near Infrared and Raman spectroscopy) providing information related to biological, physical and chemical attributes of the materials being processed

For a more comprehensive reading, Rathore et al. [31] have made a compilation of PAT and QbD activities during the past decades.

Some issues remain in the implementation of process analyzers. An appropriate location of the sensor is necessary in order to accomplish representative sampling and to

minimize the influence of the analyzer on the process. When probes are present, fouling is also a recurring issue therefore, systems with purging gases or mechanical removal of material disturbing the measurements were developed.

The PAT framework includes not only the implementation of novel sensors and analyzers, but also the utilization of other fundamental tools for understanding increasing and for implementation of risk management strategy such as, design of experiments (DoE), advanced data analysis techniques, first-principles-based process modeling and control, and fundamental material characterization together with molecular modeling [32]. Modern industrial processes typically perform a large number of measurements. It is clear that to enable understanding, monitoring and control of those processes, multivariate data analysis techniques are required. Process data (temperatures, pressures, etc.) together with real-time analyzers provide valuable information about events that may not only affect the final quality, but also give early warnings for potential equipment failure.

2.4 Multivariate Statistical Process Control

Multivariate Data Analysis (MVDA), as detailed by its name, is the analysis of multiple variables at once. The FDA in the Guidance on PAT has stated the importance of this methodology. MVDA has seen its implementation strive with the development of fast and inexpensive computers, together with powerful analytical software.

Using MVDA tools for process monitoring is often called Multivariate Statistical Process Control (MSPC) [33, 34]. MSPC allows monitoring the wellness of both product and process in real-time by getting a reduced summary of all variables simultaneously, as they are collected. By the application of this concept, several industries managed not just to assure product quality, but at the same time improve performance by gathering knowledge about the process [35].

Sometimes it is not possible to characterize a product based only on end product quality measurements, since normally, not all quality properties can be measured. Also, these properties may differ, if dissimilar process trajectories are attained. In order to guarantee consistency, process conditions must be kept under control [36].

During production, various process and product variables are collected by implemented process analyzers and in-built sensors, often under the supervision of control loops. These loops utilize Proportional-Integral-Derivative (PID) and other controllers to compensate for many types of disturbances and keep the process at the defined setpoints. Despite these being able to handle many types of disturbances, there

are some types (faults) that cannot be handled in this way and therefore still need to be adequately detected, diagnosed and removed. MSPC helps to fill this gap.

2.4.1 Latent Variable Projection Methods

Latent variables (LV) projection models provide the means for obtaining fundamental process understanding by compressing information in a few variables and allowing its visual representation. They can be used for process monitoring and troubleshooting. A LV model will adequately deal with specific characteristics of process data such as low rank and high collinearity [37]. Methods such as PCA and PLS are often used to perform MSPC [38]. MSPC based on PCA or PLS is robust to redundancy, highly correlated, and noisy data, also at some extent to missing data; furthermore, it can be applied to data obtained from all modes of manufacturing (batch, continuous, etc.).

Collinear variables in a system can be combined in latent LVs describing the underlying structure of the data. LV models are generally used in the pharmaceutical industry, and can be helpful tools in acquiring process knowledge as discussed in detail by Kourti in 2006 [36].

2.4.1.1 Principal Component Analysis

The most common LV projection method is PCA [39-41]. In PCA a data matrix X is decomposed into a number of LVs (or principal components, PC) in two steps. Firstly, the first PC is found in the direction of the largest variance. Secondly, further PC's are calculated each PC under the constraint of being orthogonal (and hence independent) to the previous one. This results in a bilinear model, a product of **T** scores and **P** loading matrices (Eq. 1):

$$X = TP^T + E = t_1 p_1^T + t_2 p_2^T + \ldots + t_A p_A^T + E \quad (Eq. 1)$$

where, X is a **N x K** matrix consisting of **N** observations/samples (rows) with **K** measured X variables (columns). **M** is the number of Y variables. **T** is a **M x A** matrix and $\mathbf{P^T}$ is a **A x N** matrix, where **A** is the number of PC that were calculated. **T** and **P** consist on orthogonal and orthonormal vectors, respectively. **E** is an **M x N** matrix containing residuals, the variance which is not explained by the calculated PCs.

PCA is a technique for data compression and visualization. Each observation is attributed a score value on each PC. Plotting scores together can be utilized to reveal patterns in data such as clusters and outliers. Plotting the loadings for a PC reveals how

the original variable contributes to the PC, i.e. their collinearities. Together scores and loadings map the covariance structure of the data.

The number of components fitted in a model can be up to the number of original variables, but fitting that many principal components would annul the compression capabilities of PCA, and worse, noise would definitely be fitted. Only the PCs that capture the general variation in the data should be kept and the remaining (noise) should be left in the residuals (i.e. unmodeled).

2.4.1.2 Partial Least Squares

PLS is a regression technique, i.e. it aims at establishing a relationship between a set or predictor variables (X) and a set of responses (Y) and reveal how the later can be explained by the first [42-44]. If the number of X variables is reduced, if they are independent and contain little noise a Multiple Linear Regression (MLR) model is enough however, in most pharmaceutical problems X data is correlated. This happens for example when spectral data is analyzed. An approach to solve the collinearity problem is to regress the PCA scores from that data (which are independent) with Y, a procedure known as Principal Component Regression (PCR). This however can provide complex models since the relevance of the data compressed in the PCs might not be of relevance to the Y variables. To solve this issue PLS was finally suggested. In PLS a set of new variables/PC is calculated just like in PCA but the criteria for optimization is now the maximum covariance between X and Y data. Given the X matrix of predictor data and Y responses the model can be summarized by (Eq. 2 and 3):

$$X = tP^t + E = t_1p_1^T + t_2p_2^T + \ldots. + t_Ap_A^T + E \; (Eq.2)$$

$$Y = uQ^t + FE = u_1q_1^T + u_2q_2^T + \ldots. + u_Aq_A^T + F \; (Eq.3)$$

The information related to the observations are contained in the score matrices, **T** and **U**. Information related to variables is stored in loading matrix **P,** and the Y weight matrix **Q**. Variation in the data that is left out (noise) is on the X and Y residual matrices, **E** and **F**.

2.4.1.3 Other Approaches

Since their development, PCA and PLS have been subject to various modifications to suit multiple purposes and applications. Multiway PCA was proposed by Nomikos and MacGregor [45] to deal with the monitoring of three-dimensional batch process data, as was multiway PLS [46]. Multiblock PCA and PLS [47, 48] allowed to model datasets with

large number of variables and made the models more interpretable. In multiblock models, variables are divided into a number of blocks containing different information, the relative importance of the blocks can be defined, i.e. the weights of each block, leading to an improved interpretability of the model. Dynamic PCA and PLS methods allow to account for variable cross- and auto-correlation, i.e. process dynamics, by introducing time-lagged versions of the original variables in the X data [49-51]. Non-linear PCA and PLS algorithms have surged to deal with non-linear relationships, common in many chemicals and practical situations [52]. Johan Trygg and Svante Wold proposed the OPLS method in 2002 [53]. The difference between OPLS and a regular PLS models is that in OPLS the predictive and uncorrelated information captured by the PLS model are separated in different components [54, 55].

Other than PCA, PLS and their numerous extensions one can find in literature many other data-driven methods for building models from process data. Independent Component Analysis (ICA) [56], Artificial Neural Networks (ANN) [57, 58] and Support Vector Machines (SVM) [59-62] are some of these approaches. According to Kourti et al. [28] these fall within the class of regression methods that do not model the X-space and thus assume data to be full rank. These methods can be useful in some cases, even with process data however, they do not provide unique models, or allow for an easy interpretation. They also have limited ability to handle missing data, or test for outliers in new data. Methods for addressing the problem of utilizing future multivariate observations with missing data have been presented by Arteaga and Ferrer [63, 64]

2.4.2 MSPC for continuous and batch manufacturing

LV models are empirical, therefore, to develop a model for process monitoring, the first step is to acquire data that represents "normal" process behavior, i.e. obtain historical data from when the process was performing as expected, resulting in a product that is within specification. For the development of a monitoring model, multiple runs are performed, which should include as much variability during normal operation as possible (different seasons, different operators, etc.). Generally, the variables of interest are recorded through the production of a large number of batches under normal operation. If the objective is only to detect and diagnose process faults, a PCA model is enough, if in addition the objective is to predict process performance variables or critical quality attributes in real-time, multivariate regression models are then necessary (e.g. PLS).

MSPC for the case of continuous processing, works under the assumption that the system must be at a steady state of operation during the time of production i.e. the process mean and standard deviation remain time invariant.

During continuous processes, raw materials are continuously transformed throughout all the manufacturing steps until the final product is obtained. This is achieved by linking different processing units into a single manufacturing line [6]. Process parameters, and consequently end product quality, are kept constant over the whole production run and the system achieves, mostly after a startup period, a steady state of operation that will keep on until the process is stopped. Disregarding startup and transitions historical data collected from a continuous process is a bi-dimensional data array X with K variables measured at N time points.

BSPC is the batch approach to MSPC. Differently from the continuous case, batch processes possess a time-varying trajectory and a delimited duration. Batch process data is three-dimensional with K variables, measured at N time points for each one of B batches. To deal with these particularities of batch data, Nomikos and MacGregor in a series of landmark papers presented a new BSPM strategy suitable for three-dimensional batch data, the Multiway PCA/PLS [45, 46, 65, 66].

Another alternative for monitoring batch processes generating three-dimensional data is for example PARAFAC [67, 68]. Comparisons between different methods for multivariate statistical analysis of batch process data have been established by Westerhuis et al. [69] Chiang et al. [70] and Louwerse et al. [71].

Batch data can be made two-dimensional by unfolding. There are multiple ways to unfold batch data [72]. Westerhuis et al. [69] compare several ways to unfold the data matrix and discuss their effects for the monitoring of batch processes. Once the three-way data structure is unfolded the new two-way array can be modeled by regular PCA or PLS. The two most prevalent unfolding methods are batch-wise unfolding (BWU) and observation-wise unfolding (OWU). In BWU batch direction is preserved and the information of one batch is contained in a single row of the resulting matrix [73] while in OWU the variable direction is preserved. A critical discussion on other batch process modelling and monitoring procedures and other issues related to batch process analysis can be found in the papers of Kourti [49, 74]. All of these allow deciding on the rejection or acceptance of the product generated from a certain batch based on the information collected from the process in real-time.

Before modelling batch process data, it might be necessary to proceed to the alignment or synchronization of the trajectories. With alignment or synchronization, we look to obtain common start points for each of the phases of the run and match the shape

of the variable trajectories. Once shapes match the length of the batches is no more required to be the same. There are multiple approaches that have been reported for batch data synchronization such as the use of an indicator variable [35, 75] and dynamic time warping [76, 77]. If an indicator variable exists (or can be constructed by nonlinear transformations from other variables and/or process knowledge), the indicator variable approach is usually chosen as it is the simplest and most convenient for industrial applications.

Data other than batch process data can be modelled by the BSPM methodology. For example, data collected from transitions of continuous processes such as grade-to-grade transitions, startups, restarts, etc. It is appropriate to do so given that just as in a batch process these transitions possess a time-varying trajectory and delimited duration. A detailed discussion can be found in references [49, 78]. Several industrial applications of LV projection methods for multivariate statistical process control of transitions have been reported [79].

2.4.3 On-line Implementation of a MSPC LV model

Before being implemented in an on-line application a monitoring model should be comprehensively tested off-line and challenged by utilizing historical data with faults in order to assure that they can be detected. The performance of the new LV model should be compared with the previously employed monitoring scheme by running the two in parallel for some time. When an operator is familiar and comfortable with the new monitoring approach the shift to the new methodology can be completed. During real-time monitoring of processes the data being acquired is projected on the model for the calculation of the scores and the detection of deviations is from then on possible by means of different control charts where sudden changes, slow process drifts, etc. can be readily observed. Examination of loadings and/or score contribution plots provide the means for interpretation of the observed deviations and changes. Commonly, Hotelling's T^2 and Squared Prediction Error (SPE) are charted. SPE can be also called as Q residuals statistic. Alternatively, the Distance to the model in the X-space (DModX) is charted [33]. Control charts are essential tools to detect process upsets, equipment malfunctions, or other special events as early as possible and then to find and remove the factors causing those events.

2.4.3.1 Multivariate control charts

A monitoring statistic to be charted should be helpful in the identification of the process trends. When out-of-control moments occur, the cause for the deviation from normal behavior should be easy to determine resorting to contribution plots. Additionally, small process changes should be detected.

Various multivariate statistics and charts have the potential to be utilized for monitoring new process runs or batches during real-time being Hotelling's T^2 and DModX the most common. These two charts (Hotelling's T^2 and DModX are complementary indices; together they provide a picture of the state of the system. Hotelling's T^2 Charts are used to detect deviations from normal operation that are explained by the current model and within the overall variability but represent unusually high variation compared to the average process behavior. This means the observation is projected into the model plane but far from its center. The Hotelling's T^2 statistic for scores derived from latent variables models for *A* latent variables is calculated as follows (Eq. 4):

$$T_j^2 = \sum_{a=1}^{A} \frac{t_{j,a}^2}{s_a^2} \quad (Eq.4)$$

In (Eq. 4) s_a^2 is the estimate of variance of the corresponding latent variable a.

The Q residuals statistic is used for process deviation detection when events are not explained by the model. When these values fall out of the control values this means that there is a change in the underlying correlation structure of the data and the observation falls out of the normal variability observed in the reference runs utilized for model development (Eq.5).

$$Q_j = \sum_{k=1}^{k} (x_{j,k-} \hat{x}_{j,k})^2 \, (Eq.5)$$

Alternatively, the DModX can be calculated (Eq 6.):

$$DModX_{abs} = \sqrt{\frac{\sum_{k=1}^{k} (x_{j,k-} \hat{x}_{j,k})^2}{K - A}} \, (Eq.6)$$

where, $\hat{x}_{j,k}$ is the data estimation using the model with *A* components for sample *j* and variable *k*. DModX values are generally normalized by dividing them by the pooled relative standard deviation of the model in the X space s_0 . (Eq.7).

$$s_0 = \sqrt{\frac{\sum \sum_{k=1}^{k} (x_{j,k-} \hat{x}_{j,k})^2}{(N - A) * (K - A)}} \, (Eq.7)$$

where N represents the number of observations included in the model.

Moreover, score control charts allow monitoring the process performance at each model dimension separately differently from Hotelling's T^2 which allows monitoring all the model dimensions over the course of a batch run by using a single statistic calculated from all scores [33, 37].

In general, Shewhart type control charts are utilized. These charts use the information only from the current sample and they are relative insensitive to small and moderate shifts in the mean vector. Multivariate Cumulative Sum (CUSUM) control charts, and Multivariate Exponentially Weighted Moving Average (EWMA) control charts were developed to overcome this problem [33]. The calculation of the control limits for the charts is discussed in reference [28].

Minimizing the amount of false alarms is very important given that if during the process alarms occur very frequently, operators start disregarding the control system or even end up switching it off. In practice, control limits can be adapted to the current knowledge about the process and its characteristics in order to reduce the amount of false alarms. A method for false alarm reduction was proposed by Chen in 2010 [80].

2.4.3.2 Contribution Plots

In classical SPC, where only quality variables are monitored, it is up to process operators to diagnose a probable cause for an out-of-control event using their process knowledge and a one-at-a-time inspection of the process variables time-series charts. Alternatively, when PLS or PCA models based on process data are used to construct the multivariate control charts, they provide the user with the tools to diagnose assignable causes.

If the points in a process are within their respective limits everything is in order. If an out-of-limit point occurs, it means that some variables or a variable is deviating from the historical average behavior. To inspect which variables are contributing to the inflated statistic, contribution plots can be derived and will help clarifying the underlying cause. Miller et al. [81] and MacGregor et al. [47] were the ones to propose the idea of contribution plots for the first time. Other authors as Westerhuis et al. [82] also addressed this kind of plots.

The variable contribution plots provide a powerful tool for fault identification. However, care should be taken about its interpretation as it will help designate a variable or a group or variables that will numerically contribute to the out of the control signal but causality cannot be established. Variables with high contributions simply reveal the

signature of the fault, but it is the role of the operator or process engineer with knowledge about the process in question to try and attribute a probably cause and proceed to correct it in case of necessity.

2.4.4 Advantages of MSPC

Multivariate charts have superior detection capabilities when compared to univariate charts. Experience in industrial cases of MSPC reveals that 'in most cases in practice, changes in the covariance structure precede detectable deviations from nominal trajectories [36]. This indicates that in the event of an upset the correlation structure among the variables is broken in first instance before the problem becomes more evident and a clear deviation of the variables from their normal trajectories is observed. There have been reports of cases where a process upset changes the correlation between variables not causing the variables themselves to differ significantly from their normal trajectory. These cases are rare but they can continue unnoticed for very long periods until they are finally detected generally due to a complaint [83].

2.4.5 MSPC Model Maintenance and Update

LV model maintenance is often underrated and even forgotten about. However, a continued evaluation of the control system's behavior is vital and should it be necessary action has to be taken in the direction of ensuring proper performance. If a monitoring strategy is no longer found suitable for a specific application after a period of time after its implementation it means that it has become invalid due to the introduction of new sources of variability in the process that need now to be taken into account. Off-line updating or even rebuilding of models may be necessary in order to ensure that the empirical models maintain a high degree of fidelity to the process [84].

2.4.6 Examples from Industry Practitioners of the use of LV models for MSPC/BSPC

Since their appearance numerous examples of the application of the concepts of MSPC in various industries have surged, for both batch and continuous applications. **Table 2**. presents a literature summary of some of these applications that make use of PCA, PLS and their multiway extensions for monitoring purposes. Many reports can be found about of fermentation processes. Metallurgical and chemical are two other industries where the use of MSPC is often reported.

In the pharmaceutical industry the amount of case-study reports for process overview, monitoring and control has tremendously increased after the introduction of the QbD initiative and of the PAT framework [85]. Examples of these applications for process monitoring have been extensively reviewed by different authors [86, 87].

Nevertheless, few reports exist for MSPC/BSPM of pharmaceutical production processes. In 2009, Garcia-Muñoz et al. proposed a real-time process monitoring and fault detection tool for the continuous quality assurance of a spray drying application [88]. In 2011, Burggraeve et al. developed a BSPM methodology for a fluid bed granulation process based on in-line particle size distribution measurements (using spatial filter velocimetry) combined with continuous product temperature registration [89].

In 2011 Rosas et al. [90] proposed an approach for determining the composition of a pharmaceutical gel and assess the temporal changes in major physical factors affecting the quality of the product (specifically, viscosity and pH) using NIR spectroscopy. Kona et al. (2013) [91] brought together in-line NIR spectroscopy and temperature and relative humidity data for developing PLS and PCA models allowing the statistical process monitoring of fluid bed granulation. In 2014 Barla et al. [92] proposed a PCA method to obtain qualitative information during real-time monitoring of a fluid bed dryer using NIR spectroscopy. They also proposed a PLS methodology for real-time monitoring of moisture content using the same NIR data. Sarraguça et al. have recently proposed a batch statistical process control approach based on PCA to monitor a cocrystallization process between furosemide and *p*-aminobenzoic acid [93].

Given the much-reduced number of reports of the monitoring of processes by means of LV models in pharmaceutical manufacturing applications in comparison to the other industries one can recognize that these tools are still to be explored by the pharmaceutical industry and can be of great added value in the near future.

Table 2 – Summary of publications regarding the application of Multivariate statistical process control for different applications

Field	Application	Multivariate Monitoring based on	Reference
Alarm systems	Early fire warning detection	PCA	[94]
Analytical methods	Chromatography	PCA	[95]
Analytical methods	HPLC	Window Factor Analysis (PCA extension)	[96]
Analytical methods	Particle size analysis in wet processes	PCA	[97]
Analytical methods	Chromatography	PCA	[98]
Analytical methods	Chromatography	PCA	[99]
Automotive industry	Autobody Assembly Process	PCA	[38]
Biodiesel production	Monitoring transesterification reactions	PCA PLS	[100]

Table 2 – Summary of publications regarding the application of Multivariate statistical process control for different applications

Field	Application	Multivariate Monitoring based on	Reference
Bioprocessing	Fed-batch Fermentation	PCA PLS Neural Networks	[101]
Bioprocessing	Batch Fermentation	PCA	[102]
Bioprocessing	Fermentation	Multiway PCA Multiway PLS	[103]
Bioprocessing	Fed Batch Fermentation	Multiway PCA	[104]
Bioprocessing	Fed Batch Fermentation	PCA PLS	[105]
Bioprocessing	Continuous Stirred Tank Reactor	PCA	[106]
Ceramic industry	Liquid fed ceramic melting	PCA	[107]
Ceramic industry	Production of melt polycarbonate	PCA	[108]
Chemical industry	Continuous product recovery	PLS	[35]
Chemical industry	Titanium dioxide manufacturing	PLS	[109]
Chemical industry	Electrolysis	PCA	[110]
Chemical industry	Fluidized bed reactor	PCA PLS Canonical Variates Analysis	[72]
Chemical industry	Alcohol precipitation	Multiway PCA	[111]
Chemical industry	Resin Production	PLS	[72]
Culture cell monitoring	Chinese hamster ovary cell culture	PCA	[112]
Culture cell monitoring	Complex cell cultures	Multiway PCA	[113]
Food industry	Sugar crystallization	Batch Dynamic PCA Moving Window PCA Batch Observation Level Analysis Time-varying State Space Modelling	[114]
Geoengineering	Carbon Dioxide Removal	PCA	[115]
Metallurgical industry	Grinding	PCA	[116]
Metallurgical industry	Steel casting	PCA Multiway PCA	[117]
Metallurgical industry	Continuous slab casting	PCA	[118]
Metallurgical industry	Desulfurization	Adaptative PCA	[118]
Metallurgical industry	Sulfite pulp digestion	PCA	[118]
Metallurgical industry	Continuous slab casting	Multiway PCA	[119]
Metallurgical industry	Aluminum smelting	Multiway PCA	[120]
Nanomaterials	Ultrasonic Attenuation in Nanomaterial Processing	PCA	[121]
Paper industry	Paperboard manufacturing	PCA	[118]
Paper industry	Paperboard manufacturing	PCA	[122]
Petrochemical industry	Petroleum refining	PCA Multiblock PCA Recursive PCA	[123]
Pharmaceutical industry	Crystallization and solid state analysis	PCA PLS	[124]
Pharmaceutical industry	Cocrystallization	PCA	[125]
Pharmaceutical industry	Crystallization	PCA PLS	[124]
Pharmaceutical industry	Fluid bed granulation	PLS	[89]
Pharmaceutical industry	Fluid bed granulation	Multiway PCA	[91]

Table 2 – Summary of publications regarding the application of Multivariate statistical process control for different applications

Field	Application	Multivariate Monitoring based on	Reference
Pharmaceutical industry	Gel manufacturing	PLS	[90]
Pharmaceutical industry	Fluid bed drying	PCA PLS	[92]
Pharmaceutical industry	Spray-drying	PLS	[88]
Polymer industry	Emulsion polymerization	Multiway PCA Multiway PLS	[75]
Polymer industry	Semi-batch polymerization	Multiway PLS	[35]
Polymer industry	Injection molding	Multi-stage sub PCA	[126]
Polymer industry	Injection molding	Dynamic PCA	[127]
Polymer industry	Polypropylene catalyzer reaction	PCA Independent Component Analysis (ICA) PCA with KDE (KPCA) ICA with KDE (KICA)	[128]
Printing	Fault Diagnosis Method of Feeding Mechanism in Printing Machine	PCA	[129]
Processing industry	Batch product drying	PLS Multiway PLS	[130]
Safeguarding pharmaceuticals, diagnostic and clinical safety	Quality Monitoring of Isoniazid and Rifampicin	PCA	[131]
Safeguarding pharmaceuticals, diagnostic and clinical safety	Monitoring pharmacokinetics, clinical and pharmacological studies	PCA PLS	[132]
Safeguarding pharmaceuticals, diagnostic and clinical safety	Purity analysis of biopharmaceuticals	PCA	[133]
Safeguarding pharmaceuticals, diagnostic and clinical safety	Monitoring captopril stability	PCA	[134]
Structural assessment	Damage identification	PCA Multiway PCA Multiway PLS	[135]
Water supply	Detect anomalous behaviors in a water supply system	PCA	[136]
Water supply	Burst detection in water networks	PCA	[137]
Waste management	Wastewater treatment	Multiway PCA PLS	[138]
Waste management	Wastewater treatment	PCA	[139]
Waste management	Wastewater treatment	Adaptative PCA	[140]
Waste management	Wastewater treatment	PCA	[141]
Waste management	Solid waste moving grate-type incineration	PCA PLS	[142]
Wood industry	Wood Pelletizing	PCA PLS	[143]

Chapter 3

Particle sizing measurements in pharmaceutical applications: comparison of in-process methods versus off-line methods[1]

3.1 Abstract

It has been previously described that when a sample's particle size is determined using different sizing techniques, the results can differ considerably. The purpose of this study was to review several in-process techniques for particle size determination (Spatial Filtering Velocimetry, Focused Beam Reflectance Measurements, Photometric Stereo Imaging and the Eyecon™ technology) and compare them to well-known and widespread off-line reference methods (laser diffraction and sieve analysis). To start with, a theoretical explanation of the working mechanism behind each sizing technique is presented, and a comparison between them is established. Secondly, six batches of granules and pellets (i.e. spherical particles) having different sizes were me-asured using these techniques. The obtained size distributions and related D_{10}, D_{50} and D_{90} values were compared using

[1] This chapter has been adapted: from: A.F.T. Silva, A. Burggraeve, Q. Denon, P. Van der Meeren, N. Sandler, T. Van Den Kerkhof, M. Hellings, C. Vervaet, J.P. Remon, J.A. Lopes, T. De Beer, Particle sizing measurements in pharmaceutical applications: Comparison of in-process methods versus off-line methods, European Journal of Pharmaceutics and Biopharmaceutics, 85 (2013) 1006-1018 (http://dx.doi.org/10.1016/j.ejpb.2013.03.032).

the laser diffraction wet dispersion method as reference technique. As expected, each technique provided different size distributions with different D values. These dissimilarities were examined and explained considering the measurement principles behind each sizing technique. The particle property measured by each particle size analyzer (particle size or chord length) and how it is measured as well as the way in which size information is derived and calculated from this measured property and how results are presented (e.g. volume or mass distributions) are essential for the interpretation of the particle size data.

3.2 Introduction

Building quality into pharmaceutical products is the leading purpose of the PAT initiative [9]. Particle size is a critical quality parameter in a number of pharmaceutical unit operations such as pre-mixing/mixing, granulation, drying, milling, roller compaction, spray-drying, coating and compression. An adequate particle size distribution (PSD) is essential to ensure optimal manufacturability which will have an important impact on the end product's safety, efficacy and quality. Therefore, monitoring and controlling particle size via in-process particle size measurements is essential to the pharmaceutical industry.

The application of in-process particle sizing tools for the assessment of the influence of process and formulation parameters upon critical product quality attributes has been studied for several pharmaceutical processes such as fluid bed granulation [89, 144, 145], hot melt granulation [146], spheronization [147] and crystallization [148-150]. However, differences between the measurement mechanisms and principles of the particle size analyzers (both off-line and in-process) make the direct comparison between them a challenging task [148, 151]. The aim of this study is to review different in-process particle sizing techniques and compare them to acknowledged off-line techniques (laser diffraction (LD) and sieve analysis). To establish this comparison six batches of granules and pellets (i.e. spherical particles) having different sizes were measured with the different equipment. The evaluated in-process techniques include Focused Beam Reflectance Measurements (FBRM), Spatial Filtering Velocimetry (SFV), Photometric Stereo Imaging (PSI) and the Eyecon™ technology. Table 3 provides a comparison between the assayed equipment. It discloses the underlying theoretical assumptions behind each instrument's measurement mechanism, unveils the way in which size is acquired and presented by each instrument, describes their applicability, known capabilities and drawbacks. The choice of an appropriate analyzer for measuring particle size in a specific case has to take into consideration these listed characteristics. In an industrial environment, when a new particle size analyzer is implemented in a process environment, an often-executed

procedure is to attempt to correlate the data from the traditionally used off-line analyzer with the data from the new in-process analyzer. However, due to the different measurement principles behind each sizing technique, it is obvious that this is not an accurate and reliable procedure as mostly very different particle properties are measured by each sizing technique, hence providing uncorrelated results. A particle size distribution is usually depicted by a histogram where the size-related property measured by the analyzer (total particle volume, number of particles or counts, total particle length, total particle area, etc.) is plotted as a function of demarcated size classes. D values are parameters often used in the characterization of a PSD, a D_i value of x indicating that particles with a size smaller or equal to x account for $i\%$ of the measured size-related property.

Table 3 – Comparison between the different studied particle size analyzers (CL – chord length; CLD – chord length distribution; N/A – non-applicable; PS – particle size; PSD – particle size distribution; LD – laser diffraction; SFV – spatial filtering velocimetry; PSI – photometric stereo imaging).

Instrument	Mastersizer™ S (LD)	Sieve Analysis	Parsum™ IPP70 (SFV)	FBRM™ C35	FS3D™ (PSI)	Eyecon™
Assumptions	Mie's theory: Assumes particles to be spherical | Assumes scattered light is measured before it is re-scattered by other particles | Optical properties of the particles and the medium surrounding them are supposed to be known [13]	Particles will pass through the mesh when the second largest dimension is less than the mesh size, i.e. at least two dimensions of the particle must be smaller than the sieve size	No assumptions about particle shape are made	No assumptions about particle shape are made | Particle velocity is small compared to the laser rotational velocity [16, 29, 30]	Samples are positioned against a straight glass and therefore the surface is assumed to be approximately straight. Linear integration in a horizontal direction which allows the obtention of a 3D surface. Peaks on this surface are assumed to be particles	An ellipse is fitted to the particle edges in order to obtain an average particle diameter | Assumes particle as being spherical to allow the calculation of its (relative) mass from the average diameter, this mass is used in the calculation of D values
Size Distribution type	Volume-based PSD	Mass-based PSD	Volume or number-based PSD (obtained by conversion from a CLD)	Chord length distribution (possible to apply different weighting methods)	Volume-based PSD	Number-based PSD
PS Interval	0.05-3500 µm	> 38 µm [10]	50-6000 µm	3-3000 µm	> 20 µm (maximum size depends on the performed calibration)	50-3000 µm
Particle velocity	N/A	N/A	0.01 – 50m/s	N/A	N/A	N/A
Destructive	Sample can be retrieved but sample dispersion by means of pressurized air may cause particle breakage [3]	Yes	No	No	No	No
In-process measurements	Not with Mastersizer™ S. Possible with some equipments (Insitec™ by Malvern, UK; Mytos ™ by Sympatec™, Germany) but	N/A	Yes (bench-top version also available)	Yes	On-line measurements are possible with appropriate feeder (however the measured sample has to be static)	Yes (bench-top version also available)

Table 3 – Comparison between the different studied particle size analyzers (CL – chord length; CLD – chord length distribution; N/A – non-applicable; PS – particle size; PSD – particle size distribution; LD – laser diffraction; SFV – spatial filtering velocimetry; PSI – photometric stereo imaging).

Instrument	Mastersizer™ S (LD)	Sieve Analysis	Parsum™ IPP70 (SFV)	FBRM™ C35	FS3D™ (PSI)	Eyecon™
	difficulties on presenting the sample in the appropriate concentration [11]					
Suitable for	Non-fragile particles (powders or liquid suspensions or emulsions)	Powder or granular solid particles > 38µm	Solid particles suspended in an air stream	Solid particles suspended in a liquid or in an air stream	Solid particles	Solid particles
Advantages	Ease of use | Little maintenance | Rapid measurements | Highly repeatable | Allows background subtraction | No calibration required [13]	Cheap	Ease of use | Little maintenance | Rapid measurements | No calibration required | Shape is taken into consideration [18, 21] | Probe fouling is prevented by means of a pressurized air system | Clip-in accessories are available for measurements under difficult process conditions	Ease of use | Little maintenance | Rapid measurements | Shape is taken into consideration | In this model a window scrapper allows measurements in highly concentrated particle systems	Ease of use | Little maintenance | Rapid measurements | Provides particle size alongside with important morphological information | Able to image overlaying, wet particles and extract PSD	Ease of use | Little maintenance | Rapid measurements | Provides both size and morphological information under dynamic conditions
Disadvantages	Relatively large amount of sample required for dry dispersion method (depending on the size of the particles)	Cohesive and agglomerated materials are difficult to measure | Low resolution due to the limited amount of sieves that can be fitted| Measurement times and operating methods have an influence on the results | Particle shape has a big influence on the results | Laborious and time-consuming| Requires large amount of sample [10]	Property being measured is CL and not PS (conversion to PSD by the instrument's software) | Not suitable to measure sizes < 50 µm | Probe may be susceptible to fouling [20] |Accessories are adequate only within a particle size range, if particles are outside this range they are not measured thus providing biased results	Property being measured is CL and not PS [30, 45, 48, 49]	Samples have to be static |Shades caused by irregularities in the surface of a particle may trick the instrument in detecting multiple particles [42] | Transparent particles cannot be measured | Coverage of the measurement window by fines at the moment of sampling may lead to particle size underestimation [42, 47]	At the moment, no information available about performance during in-line measurements

3.2.1 Off-line particle sizing methods

3.2.1.1 Laser Diffraction

LD is the most applied technique for the particle size measurement of pharmaceutical powders and granules. It can be used as an in-process method [152] or as an off-line method. A dispersed sample passes through a beam of monochromatic light causing light scattering which is measured as a function of scattering angle by a multi-element detector. As the scattering pattern, i.e. scattered intensity as a function of scattering angle, is largely particle size dependent, it follows that particle size information can be extracted from the experimentally determined pattern. Older instruments mainly rely on the Fraunhofer approximation to derive particle size information from the scattering pattern, while recent LD particle size analyzers are based on Mie's theory [151]. The Fraunhofer approximation is based on a number of assumptions: it assumes that particles are opaque discs, that light is scattered at only narrow angles and that all particle sizes scatter with the same efficiency. Furthermore, it does not take into consideration the optical properties of the measured material, and therefore its use is recommended when measuring mixtures of different materials. Differently, Mie's theory predicts the scattering intensity induced by particles, irrespective of the fact whether they are transparent or opaque. It is based on the assumptions that the measured particles are spherical, that the dispersion is dilute, so that light is scattered by one particle and detected before it interacts with other particles, that the optical properties of the particles and the medium surrounding them are known and that particles are homogeneous i.e. uniform in composition. Nowadays, the ISO13320 standard for LD particle size analysis acknowledges the superiority of Mie's theory [153, 154]. LD particle size analyzers that use Mie's theory (e.g. Mastersizer™ S) base their particle size calculation on the assumption that particles are spherical, which is rarely true. This is a solution to deal with the fact that the only shape that can be described by a single dimension is the sphere. LD results are generally presented as a volume-weighted particle size distribution. Thus, in LD results reporting that the median value (D_{50}) of a volume-based PSD is 100 μm means that particles with a size up to 100 μm account for 50% of the measured sample volume. Alternatively, a number-weighted distribution can be extracted, depending on the analyzer's software.

3.2.1.2 Sieve analysis

Before the introduction of LD, sieving used to be the most commonly applied sizing method, and it is still widely used for the determination of particle size because of its inexpensiveness. It is described in the European Pharmacopoeia [155] that sieve size is the "size of the aperture measured perpendicular to the wire through the center of the opening". The mass of material that is retained on a specific sieve is weighted and presented as a percentage of the total assayed material. Therefore, a mass-based PSD is generated. The results are generally presented as a cumulative mass distribution. In this case a median (D_{50}) of 100 µm indicates that 50% of the total weight of the measured material is constituted by particles that would pass through a sieve with 100 µm apertures. It is acknowledged that for a particle to pass through a sieve it must have two dimensions smaller than the sieve size. This is why it can be assumed that sieve analysis separates particles according to their second largest dimension. Some of the described disadvantages of sieve analysis are as follows: test sieves require regular care in order to maintain their performance, their cleaning must be careful as vigorous brushing may distort sieve openings, it is not possible to perform sieve analysis on sprays or emulsions, measurement of dry powders with sizes under 38 µm is very difficult as electrostatic charges may cause loss of material (wet sieving may be a solution but this technique provides very poor reproducibility and is difficult to carry out) and cohesive or agglomerated materials are problematic to measure as they form aggregates that will not pass through the sieve's aperture [151, 156]. Sieve analysis also requires a relatively large amount of sample and, as a consequence, is not appropriate for costly materials or materials of which only small quantities are available. Samples can be eroded due to attrition during the analysis making sieving unsuitable for these materials. Measurement times and operating methods (e.g. shaking) need to be standardized as the longer the measurement is performed, the smaller the obtained particle size is as particles have time to orient themselves to fall through the sieve. This is particularly important when dealing with odd-shaped particles which are difficult to sieve and may generate peculiar results. For instance, measuring the particle size of needle-like or rod-like particles by means of sieve analysis might not be the best choice. Additionally, there is an increase in the risk of particle erosion as sieving time increases. These and further disadvantages of this method are described in Table 3.

3.2.2 In-process particle sizing method

3.2.2.1 Methods based on chord length measurements

There are in-process particle size analyzers that measure chord length instead of actual particle size such as SFV and FBRM. A particle's chord length can be defined as a geometric line segment whose endpoints both lie on the surface of the particle. These analyzers utilize a laser beam that crosses the particle randomly acquiring a chord length. The number of times a given chord length is measured takes the form of a probability density function. In case of spherical particles, the diameter is the largest chord possible and the probability of the measured chord length is independent of the particle orientation towards the laser beam (Figure 3 1.) while for irregular and odd-shaped particles, shape and orientation will influence the measured chord lengths (Figure 3 2a. and 2b.). Hence, the chord length distribution (CLD) depends on both the PSD and the particle shape. Presenting the results as particle size is easier to interpret than chord length as particle size is often directly related to product quality, and it allows the comparison to particle size measured by other instruments [157]. Both SFV and FBRM utilize a laser beam for their measurements: SFV calculates the chord length from the shadows cast by the particles that cross the laser beam, FBRM calculates it from the laser light that is reflected back from the particle and propagated back through the probe.

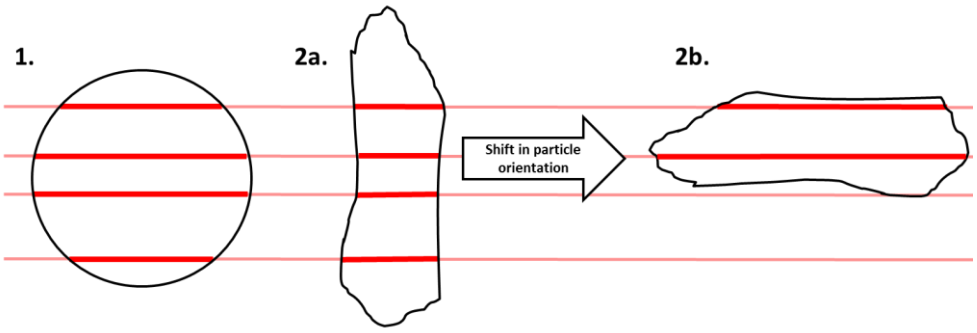


Figure 3 – Examples of the measured chord length (bold line) when a laser beam crosses (1) a spherical particle and (2a. and 2b.) an irregular particle in different positions – illustration of the effect of particle orientation on the obtained chord length.

Spatial filtering velocimetry

A system based on the SFV principle is the Parsum™ IPP70 probe which was utilized in this study. The working principle of the Parsum™ IPP70 SFV probe is presented in Figure 4. When passing in between the two sapphire windows of the probe, the laser beam hits the particles. These particles cast then a shadow on a detector array of optical fibers generating two burst signals (burst a and burst b). The difference between

these two bursts is obtained ("BURST") and its frequency calculated. This frequency (f) is then multiplied by the spatial filter constant (g), which corresponds to the distance between the detector arrays, and the particle's velocity is obtained (v). When the particle travels through the probe, a secondary signal ("PULSE") is also acquired by a single optical fiber, and the duration of this pulse is measured (tp). The chord length (x) of the particle is then calculated by multiplying the particle's velocity with the pulse signal's duration [158-160]. The Parsum™ IPP70 system is able to report size after converting the raw CLD to a number or volume-based PSD performed by an algorithm in the system's software. Some of the algorithms that have been described to convert from a CLD to a PSD are addressed later in this section. Furthermore, the user is allowed to define the different size classes. In this way, the percentage of particles with sizes in-between the user-defined values is calculated. Fouling is a recurrent problem during in-process measurements and for that reason the sapphire windows of the Parsum™ IPP70 probe are kept clear by feeding compressed air through the probe itself, though this is not always efficient [161]. Additionally, a range of different clip-in accessories is available. Two different flushing cells: SZ11 (an open flow cell slit) and SZ20-4 (a cell with a front side aperture of 6mm), both designed to protect and keep the probe's windows clear. According to the manufacturer's specifications, the SZ11 is appropriate for the measurement of free-falling particles with sizes between 100 and 4000 µm and a very low percentage of fines, whereas the SZ20-4 is suitable for measuring free-falling particles with sizes from 50 to 2500 µm and a low to average content of fines. A disperser accessory with a ring injector, diluter, and a back flush function aperture of 4mm (D23) with both an external and internal air connection is also available and is particularly fit for the measurement of small particles (50 to 2000 µm), especially in processes with high particle concentrations in the measurement volume, i.e. high particle loadings. Parsum™ IPP70's software enables the monitoring of particle loading during measurements expressed as a percentage of the measurement volume. By the use of the disperser D23 particles are accelerated, and consequently the distances between them become larger ensuring that they are presented to the instrument's detector in a suitable concentration for an accurate measurement. The choice of the appropriate accessory to utilize depends on the characteristics of the particles being analyzed. For instance, the measurement of a sample containing particles outside the appropriate size range for a certain accessory will result in biased results. SFV has already been suggested for particle size monitoring in fluidized bed processes [144], mixing and coating, high-shear wet granulation, dry granulation and spray drying [162].

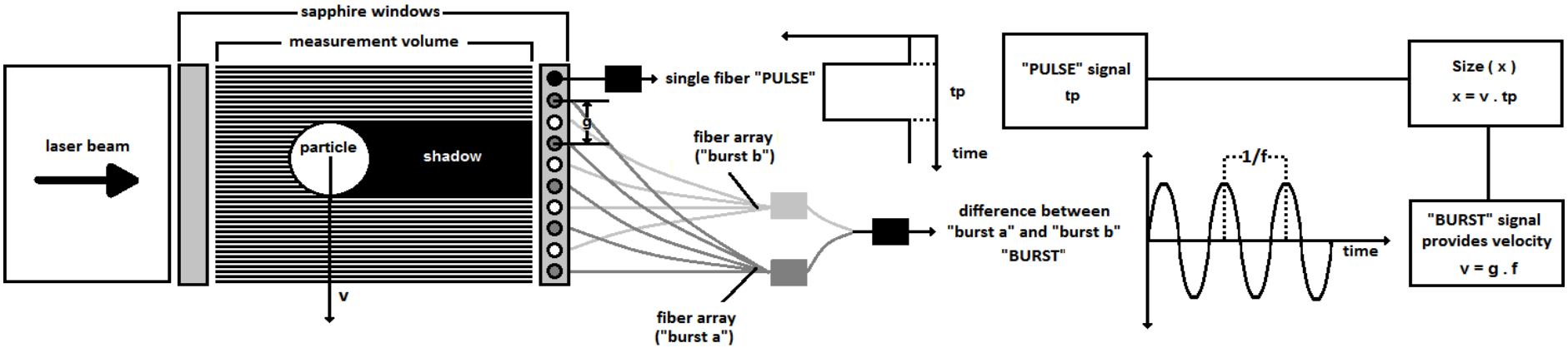


Figure 4 – Working principle of the Parsum™ IPP70 probe.

Focused Beam Reflectance Measurements

FBRM is another process analytical tool designed for measuring chord lengths. It measures the light that is reflected and propagated back through the probe when a tightly-focused laser beam, rotating at a high speed (2 - 8 m/s), hits a particle (Figure 5). The chord length is then calculated by multiplying the duration of reflection with the laser beam's scan speed.

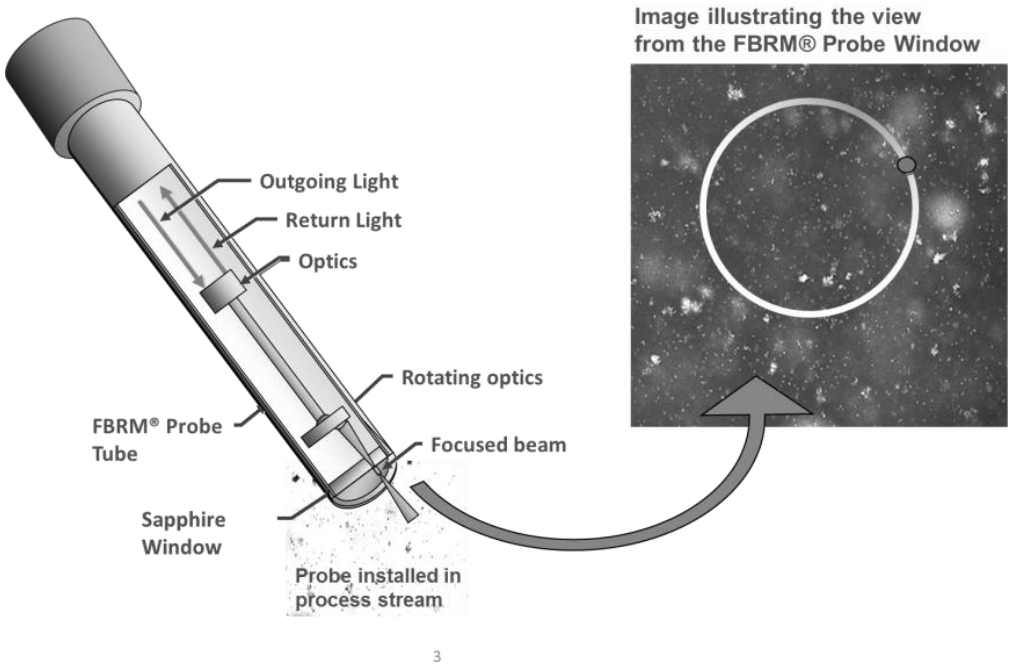


Figure 5 – Working principle of the FBRM™ technology [163].

FBRM has already been successfully applied for suspensions and crystallization processes [148-150, 164-166] and has also been studied for fluid bed granulation in comparison with other PAT tools [167]. Polymorphic transition monitoring [168], control of particle disruption [169], and solubility measurements [170] are some other applications where the use of FBRM has already been reported. FBRM™ C35 measurements can be performed in highly concentrated particle systems as a scraping system is installed on the probe's sapphire window, keeping it clean and preventing probe fouling during in-process measurements. FBRM™ C35 is a count-based technique which means that the sizing results are presented by the FBRM™ C35 software (iC FBRM™) as a number-based chord length distribution (number of particles measured within a chord length class). This software also

allows the extraction of D values from these distributions and of size (chord length) classes. As mentioned previously, size results are usually presented as a number, length, area or volume-based distribution. However, the FBRM™ C35 system results are presented as a raw chord length frequency distribution and can be transformed into a 1/length-weighted, length-weighted, square-weighted or cubic-weighted chord length frequency distribution. The weighing method to use depends on the aim of the measurement. If there is the necessity of detecting slight changes in the fraction of smaller particles, no weighing or length-weighing will emphasize these rather than the larger ones. On the other hand, if the interest lies on detecting small changes in the larger particles square and cubic weighing emphasize the coarser particles at the same time making the detection of changes in the smaller size range more difficult. It is described that the raw chord length data are similar to a length-based PSD since the probability of a certain chord length being detected is proportional to the linear dimension of a particle.

SFV and FBRM both measure chord length, but their measuring principles differ substantially. For the FBRM™ C35 system, it is necessary for particles to flow over the sapphire window. Particles that are positioned a few hundred micrometers away from the sapphire window will most likely not be measured, hence making the placement (implementation in the process environment) of the FBRM™ C35 probe of utmost importance. Dispersing and measuring the sample in a liquid in which it is insoluble is a highly suitable solution for the FBRM™ C35 system. However, the sample has to be diverted from the process and is not reusable. Parsum™ IPP70 or any other SFV system cannot be applied in suspensions. FBRM™ C35 is capable of measuring smaller particle sizes (i.e. 3 – 3000 µm) while SFV is adequate for the measurement of particle systems sized 50 – 6000 µm. The main advantages of using these systems include the fact that no calibration is needed and the capability of measuring in-line at high particle concentrations (loadings) due to the use of purging systems [171, 172].

As a particle chord length is not identical to the generally used particle size, several authors have presented their solutions to express the relationship between PSD and CLD. The Parsum™ IPP70's software performs this CLD-PSD conversion itself, while for the FBRM™ C35, the results are expressed as chord lengths. The easiest way to convert a CLD into its corresponding PSD is by developing a PSD-CLD model to calculate CLD corresponding to a known PSD and shape and afterwards invert it to obtain a PSD from the CLD (CLD-PSD model) [157]. For two-dimensional spherical particles, the translation from

PSD to CLD is based on different methods such as the probability apportioning method and Bayes' Theorem [173-177]. The probability apportioning method can also be used to calculate CLD from PSD for two-dimensional ellipsoidal particles [177-180]. For non-spherical two-dimensional and three-dimensional particles, little has been described [177, 181]. In 2001, Langston and Jones [181] presented a method in which for a certain PSD of non-spherical particles, the chord length probability distribution is determined by simulating random cuts in the particles. This method is highly dependent on assumptions made during the calculation, and the resulting data is not accurate. On the other hand, Ruf et al. [172] presented another procedure in which, for a three-dimensional ellipsoidal particle, the chord length probability distribution is obtained from two-dimensional projections at every orientation. The conversion of a PSD from a CLD is an inversion problem and the most utilized methods to solve this problem include the Least Squares and Constrained Least Squares algorithms [173, 175, 176, 179, 180, 182]. However, these might provide negative numbers of particles when the CLD measurements are noisy, and therefore, an interactive apportioning method utilizing Bayes' Theorem has been developed to overcome this limitation [173, 176]. For most processes, however, a good precision is often more important than accuracy as the interest relies on the monitoring of process dynamic changes such as particle shape and/or concentration of the suspensions [157, 172].

3.2.2.2 Photometric Stereo Imaging

Another studied technique was Photometric Stereo Imaging. The Photometric Stereo Imaging unit Flashsizer3D™ (FS3D™) consists of a monochrome CCD camera connected to a metal cuvette with a glass window and a computer. The tool is equipped with a sampling unit that allows on-line measurements. Two light sources, positioned relative to each other at an angle of 180°, illuminate the sample, and two digital images of the sample are obtained. A grey-scale value between 0 (black) and 255 (white) is attributed to each individual pixel, and the shading effects expose the topography of the surface (Figure 6).v

The gradient fields are subjected to line integration in a horizontal direction to obtain a 3D surface. This surface is assumed to be approximately straight as the samples are placed against a straight glass surface during measurement. Therefore, peaks on the 3D surface are assumed to be particles, and the projected volume-based (V) particle size is then calculated from the area of the peaks in the *xy* direction:

$$d = \sqrt{(area)}\,.\,c \quad (\mathrm{Eq.}\,8)$$

$$V = d^3 \ (\mathrm{Eq.}\ 9)$$

where *d* is the diameter of the particle, *area* is the area of the peaks and *c* a calibration constant, calibrated by default with pellets [183]. If the shape of the particles to be measured differs significantly from spherical, the calibration constant can be changed accordingly to the particles to be measured [26]. This imaging unit allows the acquisition of a volume-based PSD and related D values of the particles captured in each image. The size classes can be defined by the user. The FS3D™ system has been used for the measurement of powders [184] , granules [183], and pellets [147] showing the potential of this technique as a fast particle size analyzer for various types of material.

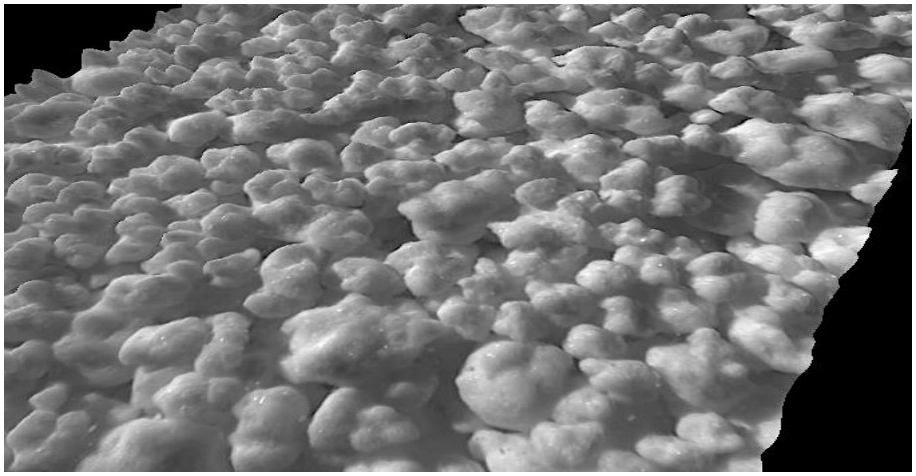


Figure 6 – A typical example of a surface visualized in 3D that is used in particle sizing with the photometric stereo imaging approach.

3.2.2.3 Eyecon™

The Eyecon™ particle sizing technology was also tested. This is a very recent 3D-imaging system that allows the determination of the PSD for moving particles using a flash imaging technique (Figure 7).

The equipment can either be used off-line or in-process. During measurements, a powerful short light pulse is created and provided that the particle movement during this pulse is negligible a sharp image without blurring is captured. The particles are illuminated with red, green, and blue LEDs from different angles. The color on the surface of the particle is captured in an image, and for each individual pixel, a map of the surface height is built. Furthermore, using image gradient data an ellipse is fitted on the particle edges and its

maximum and minimum diameters are obtained. These are used to calculate the average aspect ratio (AAR) of particles as an indicator of their sphericity by means of the following equation:

$$AAR = \frac{D_{max}}{D_{min}} \ (\text{Eq. } 10)$$

where D_{max} represents the maximum measured diameter and D_{min} the minimum measured diameter. Also, the average diameter can be assessed according to the following equation:

$$d = \frac{D_{max} + D_{min}}{2} \ (\text{Eq. } 11)$$

Posteriorly, the particle is modeled as a sphere and its mass is obtained by means of the following equation:

$$Mass = \frac{\pi . d^3 . \rho}{6} \ (\text{Eq. } 12)$$

where ρ represents the density of the particles. As it is an unknown value, all the particles are assumed to have the same density, and therefore, it is a constant that can be eliminated, as can π and *6* and Eq. (13) is obtained:

$$Mass = \mathrm{d}^3 \ (Eq. 13)$$

The obtained mass value is then a relative mass value, not a true mass. Each captured image is analyzed by Eyecon™ resulting in a group of ellipses. Results can either be computed using only the current image or also include data from previous images and are presented as a histogram.

The D values are calculated by ordering particles in order of ascending relative mass. Firstly, the total mass is computed, and then, an iterative algorithm adds up starting with the smallest of the particles. As the running total reaches 10%, 25%, 50%, 75% and 90% of the total mass, the diameter of the last added particle is recorded as being the D_{10}, D_{25}, D_{50}, D_{75} and D_{90} diameter, respectively.

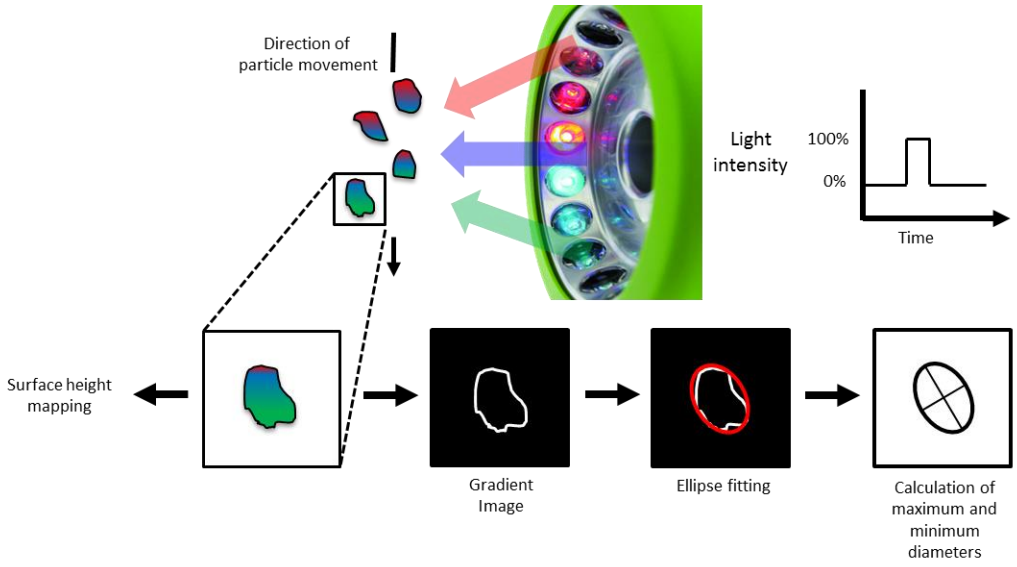


Figure 7 – Working principle of the Eyecon™ equipment.

3.3 Materials and methods

3.3.1 Materials

Six different batches of particles were used in this study. Three of them were granules of different sizes prepared in a laboratory-scale fluid bed granulator (GPCG 1, Glatt, Binzen, Germany). These batches consisted of 700 g of dextrose monohydrate (Roquette Frères, Lestrem, France) and 277.5 g of unmodified maize starch (Cargill Benelux, Sas van Gent, The Netherlands) and were granulated with an aqueous binder solution of 20 g HPMC (Dow Chemical Company, Plaquemine-LA, USA) and 2.5 g Tween 20 (Croda Chemicals Europe, Wilton, United Kingdom), sprayed as a 4% (w/w) solution. The three granulations were performed varying the process parameters: inlet air temperature during the spraying phase, spray rate and inlet air temperature during the drying phase, in order to obtain batches with different granule sizes (Table 4). The remaining three batches consisted of commercially available microcrystalline cellulose spherical pellets of different sizes commonly known as Cellets™ (Cellets™, Pharmatrans Sanaq Pharmaceuticals, Basel, Switzerland). The selected pellet sizes were Cellets™ 350 (350 to 500 µm), Cellets™ 500 (500 to 710 µm) and Cellets™ 1000 (1000 to 1400 µm). Images of the three granule and three pellet batches were obtained using the FS3D™ equipment (Figure 8).

Table 4 – Process parameters varied on the performed granulations.

	Inlet air temperature during the spraying phase (°C)	Spray rate (rpm)	Inlet temperature during the drying phase (°C)	Median particle size (D_{50}) obtained with Mastersizer™ S (µm)
b1	30	22	50	193
b2	30	36	70	383
b3	40	29	60	207

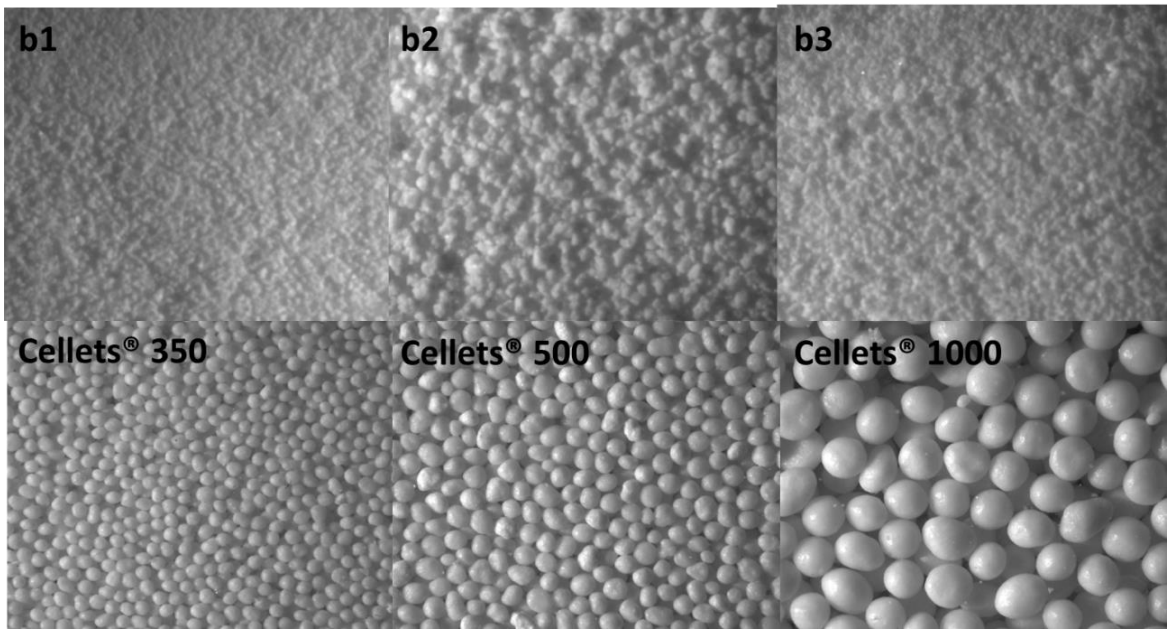


Figure 8 – Pictures of the assayed granules and Cellet™ batches taken with the FS3D™ equipment.

3.3.2 Methods

3.3.2.1 Off-line methods

Malvern Mastersizer™ S

Samples from each batch were measured twice using the LD equipment (Mastersizer™ S long bench, Malvern Instruments, Malvern, UK) by means of three different methods: dry dispersion, wet dispersion, and free fall experiments. In all cases, the 1000F lens was utilized, the particle size analysis of each sample was performed using 10000 sweeps and the obtained particle obscuration was comprehended between 10 and 30%.

In the dry method, the MS-64 sample dispersion analyzer was utilized. Two air stream pressures to aid with sample dispersion were tested (1 and 3 bars). For the wet method, Miglyol 812 (Fagron, Capelle aan den IJssel, The Netherlands) was chosen as dispersant given that both types of particles are insoluble in this liquid. The diluted dispersion was recirculated from the small volume sample adaptor through a flow cell using a peristaltic pump. In the free fall method, the sample was fed by a vibratory feeder (DR100, Retsch, Haan, Germany) to an in-house free fall controlled flow unit. At the end of this unit a vacuum cleaner (GS80, Nilfisk, Brøndby, Denmark) was placed to collect the sample and so avoid repeated measurements of the same particles.

PSD was estimated using standard reference indexes and by way of an algorithm based on Mie's theory, provided with the diffractometer. The granule batches were analyzed as polydisperse and pellet batches as monomodal. When measuring the pellet particles with sizes around 1000 μm and larger (Cellets™ 100) via the wet dispersion method, the background signal was already very high due to the combination of the lens and the use of Miglyol as the dispersant. In order to overcome this experimental difficulty, the signal at the first channels was discarded with minimal influence in the accuracy of the results.

Sieve Analysis

Sieve analysis was performed, in triplicate, on 20 g of sample from each batch. Nine sieves with mesh sizes of 2000, 1400, 1000, 500, 315, 250, 180, 100 and 50 μm were stacked. A collector pan was placed below the sieve with the smallest mesh size. The samples were placed on the top sieve (2000 μm) and a lid was placed on it. The assembly was vibrated on an automatic sieve shaker (VE 1000, Retsch, Haan, Germany) for 5 min with an amplitude of 2 mm. Such gentle conditions were chosen to prevent breakage of the granule samples. After shaking each sieve was weighted individually and the mass percentage of material retained on each sieve was calculated.

3.3.2.2 In-process methods

Parsum™ IPP70

The six different particle batches were fed to the measurement zone of the SFV probe (Parsum™ IPP70; Gesellschaft für Partikel-, Strömungs- und Umweltmesstechnik, Chemnitz, Gemany) by means of a vibratory feeder (DR100, Retsch, Haan, Germany) to simulate in-line measurements of a process's particle flow stream. The disperser accessory with a ring injector (D23) was employed during the size determination of the three granule batches in order to facilitate the measurement of the smaller particles. This cell was operated with an internal (15 L/min) and external (3 L/min) pressurized air connection. The open flow cell (SZ11) was used for the measurement of the pellet batches, working with a pressurized air stream of 4 L/min. SFV measurements were taken every 10 s for a period of 5 min. Six replicate measurements were performed for each batch, on six different days. After the analysis of the acquired size distributions and D values, it was found necessary to perform an extra series of measurements where the particle loadings were controlled (see section 5.3.3 Results and discussion). Six replicate measurements of each batch were performed. Data was acquired every 10 s for a period of 5 min.

FBRM™ C35

A small quantity of granules (2.5 g) or pellets (3.5 g) was added to a beaker containing 40 ml of Miglyol 812 (Fagron, Capelle aan den IJssel, The Netherlands). The FBRM probe (FBRM™ C35, Mettler-Toledo AutoChem, Columbia, MD, United States) was immersed in the suspension at an angle of approximately 45˚ to allow optimum sample presentation [185]. A magnetic stirrer was used to gently agitate the suspension without breaking down the granules. FBRM measurements were performed every 10 s, during a period of 3 min. The six batches were measured in triplicate, on three different days. The size information was extracted through the iC FBRM™ 4.0 software (Mettler-Toledo AutoChem Inc., Columbia, MD, United States).

FlashSizer™ 3D

The particle size of the six batches was assessed six times on six different days. The samples were filled into a petridish and positioned on top of the imaging instrument's

(Flashsizer 3D™, FS3D™, iPAT Ltd., Turku, Finland) glass window. Two digital images were captured during each measurement and combined to obtain a 3D surface from which the relevant particle size information was calculated.

Eyecon™

The Eyecon™ 2D and 3D particle imager (Eyecon™, Innopharma Labs™, Dublin, Ireland) was used off-line. As particle movement during the light pulse is negligible, in theory, results obtained off-line will be similar as when the measurements were performed in-process. One sample was collected from each batch and placed on a petri dish. Twenty images were taken from each individual sample, and an average of the PSD parameters of interest was calculated.

3.3.2.3 Acquired particle size parameters

As mentioned previously, from sieve analysis, mass-based sieve fractions were obtained. On the other hand, FS3D™, Parsum™ IPP70 and FBRM™ C35 equipments allowed the definition of size or chord length classes. Sieve fractions were first chosen for sieve analysis and afterwards the same sieve sizes were introduced in each instrument's software (with the exception of Eyecon™ where this feature is not available). The selected sieves (size classes) 50 µm, 100 µm, 180 µm, 250 µm, 315 µm, 500 µm, 715 µm, 1000 µm, 1400 µm and the related D values D_{10}, D_{50} and D_{90} were acquired from each individual instrument's software. For sieve analysis, D values were calculated from the sieve distributions by linear interpolation of the obtained cumulative mass percentage curve, while for the other techniques, the instrument's software directly provided these results.

3.4 Results and discussion

The comparison between the particle size information obtained from the different studied particle sizing techniques is not straightforward since each technique is unique in its way of measuring and calculating size. Not only the underlying measurement method might have an influence on the acquired particle size (or chord length), but also the algorithms used for obtaining particle size information from the measured particle properties and the way in which size results are presented are essential. Sieving results are mass-weighed, while

volume-weighed data are obtained from Mastersizer™ S and FS3D™. Eyecon™'s D values are based on relative mass values. The chord length measurements performed with Parsum™ IPP70 are converted by the instrument's software to a volume-weighed particle size distribution. Square-weighted chord length distributions were obtained through FBRM™ C35 as it has been previously described by Heath et al. (2002) [163] that this data presents the best agreement with the particle size distribution obtained from LD measurements. FBRM measures the first diameter weighing of the chord distribution and it is then effectively a cube (volume) weighing which is comparable to the volume-based distribution obtained from laser diffraction. The sieve size distributions of each batch are represented in Figure 9a to f and the related D values in Figure 10.

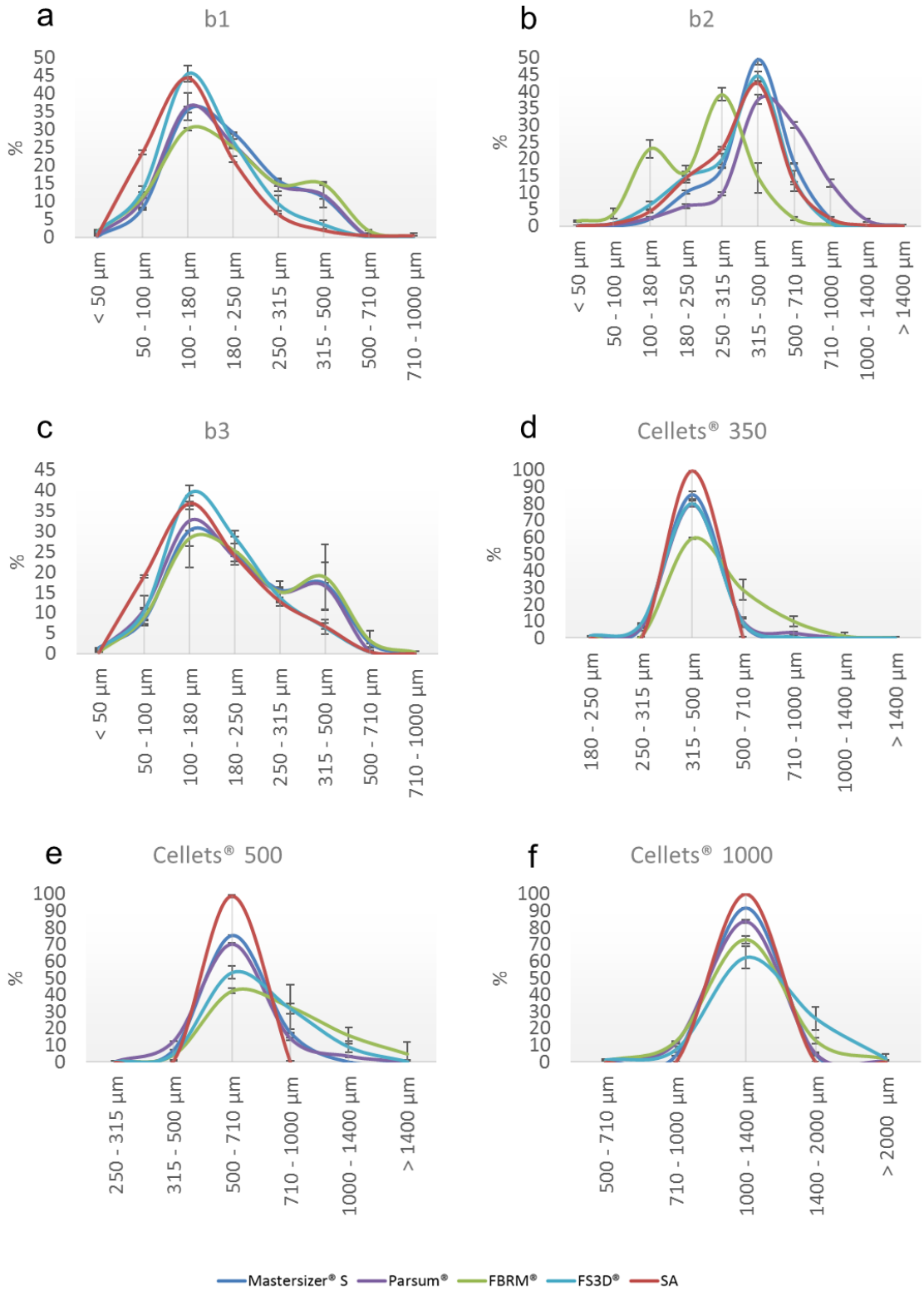


Figure 9 – Particle size distributions of the assayed batches obtained with the different equipment.

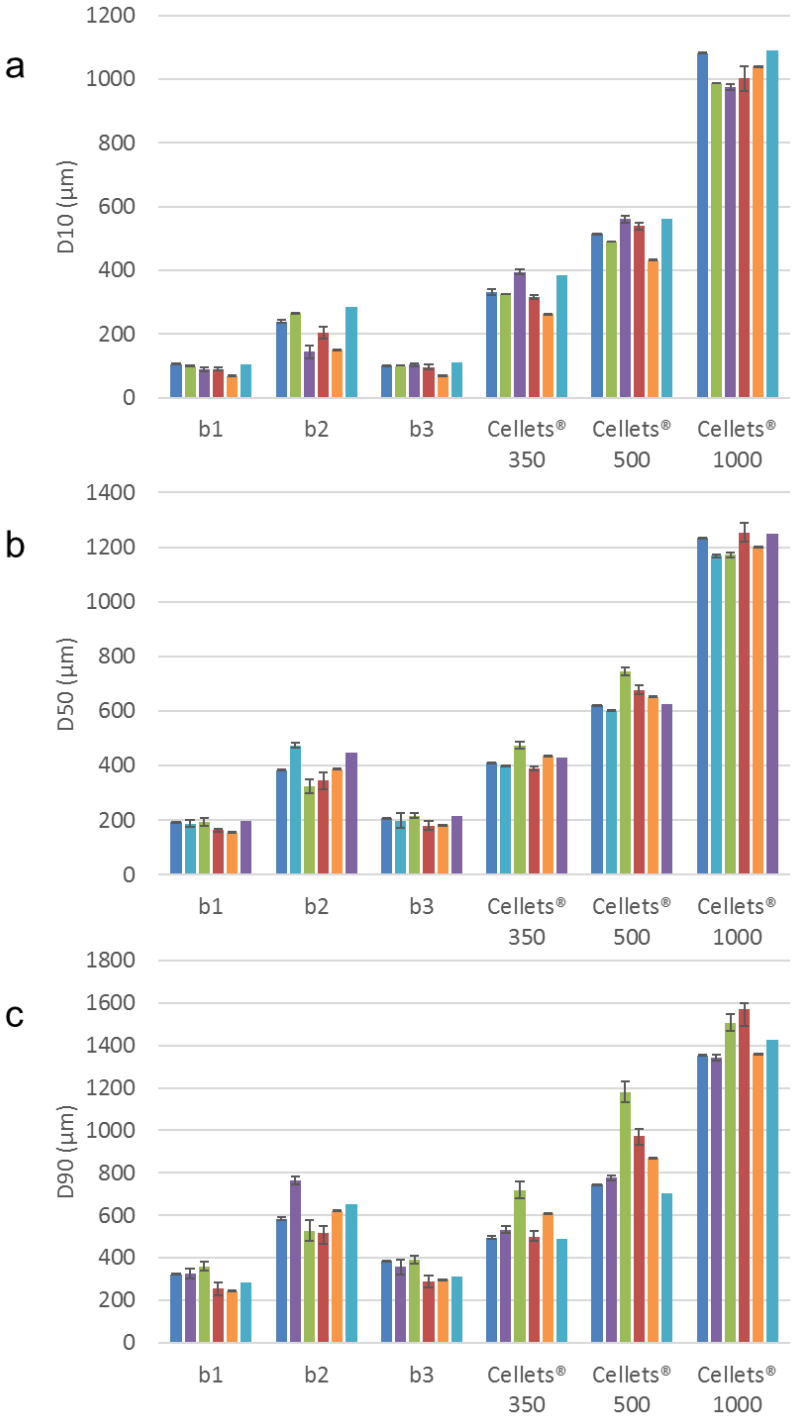


Figure 10 – D values of the assayed batches obtained with the different methods.

3.4.1 Choice of the reference method (Mastersizer™ S)

Laser diffraction is often utilized as a reference for comparison with other sizing techniques [153, 186] . Figure 11 displays the D_{50} value of the six studied batches, obtained with Mastersizer™ S using the different measurement methods.

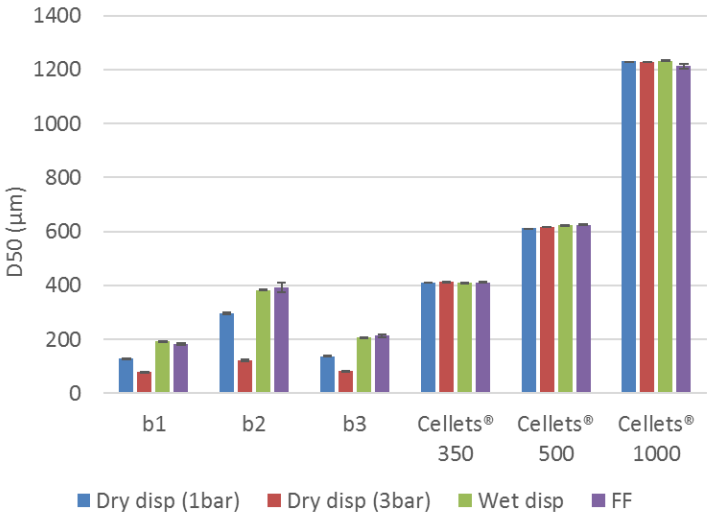


Figure 11 – Comparison between the D_{50} values of the assayed batches obtained with the Mastersizer™ S equipment by means of the dry dispersion (Dry disp), wet dispersion (Wet disp) and free fall (FF) methods.

It was observed that particle sizes obtained from dry dispersion measurements of the granule batches were significantly smaller than the ones attained from the wet dispersion and free fall methods. When air pressure was augmented from one to three bar, this reduction in size was more pronounced. Therefore, we believe that the use of pressurized air as the dispersing agent damaged the fragile granules. In contrast, for the pellet batches (particles with a low friability) no meaningful size differences were observed between all methods. Measurements in air dispersion and free fall presented similar results for all batches. All in all, the results obtained for all batches for wet dispersion, and free fall experiments are very similar but the wet dispersion method presented a better precision. For that reason, Mastersizer™ S wet dispersion method was used as the reference sizing technique for the rest of this study.

3.4.2 In process methods and sieve analysis versus reference method

3.4.2.1 Parsum™ IPP70

Both Mastersizer™ S and Parsum™ IPP70 present their results as volume-based particle size distributions. Nevertheless, it is important to always keep in mind that, differently from Mastersizer™ S, Parsum™ IPP70 does not measure particle size but chord, converting it afterwards into a PSD by means of the instrument's software. The comparison between the D values obtained from the initial Parsum™ IPP70 measurements and from the Mastersizer™ S measurements is depicted in Figure 12.

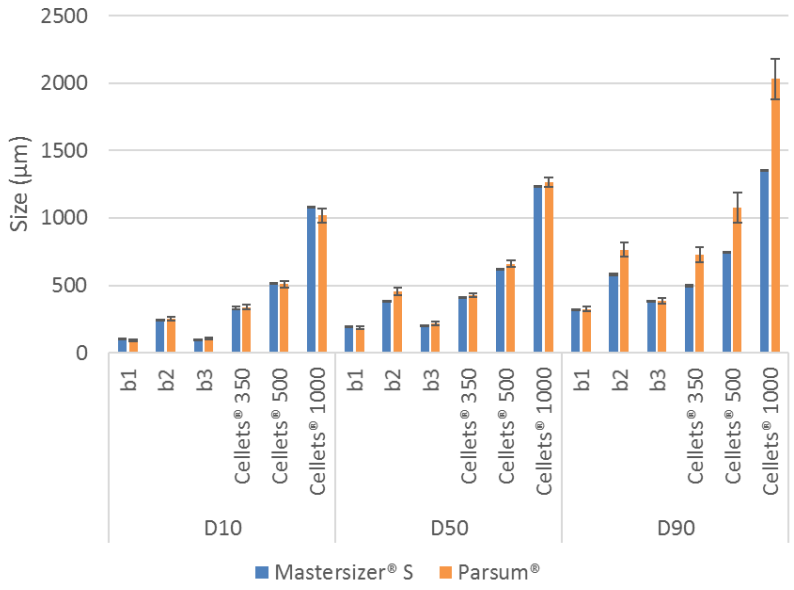


Figure 12 – D values of the assayed batches obtained with Mastersizer™ S and Parsum™ IPP70.

The D values attained with both techniques, Parsum™ IPP70 and Mastersizer™ S, for granule batches b1 and b3 (the batches with the smallest granule size) are perfectly in agreement. Differently, for batch b2 particularly the obtained D_{90} value was larger for Parsum™ IPP70 than for Mastersizer™ S. Looking into the results for pellet batches, it is observed that the D_{90} values are also larger when measured with Parsum™ IPP70 and that this difference is more obvious with a larger particle size. All things considered the D_{90} values obtained with Parsum™ IPP70 are mostly larger than the values obtained with Mastersizer™ S which strongly suggests that these results could be a result of particle coincidence. If

occasionally two particles cross the laser beam at the same time, the Parsum™ IPP70 probe is not able to make the distinction between the two shadows of the different particles hence detecting as one large particle. Therefore, a falsely increased number of larger particles might be reported hence shifting the distribution towards larger sizes. The average particle loadings registered for this initial set of measurements (Figure 12) are depicted in the first column of Table 5. It is observed that there are high loadings, which means that a high percentage of the measuring volume is at any time occupied with particles and, as these particles might overlap, the probability of errors by coincidence of particles is higher. To prove that high particle loadings might cause size overestimation in this system, two different measurements of each batch with a different average particle loading were performed. The different loadings per batch are depicted in the second column of Table 5. The obtained corresponding D values are portrayed in Figures 11a and 11b. Batches b1 and b3 did not reveal a significant influence due to the increase in particle loading (Figure 13).

Table 5 – Average particle loadings for each batch in the performed Parsum™ IPP70 measurements.

	Average loadings of the first set of measurements (%)	**Low/high average loadings of the repeated measurements (%)**
b1	6.84	3.65/11.38
b2	6.79	2.94/9.76
b3	7.09	2.6/9.05
Cellets™ 350	16.57	3.1/13.7
Cellets™ 500	16.09	2.6/16.71
Cellets™ 1000	17.7	2.2/17.43

However, as the size of particles increases (batch b2 and pellet batches, Figures 11a and 11b), and with the increase of particle loading, the overestimation becomes more noticeable, especially for the D_{90} values. Particle coincidence due to high loadings is a complication which needs to be taken into account when performing measurements with the Parsum™ IPP70 system. In the instrument's software, several settings can be altered in order to prevent this type of errors. For instance, a maximum loading level can be established and if at a certain time the particle loading exceeds the set maximum value, these data are not recorded. Also, a search for coincidence and removal can be activated by defining a coincidence level. This coincidence level is a user-set percentage of the highest size class of

the acquired number distribution and, notionally, it should be as low as possible in order to achieve the maximum sensitivity.

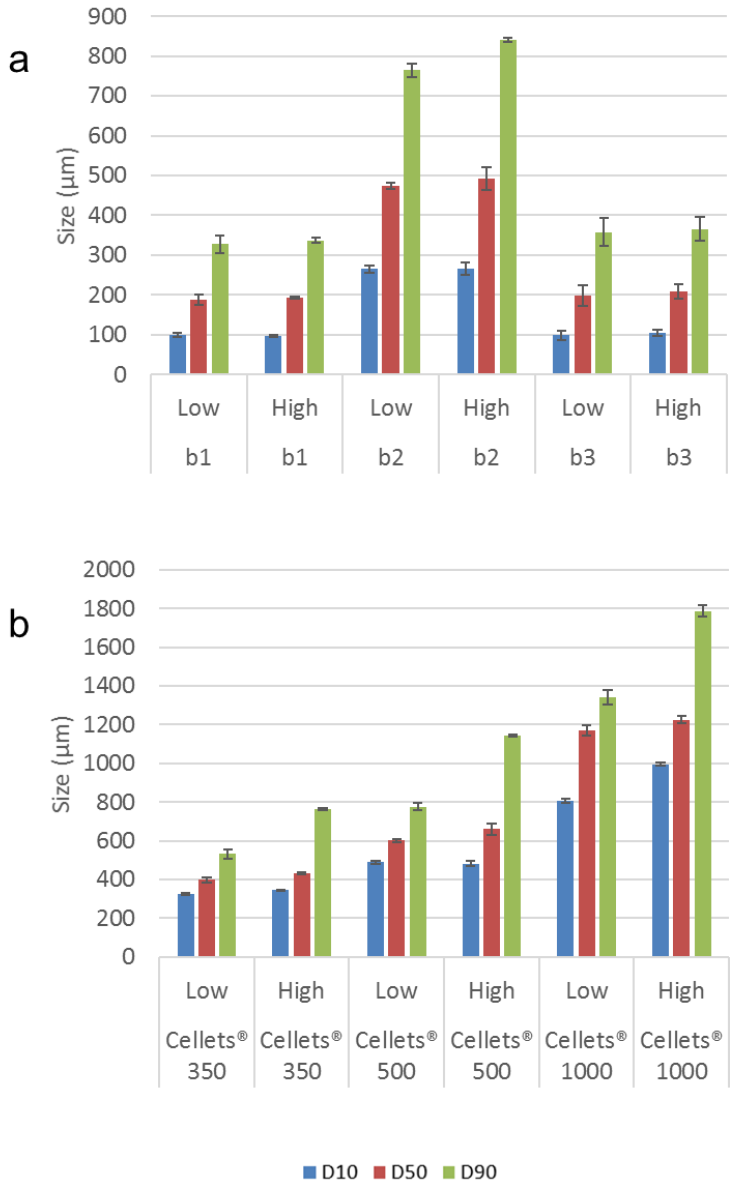


Figure 13 – D values of the Parsum™ IPP70 measurements performed at different particle loadings (see Table 5 for further information on loadings).

Concerning the Parsum™ IPP70 equipment itself, especially in processes with small particles and high particle loadings, the use of the in-line disperser D23 is important and recommended in order to keep the loading low enough avoiding coincidence errors. Finally, Parsum™ IPP70 and Mastersizer™ S were compared utilizing the data acquired with the

lowest recorded particle loading (Figure 9a to f and Figure 10a to c) to discard the differences that could arise from coincidence. Batches b1 and b3 presented similar size distributions when measured with both instruments (Figure 9 a and c) and, hence, also very similar D values (Figure 10a to c). Regarding batch b2, all obtained Parsum™ IPP70 D values were larger compared to the corresponding Mastersizer™ S values (Figure 10a to c). In the size distribution, a shift of the Parsum™ IPP70 distribution towards larger sizes is observed when comparing to the Mastersizer™ S's distribution (Figure 9b). This shift is most probably related to a sampling difficulty rather than to an instrumental dissimilarity. Batch b2 presented a very broad size distribution and it was possible to observe that the larger particles were the first to be directed from the feeder to the probe's measurement zone and into the aluminum tray, while some of the batch's fines were retained in the feeder not being measured. Concerning Cellets™ 350 the D values (Figure 10a to c) and size distributions obtained with Parsum™ IPP70 and Mastersizer™ S (Figure 9d) are similar. Regarding Cellets™ 500 and Cellets™ 1000 size distributions from Parsum™ IPP70 and Mastersizer™ S are identical (Figure 9e and f) as are the D_{90} values (Figure 10c). However, D_{10} and D_{50} are slightly larger for Mastersizer™ S (Figure 10a and b), specially for Cellets™ 1000.

3.4.2.2 FBRM™ C35

The size distribution curves of the batches with the smallest granules, b1 and b3, obtained with Mastersizer™ S (volumetric particle size) and with FBRM™ C35 (square weighted chord length) (Figure 9a and c) and the D values obtained from both techniques are very similar (Figure 10a to c). This was expected as volume-based PSDs measured by LD, and square-weighted CLDs are predicted to be in good agreement for spherical particles [163]. However, in comparison to Mastersizer™ S, the PSD measured using FBRM™ C35 of the granule batch b2 is smaller (Figure 9b) as are the D values (Figure 10a to c). When looking at the obtained particle size distributions (Figure 9b), a major shift towards smaller sizes is visible. The agitation needed and used to keep the particles in suspension (see materials and methods) might have been responsible for their breakage. Pellet batches Cellets™ 350 and Cellets™ 500 were overestimated by FBRM™ C35 when comparing to Mastersizer™ S (Figure 9d and e) resulting in larger D values, particularly D_{90} (Figure 10a to c). It is possible that when performing the measurements for this batches particles were too close together causing a phenomenon of chord concatenation i.e. an error by coincidence of particles. This type of error occurs when two particles cross the laser beam so close together

that the analyzer cannot individualize them and counts two particles as one. For batch Cellets™ 1000, the distributions from both analyzers are in good agreement though, as expected, the square-weighted CLD obtained from FBRM™ C35 is wider than the volume-based PSD acquired with Mastersizer™ S (Figure 9f) resulting in a smaller D_{10} and a larger D^{90} (Figure 10a and c). One must always keep in mind that no perfect correspondence between an unweighted or weighted CLD and PSD is achievable and converting from CLD to PSD (as in the case of the Parsum™ IPP70 equipment) by assuming a certain particle shape might be a suitable solution for simultaneous comparison of results from instruments that measure chord length with instruments that determine particle size.

3.4.2.3 Flashsizer™ 3D

Both FS3D™ and Mastersizer™ S present their results as volume-based PSDs though they possess distinct underlying methods for calculating the particle size. In comparison to the Mastersizer™ S results, the particle size of granule batches measured with FS3D™ was underestimated (Figure 9a to c), presenting slightly smaller D_{10}, D_{50} and D_{90} values (Figure 10a to c). The irregularities of the particles that present a rough surface (as is the case for the granules), may cast shades which are interpreted by the instrument as being the edges between two particles, consequently causing particle size underestimation. This phenomenon has been previously described [184]. Another possible explanation for the observed underestimation is that, when there is a relatively large amount of fines, during the sampling procedure these can cover the measurement window preventing the measurement of the larger particles [187]. Cellets™ 350 size distributions (Figure 9 d) and D values (Figure 10a to c) for both instruments were in good agreement but, for the Cellets™500 and Cellets™ 1000, an overestimation of the larger particles was observed in comparison with Mastersizer™ S (Figure 9 e and f) with significantly larger D_{90} values obtained for FS3D™ (Figure 10a to c). An explanation for this phenomenon is that as particle size increases fewer, and fewer particles are measured per image, and statistically, a reduced number of large particles can greatly influence a volume-based particle size distribution shifting it towards the larger end of the distribution.

3.4.2.4 Eyecon™

Eyecon™ and Mastersizer™ S use different measurement principles and algorithms for the calculation of particle size. From Eyecon™, it is possible to extract a number-based size distribution (number of particles versus the average diameter) divided into several classes. However, the user cannot set these classes himself and, consequently, it was not possible to plot the results altogether with the size distributions obtained with the other equipments. Therefore, the comparison between Eyecon™ and Mastersizer™ S is established regarding the acquired D values (Figure 10a to c). It is difficult to affirm if the acquired D values are significantly different, as the Eyecon™ does not allow the presentation of the standard deviation values between the several measurements of a single sample. Nevertheless, assuming that the standard deviation between measurements with this technique is approximately as large as the standard deviations obtained with the other techniques, it is possible to observe a good agreement between the D_{10} and D_{50} values measured for the smaller granule batches b1 and b3 (Figure 10a and b). Differently, the D_{90} values obtained with Eyecon™ are slightly smaller than the ones obtained with Mastersizer™ S (Figure 10c). On the other hand, the granule batch with the largest size, batch b2, had its size overestimated by Eyecon™ (Figure 9b) presenting larger D_{10}, D_{50} and D_{90} values (Figure 10a to c). These differences of granule size may arise from the fact that, if particles are not conveniently separated, Eyecon™ is not capable of individualizing them and provides erroneous results. This demonstrates the importance of a good sampling procedure. Granules need to be efficiently separated to allow the correct identification of the individual particles, hence avoiding errors during the measurement, especially during off-line measurements. During on-line measurements, the distance created between the particles due to their movement should be enough for an accurate detection. In contrast, as pellets are spherical particles with a smooth surface Eyecon™ easily can identify the well-defined particle edges successfully individualizing them. The D values for the pellet batches (

Figure 10Figure 10a to c) obtained with Eyecon™ and Mastersizer™ S revealed a good agreement and are comparable.

3.4.2.5 Sieve Analysis

Sieve analysis is a mass-based technique while Mastersizer™ S is volume-based. Assuming that all particles have the same density, there should not be significant differences

between the size data obtained with both techniques. When looking at both size distributions (Figure 9a to f) and D values (Figure 10a to f) resulting from sieve analysis, a lack of consistency emerges, especially for pellet batches. As can be observed in the size distributions from pellet batches (Figure 9 d to f) those obtained with the Mastersizer™ S are broader than those resulting from sieve analysis and, therefore, larger D_{10} and smaller D_{90} values were expected for sieve analysis in comparison to Mastersizer™ S, but this was not the case (Figure 10a and c). Commonly, D values are estimated from sieve analysis mass percentage distributions by linear regression. This method is not the most adequate as these distributions do not present a linear profile. Given this, it may become difficult to obtain reliable estimates of D values located in the left (D_{10}) or right (D_{90}) ends of the distribution curve. This is also why, from the three interpolated D values, D_{50} values appear to be the most reliable and, therefore the only one that will be discussed. In comparison to the results provided by Mastersizer™ S, the size distributions of granule batches b1 and b3 obtained with sieve analysis shift towards smaller sizes (Figure 9a and c) and, in accordance, the obtained D_{50} values are also smaller (Figure 10b). We believe that this might be a result of the erosion of granules during the analysis, even though gentle conditions were used to prevent granule breakage. Sieve analysis is not a method fit for fragile particles as the friction generated during the analysis may deteriorate the sample. Batch b2 was not as similarly affected by erosion and thus the shift towards the left end of the distribution is less evident than for the other granule batches (Figure 9b) and, also the obtained D_{50} value is not significantly different (Figure 10b). On the distributions obtained for pellet batches no shift is perceived when comparing the distributions (Figure 9d to f) though D_{50} differ slightly (Figure 10b).

3.5 Conclusions

In this work, several in-process particle sizing methods and two of the most commonly used off-line methods (sieve analysis and laser diffraction) were reviewed and compared. At first, the differences between all methods were explored theoretically and a table was made in order to facilitate the comparison between the different assessed methods. Further on, the particle size of three batches of granules and three different types of pellets was measured. The laser diffraction Mastersizer™ S (wet dispersion method) was utilized as the reference technique. Significant dissimilarities in the measured particle size were observed when comparing all the assayed techniques with the reference method. These differences were

elucidated taking into account previous knowledge about the assessed instruments and aimed to simplify the not forthright task of comparing particle size information from different instruments, exposing the reasons for the observed differences rather than finding a correlation between obtained results. The two types of particles that were tested were either homogenous and nearly perfectly spherical (pellets) or porous, almost perfectly spherical aggregates (obtained via a wet granulation method). The effect of particle shape on the estimation of particle size distribution is also of great interest and should be focus of attention if the future [188-191].

Chapter 4

Statistical process control of cocrystallization processes: a comparison between OPLS and PLS[2]

4.1 Abstract

Orthogonal partial least squares regression (OPLS) is being increasingly adopted as an alternative to PLS regression due to the better generalization that can be achieved. Particularly in multivariate BSPC, the use of OPLS for estimating nominal trajectories is advantageous. In OPLS, the nominal process trajectories are expected to be captured in a single predictive principal component while uncorrelated variations are filtered out to orthogonal principal components. In theory, OPLS will yield a better estimation of the Hotelling's T^2 statistic and corresponding control limits thus lowering the number of false positives and false negatives when assessing the process disturbances. Although OPLS advantages have been demonstrated in the context of regression, its use on BSPC was seldom reported. This study proposes an OPLS-based approach for BSPC of a cocrystallization process between hydrochlorothiazide and *p*-aminobenzoic acid monitored

[2] This chapter has been adapted from: A.F.T. Silva, M.C. Sarraguça, P.R. Ribeiro, A.O. Santos, T. de Beer, J.A. Lopes, Statistical process control of cocrystallization processes: a comparison between OPLS and PLS, International Journal of Pharmaceutics (Accepted manuscript, http://dx.doi.org/10.1016/j.ijpharm.2017.01.052)

on-line with near infrared spectroscopy and compares the fault detection performance with the same approach based on PLS. A series of cocrystallization batches with imposed disturbances were used to test the ability to detect abnormal situations by OPLS and PLS-based BSPC methods. Results demonstrated that OPLS was generally superior in terms of sensibility and specificity in most situations. In some abnormal batches, it was found that the imposed disturbances were only detected with OPLS.

4.2 Introduction

The industrial production of drug substances and dug products is very often achieved as a sequence of batch operations [192]. Statistical process monitoring proposes a strategy to detect deviations from normal process trajectories. It was firstly developed for continuous processes operating in steady state [192]. Batch processes present challenges for modelling and monitoring due to time changing dynamics, variable duration, non-linear intrinsic nature and batch-to-batch variability [49, 192]. Latent variable methods such as PCA and PLS, applied in the context of batch statistical monitoring, provide a possible way to overcome some of the difficulties. To establish a multivariate BSPC scheme, PCA or PLS methods rely on batches produced under normal operating conditions (NOC). A BSPC model structure is optimized projecting unseen NOC and abnormal batches. The goal is to ensure that the BSPC model can effectively detect abnormal conditions in non-NOC batches (true positives) while preventing NOC batches to be signaled as faulty batches (false positives). Optimization of these models relies on multiple factors: appropriate selection and pre-processing of process variables, effective selection of the number of latent variables and optimized definition of control limits and models' tolerance. Multivariate statistical control charts based on Hotelling's T^2 and squared residuals (Q) are often used to monitor the process [193].

Batch data can usually be arranged in a tree-way array encompassing the batch, process variable and time dimensions. Multiple methods can be used to analyze three-way arrays without the need of unfolding (conversion to two-way matrices). Parallel factor analysis (PARAFAC) [194] or n-way partial least squares regression (N-PLS) are two methods designed to handle n-way data [195]. However, to analyze data with PCA or PLS, the three-way matrix needs to be unfolded [196]. Unfolding can be done batch-wise (BWU), preserving the batch direction. Information of one batch is contained in a single row of the resulting matrix [73]. Alternatively, a second type of unfolding, called observation-wise unfolding (OWU), in which the variable direction is preserved, can be used. For this purpose, several

methods are available [77, 197-199]. There is an important difference between these two types of unfolding methods. In OWU, the PLS method using batch time as the response (Y-block) will yield components that best approximate the average trajectories. In BWU, the PLS method using a quality attribute of the final product as Y-block will yield components that best capture the variability among batches. In BSPC, the components of a PLS-OWU model are rotated to maximize their relationship to batch time or maturity. In PLS-OWU models, the systematic variation in the predictors (X-block) that is not related with the batch time can impair the interpretation of the model [73].

An alternative to PLS regression is the orthogonal partial least squares (OPLS) method [53]. OPLS is an adaptation of the orthogonal signal correction (OSC) method [200] and differs from PLS by removing systematic variation in **X** that is orthogonal to **Y**. The advantage is to reduce models' complexity (by concentrating the predictive information in one component) and increase interpretability due to the decreasing of confounding effects stored in model components [201].

Since it was first proposed in 2002 [53] OPLS and its extensions (e.g. OPLS-Discriminant Analysis, K-OPLS, etc.), have been utilized in distinct contexts for multiple applications [201-204]. On the other hand, a single application of OPLS for BSPC has been reported so far [73]. The study in question utilized OPLS for modeling a batch chemical hydrogenation process comparing it with PLS and PCA. OPLS demonstrated a superior ability to detect deviations and provided an easier root cause analysis for these deviations.

Pharmaceutical cocrystals are defined as a multicomponent crystalline structure formed by one drug substance and one or more coformers. These compounds have enhanced pharmaceutical properties such as solubility, bioavailability, stability, among others [205]. Cocrystallizations are typically operated in batch-mode even at the industrial scale. The ability to monitor on-line the cocrystallization process is relevant to ensure that the final product is consistently delivered within target specifications [93]. This is especially important, giving the increasingly adopted QbD paradigm for pharmaceutical processes development framed by the ICH-Q8 guideline [10].

Previously, PCA-OWU was reported as a BSPC strategy to monitor on-line the cocrystallization between furosemide and nicotinamide using near infrared spectroscopy (NIRS) [125]. In this work, a cocrystallization process between hydrochlorothiazide (HTZ) and *p*-aminobenzoic acid (PABA) by solvent evaporation was monitored with NIRS. The major aim is to investigate whether OPLS is consistently more adequate for BSPC than PLS.

4.3 Experimental

4.3.1 Materials and methods

HTZ (>98% purity), PABA (>99.5 % purity) and methanol (>99.5% purity) were acquired from Sigma-Aldrich (St. Louis, MO, USA).

Cocrystallization of HTZ and PABA (Figure 14) was performed by the solvent evaporation method using methanol. The process consists on weighting HTZ and PABA followed by dissolution in methanol at room temperature. The solution is then stirred at 150 rpm, in an orbital stirring table during 10 h also at room temperature. A total of nine batches were designed: five nominal and four abnormal batches. Nominal batches use previously optimized conditions for this cocrystallization [125]. From the nominal batches, four were chosen to calibrate models (B#1 to B#4) and one was used for testing (B#5). The abnormal batches were designed by imposing multiple disturbances as described in Table 6 (B#8 to B#11).

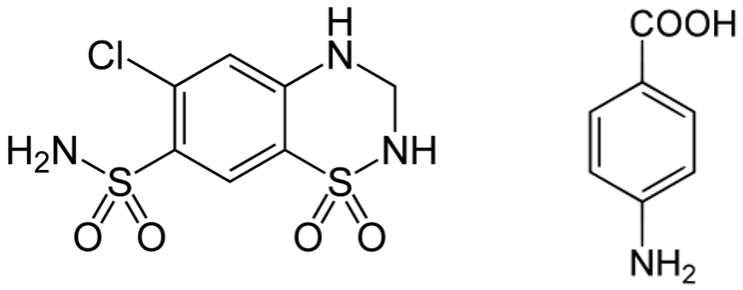


Figure 14 – Structure of a) hydrochlorothiazide (HTZ) and b) *p*-aminobenzoic acid (PABA).

Table 6 – Summary of the produced cocrystallization batches.

	Batches #	Methanol (ml)	HTZ (mg)	PABA (mg)	HTZ: PABA (molar ratio)	Type of batch	Formation of the cocrystal
Calibration	B#1 to #4	20	416.42	383.58	1:2	Nominal	Yes
Test	B#5	20	416.42	383.58	1:2	Nominal	Yes
	B#8	20	416.42	191.79	1:1	Non-nominal	No
	B#9	15	416.42	383.58	1:2	Non-nominal	No
	B#10	20	832.84	575.37	2:3	Non-nominal	No
	B#11	40	416.42	767.16	1:4	Non-nominal	Yes

4.3.2 On-line monitoring

A Fourier transform near-infrared analyzer (FTLA2000, ABB, Québec, Canada) was used to monitor on-line the cocrystallization process. The spectrophotometer is equipped with an indium-gallium-arsenide (InGaAs) detector. The measurements were made in diffuse reflectance mode using a stainless steel diffuse reflectance probe (SabIR, ThermoNicolet, Madison, USA) with a 1 cm diameter sapphire window enabling a 0.20 cm^2 illumination area. Each spectrum was acquired with a resolution of 8 cm^{-1} over a wavenumber interval between 10000 cm^{-1} and 4000 cm^{-1}. Each stored spectrum is an average of 64 scans. The instrument is controlled via the Grams LT software (version 7, ABB, Québec, Canada). A background was made before each batch, by placing a PFTE certified material (Labsphere North Sutton, NH, USA) over the probe tip. To monitor the process, the probe was set 1 cm over the cocrystallization medium in order to avoid interference with the process. One spectrum was stored every 5 min during 10 h, totalizing 121 spectra per batch. All spectra were pre-processed with the Savitzky-Golay algorithm (29 points width filter fitted with a second-order polynomial followed by a first derivative) to reduce unwanted baseline variations and standard normal variate (SNV) to compensate for scale variations. In total, a three-way array consisting on nine batches, 1556 spectral variables and 121 time points was produced.

4.3.3 Product characterization

Crystallized products were vacuum dried over 1.5 h to remove any free residual solvent. The dried products were characterized by X-ray powder diffraction (XRPD), differential scanning calorimetry (DSC) and near infrared spectroscopy (NIRS).

4.3.3.1 X-ray powder diffraction

XRPD analyses were performed on a powder diffractometer (MINIFlexII, Rigaku, The Woodlands, USA) with Cu Ka radiation (l = 1.5418 Å) operating at 40 kV/30 mA fitted with a glass sample holder. Diffraction patterns were obtained with a 0.02º (2θ) step size and 3 s per step in the range 3–45º (2θ).

4.3.3.2 Differential scanning calorimetry

DSC measurements were performed using a thermal analyzer (DSC 200 F3 Maia™, Netzsch, GmbH, Germany) with an automatic sample changer (ASC, Netzsch, GmbH, Germany). Approximately 2-3 mg were weighed in aluminum pan and then sealed. The reference pan was left empty. Heating curves for the samples were recorded with a heating rate of 10°C min^{-1} from 25°C to 300°C. The onset temperature was calculated using the software provided by the DSC equipment (Proteus, version 6.1081, Netzsch, GmbH, Germany). For each sample, three thermograms were taken and the average onset temperature considered.

4.3.3.3 Near infrared spectroscopy

The same FT-NIR spectroscopy equipment described before was used off-line for final product characterization. The difference was the use of a powder sampling accessory (ACC101, ABB, Québec, Canada) featuring a 2 cm diameter window enabling diffuse reflectance measurements on a 0.28 cm^2 illumination area. Each spectrum was acquired with a resolution of 8 cm^{-1} as an average of 64 spectra in the wavenumber range between 10000 cm^{-1} and 4000 cm^{-1}. A background was taken by using a PTFE certified material (SKG8613G, ABB, Québec, Canada). For each sample, three spectral replicates were required and the average spectrum considered.

4.3.4 Data analysis

Pre-processed NIR spectra obtained on-line organized in a three-way matrix (batch x observation x time) was firstly unfolded in a two-way structure by means of the OWU approach. After unfolding, data from batches B#1 to B#4 yielded a two-way (4x121x1556) matrix (calibration X-block). Testing data encompassing batches B#5 (NOC) and B#8 to B#11 (abnormal) were organized in a (5x121x1556) matrix (testing X-block).

Two approaches to BSPC were tested considering different models: PLS and OPLS. The response variable was the batch time (Y-block). Spectral data (X-block) and the response (Y-block) were mean-centred prior to modelling. Hotelling's T^2 and Q control charts were used to detect the deviations from the NOC. Deviations from the Hotelling's T^2 reveal differences, which can be explained by the model. On the other hand, the *Q* residuals statistic estimates the distance to model and therefore highlights differences that cannot be explained by the modelled component. By monitoring the residuals, new unexplained disturbances, different from the ones in the model, can be detected and action can be taken [83]. Both metrics are described in section Multivariate control charts (Chapter 2).

OPLS differs from PLS because it filters the variation in *X* correlated to *Y* in a single predictive component while variation uncorrelated to *Y* is captured in the orthogonal component(s). Apart from this filtering step, PLS and OPLS models for a single *Y* and the same total number of latent variables are identical and thus the *Q* statistic for both models is the same. The control charts were normalized by dividing the Hotelling's T^2 and Q *residuals* statistics by their 95% confidence limit. This procedure will lead to a control limit equal to 1. These two control charts are complementary and together they give a representation of the system [193]. A description of both Hotelling's T^2 and Q residuals statistic is found in section Multivariate control charts (Chapter 2).

The multivariate projection models based on PLS and OPLS were generated using the SIMCA 14.1 software (MKS Data Analytics Solutions, Umeå, Sweden).

4.4 Results

4.4.1 Final products characterization

The final product obtained after each batch was characterized with XRPD, DSC and NIRS. XRPD patterns of the crystals obtained from NOC batches and a prepared physical

mixture of the initial components were analyzed to investigate the actual formation of cocrystals (Figure 15). The formation a new crystalline form (a possible cocrystal), is observed by the appearance of non-existing peaks in the physical mixture, specifically at 12.9º, 14.9º, 25.4º, 26.0º, 27.3º, and 42.0º (2θ) marked in the figure by a dashed line.

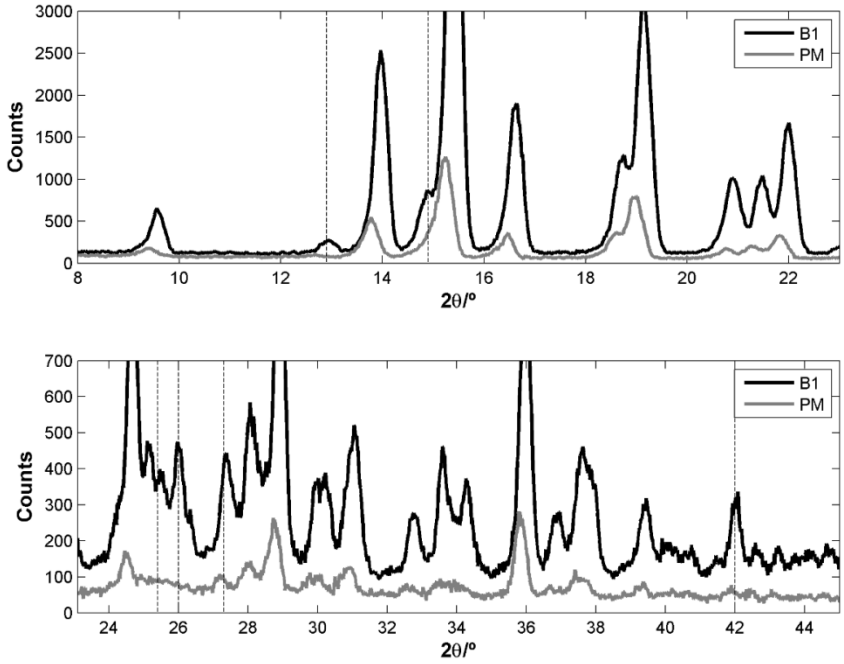


Figure 15 – XRPD patterns of the cocrystal obtained in the batch B#1 and from a physical mixture of HTZ and PABA (PM).

The proof that a single crystalline form was obtained was confirmed by DSC. The thermograms of a NOC batch final product and of the initial components are shown in Figure 16. The average onset temperature for the NOC batches final product is 175.30±0.71ºC. This value is remarkably close to that reported by Sanphui et. al. of 175.9ºC [206], and different from the onset temperature of the two initial components: 268.2±0.71ºC for HTZ and 187.1±0.23ºC for PABA.

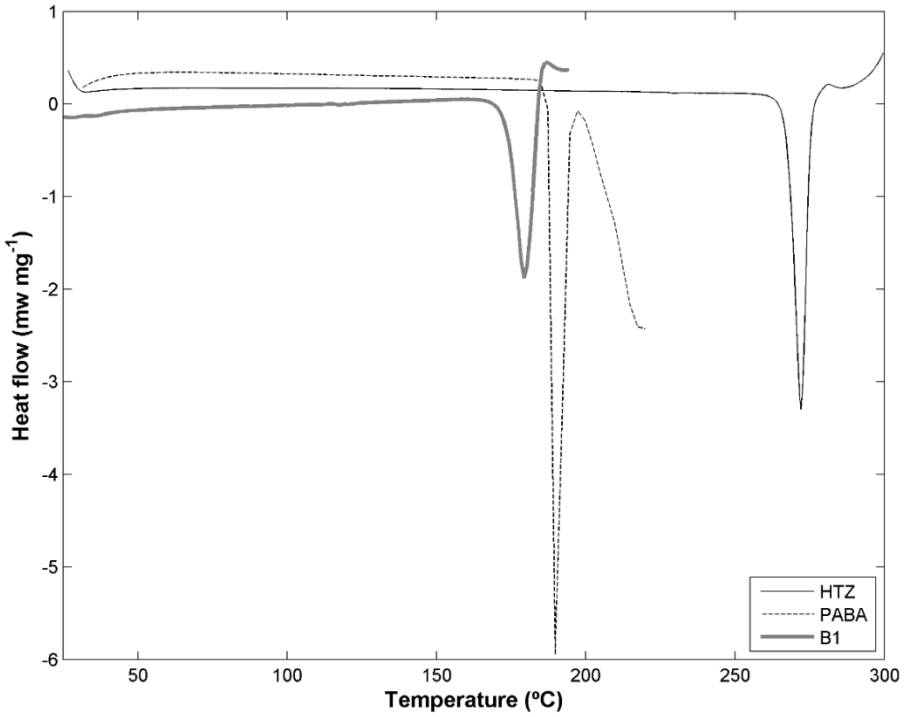


Figure 16 – DSC thermograms for the cocrystal obtained in batch B#1, pure hydrochlorothiazide (HTZ) and pure *p*-aminobenzoic acid (PABA).

The differences between NIR spectra of the crystals physical mixture and cocrystals obtained from the NOC batches are mainly visible in wavenumber regions associated with primary and secondary amides (around 6950 cm^{-1}, 6780 cm^{-1} and between 5000 cm^{-1} and 4450 cm^{-1}) (Figure 17). The wavenumber region around 5000 cm^{-1} is also associated with the carboxyl acid group. Other differences are located at 9800 cm^{-1} and 6530 cm^{-1} due to the amine group. From the NIR spectral analysis it can be established that the groups responsible by cocrystal hydrogen bonds are the sulphonamide groups of HTZ and the carboxyl and amide group of PABA as previously reported by Sanphui et al. [206].

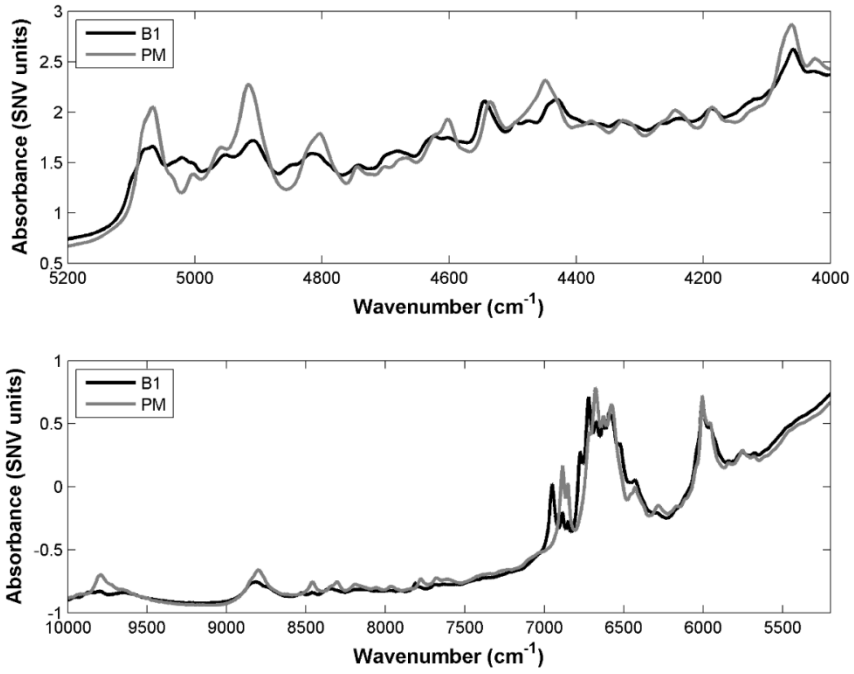


Figure 17 – NIR spectra of the cocrystal obtained in batch B#1 and from a physical mixture of HTZ and PABA (PM).

By combining the information obtained by these techniques, it can be concluded that the NOC batches produced a single crystalline phase different from the initial components, therefore a cocrystal was obtained.

These methods were additionally employed to analyze the product obtained at the end of the non-nominal batches (B#8 to B#11) revealing that batch B#11 was the only producing a cocrystal, although with an excess of PABA (supplementary material (Figure 18 to Figure 21).

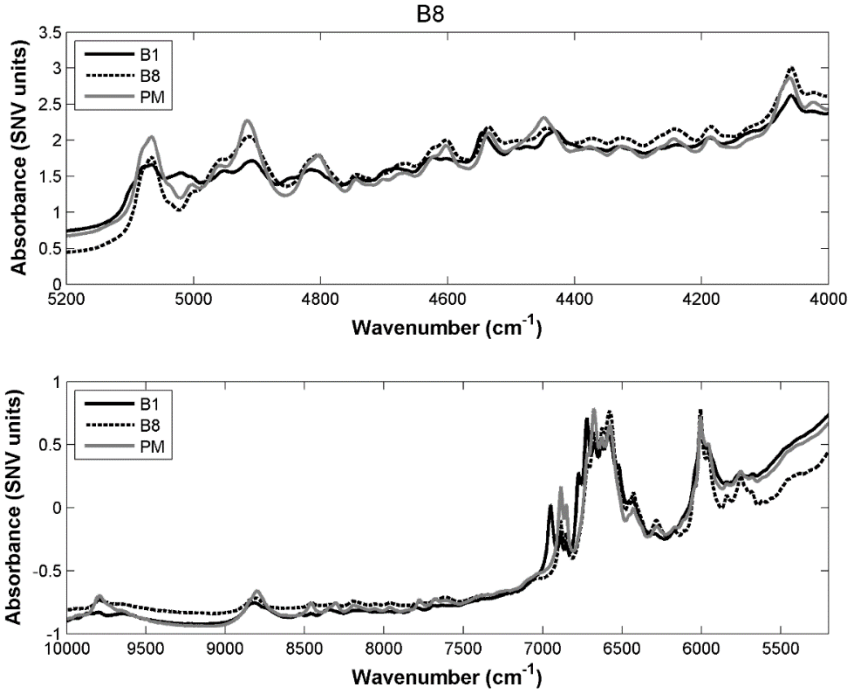


Figure 18 – NIR spectra of the cocrystal obtained in batches B#1 and B#8 and from a physical mixture of HTZ and PABA (PM).

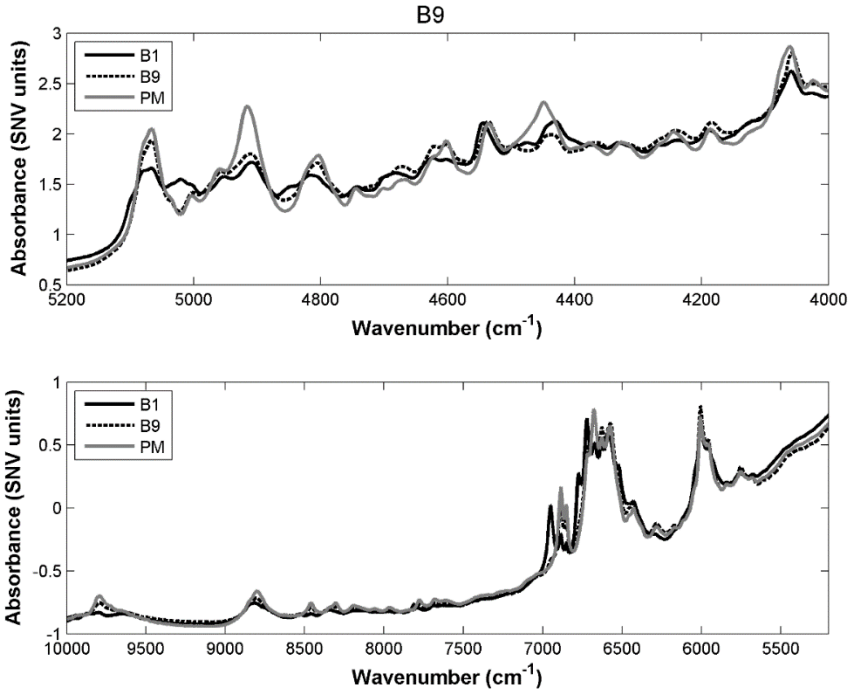


Figure 19 – NIR spectra of the cocrystal obtained in batches B#1 and B#9 and from a physical mixture of HTZ and PABA (PM).

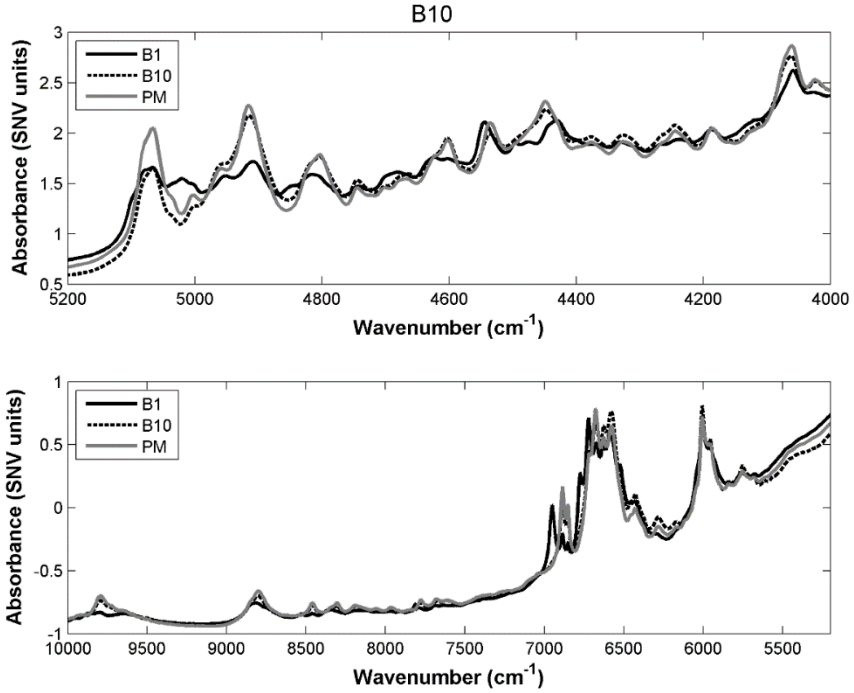


Figure 20 – NIR spectra of the cocrystal obtained in batches B#1 and B#10 and from a physical mixture of HTZ and PABA (PM).

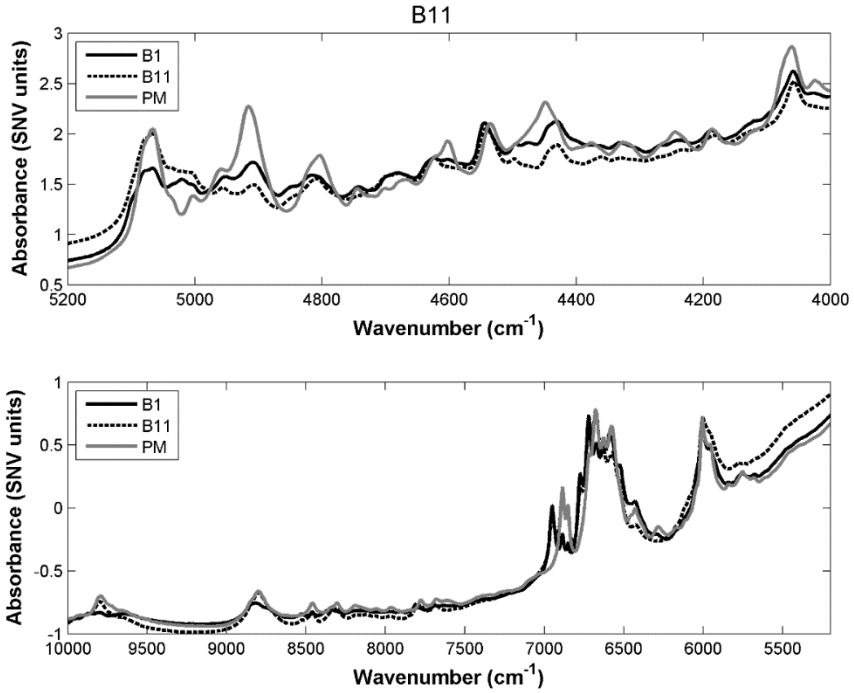


Figure 21 – NIR spectra of the cocrystal obtained in batches B#1 and B#11 and from a physical mixture of HTZ and PABA (PM).

4.4.2 Batch statistical process control

Table 7 summarizes the PLS and OPLS models developed for the monitoring of the cocrystallization process. The models' components were chosen by maximizing the variance in **Y** while minimizing the detection of false deviations (false positives) in the calibration batches. Two components were fitted on the PLS model, explaining 90.0% of the variance in **Y**. The OPLS model also contains two components, one predictive and one orthogonal. As expected the overall performance of the models is the same with the OPLS explaining also 90.0% of the variance in **Y**. The advantage in the use of OPLS comes from its capability to partition the variance in **X**. Variance in **X** correlated with **Y** is retained in the predictive component and the variance in **X** uncorrelated with **Y** is captured by orthogonal component. Therefore, in the OPLS, the predictive component explains 90% of the variance in **Y** and the orthogonal 0%. The amount of explained variance in **X** in the first PLS component and in the OPLS predictive component are similar. In both models, most of the variation related to the process evolution is already captured by the first component.

Table 7 – Summary of the PLS and OPLS models based on batches B#1 to B#4.

Model	Component	R^2X (%)	Cumulative R^2X (%)	R^2Y (%)	Cumulative R^2Y (%)
PLS	1	51.6	51.6	85.5	85.5
	2	22.1	73.7	4.5	90.0
OPLS	Predictive	48.2	48.2	90.0	90.0
	Orthogonal	25.6	73.7	0.0	90.0

It can be observed that the loading for the first PLS component is almost equivalent to the same loading obtained with OPLS. The second PLS component is very similar to the orthogonal component of the OPLS model with some exceptions in the region between 5100 cm^{-1} and 4550 cm^{-1} (Figure 22).

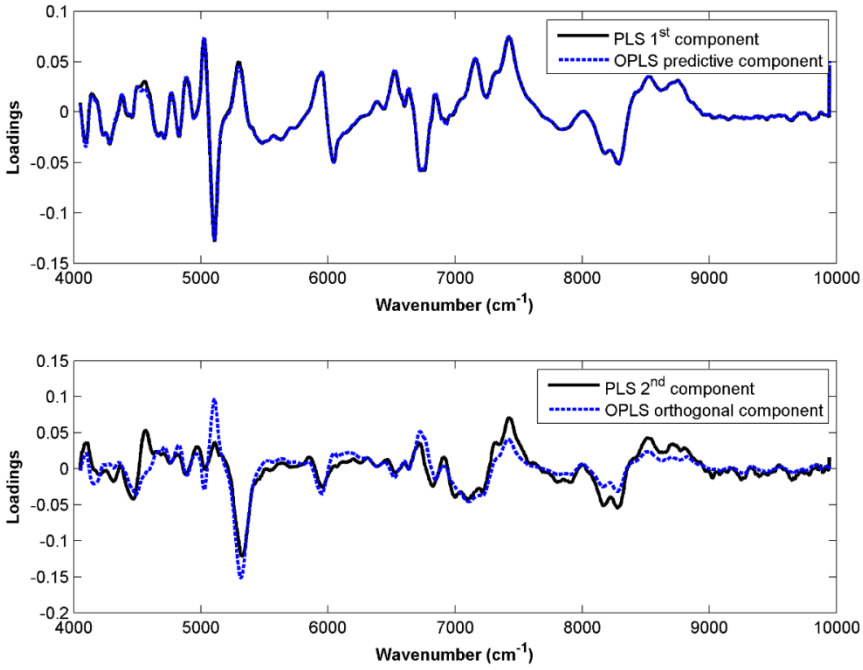


Figure 22 – Loadings for PLS and OPLS models.

The first PLS and OPLS predictive component scores (Figure 23) are similar. In both models, the first 100 min (20 spectra) are related with the rapid evaporation of methanol. The vanishing of the methyl group bands at 8500 cm^{-1} (Figure 24) proves this. After 250 min (50 spectra) there is a slight inflection on the scores that is related with the beginning of the appearance of bands related with the cocrystal, between 5000 cm^{-1} and 4000 cm^{-1} and the disappearing of methanol related bands at 7000 cm^{-1} (OH vibration) and 6150 cm^{-1} (methyl group). After 440 min the scores remain constant which means that the cocrystal is formed and there is no longer any methanol present. The scores corresponding to the second PLS component and for the orthogonal component of OPLS show the evaporation of the methanol until the bands related with the cocrystal appear after 250 min. After 440 min, the scores remain constant and the cocrystallization is finished (Figure 23). The trajectories of the PLS and OPLS scores are very similar. The slight differences that can be observed can be attributed to the fact that while in the PLS model the information regarding batch evolution is split between the two components, in the OPLS model the predictive component already captures all process-related variation (48.2%) filtering the orthogonal variation to the second component (25.5%).

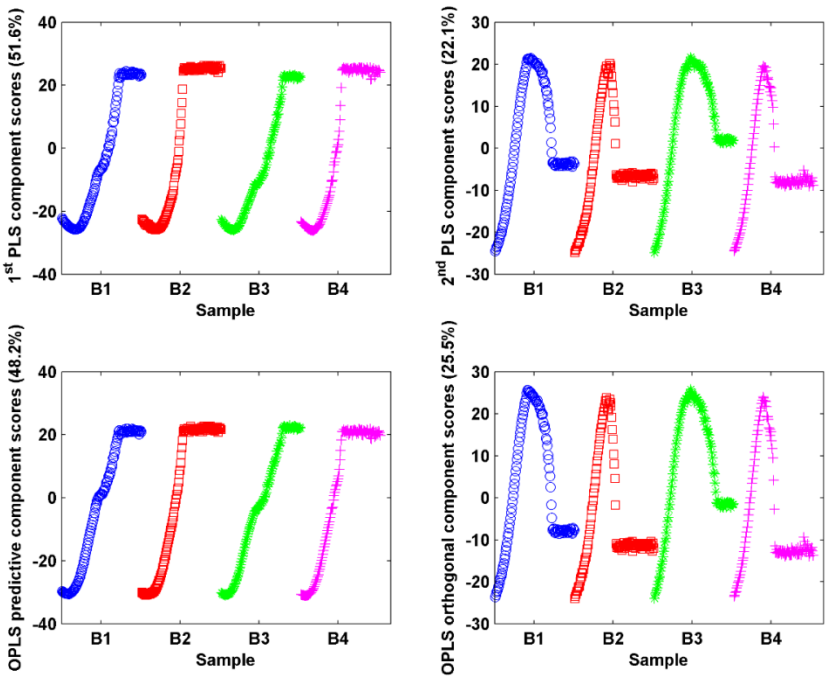


Figure 23 – PLS and OPLS scores for the four calibration batches (batches B#1 to B#4).

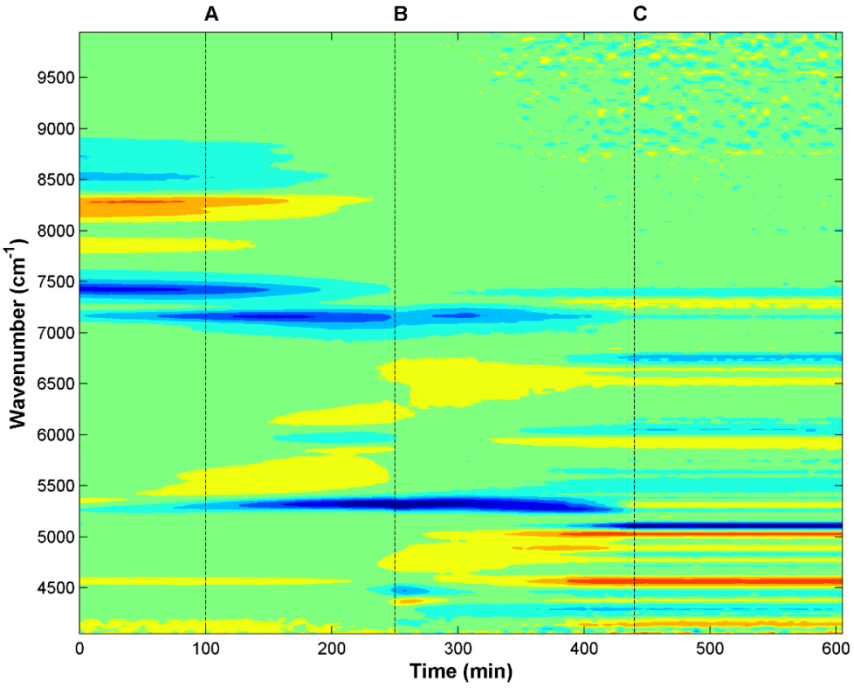


Figure 24 – Contour plot for the NIR spectra preprocessed with the Savitzky-Golay algorithm (29 points width filter fitted with a second-order polynomial followed by a first derivative) obtained for a nominal cocrystallization process (batch B#1). A – End of the rapid solvent evaporation; B – Beginning of the cocrystallization; C – End of solvent evaporation, only cocrystal present.

BSPC control charts for the Hotelling's T^2 and Q show very similar trends when considering the four calibration batches. Additionally, no false positives were detected (Figure 25 and Figure 26).

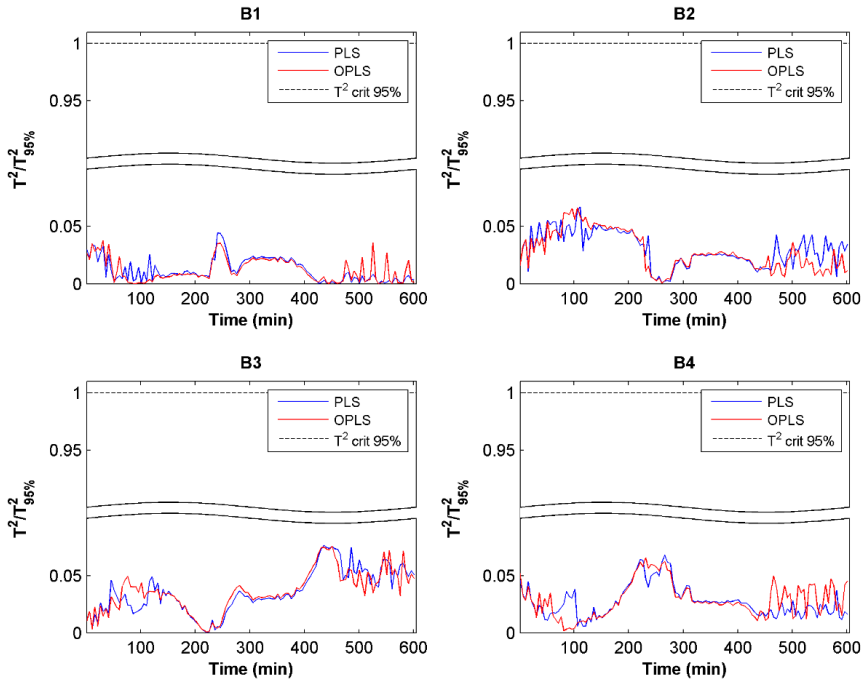


Figure 25 – Hotelling's T^2 control charts generated with PLS and OPLS for the calibration batches.

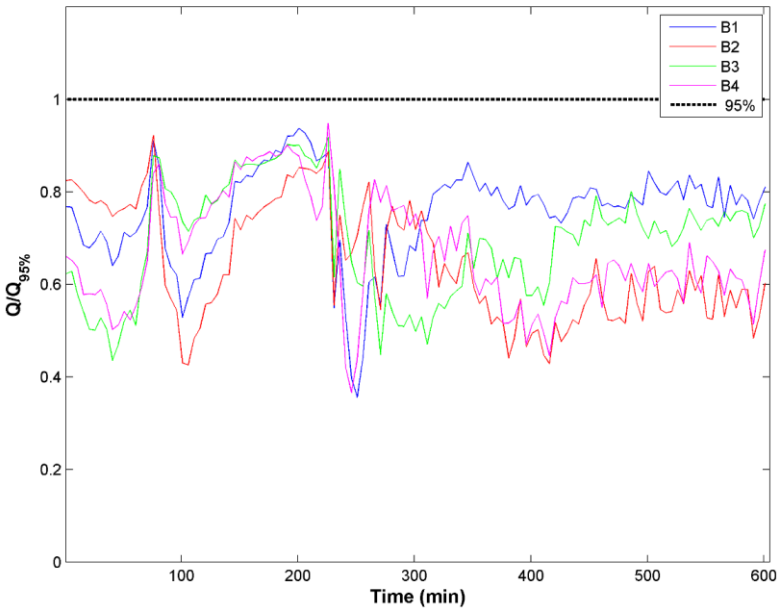


Figure 26 – Squared residuals control chart for the calibration batches (PLS/OPLS).

Comparing the PLS and OPLS scores for the test batches the same conclusions can be drawn as for the calibration batches (Figure 27). The scores for the first PLS component and the scores for the predictive OPLS component are similar for the five test batches as well as the scores for the second PLS component and for the orthogonal OPLS component. The evolution of batch B#11 is the most dissimilar in both components compared with the NOC batches since the scores of both components capture the methanol evaporation and this batch has the highest content of methanol (40 ml).

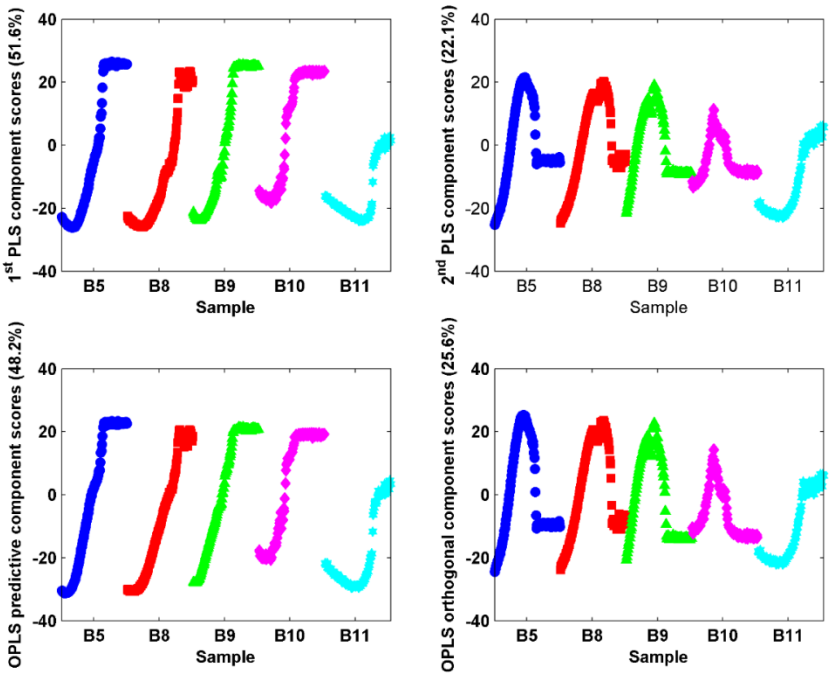


Figure 27 – PLS and OPLS scores for the testing batches (B#5 and B#8 to B#11).

In Figure 28, the Hotelling's T^2 control chart for the test batches for the PLS and OPLS models is shown. Batch B#5, as expected, is always below the control limit. However, differences between the Hotelling's T^2 from the PLS and OPLS models can be seen especially at the beginning and end of the process. For batch B#8 (molar ratio of 1:1), after 400 min, i.e. after the solvent evaporated, the process is out-of-control in the OPLS but not in the PLS control chart. For this batch, the final product is not the cocrystal but a mixture between HTZ and PABA, so it is expected that the Hotelling's T^2 for the final product will be signaled as out-of-control.

The OPLS model is capable of capturing this deviation while the PLS was not. Batch B#9 is out-of-control in the beginning of the process in both PLS and OPLS control charts.

This corresponds to the first phase of the process when solvent is evaporating. This batch started with less solvent (15 ml) than the nominal runs which explains the deviation. The final product of this batch was also a mixture between the two components and not the cocrystal (Figure 19). However, both models failed to detect this deviation in the final part of the process. For batch B#10, the amount of solvent is the same as in the nominal batches (20 ml), however the initial mass is higher (molar ratio 2:3), leading to undissolved particles present in the beginning of the process and therefore to differences in the NIR spectra due to scattering effects. This effect can be seen in the Hotelling's T^2 control chart in which the first part of the process is out-of-control for both models. However, in the OPLS case there is a part of the process' that is out-of-control (around 500 min of process time), which is an indication that the final product was not a cocrystal. B#11 has the highest amount of solvent (40 ml) and a molar ration of 1:4. The batch product was the cocrystal with an excess of PABA. For both models, out-of-control signals were detected in the beginning and end of the process, although more evident for OPLS.

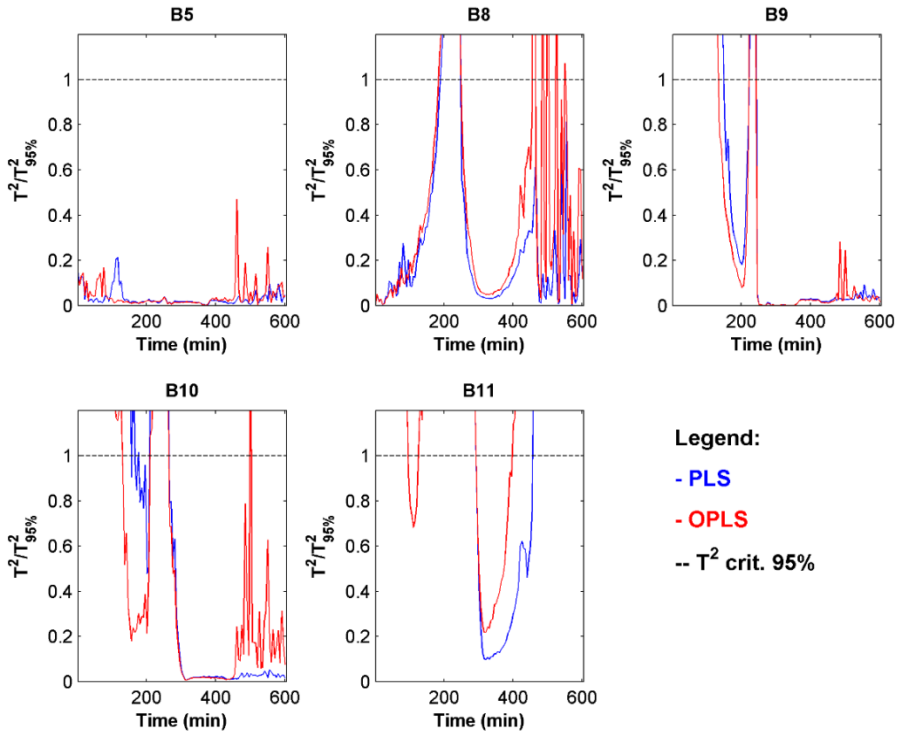


Figure 28 – Projection of the testing batches on the normalized Hotelling's T^2 control charts generated with PLS and OPLS models.

Summarizing, after the solvent evaporation, the PLS model gave false positives for batches B#8, B#9 and B#10. For the OPLS model, a false positive was detected for batch B#9. Therefore, it can be concluded that the two models yield different results for batches B#8 and B#10 especially at the end, for which the OPLS is above the 95% limit.

The Q control chart (Figure 29) is the same for both models as for a single Y the PLS and OPLS models' solutions are identical. Before solvent evaporation (<400 min), batches B#5 and B#8 are in control since they have the nominal amount of solvent. Batch B#10 has also the nominal amount of solvent but higher amount of solids, which lead to a large variation of the NIR spectra leading to the batch being out of control in the beginning of the process. For batch B#9 and B#11 both of them had non-nominal amounts of solvent and therefore are out-of-control in this part of the process. In the end, only batch B#11 is out-of-control (as it should). This means that false positives were obtained for batches B#8 B#9 and B#10.

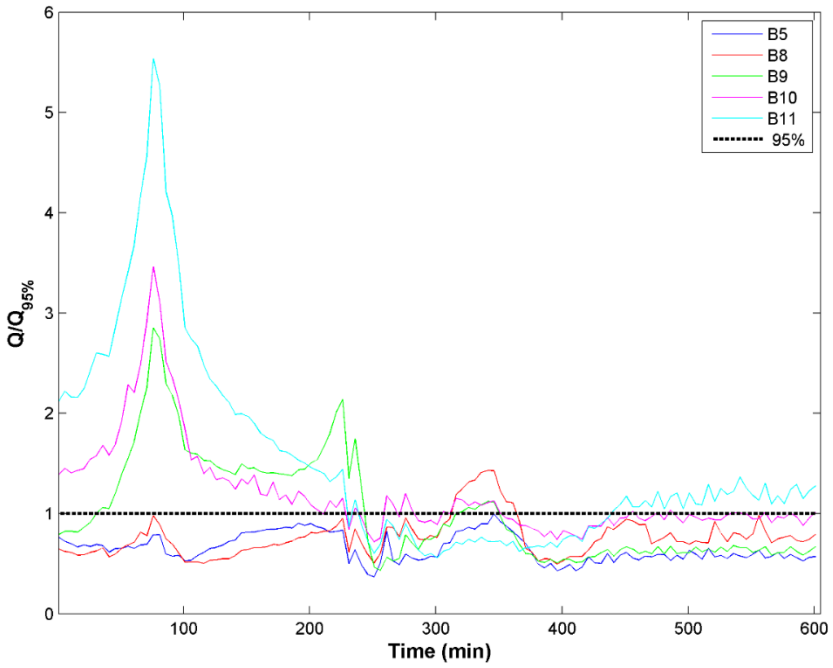


Figure 29 – Projection of the testing batches on the normalized squared residuals control charts generated with PLS/OPLS models.

4.5 Conclusions

Both PLS and OPLS-based BSPC approaches are suitable to follow the evolution of the cocrystallization of hydrochlorothiazide (HTZ) and *p*-aminobenzoic acid (PABA) monitored with NIR spectroscopy. Both approaches can detect deviations from normal trajectory when non-nominal batches are projected into the model. Regarding the Hotelling's T^2 control chart in the end of the process the PLS model gave false positives for batches B#8, B#9 and B#10. For the OPLS model, a false positive was obtained for batch B#9. The main difference between the two models were for batches B#8 and B#10. The deviations were only detected in the OPLS model based control charts. OPLS also demonstrated to be more sensitive to the imposed disturbances. This indicates that the use of OPLS resulted in an improved Hoteling's T^2 statistic calculation that allowed a better detection of process disturbances.

The Q control chart is the same for both models so no comparison can be made. However, this chart shows differences mainly in the beginning of the process before the solvent complete evaporation. Towards the process end, the Q control chart shows false positives for B#8, B#9 and B#10.

Chapter 5

Understanding and supervising a continuous pharmaceutical twin-screw granulation and drying process

5.1 Continuous granulation

Approximately 70 to 80% of all pharmaceutical preparations are tablets. They are inexpensive and easy to manufacture and they deliver an accurate dose of API. There is also a high patient compliance [207, 208].

The most convenient way to make a tablet is by direct compression of the powder blend. This is a rather simple process which does not involve any liquid addition. However, 80% of pharmaceutical ingredients are not suitable for direct compression due to insufficient flowability, tabletability and homogeneity [208]. To avoid content uniformity issues the powder mix must be homogeneous. Furthermore, to ensure an uniform die filling during high-speed tableting an excellent flowability is essential. Flowability can be improved by resorting to agglomeration techniques such as granulation. Granulation can increase bulk density, flowability and solubility. It also reduces the risk of size segregation and dust formation [209]. Granulation can be defined as a particle size enlargement process whereby small powder particles are gathered into larger, permanent structures in which the original particles can be distinguished [210]. There are two types of granulation, wet and dry [211]. In wet granulation

a binder liquid is introduced onto agitated powder particles binding these together through a combination of capillary and viscous forces [212]. The wet granules are then dried and more permanent bonds are established. The solid bridges formed by hardening of binders and the crystallization of dissolved particles are responsible for the strength of a granule. Dry granulation is most appropriate to moisture sensitive drugs as it is achieved by compaction of a powder mass which is then crushed and fractionated. Granulation is easily transferable to a continuous process as blending and feeding can be performed continuously [213].

Vervaet and Remon reviewed several continuous wet granulation techniques [214]. According to these authors continuous granulation is easy to automate, it requires a reduced handling of material and in case of a deviation from normal operation, the faulty material can be discarded instead of the entire batch being destroyed or reworked. This results in major savings.

Spray-drying, roller compactors, fluid bed granulation, instant agglomerators, extrusion and semi-continuous granulators are some of the options presented by these authors.

During spray-drying solutions, suspensions or emulsions are atomized in a drying chamber. The liquid evaporates yielding powder particles. This is a continuous process but non-agglomerated particles are obtained or, in the best scenario, loosely-bounded agglomerates. These products are not suited for tabletting without further agglomeration due to they poor flowability. A schematic overview of the spray drying process is shown in Figure 30.

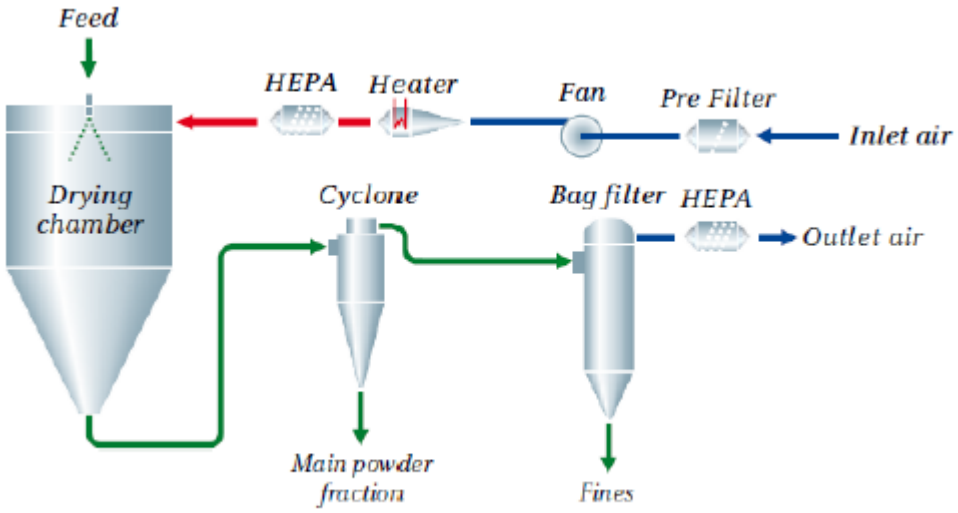


Figure 30 – Schematic overview of the spray drying process [courtesy GEA Pharma Systems].

Roller compaction, is inherently continuous. However, its most noticeable limitation is being a dry granulation technique which relies only on the material compressibility. A lot of fine material is generated while milling to the desired particle size. The used of recycling unit for fines is not ideal as it is known to cause uniformity issues [214].

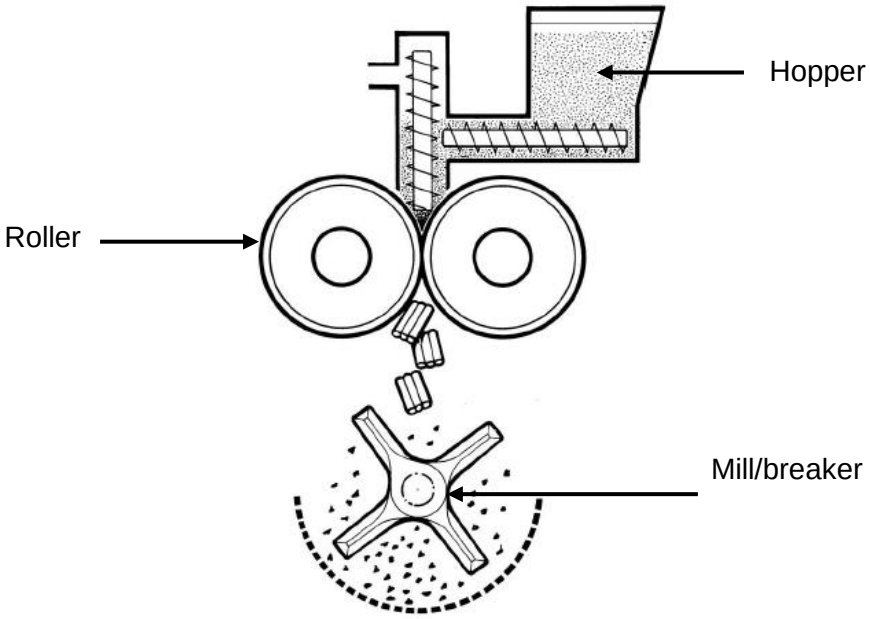


Figure 31 – Schematic representation of a roller compactor. Adapted from [215].

The most widespread continuous fluid bed granulators work horizontally (e.g. Glatt GF series, Niro Vibro-Fluidizer, Heinen Drying Technologies) and consist of different functional zones where feeding, mixing, spraying, drying, cooling and discharging are achieved. These equipments were not specifically designed for pharmaceutical use and they are used mainly in chemical, dairy and food industry (Figure 32) [211]. In the pharmaceutical industry, they are not popular for the production of solid dosage forms as they can only process high volumes (20 kg up to several tons per hour) and are unable to operate at lower volumes.

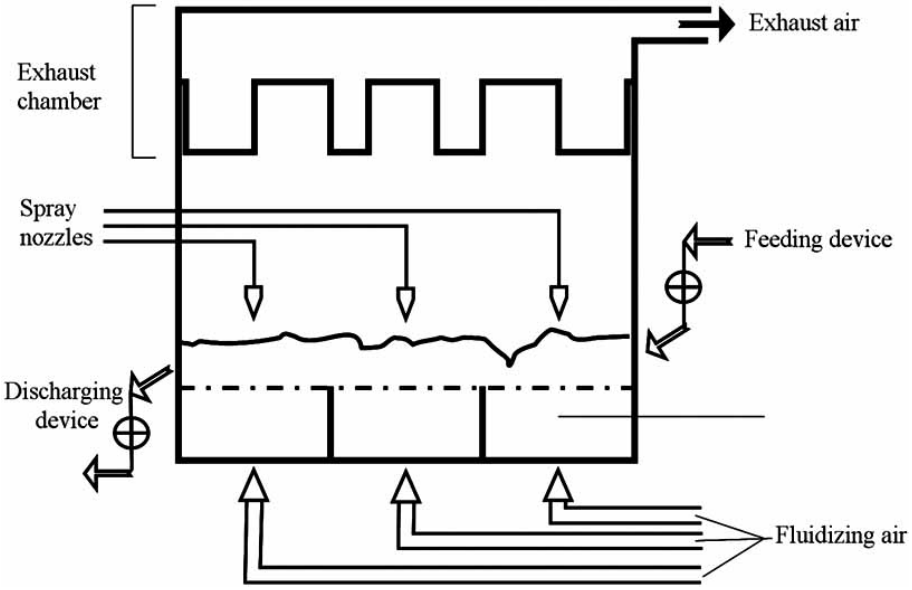


Figure 32 – Horizontal continuous fluid bed granulator [216].

Glatt developed the AGT-Series fluid bed to cope with the flow limitation by developing a system able to operate in a broader range of production rates. The material is confined to a limited space and there is a continuous discharge of granules through a round-shaped outlet at the bottom of the screen (Figure 33). The existence of a counter current air flow through the pipe at the centre of the bottom plate ensures that only particle agglomerates exit the fluid bed. A major drawback of this system is the long and uncontrollable residence time of granules.

Continuous granulators are able to create granules in seconds. These systems continuous mix powder and liquid using high speed mechanical agitation. The amount of powder bring processed is therefore very limited reducing produced waste but also limiting product throughput. Some examples of these systems are the Nica M6 mixer/granulator (Ivarson mixer) and the Schugi Flexomix (Hosokawa) (Figure 34).

Twin-screw granulation in an extrusion process where the shear of mixing paddles inside a barrel causes wetted powder to agglomerate. Differently from conventional extrusion, no die is placed at the exit of the barrel and the material is discharged as granules. The barrel is short as this process is effective. By varying a number of process variables such as barrel temperature, powder and liquid feed rate, screw configuration, screw speed, etc., granules with very different characteristics at a desirable throughput can be obtained. This is therefore, a very flexible process, which allows both developmental work and production to be performed in the same apparatus. Monitoring and controlling of process parameters is important to keep the process in a state of control. A key advantage of twin-screw granulation is the flexibility in design and throughput. A disadvantage toward fluid bed granulation techniques is that extrusion yields wet granules this require a continuous drier.

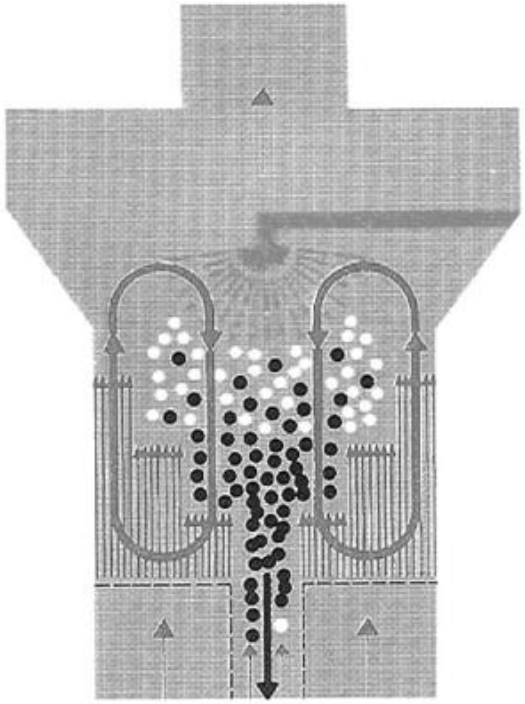


Figure 33 – Glatt AGT continuous fluid bed granulator [214].

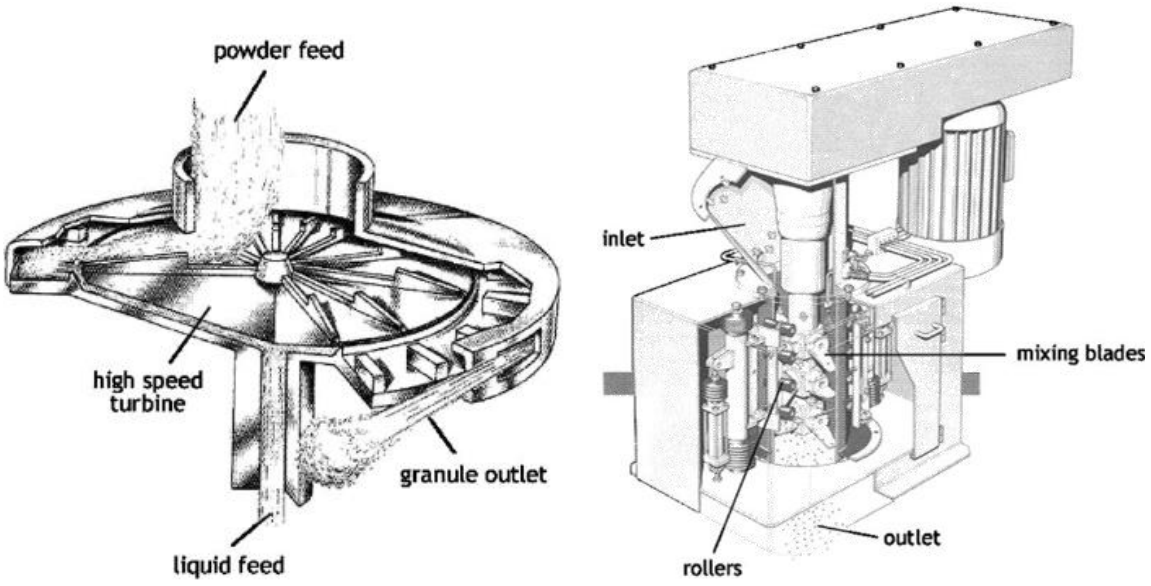


Figure 34 – A Nica M6 mixer/granulator (Ivarson mixer) on the left and a Schugi Flexomix (Hosokawa) on the right [214].

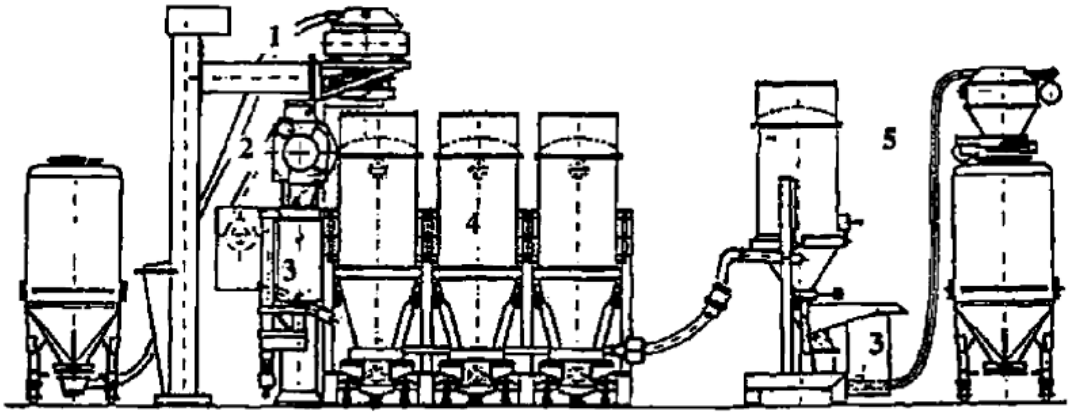


Figure 35 – Glatt Multicell continuous granulation system (Glatt AG, CH-4133 Pratteln): 1. dosing unit, 2. high-shear mixer, 3. sieve, 4. fluid bed, 5. Pneumatic transport system [217].

A truly continuous granulation line is the ConsiGma™-25 by GEA Pharma Systems, Collette™ (Wommelgem, Belgium) (Figure 36). This is one of the first commercially available fully continuous manufacturing lines. The ConsiGma™-25 consists of three main modules: a continuous twin-screw granulator, a six-cell segmented fluid bed dryer and a product control unit. Attached to the granulator there are also a powder dosing unit (i.e. a feeder) and a liquid addition module. In addition, a Modul™ rotary tablet press and an Omega™ coater can be

attached to the product control unit, at the end of the line. The continuous pharmaceutical granulation and drying work conducted in the ambit of this thesis was completed on two similar installations of the ConsiGma™-25 system, and focuses only on the granulator and dryer units.

A ConsiGma™-1, is an easily transportable, reduced version of the ConsiGma™-25 line which facilitates research and development (Figure 37). An identical granulator barrel as the ConsiGma™-25 line is linked to the equivalent to a single segment of the six-cell segmented fluid bed dryer from the ConsiGma™-25 line. Drying is performed batch-wise because only one cell is present, but due to its similarity to the larger scale line, it allows short granulation runs ideal for the earlier research and development work.

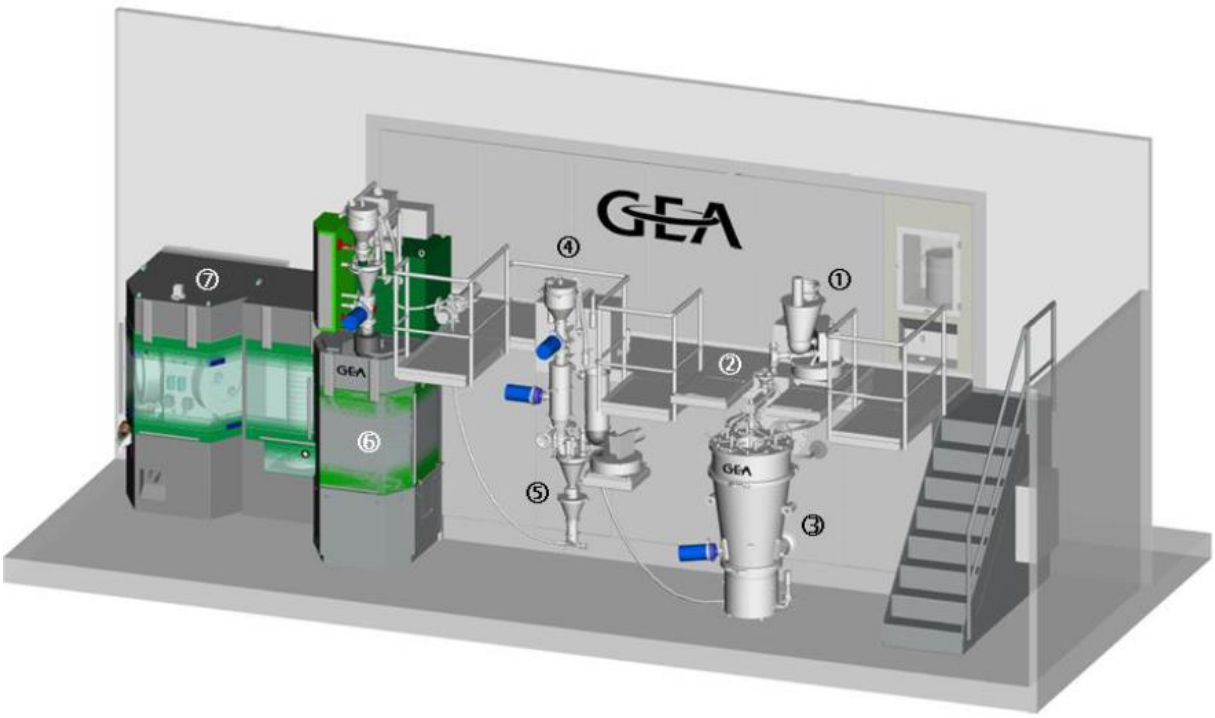


Figure 36 – ConsiGma™-25 continuous from powder-to-tablet line: 1. Powder dosing unit; 2. Twin-screw granulator; 3. Six-cell segmented fluid bed; 4. Product control unit; 5. Blender (external phase); 6. Tablet press; 7. Tablet coater (Courtesy of GEA Pharma Systems).

Figure 37 – ConsiGma™-1 with: 1. Powder dosing unit, 2. Twin-screw granulator, 3. Fluid bed dryer [Courtesy of GEA Pharma Systems].

The MODCOS™ line is also based on coupling twin-screw granulation with fluid bed dryer and a tablet press. It is the result of a collaboration Glatt with Thermo Fisher™ Scientific and Fette™. The Böhle Conti Granulator BCG (by Böhle™ in collaboration with Korsch™) is also another line that merged continuous screen granulation and fluid bed drying.

Granucon™ is composed by a ringlayer mixer and a horizontal fluid bed dryer. This is a design by Lödige™. Lödige™ implemented a screw in the dryer to narrow down the wide retention times of granules in the dryer, characteristic of continuous fluid bed dryers.

5.2 ConsiGma™-25 continuous line

The ConsiGma™-25 (GEA Pharma Systems, Collette™ Wommelgem, Belgium) continuous production line contains three units: a continuous twin-screw high-shear granulation module, a six-cell segmented fluid bed dryer, and a product control unit (Figure 38) [161, 218]. Additionally, this type of line also allows for additional blending, tableting and coating by coupling a tablet press and coater, however tableting was not in the scope of the work described in this dissertation. This system continuously generates univariate data (e.g., temperatures, pressures, etc.) during operation acquired from sensors implemented at different locations.

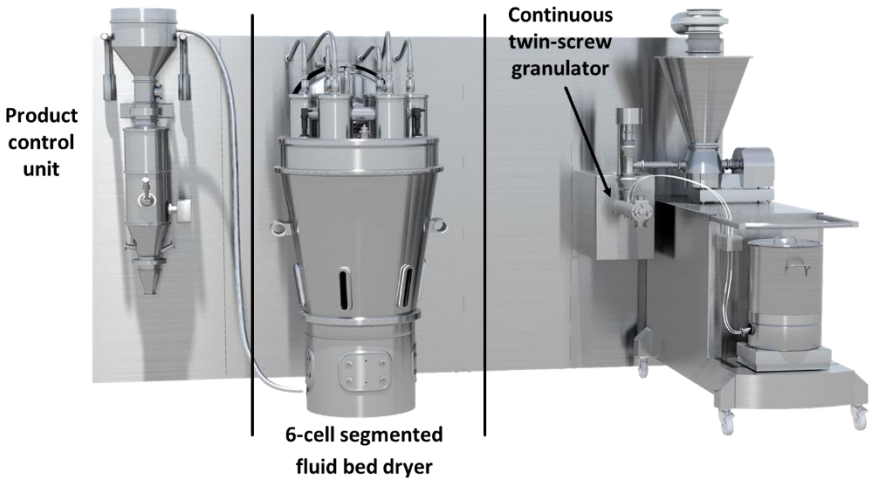


Figure 38 – ConsiGma™-25 continuous manufacturing line granulator, dryer and product control unit modules.

5.2.1 Continuous twin-screw granulation

The granulation module has a powder dosing unit, a liquid addition module and the granulator itself (

Figure 39).

The powder dosing unit feeds the dry premix from a container. Alternatively, several feeders can act simultaneously, each one dosing a dry ingredient to a continuous feeder which homogenizes the preblend before it enters the granulator. The liquid addition module adds the granulation liquid at a predefined rate via two injection nozzles into the granulator barrel working segment by means of two peristaltic pumps. The powder dosing unit and the liquid addition unit both work under the loss-in-weight principle.

The granulator consists of a twin-screw co-rotating system (

Figure 40). The screws are modular and allow for different types processing thorough the barrel. The feeding segment consists of conveying elements which transport the powder through the barrel and into the working segment. In the working segment, kneading elements intensively mix it with the granulation liquid. Screw speed and barrel temperature are user-set. A temperature control unit is responsible to cool or heat the granulator barrel medium (water) to the chosen granulator barrel temperature setpoint. The temperature of the granulator is monitored, however, it is not controlled and there is no feedback system toward the temperature control unit.

The granulation barrel is short, and therefore the residence time is also short, which is an advantage during development as a minimal amount of material is necessary to screen the different settings. After granulation, wet granules are transported via a vacuum transfer line or gravimetrically to the fluid bed dryer.

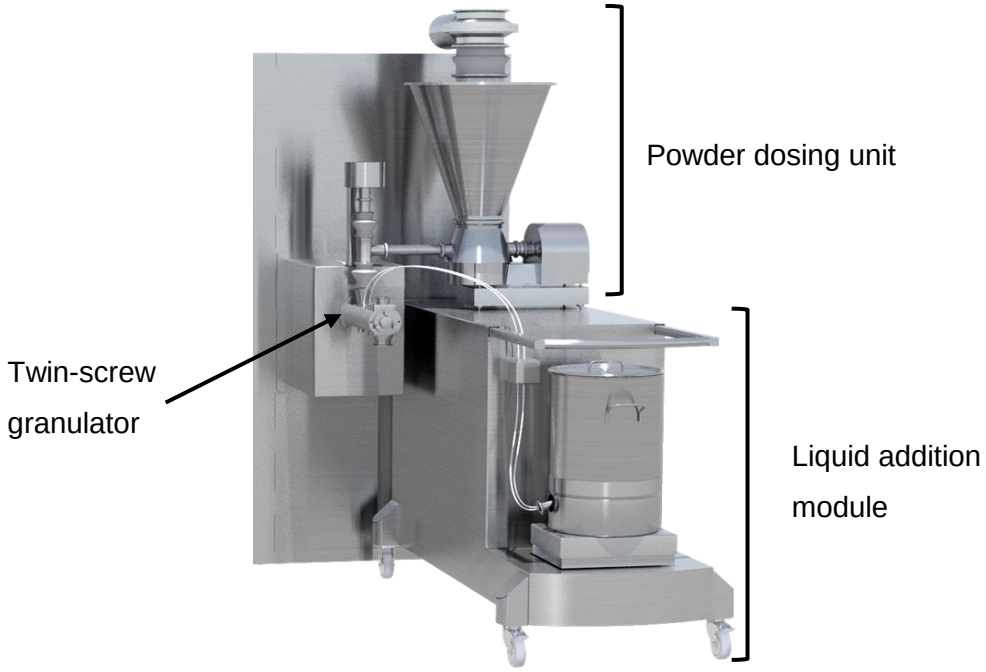


Figure 39 – Continuous granulation module (Courtesy of GEA Pharma Systems).

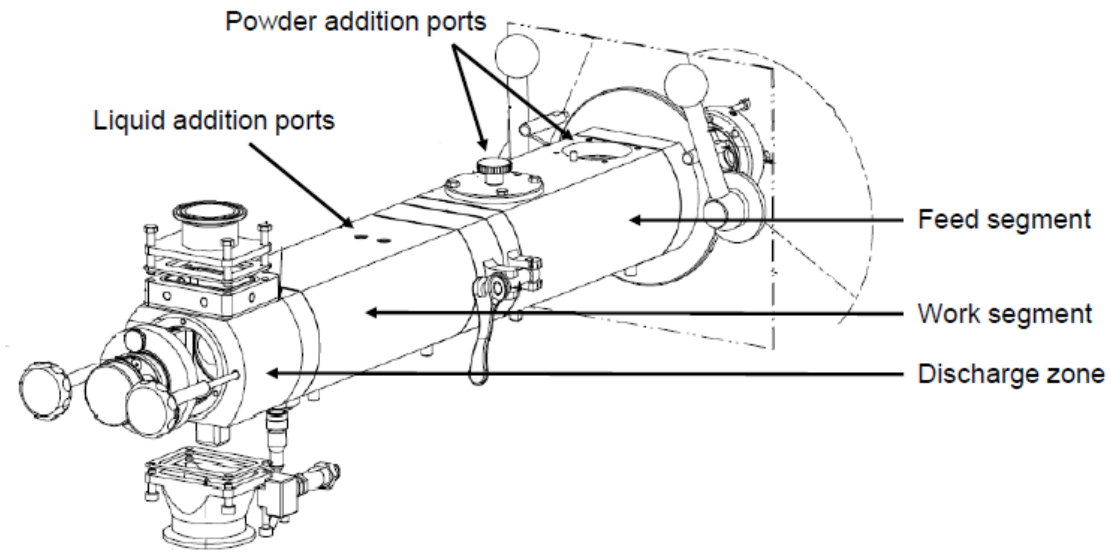


Figure 40 – Detail of the GEA continuous twin-screw granulator (Courtesy of GEA Pharma Systems).

5.2.2 Six-cell segmented fluid bed dryer and product control unit

The segmented dryer (Figure 41) works under the fluid bed principle. The drying chamber is divided in 6 identical cells which are sequentially filled and discharged one after the other ensuring a continuous flow of incoming wet granules and outcoming dry granules (Figure 42). The dryer air temperature (°C), humidity (%RH) and air flow (m^3) are user-defined while the actual temperature inside the individual cell is monitored. When the cell is filled the temperature inside the cell corresponds to temperature of the product inside. The material inside a cell ranges from 0.5 kg to a maximum of 1.5 kg (for a material throughput of 25 kg/h). The cells dry for a predefined period, and after discharge, remain inactive until they are filled again. To fit the needs of developmental work, the number of cells to be filled can be chosen in order to minimize the material used. An air handling unit, prepares the drying air to enter the dryer according to the defined setpoints. A push/fan and fan/blower systems regulate respectively, air flow and pressure within the dryer. Several HEPA filters were placed in the air outlets of the dryer and product control therefore avoiding that particulate material exits the system.

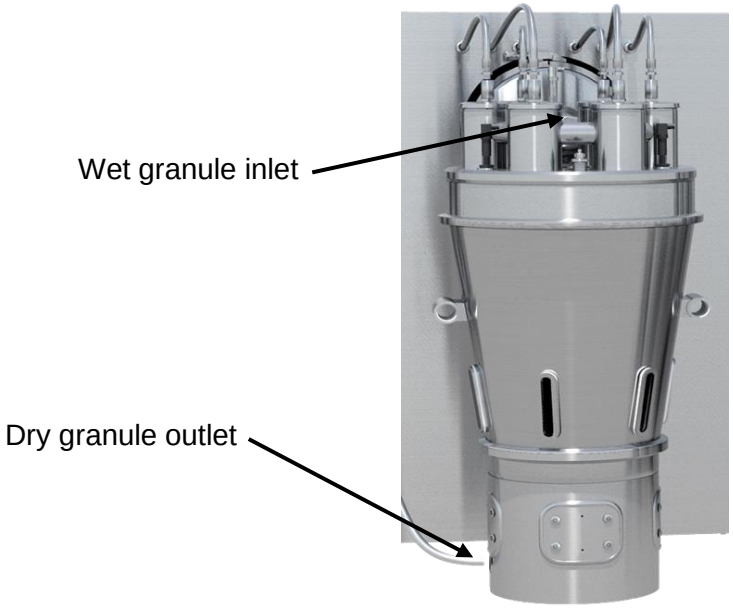


Figure 41 – Six-cell segmented fluid bed dryer (Courtesy of GEA Pharma Systems).

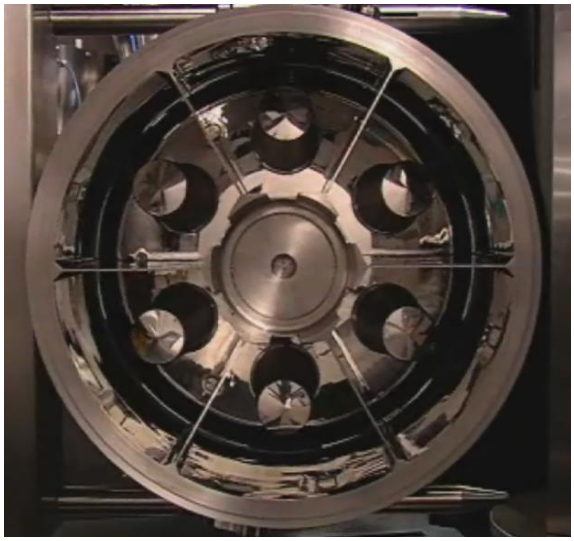


Figure 42 – View from the inside of the six-cell segmented fluid bed dryer (Courtesy of GEA Pharma Systems).

5.2.3 Product control unit

The material discharged from the cells goes into a product control unit (Figure 43). In the product control unit, there is the possibility to install various probes for PAT monitoring and control purposes. A mill is placed at the exit of the control unit but it was not used for the specific purpose of the work included in this thesis. After the product control unit, as tableting was not in scope, the granules exit the system.

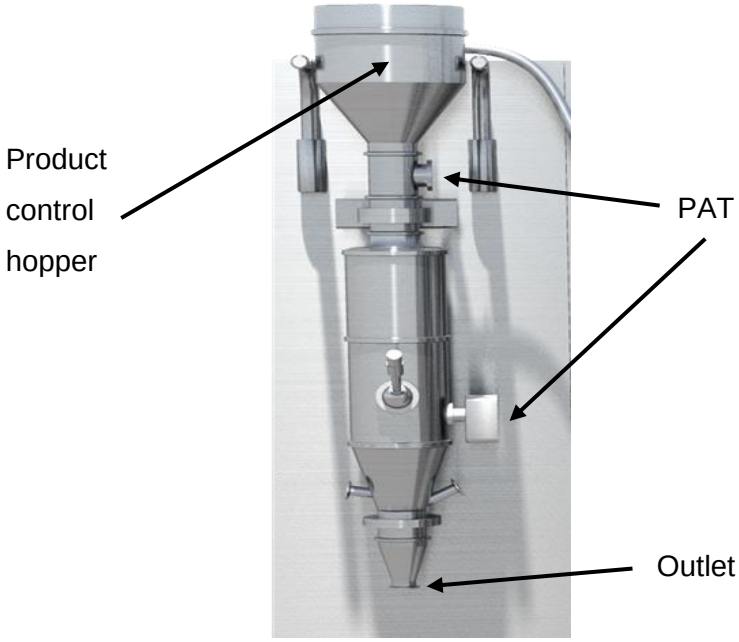


Figure 43 – Product control unit of the ConsiGma™-25 system.

5.3 Understanding and monitoring a continuous pharmaceutical twin-screw granulation and drying process using multivariate data analysis – part I[3]

5.3.1 Abstract

The present work is the first part of a study which focuses on acquiring an in-depth knowledge about a granulation and drying process performed on a ConsiGma™-25 continuous manufacturing line. During operation, the continuous line logs multiple univariate process variables, hence generating a large amount of data. Process variables are related to process parameters which impact final product quality. Ensuring that process variables are in control, reduces the variability on the final product. Three identical five-hour continuous manufacturing runs were performed. Multivariate data analysis tools, more specifically, latent variable modeling tools such as PCA were utilized to extract information from the generated datasets unveiling process trends and drifts. Furthermore, a statistical process control strategy based on the simultaneous application of MSPC and BSPC concepts is presented. The present study focuses on the use of PCA for MSPC as a possible solution for the modelling and monitoring of variables at a steady state i.e. which do not present time-relevant trends. A second part of this work will focus on the BSMP modeling of variables with time-related trends and on the process startup periods of the different units, for which process time is a relevant feature.

[3] This chapter has been adapted from: A.F.T. Silva, J. Vercruysse, M.C. Sarraguça, J.L.F. Costa Lima, C. Vervaet, J.P. Remon, J.A. Lopes, T. De Beer, Understanding and monitoring a continuous pharmaceutical twin-screw granulation and drying process using multivariate data analysis – part I, (Submission in process).

5.3.2 Materials and methods

5.3.2.1 Industrial case study

A pharmaceutical dry premix containing two API's, powdered cellulose, maize starch, pregelatinized starch and sodium starch glycolate was granulated with distilled water in the ConsiGma™-25 system. All pharmaceutical powders were provided by Johnson&Johnson (Janssen-Cilag, Italy). Three identical long continuous manufacturing runs (five hours each) were completed.

5.3.2.2 Continuous manufacturing line

The experiments in question were performed in a ConsiGma™-25 line located at GEA Pharma Systems (Wommelgem, Belgium). In this particular assembly granules are transported gravimetrically between the different subunits. The continuous production line ConsiGma™-25 (GEA Pharma Systems, Collette™ Wommelgem, Belgium) was described in-depth in subsection ConsiGma™-25 continuous line (ConsiGma™-25 continuous line).

During these runs, the powder dosing unit fed the dry premix to the granulator at a speed of 20 kg/h. The liquid addition module fed the granulation liquid, i.e. water, at a rate of 50 g/min. The screw speed and barrel temperature were set at 900 rpm and 25°C, respectively. The 6 cells were filled for 180 s sequentially after each other ensuring a continuous operation. The drying time was set to 790 s in total, including the 180 s cell filling time. During the last seconds of drying, granules were gravimetrically discharged to the product control unit after which the cell remained inactive for 290 s until new wet granules were introduced for drying. After being in the product control unit, granules exited the system. The air handling unit, responsible for preparing the air entering the dryer according to the defined setpoints, contains a dehumidifier inside which unit removes moisture from the incoming air. In addition, a push/fan system regulated the air flow inside the dryer while a fan/blower regulated the pressure. Both the air exiting the dryer and the air leaving the product control unit had to pass through different HEPA filters in order to avoid carrying particulate material out of the system.

5.3.2.3 In-process measurements

The relevant univariate process parameters, continuously measured and logged each second in multiple locations of the ConsiGma™-25, are summarized in **Table 8**. These logged parameters include both user-set variables (setpoints) and other (not set but only measured) logged variables (open loop variables). Even though setpoints are pre-defined, they still vary around their set value due to disturbances. These are then automatically corrected by the system itself through independent proportional–integral–derivative (PID) controllers.

Table 8 – Parameters logged by the ConsiGma™-25 during processing; ● - Variables included in the PCA model for process visualization of the granulator; ■ - Variables included in the PCA model for process visualization of the dryer and product control unit; ◘ - Variables included in the PCA model for process visualization of all units together; ◆ - Variables included in the MSPC model (PCA) for monitoring the continuous granulator ✱ - Variables included in the MSPC model (PCA) for monitoring the dryer and included.

Description	**Units**	**Type**	ConsiGma™-25 unit
Mass flow - Granulation liquid ●◘◆	g/min	Setpoint	Liquid addition module (granulator)
Mass - Flow powder dosing ●◘◆	kg/h	Setpoint	Powder dosing unit (granulator)
Speed - Granulator screws ●◘◆	rpm	Setpoint	Granulator
Temperature sensor - Granulator barrel ●◆	°C	Setpoint	Granulator
Power - Granulator drive ●◘◆	W	Measured	Granulator
Temperature sensor - Granulator barrel refrigeration liquid inlet ●◘◆	°C	Measured	Granulator
Torque sensor – Granulator ●◘◆	Nm	Measured	Granulator
Flow sensor - Wet granule transfer ■◘	m³/h	Setpoint	Wet granule transfer line (dryer)
Pressure sensor - Differential pressure over the wet transfer line ■◘	mbar	Measured	Wet granule transfer line (dryer)
Pressure sensor - Wet transfer line ■◘	mbar	Measured	Wet granule transfer line (dryer)
Temperature sensor - Air handling unit filter ■◘✱	°C	Measured	Air handling unit (dryer)
Temperature sensor - Dehumidifier air outlet ■◘✱	°C	Measured	Air handling unit (dryer)
Humidity sensor - Dehumidifier air outlet ■◘✱	% RH	Measured	Air handling unit (dryer)
Speed control –Fan/blower ■◘	%	Measured	Fan/Blower system (dryer)
Speed control - Push/fan ■◘	%	Measured	Fan/Blower system (dryer)
Flow sensor - Dryer air ■◘	m³/h	Setpoint	Dryer
Humidity sensor - Dryer air inlet ■◘✱	% RH	Setpoint	Dryer
Temperature sensor - Dryer air inlet ■◘✱	°C	Setpoint	Dryer
Humidity sensor - Dryer air outlet ■◘✱	% RH	Measured	Dryer
Pressure sensor – Atmospheric ■◘	mbar	Measured	Dryer
Pressure sensor - Differential pressure over the dryer filters ■◘	m bar	Measured	Dryer
Pressure sensor - Dryer air inlet ■◘	mbar	Measured	Dryer

Table 8 – Parameters logged by the ConsiGma™-25 during processing; ● - Variables included in the PCA model for process visualization of the granulator; ■ - Variables included in the PCA model for process visualization of the dryer and product control unit; ◘ - Variables included in the PCA model for process visualization of all units together; ◆ - Variables included in the MSPC model (PCA) for monitoring the continuous granulator ✱ - Variables included in the MSPC model (PCA) for monitoring the dryer and included.

Description	**Units**	**Type**	ConsiGma™-25 unit
Pressure sensor - Dryer air outlet (afteHEPA filter) ■◘	mbar	Measured	Dryer
Pressure sensor - Dryer air outlet (before HEPA filter) ■◘	mbar	Measured	Dryer
Pressure sensor - Pressure dryer top ■◘	mbar	Measured	Dryer
Temperature sensor - Dryer air outlet ■◘	°C	Measured	Dryer
Temperature sensor - Temperature dryer cell 1 ■◘	°C	Measured	Dryer
Temperature sensor - Temperature dryer cell 2 ■◘	°C	Measured	Dryer
Temperature sensor - Temperature dryer cell 3 ■◘	°C	Measured	Dryer
Temperature sensor - Temperature dryer cell 4 ■◘	°C	Measured	Dryer
Temperature sensor - Temperature dryer cell 5 ■◘	°C	Measured	Dryer
Temperature sensor - Temperature dryer cell 6 ■◘	°C	Measured	Dryer
Speed control - Vacuum pump ■◘	%	Measured	Vacuum pump (dryer)
Pressure sensor - product control unit (after HEPA filter) ■◘	mbar	Measured	Discharger
Pressure sensor - product control unit (before HEPA filter) ■◘	mbar	Measured	Discharger

5.3.2.4 Multivariate data analysis

PCA is a widely used latent variable method, to extract information from large multivariate data-sets, and to compress this data enabling better process visualization [219]. In this work, PCA was used for multiple purposes as described below. All used data sets were first mean centered and scaled to unit variance. All of the below described PCA models were developed and evaluated using SIMCA™ 14 (MKS Data Analytics Solutions, Umeå, Sweden). For all PCA models an R^2 and a Q^2 value were calculated. R^2 corresponds to the variation in the data that is captured by the model while Q^2 is the variation predicted by the model according to cross validation. For cross-validation every seventh observation was assigned to a cross-validation group, each group being excluded at a time in each cross-validation round. Components were added to the models until both R^2 and Q^2 were maximized. Adding more components to the model would result in overfitting.

Process visualization

PCA models were fit on the data from the three runs individually for the granulator unit data, the dryer data (including transfer lines to and from the dryer) and, the product control unit data, with the purpose of unveiling process trends and existing correlations between the logged variables within these units. PCA models were also fitted for each performed five-hour run, using the data from all units together, with the purpose of unveiling correlations not only between the variables logged within the same unit, but also between the variables from different units.

Multivariate statistical process control approach

As described earlier, the ConsiGma™-25 is an overall continuous process where some sub-steps operate in a continuous way (granulation) while other sub-modules operate semi-continuously (drying and product control modules). The segmentation of the fluid bed dryer in six identical cells ensures a continuity of the material flow entering and exiting the dryer by filling and discharging the cells sequentially. Material within each cell can be perceived as a "mini-batch". The proposed approach combines a MSPC modelling approach for the variables at a steady state without any time-relevant trends, and BSPC for those variables where time-related features are present and important to monitor. The present study focuses only on the MSPC part of the monitoring.

Continuous granulator

During continuous operation, the granulator is expected to run in steady state and two-dimensional data is obtained by the monitoring of variables (J) as function of process time (K). A reference PCA model was built based on historical data from the continuous granulator (including powder dosing and liquid addition units, see Table 8 for further details) while running at steady state (startup and shutdown phases were excluded). The data of two of the three performed runs (runs 1 and 2) were included in the model (reference runs) while the data from run 3 was utilized to test the model's monitoring performance.

Dryer and product control unit

The dryer and product control units operate semi-continuously. The package of granules discharged from a cell at a certain moment can be considered as a "mini-batch" since the cells operate independently and granules are not exchanged between them. Despite this, it is important to mention that the six cells physically communicate with each other since they are open at the top and therefore, variations in pressure, air humidity and inlet air temperature affect all cells.

Humidities and temperatures measured at locations that are common to all cells (outlet of the dehumidifier, inlet and outlet of the dryer) are variables that are expected to remain constant during the full processing time (two-dimensional). Hence, despite the fact that the drying process occurs independently in each cell, these variables can be monitored using a simple MSPC PCA based model. A reference PCA model was built for these logged variables (Table 8) from runs 1 and 2 (reference runs), while the data from run 3 was utilized to test the model's monitoring performance. The first 6480 s of each run where excluded from this data set due to a lower temperature of the air leaving the dehumidifier unit and a higher humidity and temperature of the air leaving the dryer during this period.

Fault Detection and Diagnosis

The individual scores of the calculated principal components of each model were plotted in Shewhart control charts providing a useful visual representation of the process evolution [220]. Control limits for the score values were set at + and – 2 times the standard deviation (i.e. within a 95% confidence interval) corresponding to a warning limit at + and - 3 times the standard deviation (i.e. within a 99% confidence interval) corresponding to an action limit. It is always possible to adjust the control limits however generally these limits are utilized.

Hotelling's T^2 and normalized DModX control charts were also utilized. These statistics are thoroughly described in section (Chapter 2). For all models the control limits in the Hotelling's T^2 control charts were set at + and – 2 times the standard deviation (i.e. within a 95% confidence interval) corresponding to a warning limit and + and - 3 times the standard deviation (i.e. within a 99% confidence interval) corresponding to an action limit). On the other hand, the DModX control limits were set at D_{crit}. The significance level for D_{crit} was set to 0.95, meaning 95% of the observations in the model have a DModX value below D_{crit}. Observations

with a DModX larger than twice the D_{crit} are considered moderate outliers. When any of the above described control charts give an out-of-control signal at a certain time point, then contribution plots were calculated to unravel the variables contributing to that deviation, hence providing a good insight into the causes of the process deviations [82].

5.3.3 Results and discussion

5.3.3.1 Process visualization

Continuous granulator

For the three five-hour runs, a PCA model with two PC was fitted for the logged granulator data, capturing 95.5% of the variation in the data (R^2). A Q^2 of 88.0% was obtained. With this model, it is possible to distinguish 3 different process phases: 1) startup, 2) steady state and 3) shutdown phases (Figure 44). All three runs are identical.

During steady state, the score values of both PCs remain close to the origin. The PC1 scores increase during startup and decrease again at the end of the process. According to the loadings of this PC (Figure 45), the mass flow of the granulation liquid, the powder mass flow, the power, the torque, the screw speed of the granulator and the temperature inside the granulator barrel are positively correlated with the scores meaning their values increase during startup and decrease at the shut-down. On the other hand, the temperature at the inlet of the refrigeration liquid of the granulator barrel presents a negative correlation with the scores which means this value decreases during startup. This is expected since once the granulator starts operating, its temperature rises and consequently the refrigeration liquid starts entering the barrel in order to keep the temperature of the granulator barrel at its setpoint. Given this, we can conclude that PC1 depicts the operating status of the granulator, i.e. the granulator is working or not. The PC2 scores (Figure 44) are positive during startup, sit around the origin at steady state and become negative at shut-down. By looking at the loadings of this PC (Figure 45), it can be seen that the main contributing variables are the temperature at the inlet of the refrigeration liquid of the granulator barrel and the temperature inside the granulator itself, both positively correlated with the scores. The time series charts of these two variables are plotted (Figure 46a and b) for clarification. At the beginning of granulation, the temperature of the granulator barrel spikes 1°C (up to 26 °C) at the start of operation. This slight increase is due to the filling of the barrel resulting in material build up on

the barrel walls and consequent increase of friction. At this moment torque and the power of the granulator drive also increase suddenly, and then gradually until steady state is reached which takes around 2000 s. In this time frame, the barrel temperature was corrected back to 25°C by the granulator temperature control system that is responsible for the heating/cooling of the refrigeration liquid entering the barrel Afterwards, the temperature of the barrel remained constant around the set 25°C until the end of the process. When the granulator stopped operating, the barrel temperature spiked towards lower values since the material left the granulator and hence s no heat by friction was generated anymore. The temperature of the refrigeration liquid at the granulator inlet heated the barrel up again to the predefined setpoint.

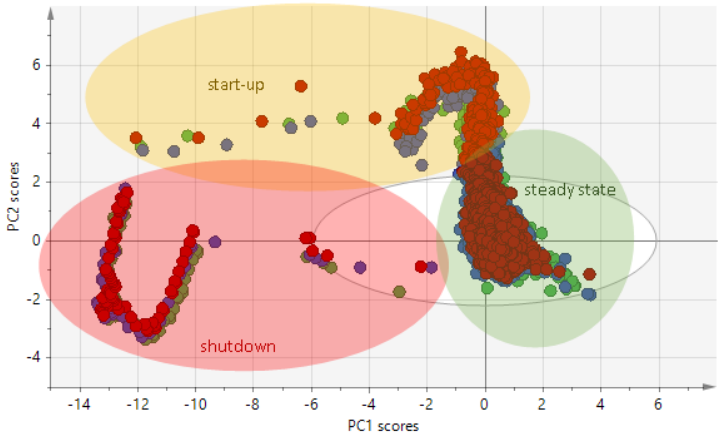


Figure 44 – Process visualization – continuous granulator - a Scores on PC1 versus scores on PC2 (Run 1 – green; Run 2 – red; Run 3 – blue).

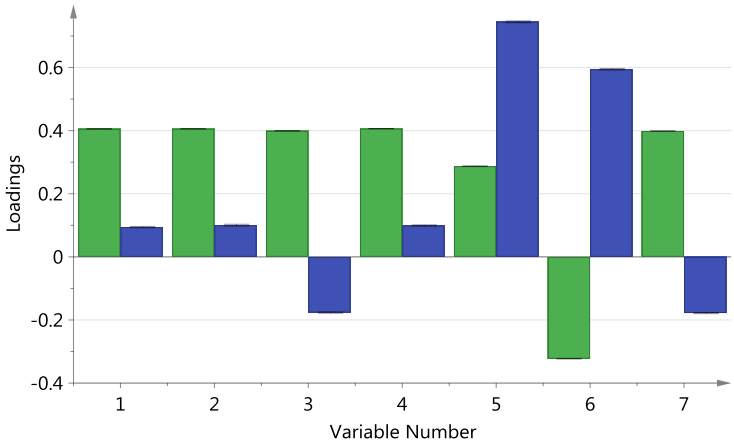


Figure 45 – Process visualization – continuous granulator - Loadings of PC1 (green) and PC2 (blue); 1 - Mass flow - Granulation liquid; 2 - Mass - Flow powder dosing; 3 - Power - Granulator drive; 4 - Speed - Granulator screws; 5 - Temperature sensor - Granulator barrel; 6 - Temperature sensor - Granulator barrel refrigeration liquid inlet; 7 - Torque sensor – Granulator.

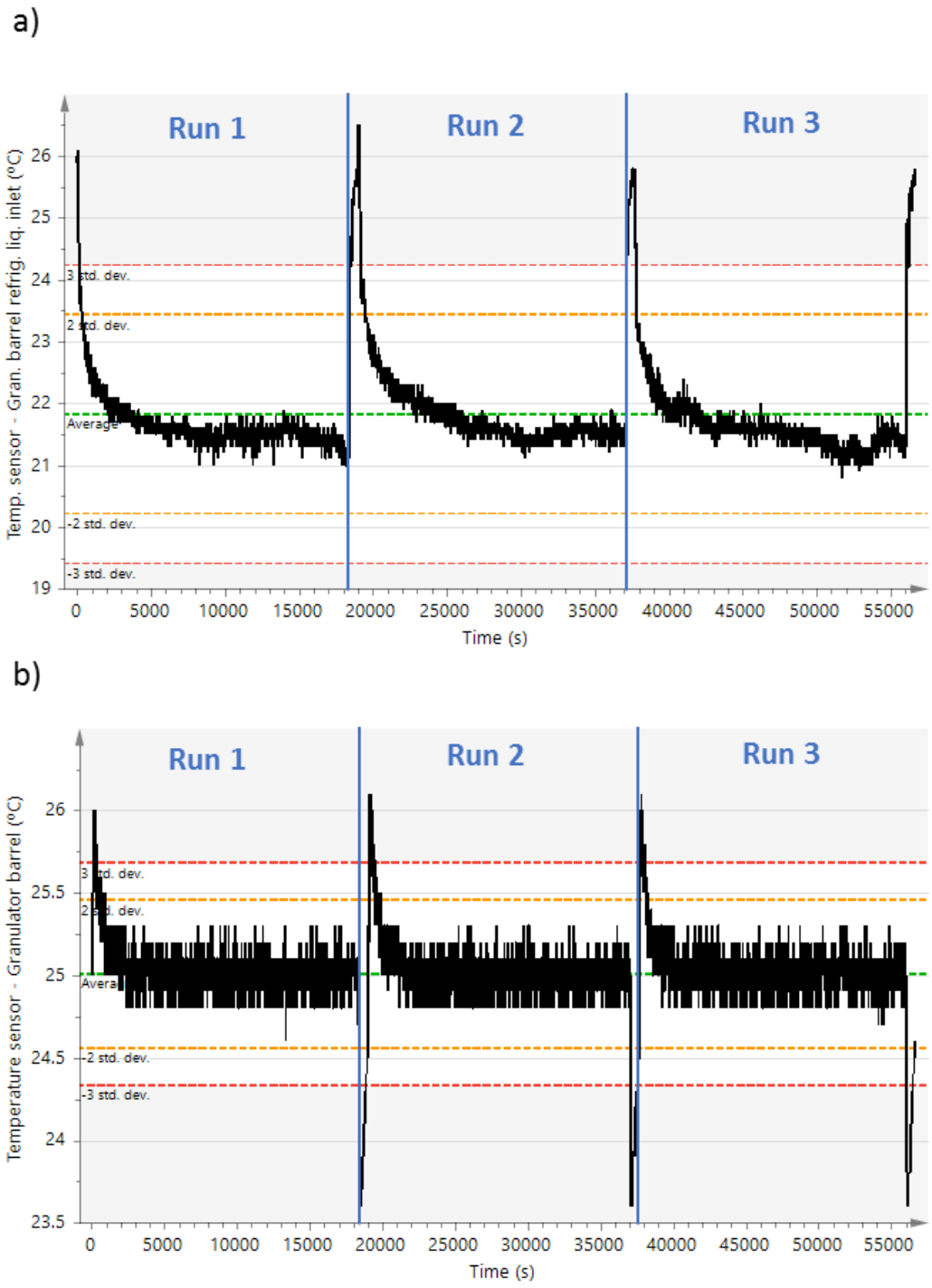


Figure 46 – Process overview – continuous granulator - Evolution during process time of the a) Temperature sensor - Granulator barrel refrigeration liquid inlet and the b) Temperature sensor – granulator barrel; (Orange dotted line – 95% confidence interval; Red dotted line – 99% confidence interval; Green dotter line – variable average value calculated from the calibration set);

Dryer and product control unit

A PCA model with 3 PCs was fitted for the logged dryer data capturing 55.1%, of the variance in the data (R^2) and with Q^2 of 44.4%.

Figure 47a shows that the PC1 scores increase as function of processing time. Figure 47d shows PC1 loadings. Many variables related to pressure contribute to this component. The differential pressure over the dryer filters is one of these variables and it increases slightly during the whole process time due to accumulation of fine material over the filters (Figure 48a). Special attention should be taken to this phenomenon as the blockage of filters might reduce fluidization and impact the drying effectiveness. The speed of the fan/blower revealed the same increasing trend as to compensate for the increasing pressure (Figure 48b). If at a certain moment filters would be completely covered with material, the speed of the fan/blower would reach 100% and still the air flow in the dryer would not reach the desired setpoint. This never happened during the presented experiments. For all runs, the temperature at the air outlet of the dryer markedly decreased during the first 1080 s. After reaching steady state, this temperature remained constant till the end of the process. Once the system is stopped, this temperature increased again to its initial values (Figure 48c). The decrease in temperature at the start of processing is probably due to the fact that during the first 1080 s the cells are filled for the first time and energy starts to be transferred to the granules for drying. This also explains why the temperature increased again when the dryer is emptied. Complementary too this, the relative humidity, measured at the same location, noticeably increased during the initial 1080 s. After the system is stopped, the humidity lowered again to the initial values (Figure 48d). Given that wet granules start being dried, the humidity of the air leaving the dryer increased since the moisture of the granules is transferred to the air. Correspondingly, when at the end of processing no granules are dried anymore, the humidity decreased.

The PC2 scores versus time show a 790 s startup period, followed by peaks at regular intervals of 180 s till the end of the process (Figure 47b). The 790 s of startup phase corresponds to the time period before the first cell is discharged. From that moment on, every 180 s a vacuum pump actuated to discharge the granules from a different dryer cell. The loadings (Figure 47d) reveal that this PC is mainly explaining variables related to the vacuum pump actuation every 180 s, i.e. to the pressure and air flow related variables inside the dryer and product control units.

PC3 (Figure 47c) explains how the temperatures and humidities at the outlet of the dehumidifier, inlet, and outlet of the dryer relate to the temperatures inside the six dryer cells and to the pressure inside the dryer (Figure 47d).

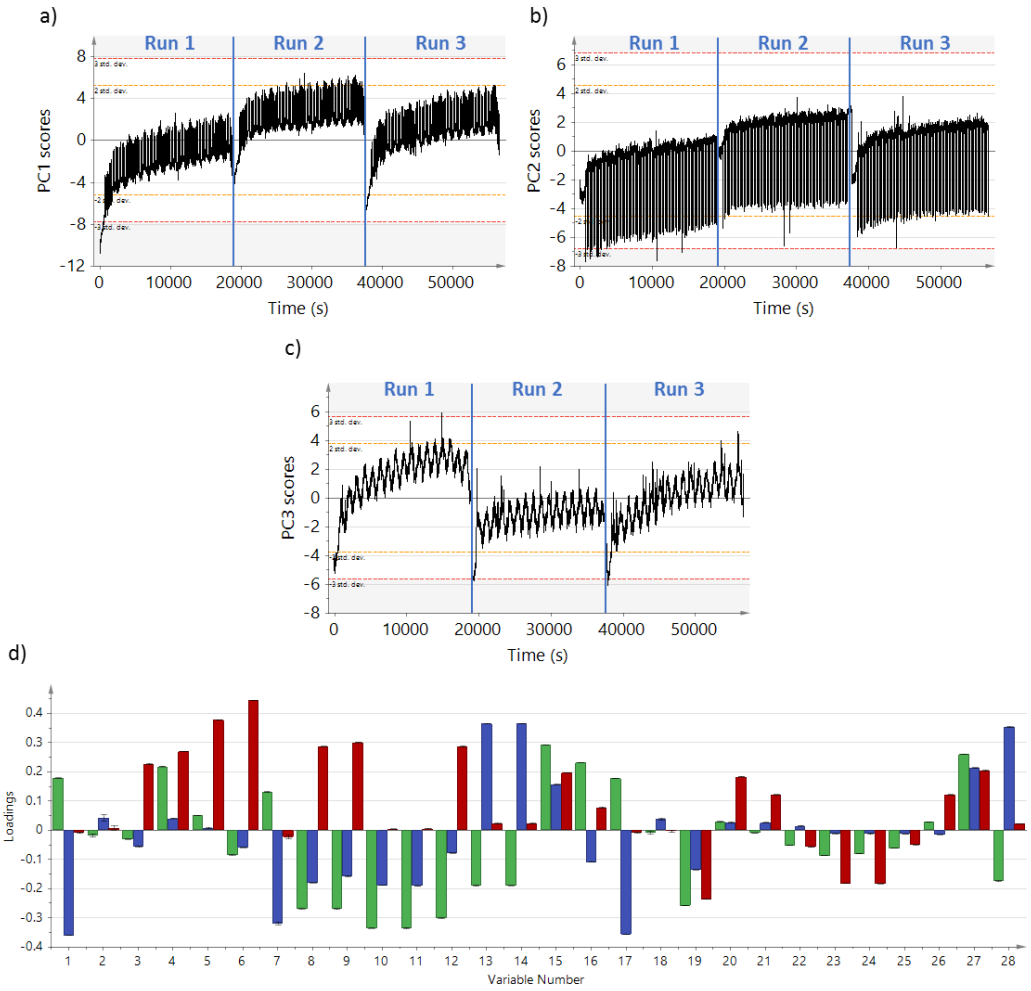


Figure 47 – Process visualization – dryer and product control units - a) Scores on PC1 versus process time; b) Scores on PC2 versus process time; c) Scores on PC3 versus process time; (Orange dotted line – 95% confidence interval; Red dotted line – 99% confidence interval); d) Loadings of PC1 to PC3 (PC1 –green; PC2 – blue; PC3 - red); 1 - Flow sensor - Dryer air; 2 - Flow sensor - Wet granule transfer; 3 - Humidity sensor - Dryer air inlet; 4 - Humidity sensor - Dryer air outlet; 5 - Humidity sensor - Dehumidifier air outlet; 6 - Temperature sensor - Dehumidifier air outlet; 7 - Pressure sensor - Differential pressure over the wet transfer line; 8 - Pressure sensor – Atmospheric; 9 - Pressure sensor - Dryer air inlet; 10 - Pressure sensor - Dryer air outlet (before HEPA filter); 11- Pressure sensor - Dryer air outlet (after HEPA filter); 12 - Pressure sensor - Pressure dryer top; 13 - Pressure sensor - product control unit (after HEPA filter); 14 - Pressure sensor - product control unit (before HEPA filter); 15 - Speed control –Fan/blower; 16 - Speed control - Push/fan; 17 - Speed control - Vacuum pump; 18 - Temperature sensor - Dryer air inlet; 19 - Temperature sensor - Dryer air outlet; 20 - Temperature sensor - Temperature dryer cell 1; 21 -Temperature sensor - Temperature dryer cell 2; 22 - Temperature sensor - Temperature dryer cell 3; 23 - Temperature sensor - Temperature dryer cell 4; 24 - Temperature sensor - Temperature dryer cell 5, 25 - Temperature sensor - Temperature dryer cell 6; 26 - Temperature sensor - Air handling unit filter; 27 - Pressure sensor - Differential pressure over the dryer filters; 28 - Pressure sensor - Wet transfer line.

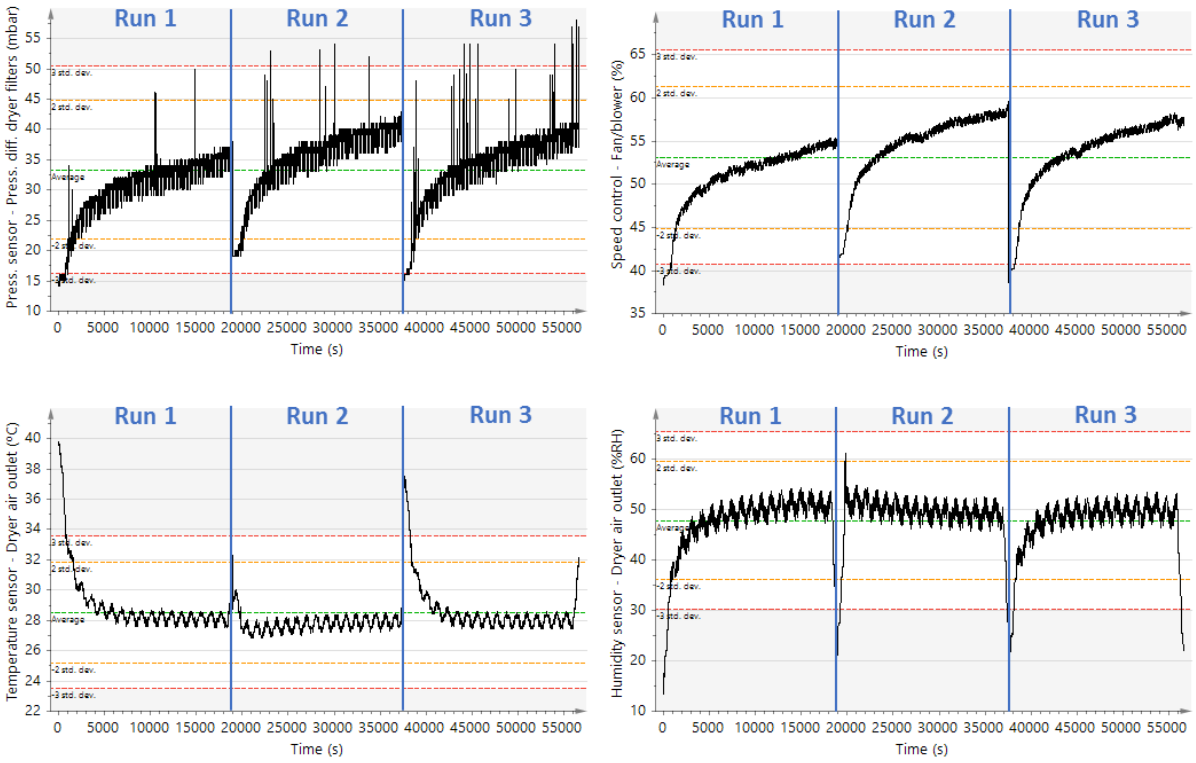


Figure 48 – Process overview – dryer and product control unit - Evolution during process time of the a) Pressure sensor - Differential pressure over the dryer filters, b) Speed control – Fan/blower, c) Temperature sensor - Dryer air outlet, d) Humidity - Dryer air outlet; (Orange dotted line – 95% confidence interval; Red dotted line – 99% confidence interval; Green dotter line – variable average value calculated from the calibration set).

All units

By fitting a PCA model on the data from all units together (granulator, dryer and product control unit), similar information was obtained as for the separate unit models, implying that inter-unit correlations between the measured variables do not exist. In the future, by implementing multivariate process analyzers during processing such as particle size analyzers and spectroscopic tools, additional production information will be obtained. This might enable the visualization of new correlations between the measured variables within and between the different units that are critical to the measured quality attributes. However, this study only aims at focusing on the univariate sensors, since no other process analyzers were implemented during the three experimental runs.

5.3.3.2 Multivariate statistical process control approach

Continuous granulator

An MSPC PCA model consisting of one principal component (explaining 32.8% of the variation in X and with an explained variance according to cross-validation (Q^2) value of 21.4% was fitted on the granulator reference data from runs 1 and 2.

The loadings for this PC (Figure 49a) reveal the contribution of 4 variables: power of the granulator drive and torque, both negatively correlated to the scores, screw speed and temperature at the inlet of the granulator jacket positively correlated to the scores. When the torque inside the barrel increases, more input power is supplied to maintain the speed of the screws which otherwise tends to decrease.

At the same time, temperature at the inlet of the granulator refrigeration liquid decreases due to the water entering inside the barrel walls which prevents its heating up. The developed MSPC model was then used for the monitoring of run 3. The aim was to test the performance of this model for the monitoring of future runs and evaluate its ability to detect deviations from the normal operation state defined by the reference runs 1 and 2. Scores, Hotelling's T^2 and DModX Shewhart control charts were plotted.

On the scores plot (Figure 49b), a few outliers are visible dispersed over the entire steady state processing time (approximately 4.5 h). By analyzing the contribution plots of these outliers (Figure 50), it can be observed that the variables contributing to the deviations are mainly increased power and torque of the granulator. In the Hotelling's T^2 control chart the same strong outliers are observed since only one component was fitted (Figure 49c). Even though these strong deviations were detected during processing, a trained operator would not feel the need to take action since normality was restored immediately in the following second in each outlier case. As alternative, it can also be defined that a deviation should only be considered after a certain number of consecutive deviations, or a percentage of deviations in a processing time interval. Being able to diagnose the detected deviation and make a decision on whether or not to react on is also an important part of an effective monitoring system. Data points with a high DModX lie further from the model plane and represent moderate outliers (Figure 49d). Overall, a significant number of observations are detected as moderate outliers in this plot probably because the reference model was built using only two reference runs which is not enough to include all possible normal variations of the process. More reference runs should be performed and included in the MSPC model.

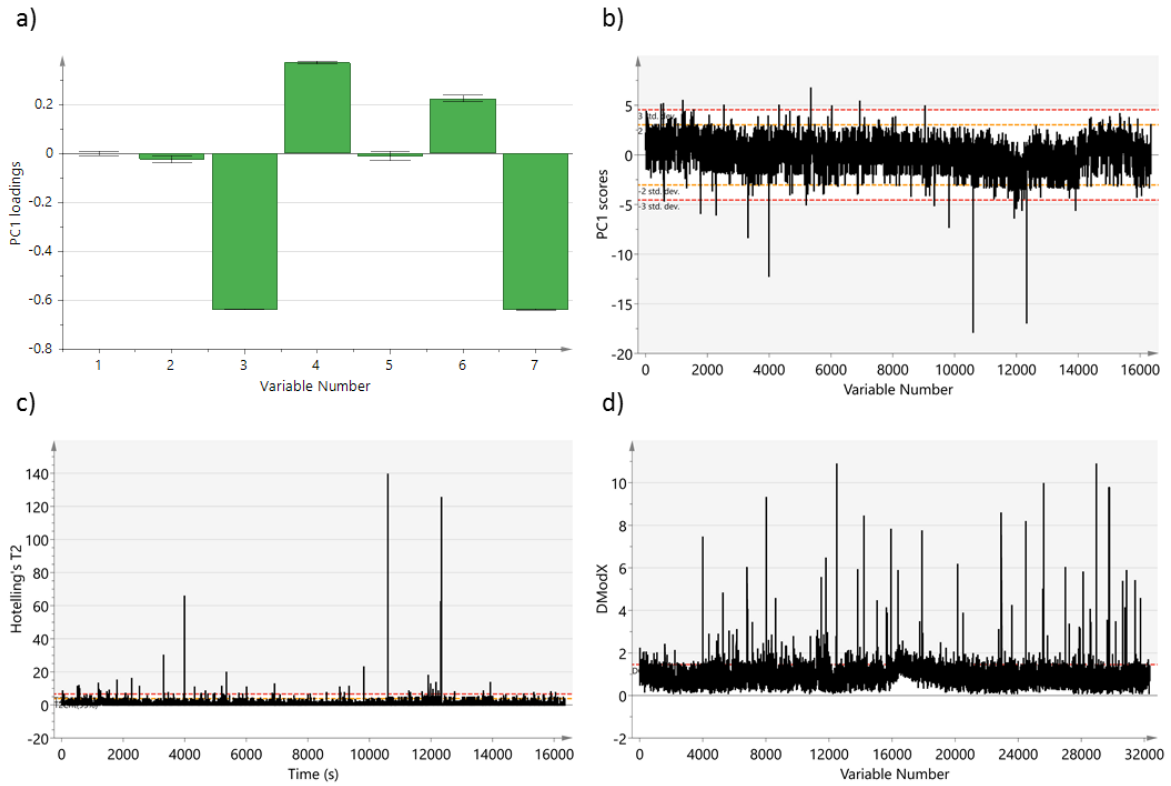


Figure 49 – MSPC approach – continuous granulator- a) Loadings of PC1; 1 - Mass flow - Granulation liquid; 2 - Mass - Flow powder dosing; 3 - Power - Granulator drive; 4 - Speed - Granulator screws; 5 - Temperature sensor - Granulator barrel; 6 - Temperature sensor - Granulator barrel refrigeration liquid inlet; 7 - Torque sensor – Granulator; b) Predicted scores on PC1 control chart (Orange dotted line – 95% confidence interval; Red dotted line – 99% confidence interval; c) Hotelling's control chart (Orange dotted line – 95% confidence limit; Red dotted line – 99% confidence limit); c) DModX control chart (Red dotted line – D_{crit}); Data from Run 3.

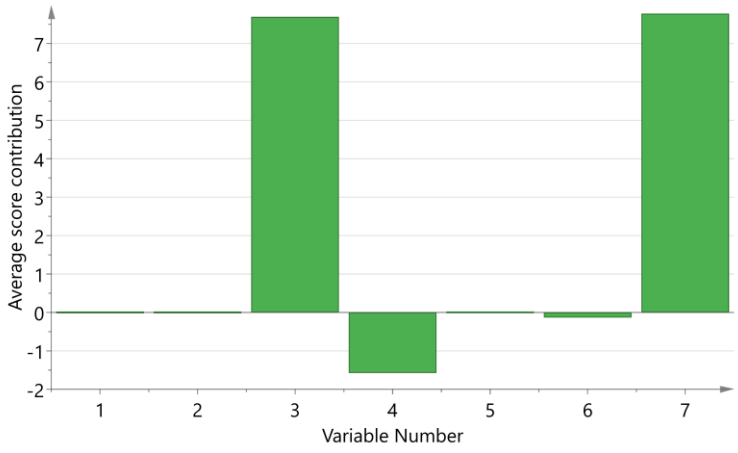


Figure 50 – MSPC approach – continuous granulator – average score contribution plot of all points outside the action limits; 1 - Mass flow - Granulation liquid; 2 - Mass - Flow powder dosing; 3 - Power - Granulator drive; 4 - Speed - Granulator screws; 5 - Temperature sensor - Granulator barrel; 6 - Temperature sensor - Granulator barrel refrigeration liquid inlet; 7 - Torque sensor – Granulator.

Dryer and product control unit

A MSPC PCA model of two PCs was developed for the steady state dryer data from runs 1 and 2 explaining 62.3% of the overall variation in the data (Q^2 of 36.9%). Adding more components to the model would result in a decrease of the Q^2 which indicates an overfitting. According to the PC1 loadings (Figure 51a), the temperature and humidity at the outlet of the dehumidifier and the humidity at the inlet of the dryer are positively correlated with the scores and with each other. This PC mainly represents the difference between the two calibration runs with all values for these variables being on average lower for run 2. In the PC2 loading plots (Figure 51a), it can be observed that the humidity at the outlet of the dryer is positively correlated to the scores and negatively correlated with the temperature at the same location. This PC represents the correlations during steady state operation of the dryer. When the air leaving the dryer carries more moisture removed from the granules, the temperature of this air is lower. Monitoring run 3 using the developed MSPC model (developed from runs 1 and 2) revealed no outliers in the PC1 and PC2 scores and Hotelling's T^2 control charts (Figure 51b to d). The DModX control chart (Figure 51e) revealed a single moderate outlier. Via a DModX contribution plot (Figure 52) it was possible to observe that at this time point the temperature at the filter of the dryer air handling unit was lower than during normal operation. However, this lasted for only a second and no corrective action was required.

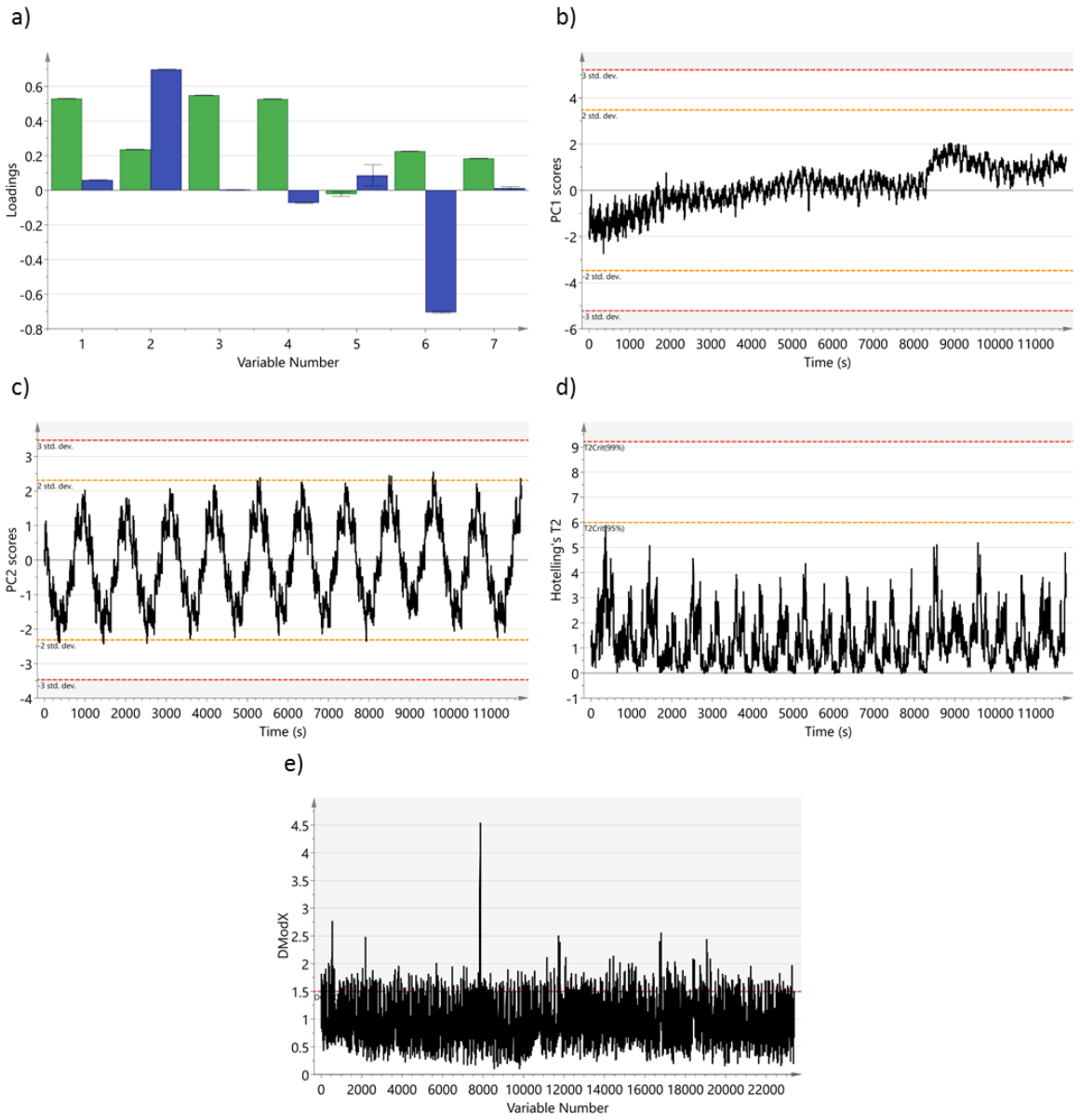


Figure 51 – MSPC approach – dryer and product control unit – a) Loadings of PC1 (green) and PC2 (blue); b) Predicted scores on PC1 control chart (Orange dotted line – 95% confidence interval; Red dotted line – 99% confidence interval) chart; c) Predicted scores on PC2 control chart (Orange dotted line – 95% confidence interval; Red dotted line – 99% confidence interval); d) Hotelling's control chart (Orange dotted line – 95% confidence limit; Red dotted line – 99% confidence limit); e) DModX control chart (Red dotted line – Dcrit); 1 - Humidity sensor - Dryer air inlet; 2 - Humidity sensor - Dryer air outlet; 3 - Humidity sensor - Dehumidifier air outlet; 4 - Temperature sensor - Dehumidifier air - outlet; 5 - Temperature sensor - Dryer air inlet; 6 - Temperature sensor - Dryer air outlet; 7 - Temperature sensor - Air handling unit filter; Data from Run 3.

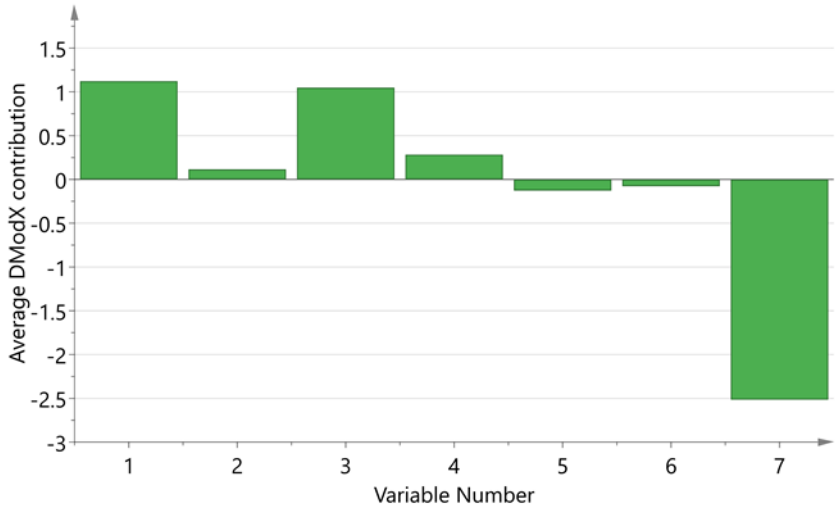


Figure 52 – MSPC approach – granulator dryer and product control unit -); Average DModX contribution plot of all points outside the action limits); 1 - Humidity sensor - Dryer air inlet; 2 - Humidity sensor - Dryer air outlet; 3 - Humidity sensor - Dehumidifier air outlet; 4 - Temperature sensor - Dehumidifier air - outlet; 5 - Temperature sensor - Dryer air inlet; 6 - Temperature sensor - Dryer air outlet; 7 - Temperature sensor - Air handling unit filter.

5.3.4 Conclusion

The PCA models on the data from the three repeated runs provided a good overview of a granulation and drying process performed in the ConsiGma-25™ system. In the PCA of the data acquired in the continuous granulator, three phases were identified: startup, steady state and shutdown. Several correlations were found between the variables measured at this location. During startup, the attrition generated by the powder filling of the barrel causes a slight increase in temperature. Torque and the power of the granulator drive also increase. The increased temperature is corrected by the temperature of the water inside the barrel. It takes 2000 s until the barrel temperature, torque and power of the granulator drive become constant around the setpoint and it remains so, until the end of the process. In the PCA of the dryer it was observed that, through the process, there is an increase in the pressure over the dryer filters due to the deposition of fines over them which may lead to an inefficient drying. To deal with this issue it is possible to set a number of "blow back" air pulses when a cell is emptied in order to prevent total blockage of the filters. In this analyses, a 790 s startup period is identified, after this startup period the actuation of the vacuum pump each 180 s is well observed in all pressure-related variables measured in the system. In the beginning of the process, the humidity at the outlet of the dryer increases and the temperature decreases due the wet granules entering the dryer, the opposite trend is observed at the end of the

process as the dryer is emptied. No correlations between variables from the granulator and variables measured at the dryer and product control units were found.

A PCA model for the MSPC of the continuous granulator was successfully developed. During the monitoring of the test run 3, several out-of-control spikes are observed in the Hoteling's T^2 and DModX control charts due to an increased power and torque, no actions were required as normal operation was immediately re-established (i.e. within 1 s). A PCA model for the MSPC of the dryer and product control unit was developed. During the monitoring of test run 3, a single moderate outlier was found in the DModX due to a temperature at the filter of the dryer air handling unit lower than during normal operation. Again, the deviation lasted for only a second and no corrective action was required. Contribution plots were found suitable to assign a probable cause for a deviation and allow a trained operator to take a corrective action to restore normal operation and prevent future complications.

5.4 Understanding and monitoring a continuous pharmaceutical twin-screw granulation and drying process using multivariate data analysis – part II[4]

5.4.1 Abstract

The present study is a second part of a work that aims at acquiring an in-depth process knowledge about a granulation and drying process performed on the commercially available ConsiGma-25™ continuous line. This second part describes not only the steady state operation of the line but also the startup of each line submodule.

Furthermore, the study focuses on the use of batch BSPC principles to model the variables logged by the system (e.g. temperatures, pressures, etc.) with a relevant time-related trajectory. This is a part of a proposed MSPC/BSPC monitoring strategy where, in addition to BSPC, MSPC based on principal component analysis is used to monitor variables varying around a steady state, i.e. that do not present a time-dependent trajectory. These logged variables are related to process parameters and can impact the product quality. Ensuring that there are no deviations from the normal process trajectories of their variables is a way of minimizing product variability and maximizing quality. A profound process understanding regarding the normal operation of the dryer and product control units as well as an in-depth description of the startup period of the different units were achieved. An approach was developed which allows to monitor the drying process, to detect and diagnose deviations from normal operation and assign possible causes for the disturbances.

[4] This chapter has been adapted form: A.F.T. Silva, J. Vercruysse, M.C. Sarraguça, J.L.F. Costa Lima, C. Vervaet, J.P. Remon, J.A. Lopes, T. De Beer, Understanding and monitoring a continuous pharmaceutical twin-screw granulation and drying process using multivariate data analysis – part II, (Submission in process).

5.4.2 Materials and methods

5.4.2.1 Industrial case study

Three repeated continuous granulations (five hours each) were completed in the ConsiGma™-25 system. A dry preblend of two API's, powdered cellulose, maize starch, pregelatinized starch and sodium starch glycolate was granulated with distilled water. Powders were provided by Johnson&Johnson, Janssen-Cilag, Italy.

5.4.2.2 *Continuous manufacturing line*

The experiments in this study were performed in a ConsiGma™-25 line located at GEA Pharma Systems (Wommelgem, Belgium). The system was previously described in subsection ConsiGma™-25 continuous line (Chapter 5). In this particular assembly, granules are transported gravimetrically between the different subunits. A powder feeder working at a constant speed of 20 kg/h fed the dry premix to the granulator. Screw speed of the granulator screws was set to 900 rpm and barrel temperature was set and 25°C. Granules exiting the granulator were transported to the six-cells segmented fluid bed dryer through a vacuum transport line. Each cell is sequentially filled after the other for 180 s. Drying time for the granules in each cell was set to 790 s in total, including the 180 s filling time. During the last seconds, granules followed to the product control unit. After emptying, the cell remained inactive for 290 s until wet granules were again loaded in the cell for drying. A mill was placed at the exit of the control unit but it was not used for the purpose of this study. The air entering the fluid bed dryer is brought to the setpoints by an air handing unit. A dehumidifier inside this unit removes moisture from the incoming air. A push/fan and fan/blower system regulated respectively air flow and pressure within the dryer. Several HEPA filters were placed in the air outlets of the dryer and product control, avoiding that particulate material exits the system.

5.4.2.3 In-process measurements

The univariate process parameters relevant to this exercise, that were continuously measured and logged each second at multiple locations of the ConsiGma™-25, are summarized in Table 9. Both user-set variables (setpoints) and other (not set but only measured) logged variables (open loop variables) are logged. Setpoints are pre-defined yet

they vary around their set value due to disturbances which are corrected by the system itself by means of independent PID controllers.

Table 9 – Parameters logged by the ConsiGma™-25 during processing; ✪ - Variables included in the OPLS model for unveiling time-dependency of variables; ◆ - Variables included in the BSPM model (PLS) for monitoring the continuous granulator startup; ◆ - Variables included in the BSPM model (PLS) for monitoring the dryer startup; ✚ - Variables included in the BSPM model (PLS) for monitoring the dryer and included in the BSPM model (PLS) for monitoring the dryer startup period.

Description	Units	Type	ConsiGma™-25 unit
Mass flow - Granulation liquid ✪◆	g/min	Setpoint	Liquid addition module (granulator)
Mass - Flow powder dosing ✪◆	kg/h	Setpoint	Powder dosing uni(granulator)
Speed - Granulator screws ✪◆	rpm	Setpoint	Granulator
Temperature sensor - Granulator barrel ✪◆	°C	Setpoint	Granulator
Power - Granulator drive ✪◆	W	Measured	Granulator
Temperature sensor - Granulator barrel refrigeration liquid inlet ✪◆	°C	Measured	Granulator
Torque sensor – Granulator ✪◆	Nm	Measured	Granulator
Flow sensor - Wet granule transfer ✪✚	m³/h	Setpoint	Wet granule transfer line (dryer)
Pressure sensor - Differential pressure over the wet transfer line ✪✚	mbar	Measured	Wet granule transfer line (dryer)
Pressure sensor - Wet transfer line ✪✚	mbar	Measured	Wet granule transfer line (dryer)
Temperature sensor - Air handling unit filter ✪✱	°C	Measured	Air handling unit (dryer)
Temperature sensor - Dehumidifier air outlet ✪✱	°C	Measured	Air handling unit (dryer)
Humidity sensor - Dehumidifier air outlet ✪✱	% RH	Measured	Air handling unit (dryer)
Speed control –Fan/blower ✪✚	%	Measured	Fan/Blower system (dryer)
Speed control - Push/fan ✪✚	%	Measured	Fan/Blower system (dryer)
Flow sensor - Dryer air ✪✚	m³/h	Setpoint	Dryer
Humidity sensor - Dryer air inlet ✪✱	% RH	Setpoint	Dryer
Temperature sensor - Dryer air inlet ✪✱	°C	Setpoint	Dryer
Humidity sensor - Dryer air outlet ✪✱	% RH	Measured	Dryer
Pressure sensor – Atmospheric ✪✚	mbar	Measured	Dryer
Pressure sensor - Differential pressure over the dryer filters ✪✚	mbar	Measured	Dryer
Pressure sensor - Dryer air inlet ✪✚	mbar	Measured	Dryer
Pressure sensor - Dryer air outlet (after HEPA filter) ✪✚	mbar	Measured	Dryer
Pressure sensor - Dryer air outlet (before HEPA filter) ✪✚	mbar	Measured	Dryer
Pressure sensor - Pressure dryer top ✪✚	mbar	Measured	Dryer
Temperature sensor - Dryer air outlet ✪✚	°C	Measured	Dryer
Temperature sensor - Temperature dryer cell 1 ✪✚	°C	Measured	Dryer
Temperature sensor - Temperature dryer cell 2 ✪✚	°C	Measured	Dryer

Table 9 – Parameters logged by the ConsiGma™-25 during processing; **O** - Variables included in the OPLS model for unveiling time-dependency of variables; ◆ - Variables included in the BSPM model (PLS) for monitoring the continuous granulator startup; ◆ - Variables included in the BSPM model (PLS) for monitoring the dryer startup; ✚ - Variables included in the BSPM model (PLS) for monitoring the dryer and included in the BSPM model (PLS) for monitoring the dryer startup period.

Description	Units	Type	ConsiGma™-25 unit
Temperature sensor - Temperature dryer cell 3 **O✚**	°C	Measured	Dryer
Temperature sensor - Temperature dryer cell 4 **O✚**	°C	Measured	Dryer
Temperature sensor - Temperature dryer cell 5 **O✚**	°C	Measured	Dryer
Temperature sensor - Temperature dryer cell 6 **O✚**	°C	Measured	Dryer
Speed control - Vacuum pump **O✚**	%	Measured	Vacuum pump (dryer)
Pressure sensor - product control unit (after HEPA filter) **O✚**	mbar	Measured	Discharger
Pressure sensor - product control unit (before HEPA filter) **O✚**	mbar	Measured	Discharger

5.4.2.4 Time series analysis

A time series can be defined as a succession of observations measured sequentially in time. Each of the variables measured by the ConsiGma™-25 can hence be represented as a time series.

In a complex system as the ConsiGma™-25, it is possible that a change or disturbance affecting one variable can impact other variable(s) at the same or different time points. Therefore, evaluating correlations between pairs of variables being one of them time-lagged (i.e. delayed) is a method to achieve better process understanding. The intention of performing this cross-correlation analysis was to later take into account the found optimal time lags in the development of the BSPC strategy in order to account for process dynamics. Variables were de-trended and correlations were calculated using the Fast Fourier Transform algorithm. All possible pairs of variables were tested either within the same sub-module or between the different sub-models of the ConsiGma™-25 system. For each variable pair, one of the two variables was kept fixed at its original time point while the other variable was shifted up to 1080 s before and after this moment. Therefore, a time lag in this context is always a measure of the delay or advance of the second variable in the pair in relation to the first variable. The 1080 s period corresponds to the maximum residence time i.e. the maximum interval since the moment the material is fed until the dry granules are discharged from the product control unit. The 1080 s are the maximum interval possible as a cell needs to be empty before being filled again. This cross-correlation analysis was performed using SIMCA™ 14 (MKS Data Analytics Solutions, Umeå, Sweden).

5.4.2.5 Multivariate data analysis

Collinear variables in a system can be combined into latent variables called principal components (PC), describing the underlying structure of the data. This approach is generally used in the pharmaceutical industry, and can be helpful in acquiring process knowledge as discussed in detail by Kourti in 2006 [193]. Latent variable methods include for example PCA, PLS regression and their various extensions such as multiblock PCA/PLS and OPLS. In this work, different latent vari(able approaches were used for multiple purposes as described below. All used data sets were first mean-centered and scaled to unit variance. All of the below described models were developed and evaluated using SIMCA™ 14 (MKS Data Analytics Solutions, Umeå, Sweden). For all PLS models a R^2 value and a Q^2 value were calculated. R^2 corresponds to the variation in the data that is captured by the model while Q^2 is the variation predicted by the model according to cross-validation. For cross-validation every seventh observation was assigned to a cross-validation group, each group being excluded at a time in each cross-validation round. Components were added to the models until both R^2 and Q^2 were maximized. Adding more components to the model would results in an overfitting.

Time dependency

In theory, a continuous system is expected to achieve steady operation where variables are unchanging in time. In many cases, this steady state is only reached some time after the continuous process has been started (startup phase). To unveil only long-term drifts, the logged data corresponding to the startup and shut-down phases of the continuous process were excluded and only data from the expected steady state phase was used to build an OPLS model. These logged variables from all units were regressed against process time (Y) to uncover the time-related variables. Table 9 overviews the variables included in this analysis. The advantage of OPLS in comparison to regular PLS is that it allows a better interpretability of the model while keeping the same predictive ability as regular PLS. In an OPLS model, information correlated to Y is separated from the uncorrelated information. While the information correlated to Y is captured in the predictive PC, the uncorrelated information is captured by the orthogonal PCs [53].

Batch statistical process control approach

The ConsiGma™-25 is an overall continuous process where some sub-steps operate continuous (granulation) while other sub-modules operate semi-continuously (drying and product control modules). Therefore, a BSPC for variables containing relevant time-related trends is proposed in a mixed solution to model and monitor this continuous manufacturing line. Variables without a time-dependence were described in section 5.3 Understanding and monitoring a continuous pharmaceutical twin-screw granulation and drying process using multivariate data analysis – part I. Moreover, BSPC is also presented as a solution for monitoring the startup phases of each module.

Dryer and product control unit

There are variables in the dryer that present time-related (i.e. drying process related) features that are of interest for statistical process control purposes. Such three-dimensional data (N batches x J variables x K time points) should be modeled using BSPC PLS based models.

The time-dependent variables in the six-cells segmented dryer are the temperatures individually measured inside each drying cell which present a singular drying trajectory. When the cell is filled with material, this corresponds to the product temperature. Furthermore, the operation of the vacuum pump for granule discharge, due to the pressure overshoot it causes, is observable in all variables related to air flow and pressures measured in the transfer lines, dryer and product control units. This pressure overshoot occurs every 180 s and cannot be monitored using the regular PCA MSPC approach since it is a time-related feature. The variables included in the development of the dryer PLS BSPC models can be found in Table 9. Individual PLS models were fitted to the data from each cell in order to develop a different model for each cell. In addition, to obtain a global model capable of monitoring all cells, the data from all cells was included in the development of an overall dryer BSPC model. Of the three available runs, runs 1 and 2 were included in the model (reference runs) while run 3 was utilized to test the model's monitoring performance. Since the first cell discharges only 790 s after the start of the process and since after that moment a new discharge of a cell happens every 180 s, the data measured during the first fill of the first four cells (first 720 s of processing) was not included in the model as it can be considered as startup of the dryer operation. The three-dimensional batch data matrices were unfolded

resulting in two-dimensional matrices where batch and time dimensions were combined, creating matrices with N×K observations (rows) and J variables (columns). PLS regression was performed relating the data matrix with a vector depicting process time indicative of the maturity of the batch [221]. This unfolding methodology was chosen since it is more sensitive to the detection of variations and deviations from normality in process monitoring [88].

Continuous granulator, dryer and product control unit startup periods

According to Kourti et al., the BSPC approach is not only useful to monitor batch processes but also for transitions such as process startup, restarts, grade to grade transitions, etc. [49, 221]. Thus, in addition to the PLS models for the monitoring of the operation of each of the dryer cells, additional PLS models were built: one for the monitoring of the granulator startup (first 2000 s of operation) and others for the monitoring of the startup of each of the dryer cells (including both time invariant and time dependent variables). For the latter, the first fill of the first four cells were used. Of the three available runs, the data from runs 1 and 2 were used to develop the models (reference runs) while the data from run 3 was utilized to test the model's monitoring performance.

Fault detection and diagnosis

All above described developed models were utilized to monitor run 3. In all cases, run 3 data was projected into the corresponding models and it was verified whether this run was within normal operation conditions.

Shewhart score control charts provide a valuable visual representation of the process evolution [220]. Control limits for the score values were set at + and – 2 times the standard deviation (i.e. within a 95% confidence interval) corresponding to a warning limit and + and - 3 times the standard deviation (i.e. within a 99% confidence interval) corresponding to an action limit. Hotelling's T^2 and DModX values can also be utilized for pinpointing deviations and summarize the variation across all PCs in the model in a single value per process time point. The equations to calculate these metrics can be found in section Multivariate control charts (Chapter 2).

Deviations that are explained by the model (i.e. that are inside the model plane) are seen in the Hotelling's T^2 values. For all models the control limits in the Hotelling's T^2 control

charts were set at + and – 2 times the standard deviation (i.e. within a 95% confidence interval) corresponding to a warning limit and + and - 3 times the standard deviation (i.e. within a 99% confidence interval) corresponding to an action limit). Alternatively, DModX represents the distance to the model plane and allows to diagnose faults not explained by the model i.e. falling out of the model plane.

Observations with high DModX values indicate for that time point a breakage in the correlation structure of the variables. DModX control limits were set at D_{crit}. For the PLS models, the DModX control limits were set at 3 times the standard deviation (99% confidence interval). DModX values were normalized by dividing them by the pooled relative standard deviation of the model in the X.

When any of the above described control charts give an out-of-control signal at a certain time point, then contribution plots were calculated to unravel the variables contributing to that deviation, hence providing a good insight into the causes of the process deviations [82].

5.4.3 Results and discussion

5.4.3.1 Time series analysis

Within the sub-units of the ConsiGma™-25, several correlated measured variables were found as represented in Table 10. Furthermore, correlated variables were also found between the dryer and the product control unit. In a complex system as the ConsiGma™-25, it is possible that a change or disturbance affecting one variable can impact other variable(s) at the same or different time points. Therefore, correlations between pairs of variables being one of them time-lagged (i.e. delayed) were also examined allowing to find the optimal time lags between variables (Table 10) for the development of the multi-model statistical process control strategy in order to account for process dynamics. To include these time-lags in the latent variable models, one of the variables should be defined as reference while all other lags are to be recalculated according to this reference. However, this was not possible since it was observed that the lags do not fit together once recalculated according to this reference. For better understanding a practical example is hereby described. According to Table 10, the pressure at the product control unit after the HEPA filter- correlates optimally with the flow of the incoming dryer air at a lag time value of 2 s (line 2 in Table 10) and the differential pressure over the dryer filters correlates optimally with the flow of the incoming dryer air at a

lag value of 2 s (line 22 in Table 10). Given this, the pressure at the product control unit after the HEPA filter and the differential pressure over the dryer filters should be optimally correlated at a lag value of 0 s while in truth the optimal lag (which maximized the correlation coefficient) found was of 4 s (line 4 in Table 10). It was also observed that the correlation coefficients found between the pairs of variables only changed slightly (few seconds) with the change in the lag value. Therefore, it was decided to keep the original time-points for all variables.

Table 10 – Correlations between the measured variables (CC - correlation coefficient).

Fixed variable 1	Lagged variables	Lag	CC
Pressure sensor - product control unit (after HEPA filter)	Pressure sensor - product control unit (before HEPA filter)	0	1.00
Pressure sensor - product control unit (after HEPA filter)	Flow sensor - Dryer air	2	-0.92
Pressure sensor - product control unit (before HEPA filter)	Flow sensor - Dryer air	2	-0.92
Pressure sensor - product control unit (after HEPA filter)	Pressure sensor - Differential pressure over the dryer filters	4	0.64
Pressure sensor - product control unit (before HEPA filter)	Pressure sensor - Differential pressure over the dryer filters	4	0.64
Pressure sensor - Pressure dryer top	Pressure sensor - product control unit (after HEPA filter)	2	0.65
Pressure sensor - Pressure dryer top	Pressure sensor - product control unit (before HEPA filter)	2	0.65
Pressure sensor - Differential pressure over the wet transfer line	Pressure sensor - product control unit (after HEPA filter)	2	-0.79
Pressure sensor - Differential pressure over the wet transfer line	Pressure sensor - product control unit (before HEPA filter)	2	-0.79
Speed control - Vacuum pump	Pressure sensor - product control unit (before HEPA filter)	2	-0.98
Speed control - Vacuum pump	Pressure sensor - product control unit (after HEPA filter)	2	-0.98
Temperature sensor - Dryer air inlet	Humidity sensor - Dehumidifier air outlet	0	0.59
Pressure sensor - Differential pressure over the wet transfer line	Pressure sensor - Differential pressure over the dryer filters	0	-0.65
Pressure sensor - Dryer air outlet (before HEPA filter)	Pressure sensor - Dryer air outlet (after HEPA filter)	0	0.66
Pressure sensor - Pressure dryer top	Pressure sensor - Differential pressure over the dryer filters	0	0.72
Pressure sensor - Pressure dryer top	Pressure sensor - Wet transfer line	0	0.73
Pressure sensor - Differential pressure over	Pressure sensor - Pressure dryer top	0	-0.87

Table 10 – Correlations between the measured variables (CC - correlation coefficient).

Fixed variable 1	Lagged variables	Lag	CC
the wet transfer line			
Pressure sensor - Differential pressure over the wet transfer line	Pressure sensor - Wet transfer line	0	-0.89
Speed control - Vacuum pump	Pressure sensor - Pressure dryer top	1	-0.63
Speed control - Vacuum pump	Pressure sensor - Differential pressure over the wet transfer line	1	0.76
Speed control - Vacuum pump	Pressure sensor - Wet transfer line	1	-0.84
Pressure sensor - Differential pressure over the dryer filters	Flow sensor - Dryer air	2	-0.63
Pressure sensor - Pressure dryer top	Flow sensor - Dryer air	2	-0.83
Pressure sensor - Differential pressure over the wet transfer line	Flow sensor - Dryer air	2	0.84
Pressure sensor - Wet transfer line	Pressure sensor - product control unit (after HEPA filter)	2	0.87
Speed control - Push/fan	Flow sensor - Dryer air	3	0.52
Pressure sensor - Wet transfer line	Flow sensor - Dryer air	3	-0.93
Speed control - Fan/blower	Pressure sensor - Dryer air outlet (after HEPA filter)	4	-0.70
Speed control - Fan/blower	Pressure sensor - Dryer air outlet (before HEPA filter)	4	-0.84
Speed control - Vacuum pump	Flow sensor - Dryer air	4	0.89
Speed control - Vacuum pump	Pressure sensor - Differential pressure over the dryer filters	8	-0.62
Pressure sensor - Differential pressure over the dryer filters	Speed control - Fan/blower	20	0.77
Pressure sensor - Differential pressure over the dryer filters	Pressure sensor - Dryer air outlet (after HEPA filter)	25	-0.56
Pressure sensor - Differential pressure over the dryer filters	Pressure sensor - Dryer air outlet (before HEPA filter)	25	-0.67
Temperature sensor - Temperature dryer cell 3	Temperature sensor - Temperature dryer cell 4	175	0.99
Temperature sensor - Temperature dryer cell 4	Temperature sensor - Temperature dryer cell 5	178	0.99
Temperature sensor - Temperature dryer cell 2	Temperature sensor - Temperature dryer cell 3	180	0.99
Temperature sensor - Temperature dryer cell 1	Temperature sensor - Temperature dryer cell 2	182	0.98
Temperature sensor - Temperature dryer cell 5	Temperature sensor - Temperature dryer cell 6	184	0.99
Temperature sensor - Temperature dryer cell 3	sensor - Temperature dryer cell 5	354	0.98
Temperature sensor - Temperature dryer cell 2	Temperature sensor - Temperature	354	0.98

Table 10 – Correlations between the measured variables (CC - correlation coefficient).

Fixed variable 1	Lagged variables	Lag	CC
	dryer cell 4		
Temperature sensor - Temperature dryer cell 1	Temperature sensor - Temperature dryer cell 3	362	0.97
Temperature sensor - Temperature dryer cell 4	Temperature sensor - Temperature dryer cell 6	362	0.98
Temperature sensor - Temperature dryer cell 2	Temperature sensor - Temperature dryer cell 5	533	0.98
Temperature sensor - Temperature dryer cell 1	sensor - Temperature dryer cell 4	538	0.96
Temperature sensor - Temperature dryer cell 3	Temperature sensor - Temperature dryer cell 6	538	0.98
Temperature sensor - Temperature dryer cell 1	Temperature sensor - Temperature dryer cell 5	716	0.96
Temperature sensor - Temperature dryer cell 2	Temperature sensor - Temperature dryer cell 6	718	0.97
Temperature sensor - Temperature dryer cell 1	Temperature sensor - Temperature dryer cell 6	901	0.95
Torque sensor - Granulator	Power - Granulator drive	0	1.00

5.4.3.2 Time dependency

By fitting an OPLS model to the expected steady state data (excluding the data corresponding to startup and shut-down) from all ConsiGma units but individually for each run, it was possible to properly identify the time correlated variables. In each individual OPLS model of the different runs, one OPLS component was fit. Thus, the model had a predictive component and no orthogonal components. The variance in the logged data (X) captured by the model (R^2X), the variance in the Y-data captured by the model (R^2Y) and the variance predicted by the model according to cross-validation (Q^2) are overviewed in Table 11.

Table 11 – Fit of the OPLS models developed for the identification of time-related variables (# PC – number of principal components, R2X – variance in the X-data captured by the model; R2Y – variance in the Y-data captured by the model; Q2 – variance predicted by the model according to cross-validation).

Run	# PC	R^2X	R^2Y	Q^2
1	1	27.0	94.7	94.7
2	1	23.1	92.1	92.1
3	1	19.6	85.3	85.3

The OPLS scores of the predictive component versus time plot for run 1 is depicted in Figure 53a)). Similar plots were obtained for the OPLS models of runs 2 and 3 and are therefore not depicted. The loadings of the first OPLS component for all developed OPLS models (i.e, for runs 1, 2 and 3) are depicted on Figure 53b)). All variables contributing to the predictive component of time are variables that follow either an increasing or decreasing trend during processing. In this plot is was possible to identify two different groups of time-related variables: (i) time-related variables that follow the same trend for the three runs and (ii) time-related variables which behave differently in the different runs. The temperature of the air at the outlet of the dehumidifier increased during time in all three runs probably due to the heating up of the stainless-steel cover of the dryer caused by the operation of the line. Despite this, there is no visible influence on the temperature of the air entering the dryer cells since at the inlet of the air handling unit, the air temperature is corrected back to the setpoint. As already reported in part I of this work, the differential pressure over the dryer filters increased in all runs due to the deposition of fine powder particles in the filters. In addition, the pressure at the air outlet of the dryer, both before and after the HEPA filter, decreases over time which is most probably also a consequence of material build up in the filters. On the other hand, the speed of the fan/blower, responsible for the control of pressure inside the dryer increases its value to compensate for the observed decrease of pressure at the outlet of the dryer air. This phenomenon, as explained before, needs attention since it can cause the drying step to be inefficient. Lastly, an overall trend observed in all runs is the decrease of the temperature at the inlet of the refrigeration liquid in the granulator. Due to the continuous operation, heat generated by the friction and movement of material inside the granulator causes an increase of the barrel temperature, which is corrected back to its setpoint by the temperature control unit. This compensation is achieved by filling the barrel with colder water which is seen in the decreasing temperature at the inlet of refrigeration liquid. The atmospheric pressure in the production room shifted during processing differently for the three runs. In run 1 it followed a decreasing trend with a maximum of 1033 and a minimum 1024 mbar. In run 2 it also slightly decreased over time but only varied between 1026 and 1022 mbar. In run out of c3 the pressure remained constant around an average of 1021 mbar (with a maximum of 1023 and minimum 1019 mbar). As a consequence, also the pressures at the air inlet, top and outlet of the dryer varied differently for all runs. The humidity at the air outlet of the dehumidifier presented different trends for all runs, since the humidity of the air inside the production room was different during the three runs. Humidities at the air inlet and

outlet of the dryer were also influenced and revealed the same trend. The temperature of the air exiting the dryer presented the opposite trend: when more moisture is present at the outlet of the dryer, its temperature lowers.

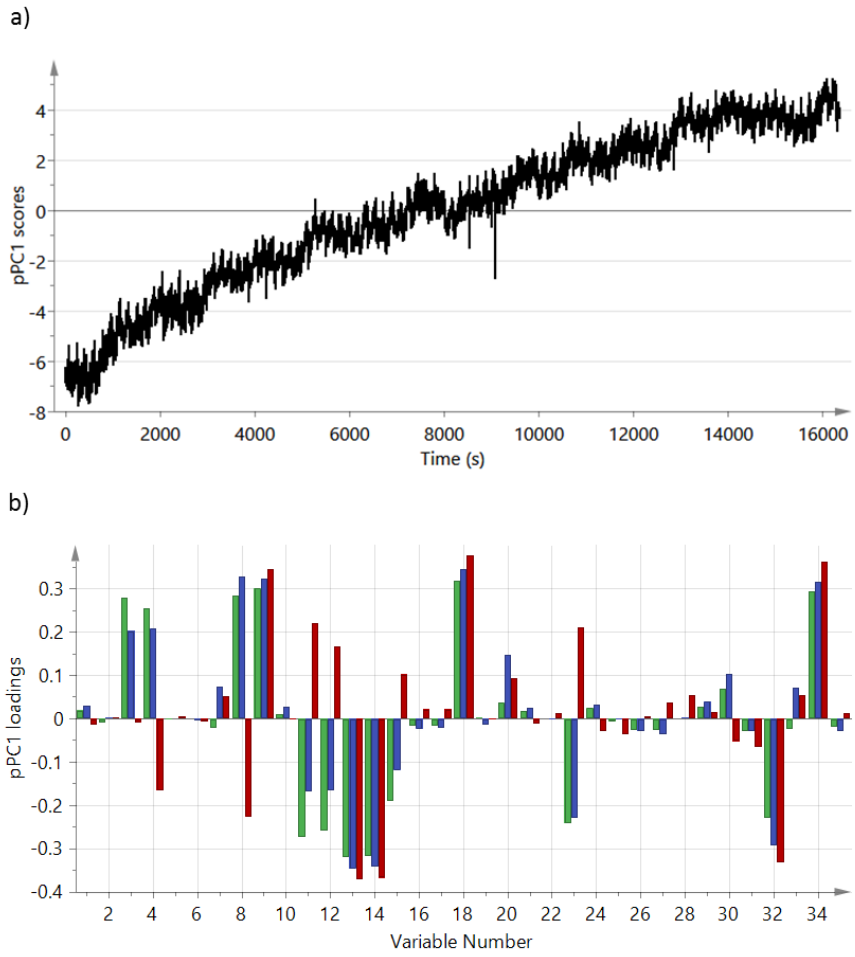


Figure 53 –Time dependency - a) Scores on predictive PC1 (Run 1); b) Loadings of PC1 (Run 1 – green; Run 2 – blue; Run 3 - red); 1 - Flow sensor - Dryer air; 2 - Flow sensor - Wet granule transfer; 3 - Humidity sensor - Dryer air inlet; 4 - Humidity sensor - Dryer air outlet; 5 - Humidity sensor - Dehumidifier air outlet; 6 - Mass - Flow granulation liquid; 7 - Mass - Flow powder dosing; 8 - Power - Granulator drive; 9 - Humidity sensor - Dehumidifier air outlet; 10 - Temperature sensor - Dehumidifier air outlet; 11 - Pressure sensor - Differential pressure over the wet transfer line; 12 - Pressure sensor – Atmospheric; 13 - Pressure sensor - Dryer air inlet; 14- Pressure sensor - Dryer air outlet (before HEPA filter); 15- Pressure sensor - Dryer air outlet (after HEPA filter); 16 - Pressure sensor - Pressure dryer top; 17 - Pressure sensor - product control unit (after HEPA filter); 18 - Pressure sensor - product control unit (before HEPA filter); 19 - Speed control –Fan/blower; 20 - Speed control - Push/fan; 21 - Speed control - Vacuum pump; 22 - Temperature sensor - Dryer air inlet; 23 - Temperature sensor - Dryer air outlet; 24 - Temperature sensor - Temperature dryer cell 1; 25 -Temperature sensor - Temperature dryer cell 2; 26 - Temperature sensor - Temperature dryer cell 3; 27 - Temperature sensor - Temperature dryer cell 4; 28 - Temperature sensor - Temperature dryer cell 5, 29 - Temperature sensor - Temperature dryer cell 6; 30 - Temperature sensor - Air handling unit filter; 31 – Temperature sensor - Granulator barrel; 32 - Temperature sensor - Granulator barrel refrigeration liquid inlet; 33 - Torque sensor – Granulator; 34 - Pressure sensor - Differential pressure over the dryer filters; 35 - Pressure sensor - Wet transfer line.

5.4.3.3 Batch statistical process control approach

Dryer and product control unit – individual cell models

Table 12 summarizes the PLS models fitted on the univariate data from the dryer and product control units involved in the monitoring of each individual dryer cell (Table 8; 6 cells = 6 models).

Table 12 – Fit of the BSPC PLS models developed for the monitoring of the individual dryer cells and product control unit (# PC – number of principal components, R2X – variance in the X-data captured by the model; R2Y – variance in the Y-data captured by the model; Q2 – variance predicted by the model according to cross-validation).

Cell #	# PC	R^2X	R^2Y	Q^2
1	2	46.9	59.1	59.1
2	2	45.1	57.3	57.3
3	2	43.4	57.0	57.0
4	2	42.4	61.5	61.5
5	2	42.3	64.5	64.5
6	2	44.7	58.4	58.4

As previously described on the section Dryer and product control unit**,** variables from both dryer and product control unit variables that cannot be monitored using the regular PCA MSPC are included in these models. According to the loadings of PC1 (Figure 54a)) the temperature inside the dryer cell is the variable that mostly contributes to this PC. Loadings of PC2 (Figure 54b)) reveal that almost all variables included in the model (temperature inside the dryer cell and several pressure-related variables) contribute to this PC with the exception of the air flow inside the wet granule transfer line and the speed control of the push/fan.

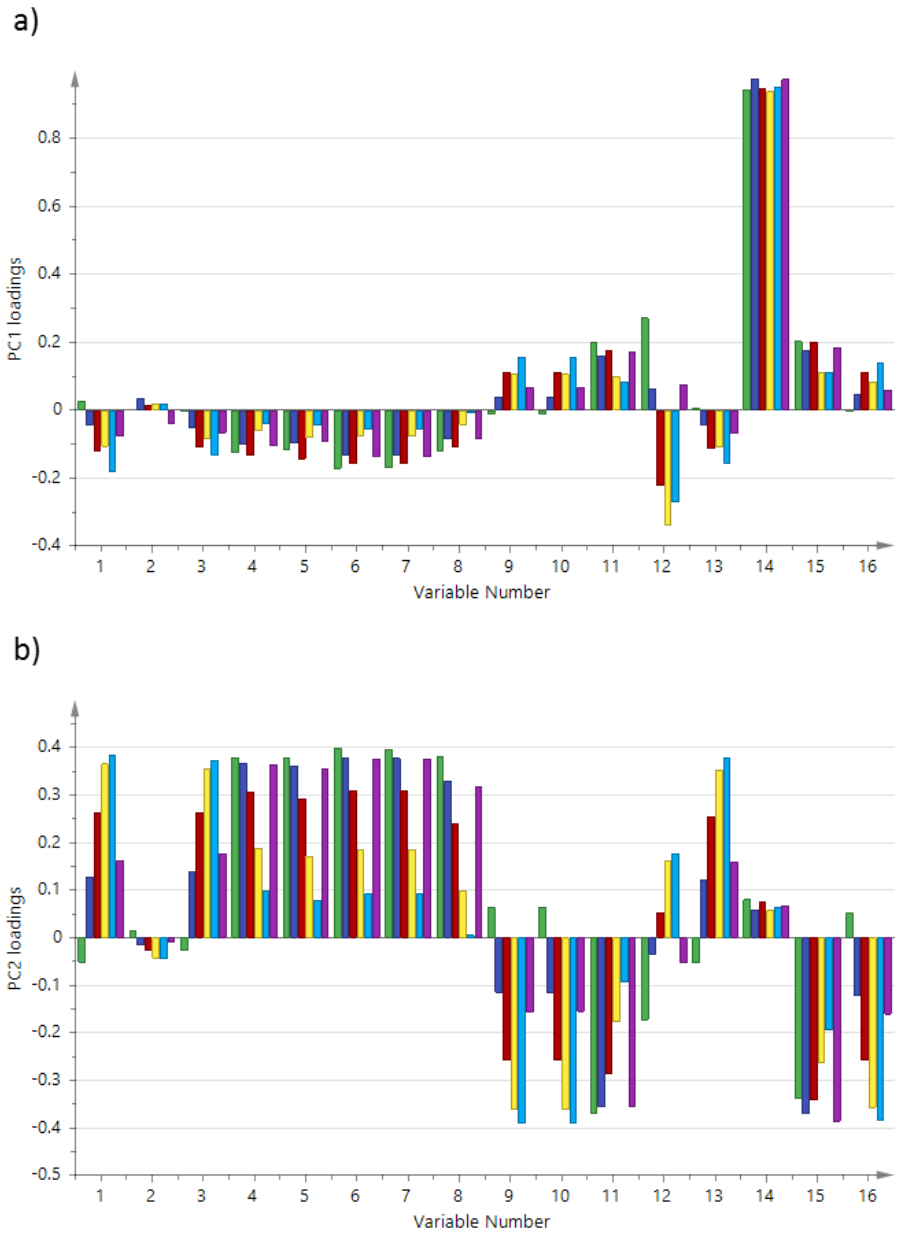


Figure 54 – Dryer and product control units – individual cell models - a) Loadings of PC1; b) Loadings of PC2 (cell 1 –green; cell 2 – blue; cell 3 – red; cell 4 – yellow, cell 5 – cyan, cell 6 - purple); 1 - Flow sensor - Dryer air; 2 - Flow sensor - Wet granule transfer; 3 - Pressure sensor - Differential pressure over the wet transfer line; 4 - Pressure sensor – Atmospheric; 5 - Pressure sensor - Dryer air inlet; 6 - Pressure sensor - Dryer air outlet (before HEPA filter); 7- Pressure sensor - Dryer air outlet (after HEPA filter); 8 - Pressure sensor - Pressure dryer top; 9 - Pressure sensor - product control unit (after HEPA filter); 10 - Pressure sensor - product control unit (before HEPA filter); 11 - Speed control – Fan/blower; 12 - Speed control - Push/fan; 13 - Speed control - Vacuum pump; 14 - Temperature sensor - Temperature dryer cell; 15 - Pressure sensor - Differential pressure over the dryer filters; 16 - Pressure sensor - Wet transfer line.

The data from each dryer cell of run 3 was projected into each individual cell model and scores, Hotelling's T^2 and DModX control charts were created. Figure 55a) to d) depicts these plots for the cell 3 model as an example. It can be observed in Figure 55a) that some of the projected PC1 trajectories from other cells deviate at times from the control limits. This is caused by the fact that even if the cells follow a similar drying temperature profile there are still slight differences from cell to cell. On PC2 no out-of-control moments were visible. Deviations from normal operation were also observed in the Hotelling's T^2 and DModX plots. According an average Hotelling's T^2 and DModX contribution plots of the moments outside the action limit (Figure 56a) and b)), the out-of-control moments are due to the pressure inside the wet granules transfer line being higher or lower than expected at the moments of the actuation of the vacuum pump. According to the manufacturer and as observed by Jurgen et al. [161] the differences between the drying temperature profiles from the different cells are small enough not to have an impact on the drying of the granules. Therefore, to simplify the monitoring strategy a global model containing information from all cells was also built.

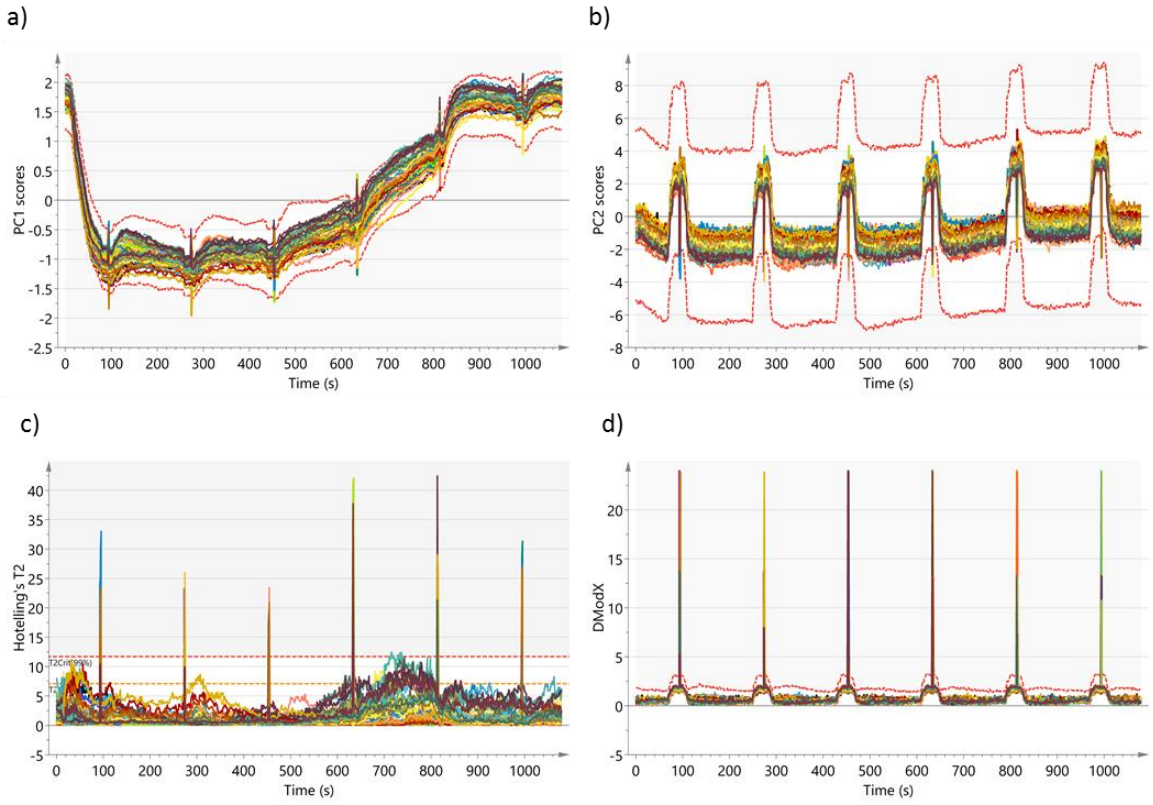


Figure 55 – Dryer and product control units – individual cell models - a) Predicted scores on PC1 control chart (Red dotted line – 99% confidence interval); b) Predicted scores on PC2 control chart; (Red dotted line – 99% confidence interval); c) Hotelling's T2control chart (Orange dotter line – 95% confidence limit, Red dotted line – 99% confidence limit); d) DModX control chart (Red dotted line – 99% confidence limit); Data from Run 3.

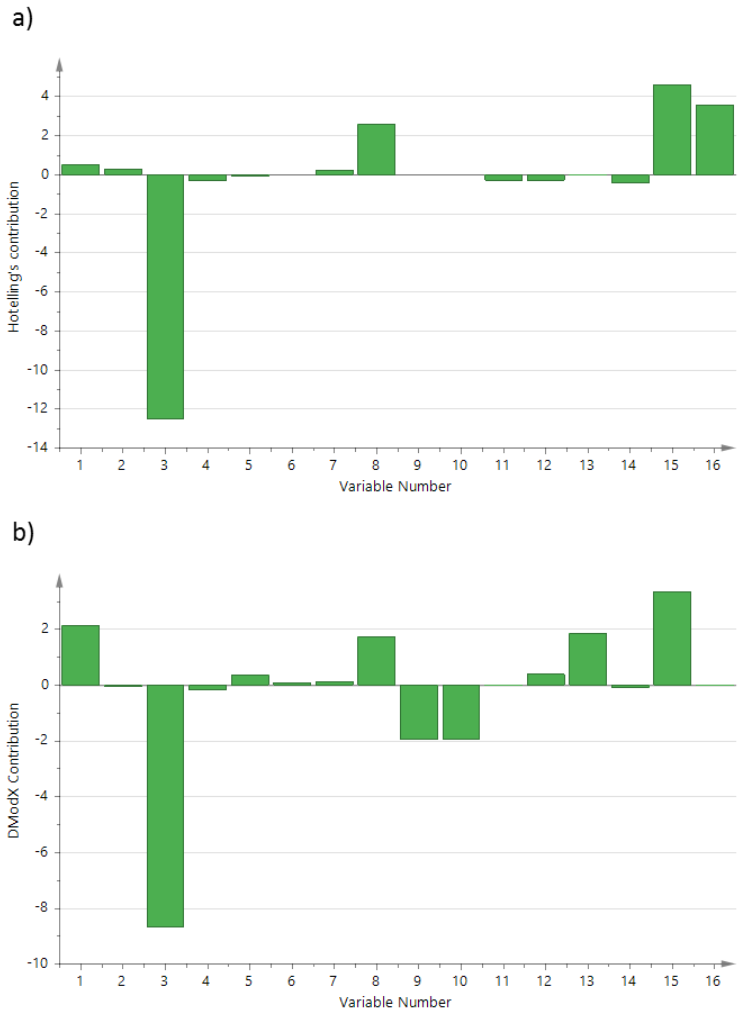


Figure 56 – Dryer and product control units – individual cell models - a) Average contribution to the out-of-control moments observed in the Hotelling's control charts; b) Average contribution to the out-of-control moments observed in the DModX control charts; 1 - Flow sensor - Dryer air; 2 - Flow sensor - Wet granule transfer; 3 - Pressure sensor - Differential pressure over the wet transfer line; 4 - Pressure sensor – Atmospheric; 5 - Pressure sensor - Dryer air inlet; 6 - Pressure sensor - Dryer air outlet (before HEPA filter); 7- Pressure sensor - Dryer air outlet (after HEPA filter); 8 - Pressure sensor - Pressure dryer top; 9 - Pressure sensor - product control unit (after HEPA filter); 10 - Pressure sensor - product control unit (before HEPA filter); 11 - Speed control –Fan/blower; 12 - Speed control - Push/fan; 13 - Speed control - Vacuum pump; 14 - Temperature sensor - Temperature dryer cell; 15 - Pressure sensor - Differential pressure over the dryer filters; 16 - Pressure sensor - Wet transfer line.

Dryer and product control unit - global model

A global BSPC model that can be applied for the monitoring of all cells was fitted on the data capturing 43.8% of the variability in the X data, 58.9% of the variability in Y and with a Q^2 value of 58.9%. The loadings of PC1 and PC2 (Figure 57a) are identical to the loadings of PC1 and PC2 described for the individual cell models (see section Dryer and product

control unit – individual cell models). The data of the operation of the dryer cells from run 3 was projected into this model and scores (Figures 4a and 4b), Hotelling's T^2 (Figure 4c) and DModX (Figure 4d)) control charts were created. Deviations from normal behavior were observed in the Hotelling's T^2 and DModX plots. Similar to the individual cell models, according to the average Hotelling's T^2 and DModX contribution plots of all out-of-control moments (Figure 57b) and c)) these happened mostly because of the pressure inside the wet granules transfer line being higher or lower than expected at the moments of the actuation of the vacuum pump.

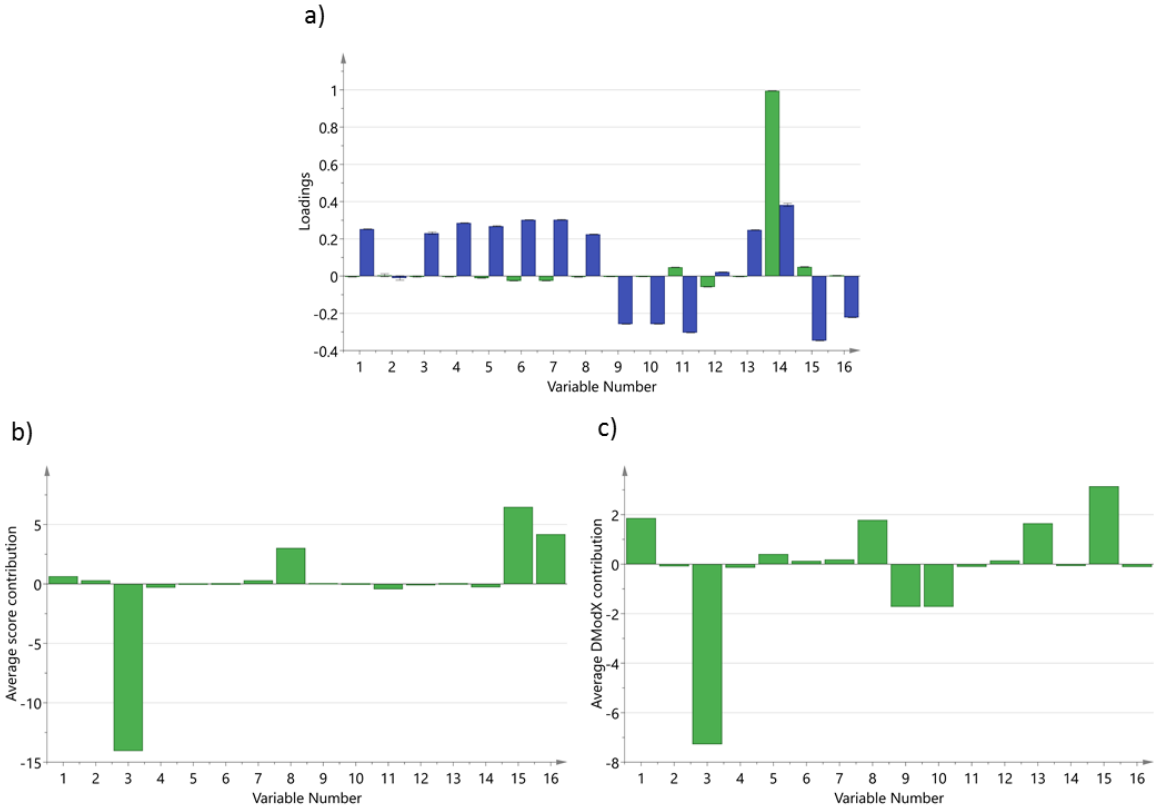


Figure 57 – Dryer and product control unit – global model - a) Loadings of PC1 (green) and PC2 (blue); - b) Average contribution to the out-of-control moments observed in the Hotelling's control charts; c) Average contribution to the out-of-control moments observed in the DModX control charts; 1 - Flow sensor - Dryer air; 2 - Flow sensor - Wet granule transfer; 3 - Pressure sensor - Differential pressure over the wet transfer line; 4 - Pressure sensor – Atmospheric; 5 - Pressure sensor - Dryer air inlet; 6 - Pressure sensor - Dryer air outlet (before HEPA filter); 7- Pressure sensor - Dryer air outlet (after HEPA filter); 8 - Pressure sensor - Pressure dryer top; 9 - Pressure sensor - product control unit (after HEPA filter); 10 - Pressure sensor - product control unit (before HEPA filter); 11 - Speed control – Fan/blower; 12 - Speed control - Push/fan; 13 - Speed control - Vacuum pump; 14 - Temperature sensor - Temperature dryer cell; 15 - Pressure sensor - Differential pressure over the dryer filters; 16 - Pressure sensor - Wet transfer line.

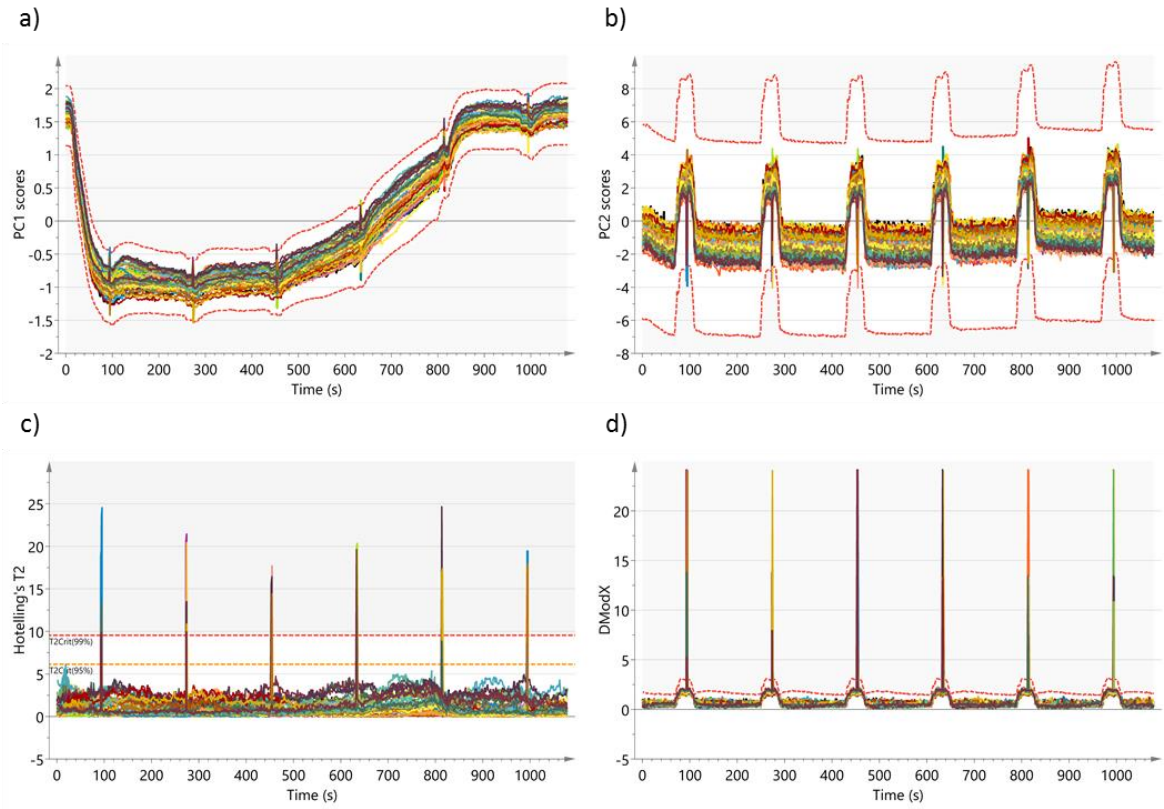


Figure 58 – Dryer and product control unit – global model - a) Predicted scores on PC1 control chart (Red dotted line – 99% confidence interval); b) Predicted scores on PC2 control chart; (Red dotted line – 99% confidence interval); c) Hotelling's T^2 control chart (Orange dotter line – 95% confidence limit, Red dotted line – 99% confidence limit); DModX control chart (Red dotted line – 99% confidence limit); Data from Run 3.

Continuous granulator startup

A BSPC model was developed to monitor the startup operation the granulator unit. Two PCs were fit capturing 81.2% of the variability in X and 80.8% of the variability in Y. Q^2 was 80.8%. The PC1 loadings (Figure 59) reveal that the variables contributing to the main source of variation are power of the granulator driver and torque, both positively correlated to the scores and negatively to the temperatures at the inlet of the granulator's refrigeration liquid and of the granulator barrel. The PC1 scores increase during the entire startup phase. Power and torque increase due to the start of operation. The temperature at the granulator barrel spikes to 26°C at the beginning of operation. This slight increase is due to material build up on the barrel walls and increase of friction. The barrel temperature returns back to its setpoint due to the water cooling down the barrel. The PC2 scores decrease during processing time. The loadings of PC2 (Figure 59) reveal that the mass flow of granulation

liquid, the powder mass flow, the screw speed and temperature of the barrel contribute to the variation captured in this PC. With the start of operation, an increase is seen in the values of these variables. Inputting the data from run 3 on the developed BSPC model shows that this third run followed the same startup since the batch control charts of scores and DModX revealed no deviation from normal operation (Figures 5a, 5b and 5c). The Hotelling's T^2 control charts could not be calculated since only 2 batches (run 1 and run 2) were used to build the two component BSPC model and therefore the critical limits could not be calculated. Calculation of these limits is only possible when the number of components is lower than the number of batches minus one.

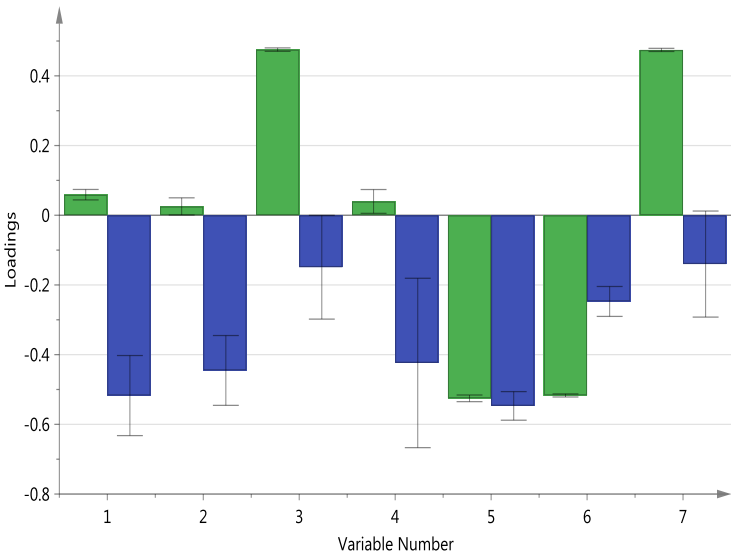


Figure 59 – Continuous granulator startup - Loadings of PC1 (green) and PC2 (blue); 1 - Mass flow - Granulation liquid; 2 - Mass - Flow powder dosing; 3 - Power - Granulator drive; 4 - Speed - Granulator screws; 5 - Temperature sensor - Granulator barrel; 6 - Temperature sensor - Granulator barrel refrigeration liquid inlet; 7 - Torque sensor – Granulator

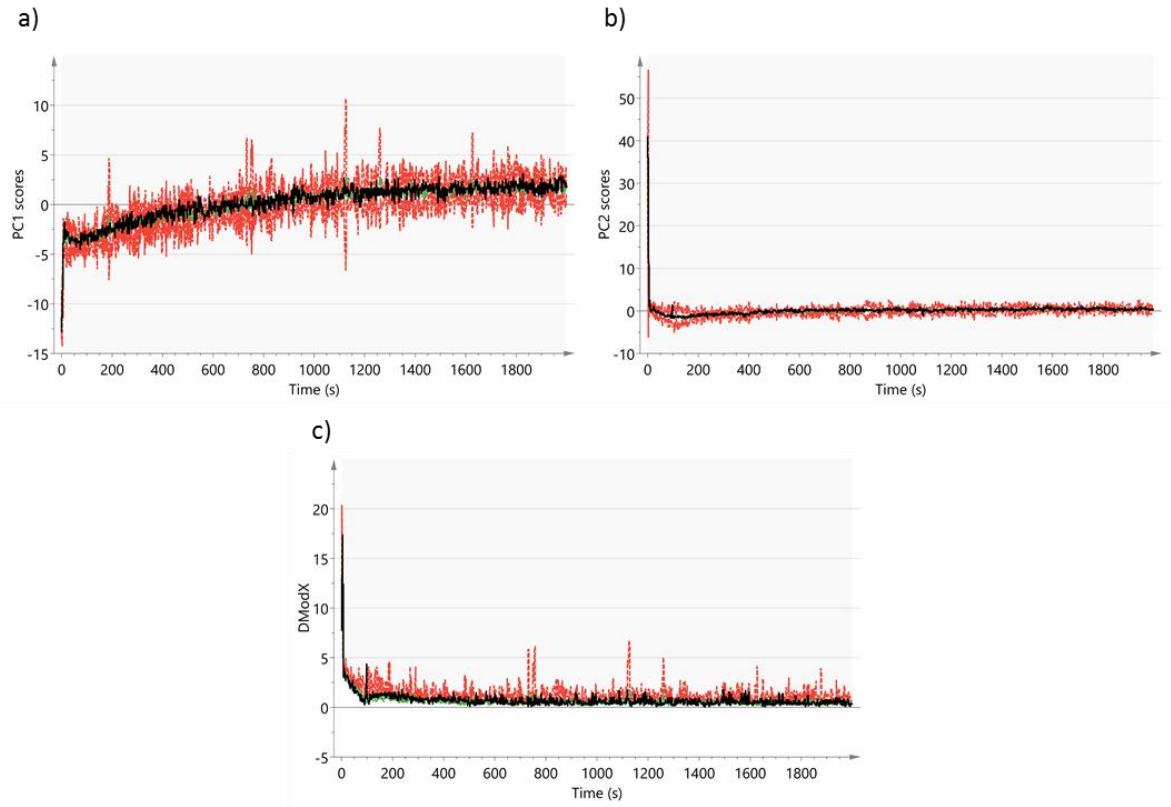


Figure 60 – Figure 5 – Continuous granulator startup - a) Predicted scores on PC1 control chart (Green – average value; Red dotted line – 99% confidence interval); b) Predicted scores on PC2 control chart; (Green – average value; Red dotted line – 99% confidence interval); c) DModX control chart (Green – average value; Red dotted line – 99% confidence limit); Data from Run 3.

Dryer and product control unit startup period

For the startup monitoring of the dryer unit, a BSPC model was developed using all data collected in each individual cell. A two-component model was fit capturing 47.4% of the variance in X, 78.3% of the variance in Y and with an R^2 value of 78.3%.

According to the loadings of PC1 (Figure 61), various variables contribute to this PC, however the differential pressure over the dryer filters and the speed of the fan/blower having the most important contributions. Both variables have loadings that are positively correlated to the scores and their scores increase during processing time. This reveals that the differential pressure over the filters increases after startup and that the speed of the fan/blower need to increase in order to maintain the pressure inside the dryer constant. As previously observed, the differential pressure over the filters of the dryer keeps increasing during the whole process time and not just during startup. The PC2 loadings (Figure 61) also include the differential pressure over the dryer filters and the speed of the fan/blower which

increases to compensate for the increase in pressure. Also, variables like the atmospheric pressure, the pressures at the inlet and top of the dryer and the speed of the fan/blower system contribute to this component.

To test the monitoring ability of the BSPC model, the run 3 data from all cells was projected on the model. According to the scores, DModX and Hotelling's T^2 control plots (Figure 62a) to d)) some moderate outliers were observed corresponding to the pressure overshoot caused by the accumulation of material over the dryer filters (see above for explanation of this phenomenon).

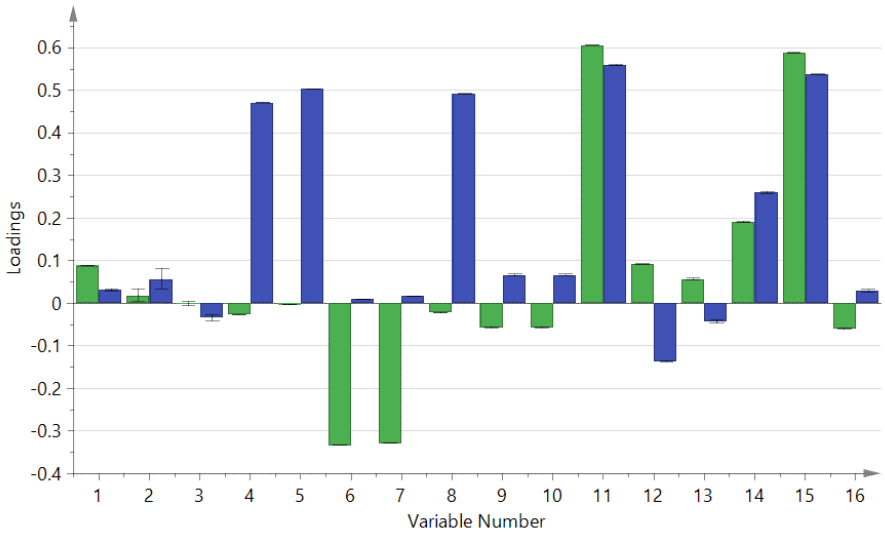


Figure 61 – Dryer and product control unit startup period; Loadings of PC1 (green) and PC2 (blue); 1 - Flow sensor - Dryer air; 2 - Flow sensor - Wet granule transfer; 3 - Pressure sensor - Differential pressure over the wet transfer line; 4 - Pressure sensor – Atmospheric; 5 - Pressure sensor - Dryer air inlet; 6 - Pressure sensor - Dryer air outlet (before HEPA filter); 7- Pressure sensor - Dryer air outlet (after HEPA filter); 8 - Pressure sensor - Pressure dryer top; 9 - Pressure sensor - product control unit (after HEPA filter); 10 - Pressure sensor - product control unit (before HEPA filter); 11 - Speed control – Fan/blower; 12 - Speed control - Push/fan; 13 - Speed control - Vacuum pump; 14 - Temperature sensor - Temperature dryer cell; 15 - Pressure sensor - Differential pressure over the dryer filters; 16 - Pressure sensor - Wet transfer line.

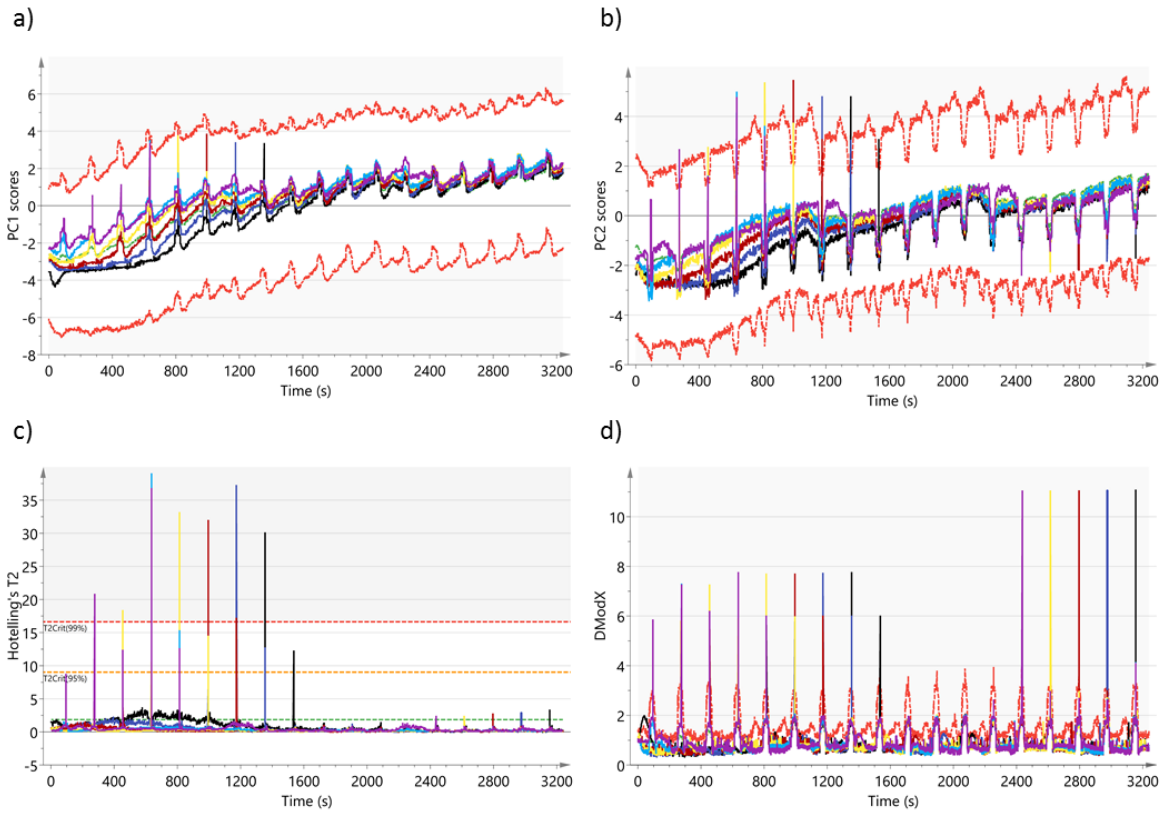


Figure 62 – Dryer and product control unit startup period - a) Predicted scores on PC1 control chart (Green – average value; Red dotted line – 99% confidence interval); b) Predicted scores on PC2 control chart; (Green – average value ; Red dotted line – 99% confidence interval); c) Hotelling's T^2 control chart (Green – average value; Orange dotted line – 95% confidence limit; Red dotted line – 99% confidence limit); d) DModX control chart (Green – average value ; Red dotted line – 99% confidence limit); Data from Run 3

5.4.4 Conclusion

The correlations between pairs of variables with and without time were examined in order to take into account the process dynamics when developing the BSPC strategy. The optimal cross-correlations were found however, the lags were not implemented since it was not possible to find a logical manner to do so, and the original time-points for all variables were kept.

An OPLS model fitted to the expected steady state data (excluding the data corresponding to startup and shut-down) and regress versus process time allowed to identify variables with a trend over process time. Three types of variables with trend were found: (i) time-related variables that follow the same trend for the three runs and (ii) time-related variables which behave differently in the different runs. These are described in depth in the extent of this study.

A PLS-based BSPC model was fitted on data from each dryer cell independently and also from the product control unit. The predicted batch trajectory from a cell other than the one used for the modelling sometimes deviates from the control limits. This occurs because even though cells are identical there are still slight differences in the drying temperature profiles from cell to cell. The differences between these profiles are considered small enough not to have an impact on the drying of the granules. A global model containing information from all cells was also built. Deviations seen in the Hotelling's T^2 and DModX plots are mostly due to the pressure inside the wet granules transfer line being higher or lower than expected at the moments of the actuation of the vacuum pump.

PLS models for the BSPC of the different unit's startup operation periods were developed. The main aim of this exercise was to obtain knowledge about the current process. Several variable correlations were found and are hereby described.

5.5 Multivariate statistical process control approach to monitor a continuous pharmaceutical twin-screw granulation and drying process[5]

5.5.1 Abstract

The present study aims at developing a multivariate statistical process control (MSPC) strategy for the innovative ConsiGma™-25 continuous tablet manufacturing line. Thirty-five logged variables encompassing twin screw high shear granulator, fluid bed dryer and a product control unit processes were used to monitor the process. The MSPC strategy was based on PCA considering data acquired under NOC for four process runs. Seven runs with imposed disturbances to the NOC were utilized to evaluate the model's monitoring performance and the ability to detect disturbances. The impact of the imposed disturbances to the process to the continuity was also evaluated using Hotelling's T^2 and Q residuals statistics control charts. The influence of the individual process variables to the statistics was also assessed by analyzing contribution plots at specific time points. Results show that the provoked disturbances were all detected in both control charts. The imposed disturbance in the granulator barrel temperature was more evident to the Hotelling's T^2 statistics. All other deviations were observable in the Hoteling's T^2 statistics but, as they could not be explained by the model, were diagnosed in the Q residuals statistic.

[5] This chapter has been adapted from: A.F.T. Silva, M.C. Sarraguça, M. Fonteyne, J. Vercruysse, F. De Leersnyder, V. Vanhoorne, N. Bostijn, M. Verstraeten, C. Vervaet, J.P. Remon, T. De Beer, J.A. Lopes, Multivariate statistical process control approach to monitor a continuous pharmaceutical twin-screw granulation and drying process, (Submission in process).

5.5.2 Materials and methods

5.5.2.1 Materials

Milled lactose monohydrate (Pharmatose™ 200M, DFE Pharma, Germany) was granulated with a 20% (w/w) PVP solution in the ConsiGma™-25 system.

5.5.2.2 Continuous manufacturing line

The experiments in question were performed in a line located at the University of Gent (Gent, Belgium). In this specific assembly, the wet granules are transported from the granulator to the top of the six-segmented fluid bed dryer by the actuation of a horizontal vacuum transport line. A second vacuum transport line takes the granules form the dryer to the product control unit. The system has been thoroughly described in section ConsiGma™-25 continuous line (Chapter 5). During operation, the continuous line logs multiple process variables including user-set and open loop variables. In the first case, independent PID controllers keep the variables around the user-defined set points. From the logged variables, thirty-five were chosen to be utilized for the monitoring strategy development (Table 13).

Table 13 – Variables logged by the ConsiGma™-25 during processing include

Variable number	Name	Description	Units	Type	Setpoint value	ConsiGma™ -25 unit
V1	Flow dry. air	Flow sensor – dryer air	m^3/h	Setpoint	360	Dryer
V2	Flow wet gran. trans.	Flow sensor – wet granule transfer line	m^3/h	Setpoint	3.6	Wet granule transfer line (dryer)
V3	Hum. dry. air in.	Humidity sensor – dryer air inlet	%RH	Setpoint	10	Dryer
V4	Hum. Dry. air out.	Humidity sensor – dryer air outlet	%RH	Measured	N/A*	Dryer
V5	Mass flow gran. liq.	Mass flow granulation liquid	g/min	Setpoint	58	Liquid addition module (granulator)
V6	Mass flow pow. dos. unit	Mass Flow – Powder dosing	Kg/h	Setpoint	25	Powder dosing unit (granulator)
V7	Pow. gran. drive	Power – granulator drive	W	Measured	N/A	Granulator

Table 13 – Variables logged by the ConsiGma™-25 during processing include

Variable number	Name	Description	Units	Type	Setpoint value	ConsiGma™ -25 unit
V8	Press. diff. dry. fil.	Pressure sensor – differential pressure over the dryer filters	mbar	Measured	N/A	Dryer
V9	Press. diff. dry. hole plate	Pressure sensor – differential pressure over the hole plate	mbar	Measured	N/A	Dryer
V10	Press. diff. HEPA fil. dry. air out.	Pressure sensor – differential pressure at the HEPA filter of the dryer air outlet	mbar	Measured	N/A	Dryer
V11	Press. diff. HEPA fil. prod. cont. unit	Pressure sensor – differential pressure over the HEPA filter at the product control unit	mbar	Measured	N/A	Product Control Unit
V12	Press. diff. HEPA fil. wet gran. trans. in	Pressure sensor – differential pressure over the HEPA filter at the wet transfer line inlet	mbar	Measured	N/A	Wet transfer line (dryer)
V13	Press. diff. wet trans. line	Pressure sensor – differential pressure over the wet transfer line	mbar	Measured	N/A	Wet transfer line (dryer)
V14	Press. dry. air in.	Pressure sensor – dryer air inlet	mbar	Measured	N/A	Dryer
V15	Press. dry. air out. after HEPA fil.	Pressure sensor – dryer air outlet after HEPA filter	mbar	Measured	N/A	Dryer
V16	Press. dry. air out. before HEPA fil.	Pressure sensor – dryer air outlet before HEPA filter	mbar	Measured	N/A	Dryer
V17	Press. dry. top	Pressure sensor – dryer top	mbar	Measured	N/A	Dryer
V18	Press. prod. cont. after HEPA fil.	Pressure sensor – product control unit after HEPA filter	mbar	Measured	N/A	Product Control Unit
V19	Press. prod. cont. before HEPA fil.	Pressure sensor – product control unit before HEPA filter	mbar	Measured	N/A	Product Control Unit
V20	Speed air hand.	Speed control –	%	Measured	N/A	Fan/blower

Table 13 – Variables logged by the ConsiGma™-25 during processing include

Variable number	Name	Description	Units	Type	Setpoint value	ConsiGma™ -25 unit
	unit fan cont.	fan/blower				system (dryer)
V21	Speed gran.	Speed – granulator screws	rpm	Setpoint	700	Granulator
V22	Speed motor pow. dos. unit	Speed – motor powder dosing	rpm	Measured	N/A	Powder dosing unit (granulator)
V23	Temp. air hand. unit	Temperature sensor – air handling unit	ºC	Measured	N/A	Air handling unit (dryer)
V24	Temp. dry. air in.	Temperature sensor – dryer air inlet	ºC	Setpoint	50	Dryer
V25	Temp. dry. air out.	Temperature sensor – dryer air outlet	ºC	Measured	N/A	Dryer
V26	Temp. dry. cell 1	Temperature sensor – temperature dryer cell 1	C	Measured	N/A	Dryer
V27	Temp. dry. cell 2	Temperature sensor – temperature dryer cell 2	ºC	Measured	N/A	Dryer
V28	Temp. dry. cell 3	sensor – temperature dryer cell 3	ºC	Measured	N/A	Dryer
V29	Temp. dry. cell 4	Temperature sensor – temperature dryer cell 4	ºC	Measured	N/A	Dryer
V30	Temp. dry. cell 5	Temperature sensor – temperature dryer cell 5	ºC	Measured	N/A	Dryer
V31	Temp. dry. cell 6	Temperature sensor – temperature dryer cell 6	ºC	Measured	N/A	Dryer
V32	Temp. gran. barrel	Temperature sensor – granulator barrel	ºC	Setpoint	25	Granulator
V33	Temp. out. tcu 1	Temperature sensor – outlet temperature control unit	ºC	Measured	N/A	Temperature control unit (Granulator)
V34	Temp. tank tcu 1	Temperature sensor – tank temperature control unit	ºC	Measured	N/A	Temperature control unit (Granulator)
V35	Torque gran.	Torque sensor – granulator	N/A	Measured	N/A	Granulator

N/A: not applicable

5.5.2.3 Production runs

Five NOC runs (Table 14) were performed by setting the user-set variables at the fixed values specified in Table 13. The total drying time of the granules inside a cell was set at 840 s (including a filling time of 180 s and a cell discharging time of 20 s). The product control unit emptying time was set at 30 s. The system takes 1080 s (6 x 180s) to finish the sequential filling of all 6 cells. For the purpose of this work, this period will be called the "fill" of the dryer. Each reference run has a total of three fills with an overall time of 3240 s. Data fed to the MSPC model require the system to be in a steady state. It was verified (data not shown) that for the first fill of each run, the process was still reaching the steady state. Therefore, for all modelling purposes, the first fill of each run was removed resulting in NOC runs with a total of 2160 s corresponding to two fills. Four of these runs were used for calibration and the one was used to test the model's performance (see Table 14).

To challenge the effectiveness of the monitoring strategy seven different runs including imposed disturbances of the NOC were performed. In each run, a different setpoint was changed when the dryer first cell was filled for the fourth time and changed back to the original value when the same cell was filled for the seventh time (Table 14). The time length of the disturbance (three fills) was chosen to ensure that the effect of the changes could be clearly identified and a new steady state could be reached.

As explained before, the first fill of each run was removed because the system was not yet at steady state resulting in runs with a total of 8640 s (eight fills).

After each run the system was cleaned. During cleaning, the granulator was disassembled and its components washed with water and allowed to dry until the next run. The dryer and product control units were thoroughly vacuum cleaned. All runs were performed in different days and by different operators to ensure maximum variability.

Table 14 – Calibration and test runs used in the MSPC model.

Name	Type of run	Purpose	Length (fills)	Variable changed	Setpoint change	Units
R1	NOC	Calibration	3	N/A*	N/A	N/A
R2	NOC	Calibration	3	N/A	N/A	N/A
R3	NOC	Calibration	3	N/A	N/A	N/A
R4	NOC	Test	3	N/A	N/A	N/A
R5	NOC	Calibration	3	N/A	N/A	N/A

Table 14 – Calibration and test runs used in the MSPC model.

Name	Type of run	Purpose	Length (fills)	Variable changed		Setpoint change	Units
F1	Non NOC	Test	9	V1	Dryer air flow	360 to 400	m^3/h
F2	Non NOC	Test	9	V32	Granulator barrel temperature	25 to 35	ºC
F3	Non NOC	Test	3	N/A	Unclean system	N/A	N/A
F4	Non NOC	Test	9	V24	Dryer air temperature	50 to 60	ºC
F5	Non NOC	Test	9	V55	Granulation liquid mass flow	58 to 66.7	g/min
F6	Non NOC	Test	9	V6	Powder dosing unit mass flow	25 to 21.7	kg/h
F7	Non NOC	Test	9	V21	Granulator screw speed	700 to 900	rpm

***N/A: not applicable**

5.5.2.4 Data analysis

Process variables of the calibration runs were organized in a three-way array (run x variable x time) that was firstly unfolded in a two-way matrix, using the observation-wise unfolding method in which the variable direction is preserved [197]. After being unfolded, the calibration runs yielded a two-way matrix with a length of 8640 lines (2160 time points x 4 runs) and 35 columns (variables). This matrix was used to calibrate the PCA [222] model. Data were scaled to unit variance prior to modeling. The number of principal components was chosen based on the lowest cross-validation error.

Hotelling's T^2 and Q residuals control charts were used to monitor the process runs. Unusually high Hotelling's T^2 (above the 95% confidence level) reveal differences which can be explained by the model. On the other hand, the Q residuals estimates the distance to model and therefore highlights differences that cannot be explained by the modelled components. It also important to identify the individual variables causing out-of-control signals in both Hotelling's T^2 and Q residuals statistics. For that purpose, contribution plots were generated and examined [82, 223, 224]. Several studies show that the relative size of the contributions relatively to the NOC batches should be examined instead of the absolute size of the contributions [82, 225]. Some authors used bootstrap as a method to determine confidence limits to contribution plots [225, 226].

In this work, in order to determine confidence limits, the variable contributions at each time point, for the calibration runs, were average and its standard deviation calculated. The confidence limit was set as twice the standard deviation.

Data analysis was performed with Matlab version 8.3 (Mathworks, Natic, MA, WA, USA) and the PLS Toolbox version 7.5 (Eigenvector Research Inc., Wenatchee, WA, USA).

5.5.3 Results and discussion

A PCA model with seven PC's encompassing 66.3% of variance was obtained. Each of the eight test runs (the NOC run and the seven runs with imposed disturbances) were projected onto the model and the Hotelling's T^2 and Q residuals statistics were calculated. The equations to calculate these metrics can be found in section Multivariate control charts (Chapter 2). Throughout the discussion, the variable number is included after the variable name to help results interpretation. A detailed analysis is given for each test run.

5.5.3.1 Run R4 (NOC)

As expected, Hoteling's T^2 values for the NOC run R4 are mostly within the control limits (Figure 63). In accordance to what happens for the calibration runs, a few values are outside the control limits, after which normal operation is promptly restored without any action being necessary. In the Q residuals control chart (Figure 63) the same trend is also observed.

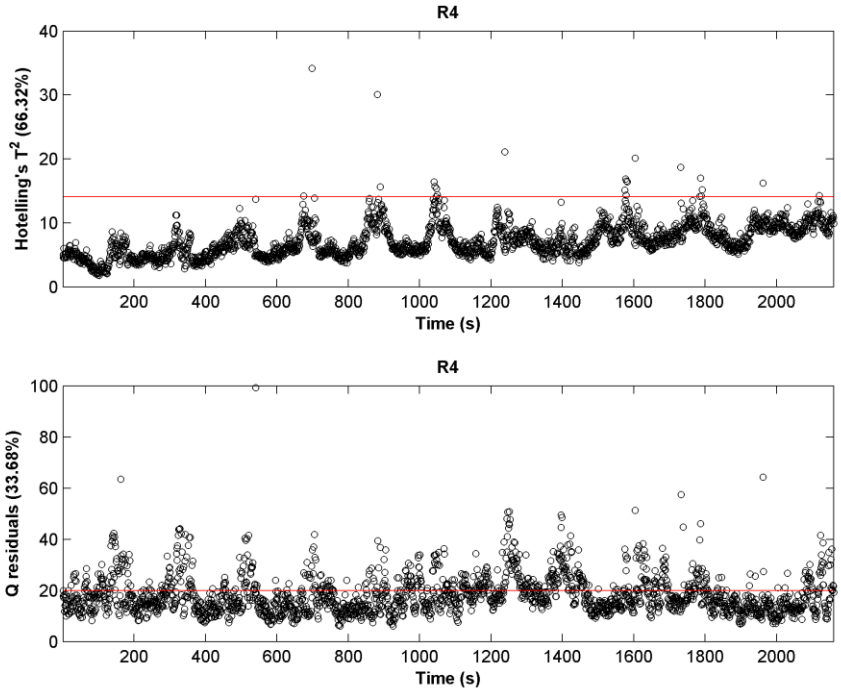


Figure 63 – Hotelling's T^2 (up) and Q residuals (down) control charts obtained by projecting NOC run R4 in the developed PCA model.

5.5.3.2 Runs with imposed disturbances

Run F1 (dryer air flow)

For run F1 the provoked changed was the increase of the dryer air flow (V1) from 360 to 400 m^3/h. The provoked change can be seen in both Hotelling's T^2 (Figure 64) and Q residuals (Figure 65) control charts with an increase and posterior decrease of the values of both statistics. A Hotelling's T^2 contribution plot versus process time for this period was generated (Figure 66) in order to clarify the variables contributing to the change. The air flow increase lead to a decrease of the humidity at the dryer outlet (V4) as did pressures at the outlet of the dryer before and after the HEPA filter (V15 and V16). On the contrary, differential pressure over the dryer filters (V8) and speed of the fan/blower (V20) increased in order to maintain the pressure in the dryer constant.

The Q residuals contribution plot (Figure 69) shows that the dryer air flow was the variable responsible for the provoked changed. Consequently, most of the measured air flows, pressures, temperatures, humidities were impacted.

A second change can be seen around 7000 s in the Hotelling's T^2 and Q residual control charts. The Hotelling's T^2 contributions during this period (7000 – 7600s) (Figure 67) show a decrease in the temperature of the granulator barrel (V32), power of the granulator drive (V7) and torque (V35). These variables were probably increased due to an increase of the friction of the material inside the granulator with the granulator walls. The cause of this phenomenon is unknown. The same information can be retrieved form the variables contribution plot for the Q residuals (Figure 68).

Another significant point if the fact that the Hotelling's T^2 chart shows deviations from the beginning of the process. This is not expected since the process is nominal for the first three fills. A contribution plot for the Hotelling's T^2 at time point 2000 s (Figure 66) shows which variables are responsible for this apparent deviation. Humidity at the dryer inlet and outlet (V3 and V4) are lower which is possibly a consequence of differences in the humidity of the production room incoming air. The differential pressure over the dryer filters (V8) and the speed of the fan/blower system (V20) responsible for the regulation of the dryer internal pressure are also increased in comparison with the calibration runs. Previous studies have shown that the differential pressure in the dryer filters (V8) generally increases over time due to gradual deposition of fine powder particles in the filters [227]. If the pressure is already high from the start, this can be an indication that filters need to be cleaned as it can lead to an ineffective drying.

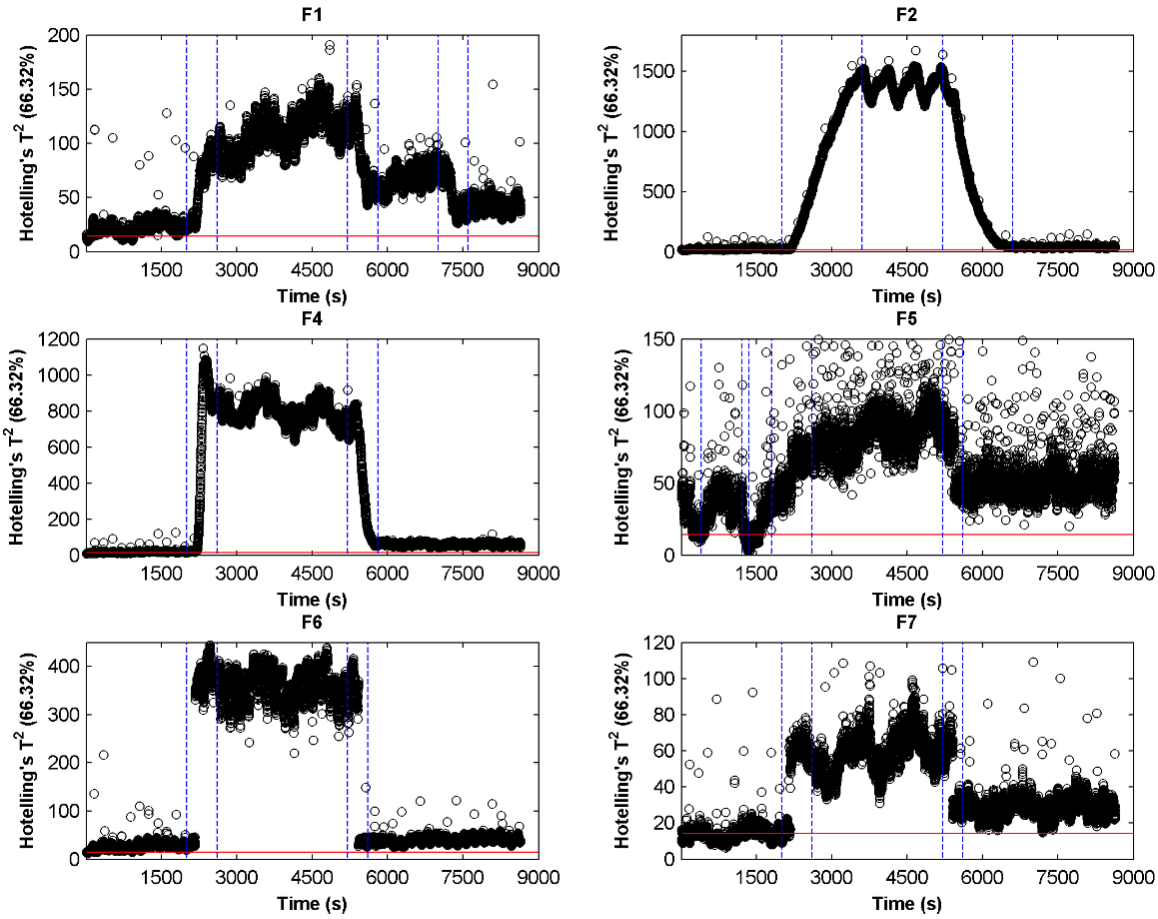


Figure 64 – Hotelling's T^2 control charts obtained projecting test runs in the developed PCA model.

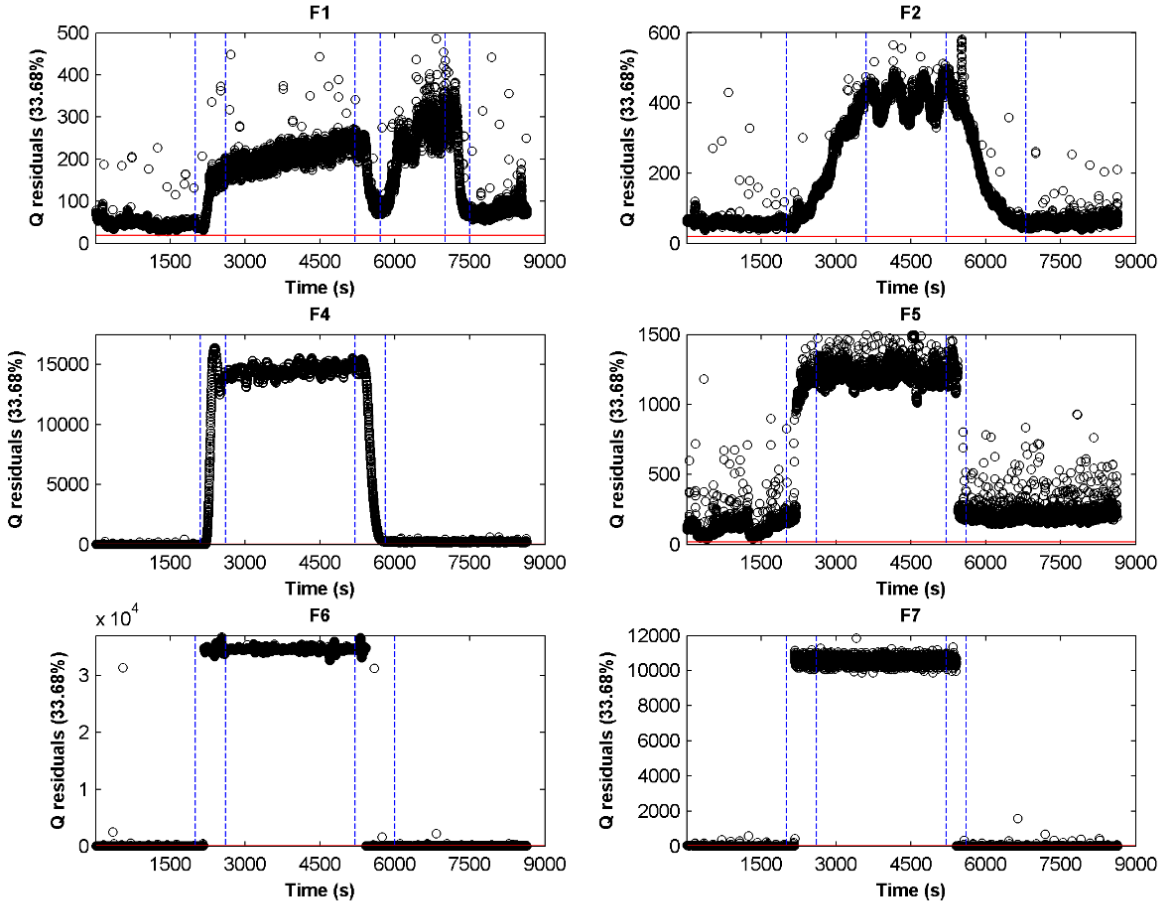


Figure 65 – Q residuals control charts obtained projecting test runs in the developed PCA model.

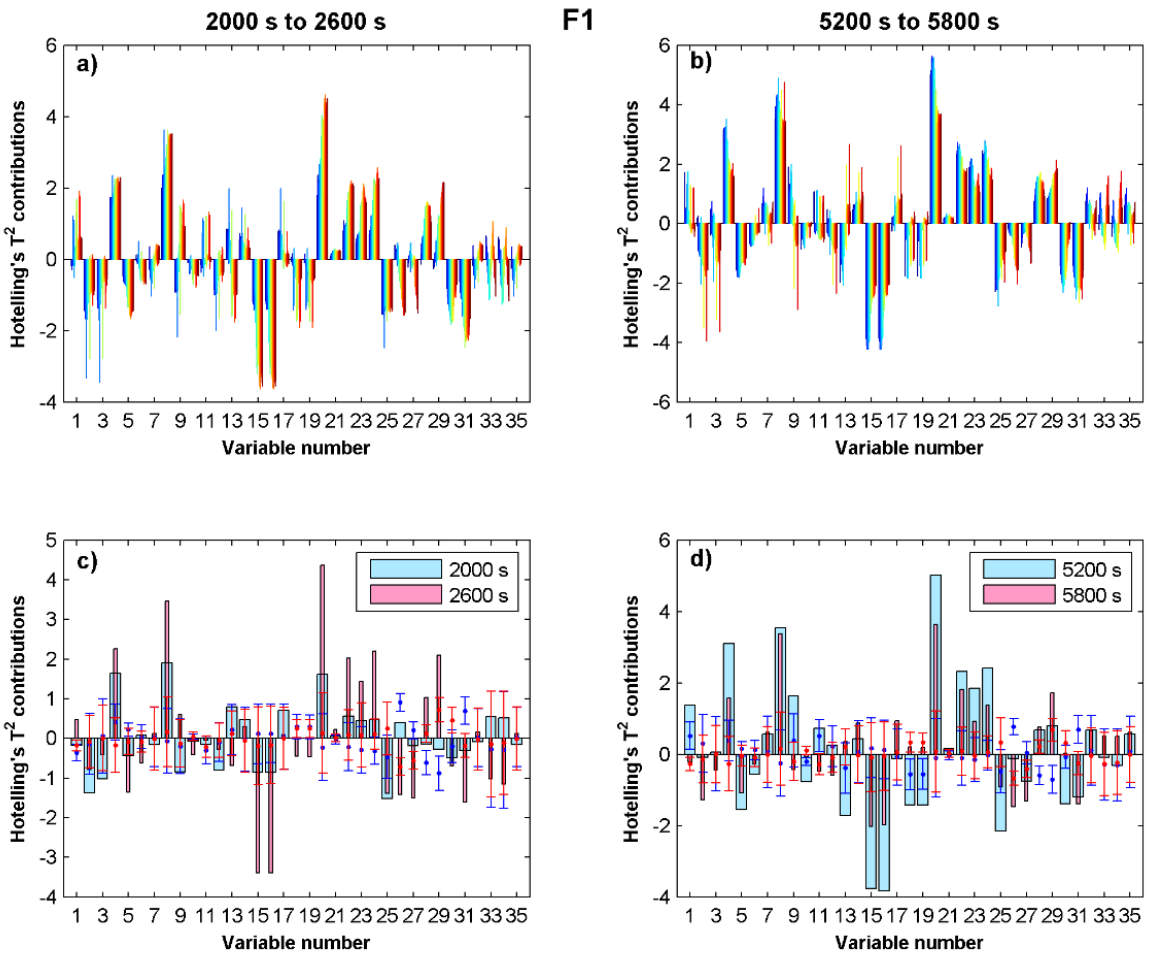


Figure 66 – Run F1 – variables contributions for the Hotelling's T^2 statistic for time: a) between 2000 and 2600 s, b) between 5200 and 5800 s, c) 2000 and 2600 s, d) 5200 and 5800 s.

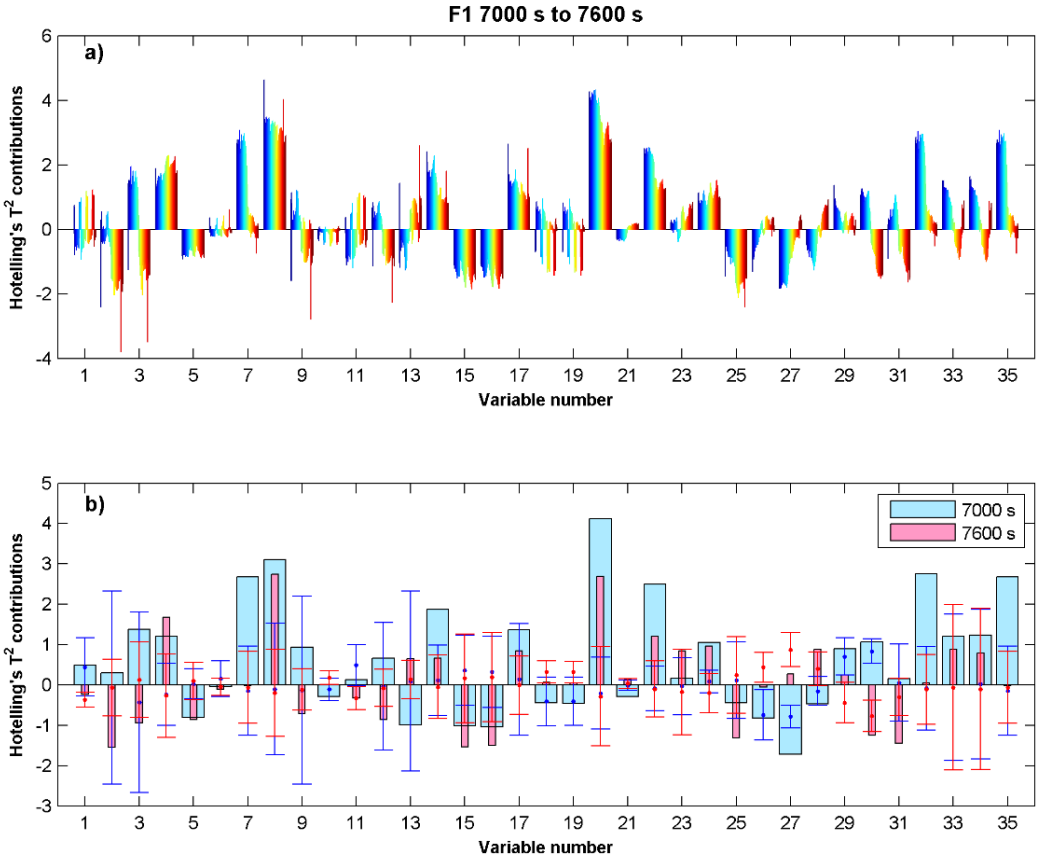


Figure 67 – Run F1 – variables contribution for Hotelling's T^2 statistic for time: a) between 7000 and 7600 s, b) 7000 and 7600 s. ´

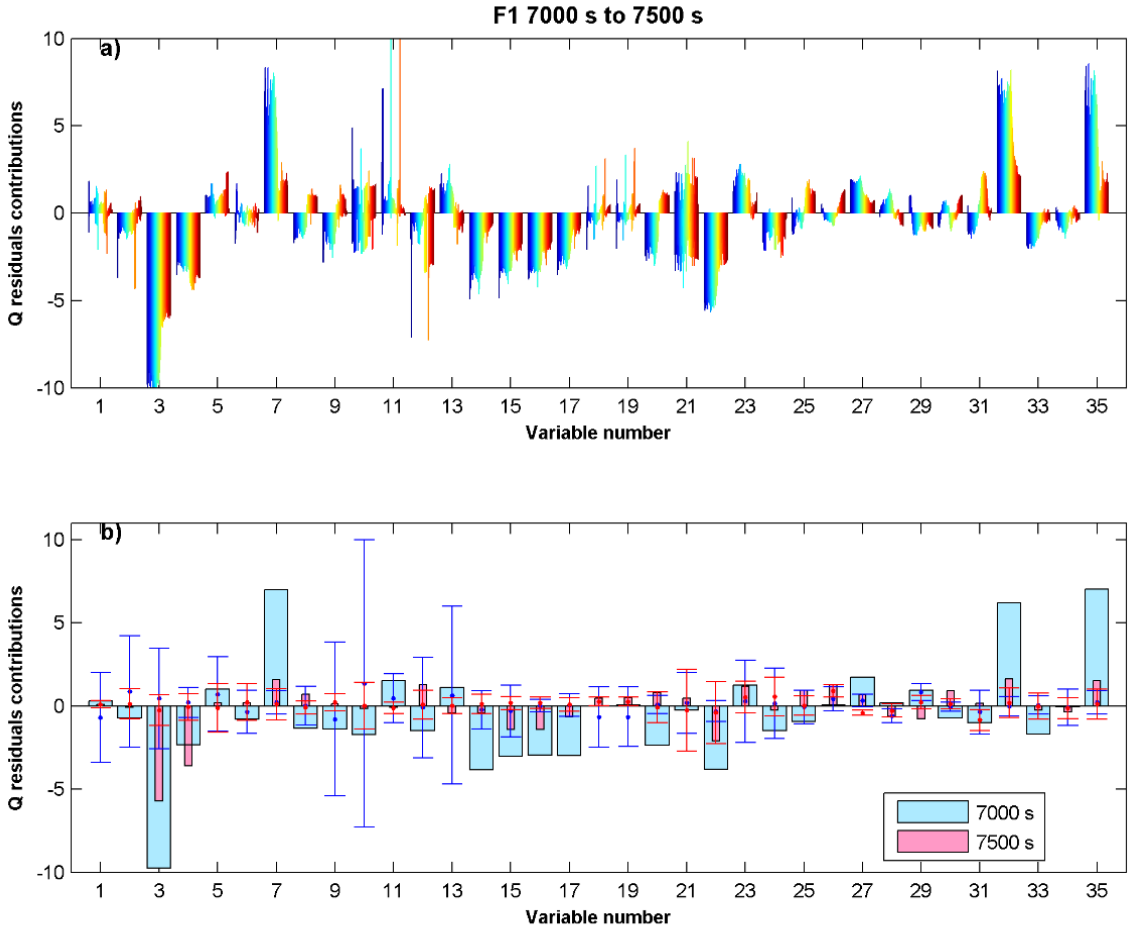


Figure 68 – Run F1 – variables contribution for Q residuals statistic for time: a) between 7000 and 7500 s, b) 7000 and 7500 s.

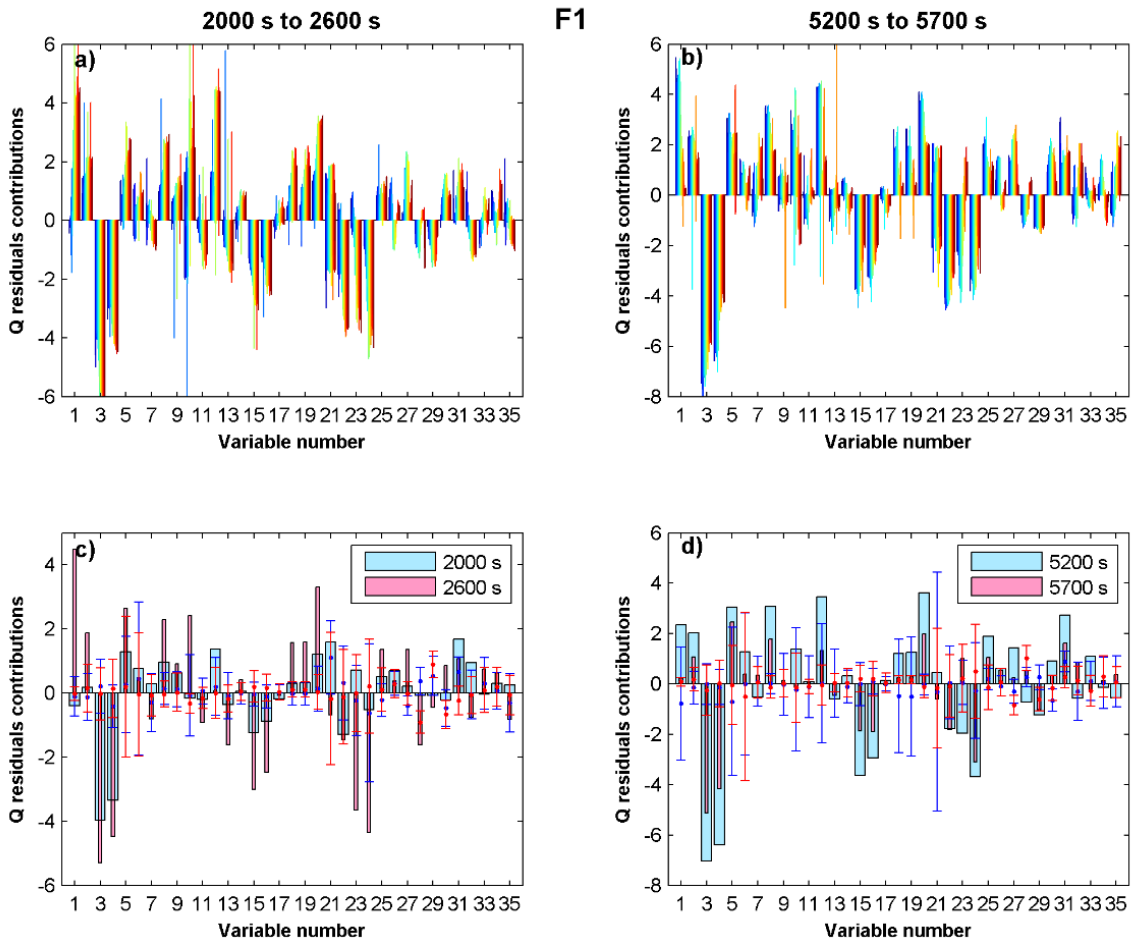


Figure 69 **–** Run F1 – variables contributions for the Q residuals statistic for time: a) between 2000 and 2600 s, b) between 5200 and 5700 s, c) 2000 and 2600 s, d) 5200 and 5700 s.

Run F2 (granulator barrel temperature)

In this run, the granulator barrel temperature (V32) was changed between 25 ºC to 35ºC. This disturbance lead to a change in the Hotelling's T^2 (Figure 64) and Q residuals (Figure 65) statistics. Contribution plots for the Hotelling's T^2 were built (Figure 70) revealing that the temperatures at the control unit outlet (V33) and tank (V34) are the main variables contributing for the deviation. With the increase of the temperature setpoint the water in the temperature control unit is heated in order to increase the temperature of the barrel causing the deviations from normal operation. Contribution plots for the Q residuals (Figure 71) shows contributions from variables associated with the granulator operation (V5, V32, V33 and V34) and also with the temperature inside the dryer cells (V27, V29, V30 and V31). The control

chart of the Q residuals also shows clearly that the process is, from the start, different from the calibration runs (Figure 65). The main variables contributing to this difference are the humidities at the dryer inlet and outlet (V3 and V4) are lower when compared to the calibration runs (Figure 71). Variations in the air humidity of the production room can be a possible explanation for these differences captured in the Q residuals statistics.

Another essential fact was that a closer inspection of the granulator barrel temperature (V32) showed that, despite the setpoint being set at 25°C, the logged average temperature varied around 30°C. When the setpoint was changed to 35°C, temperature increased to 37°C. This occurs because the granulator barrel temperature is controlled by a temperature control unit, where the medium that enters the barrel (water) is cooled down or heated to the setpoint value in order to change the temperature of the barrel itself. However, there is no feedback control implemented toward the temperature control.

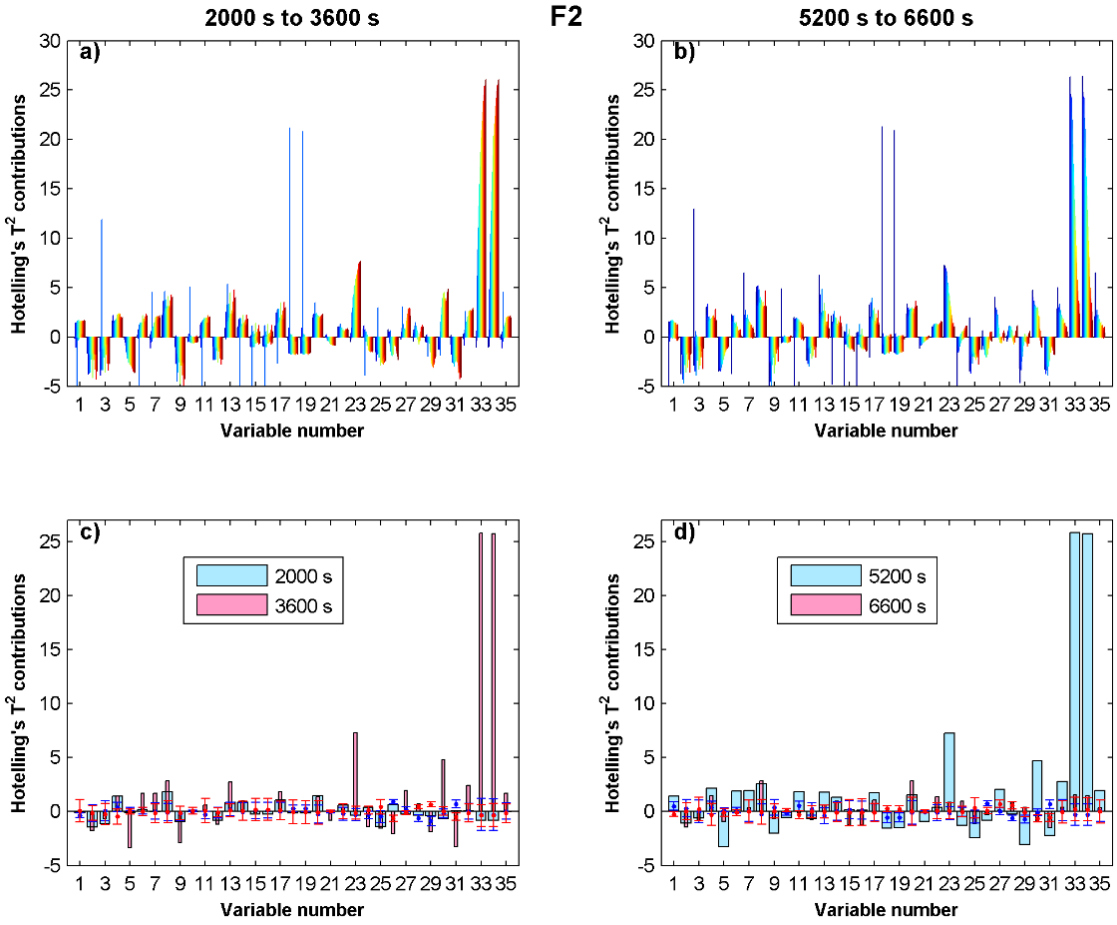


Figure 70 – Run F2 - variables contributions for the Hotelling's T^2 statistic for time: a) between 2000 and 3600 s, b) between 5200 and 6600 s, c) 2000 and 3600 s, d) 5200 and 6600 s.

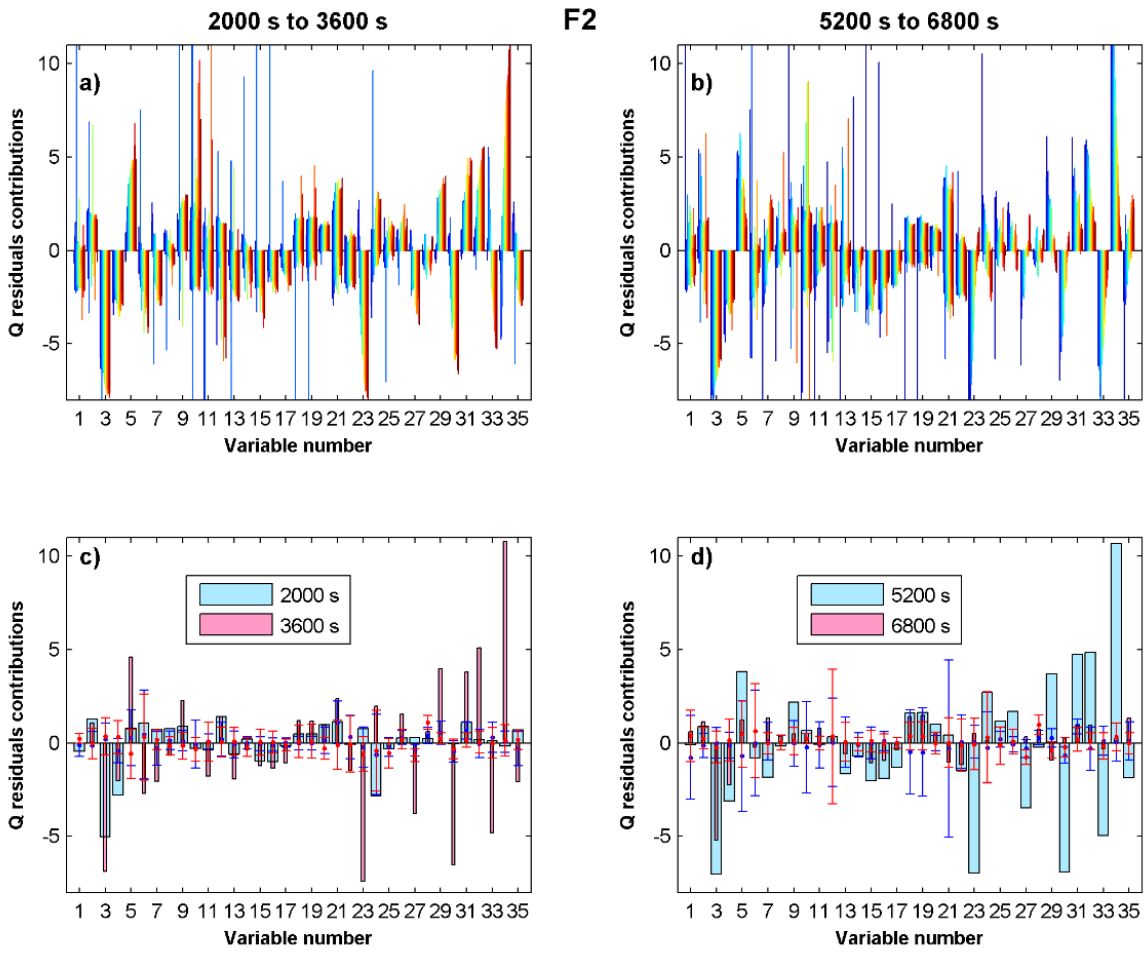


Figure 71 – Run F2 - variables contributions for the Q residuals statistic for time: a) between 2000 and 3600 s, b) between 5200 and 6800 s, c) 2000 and 3600 s, d) 5200 and 6800 s.

Run F3 (unclean system)

In this run, there was not a provoked disturbance, the difference between this run and the NOC runs is that the system was no cleaned after the previous run.

Hoteling's T^2 values for run F3 are within the control limits for the most of the process (Figure 74) and the few values outside the control limits thorough the process become within control without any action being taken. Q residuals control chart (Figure 74) also shows the same trend.

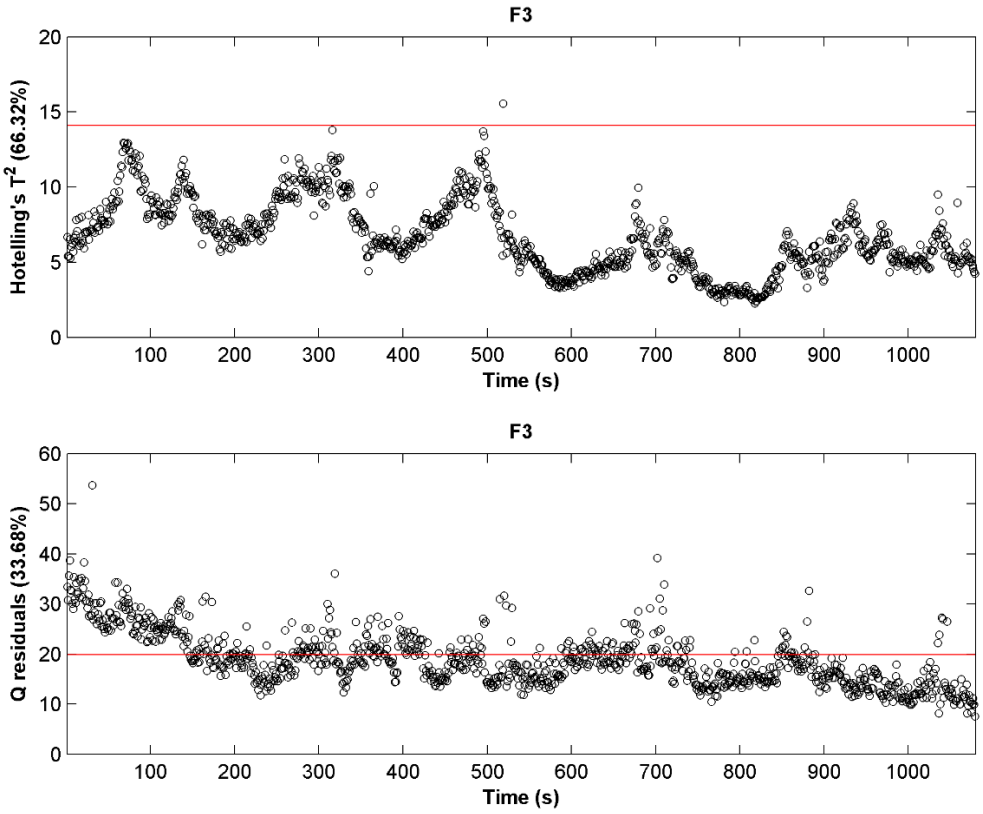


Figure 72 – Hotelling's T^2 (up) and Q residuals (down) control charts obtained projecting test run F3 in the developed PCA model.

Run F4 (temperature dryer air)

The temperature of the air inside the dryer was changed between 50 ºC and 60 ºC. Both Hotelling's T^2 (Figure 64) and Q residuals (Figure 65) statistics reflect the provoked changed. The contribution plots for the Hotelling's T^2 (Figure 73) shows a large contribution from multiple variables related to the dryer and product control units (temperatures, humidities, air flow, pressures). The contribution plots for the Q residuals (Figure 74) indicate that the main variable contributing to the change is the setpoint variable (V24).

After the setpoint changed to the reference value the Hotelling's T^2 statistics did not return to the initial values (Figure 64). The variables that contribute to this fact are mainly the temperature of the granulator barrel (V32) torque (V35) and power of the granulator drive (V7) (Figure 73). The reason is probably due to an increase of the friction of material in the granulator walls.

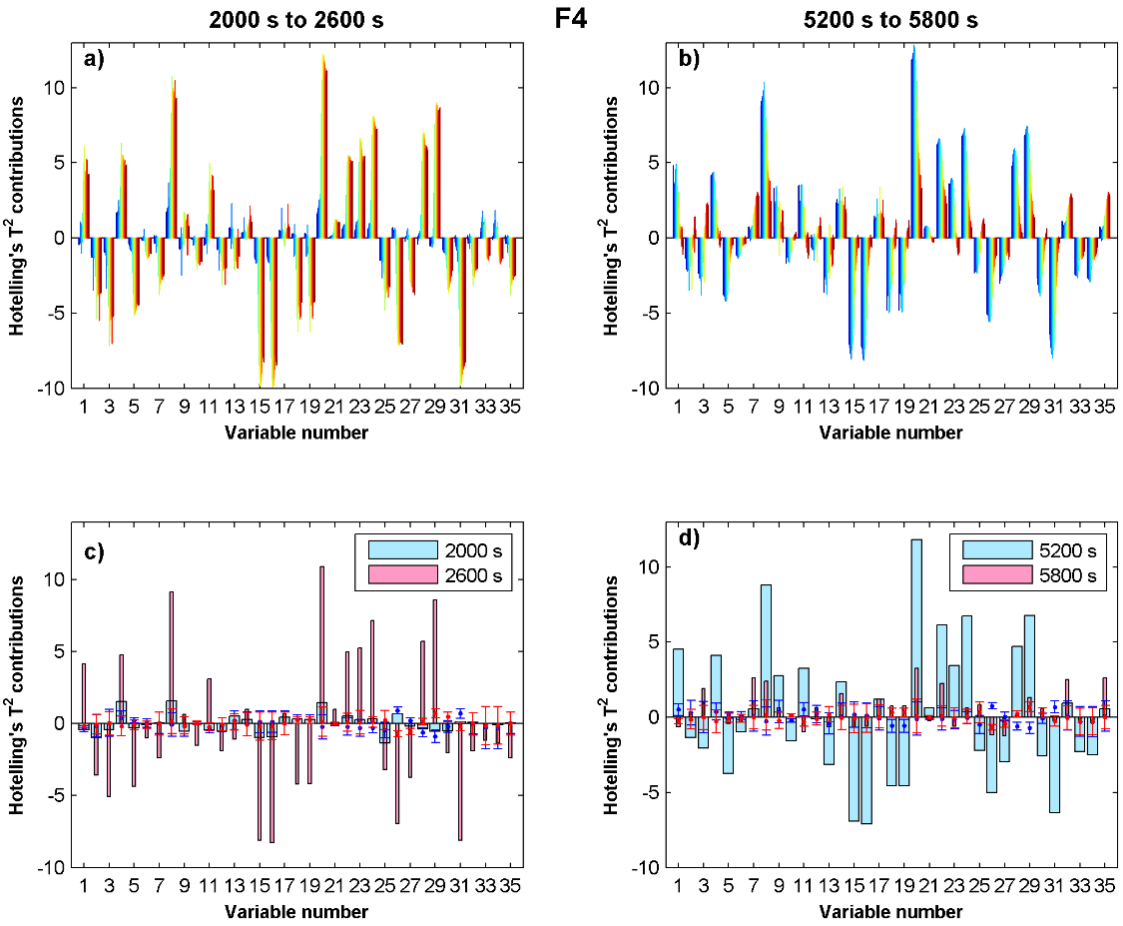


Figure 73 – Run F4 - variables contributions for the Hotelling's T^2 statistic for time: a) between 2000 and 2600 s, b) between 5200 and 5800 s, c) 2000 and 2600 s, d) 5200 and 5800 s.

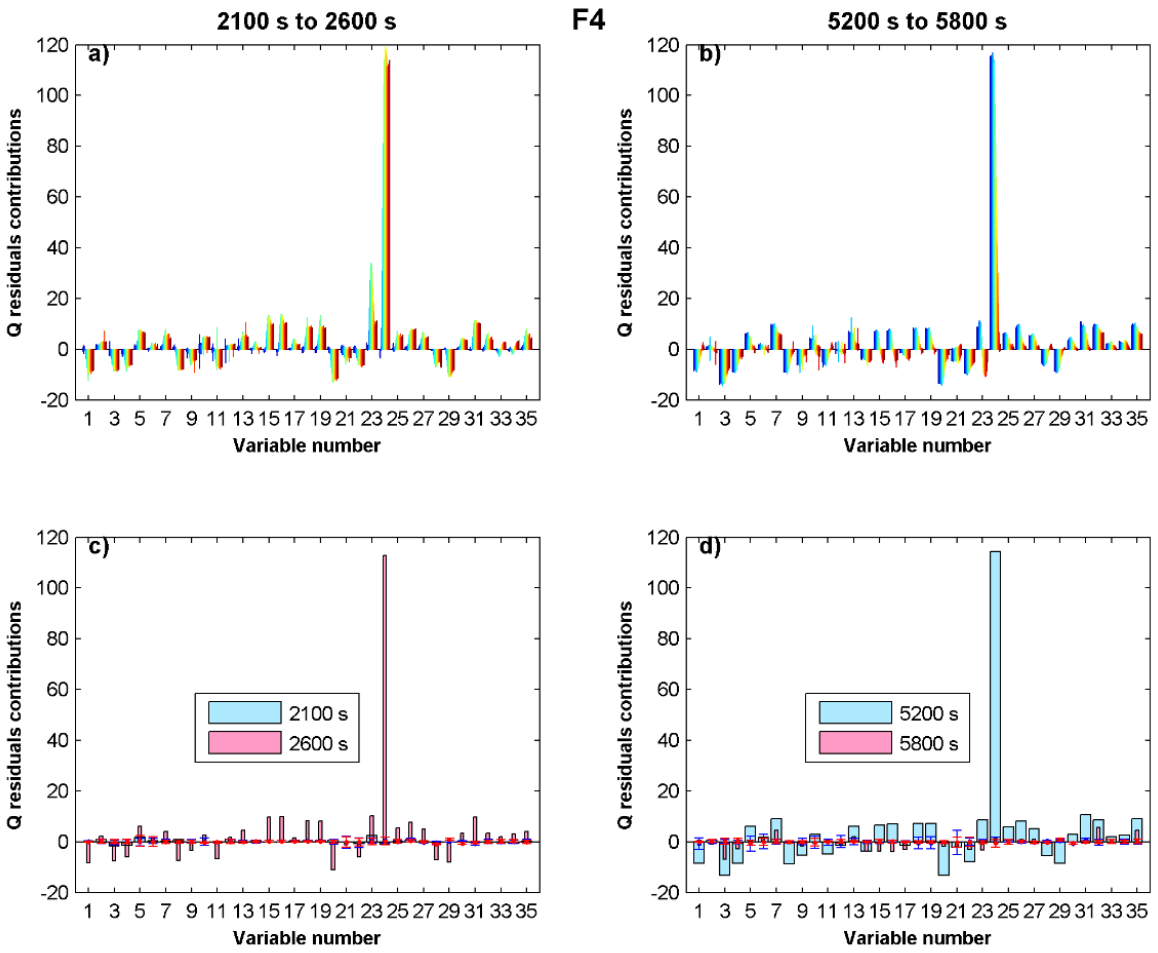


Figure 74 – Run F4 - variables contributions for the Q residuals statistic for time: a) between 2100 and 2600 s, b) between 5200 and 5800 s, c) 2100 and 2600 s, d) 5200 and 5800 s.

Run F5 (granulation liquid mass flow)

Another setpoint changed was the granulation liquid mass flow (V5) from 58 to 66.7 g/min. As expected both Hotelling's T^2 (Figure 64) and Q residuals statistics reveal this change (Figure 65).

In the variable contribution plots (Figure 65) the variables that contribute to the change in the Hotelling's T^2 values are mainly the pressures in the HEPA filters (V18 and V19) and the temperatures of the granulator temperature control unit (V33 and V34).

The Q residuals contribution plots (Figure 76) show that the mass flow of the granulation liquid (V5) was deviating from normal operation.

In the beginning of the process both statistics show an abnormal behavior (Figure 64 and Figure 65). The variables that contribute to this behavior are mainly the power of the granulator drive (V7), temperature of the granulator barrel (V32) and torque (V35) (Figure 77). As explained before, this can be due to the friction between material and granulator wall. At the end, the process is still out-of-control according to the Q residuals (Figure 65). A contribution plot at time point 5600s (Figure 76) shows that several pressures, temperatures and humidites contribute to the deviation from the calibration runs' behavior. The granulator drive powder (V7), granulator barrel temperature (V32) and torque (V35) are also visibly out-of-control.

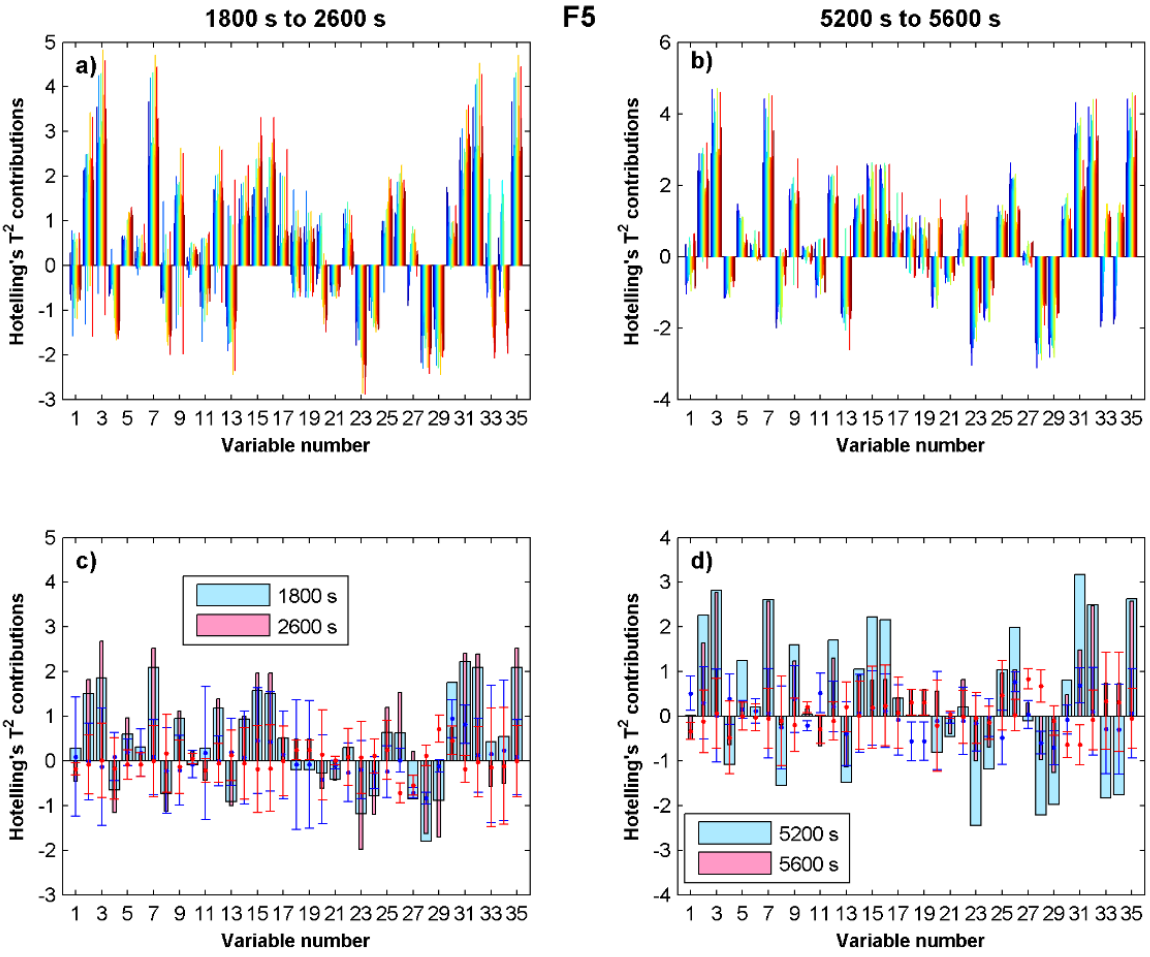


Figure 75 – Run F5 - variables contributions for the Hotelling's T^2 statistic for time: a) between 1800 and 2600 s, b) between 5200 and 5600 s, c) 1800 and 2600 s, d) 5200 and 5600 s

.

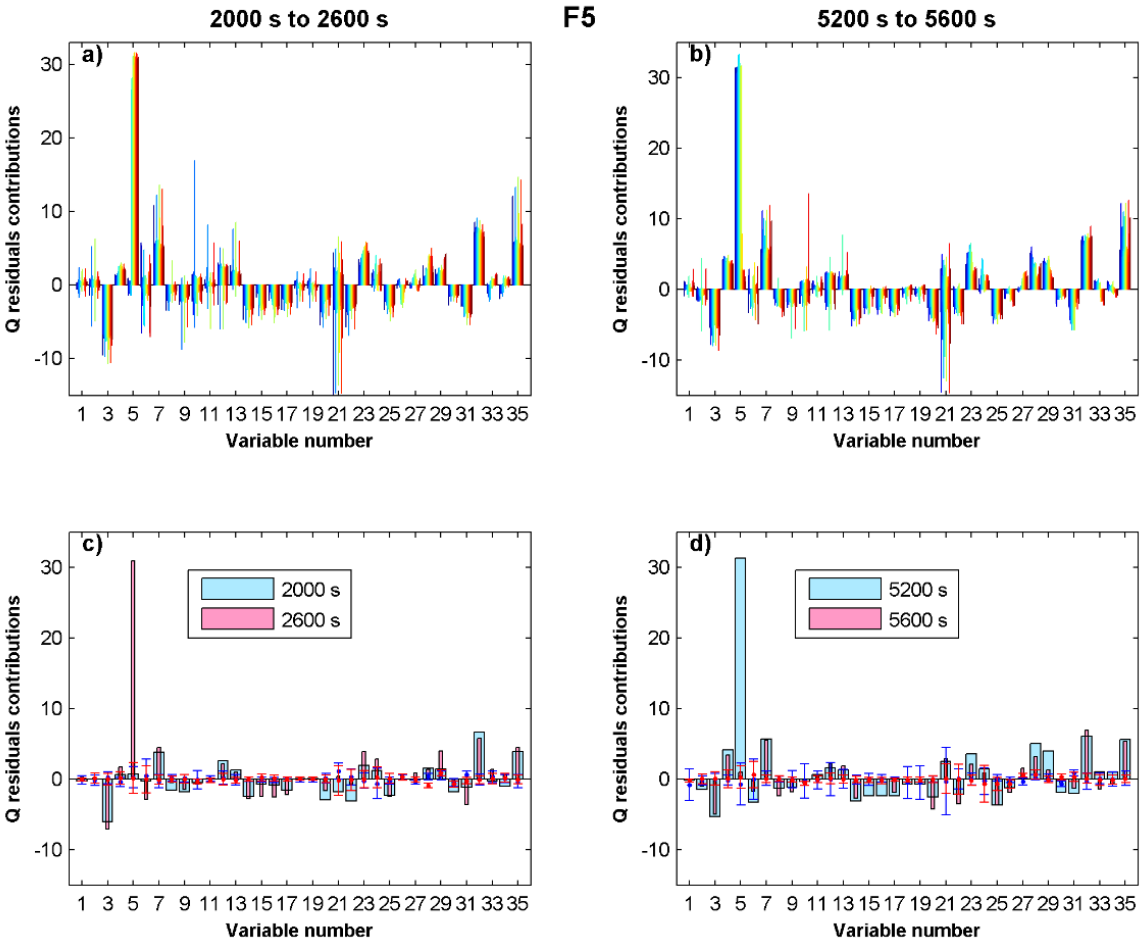


Figure 76 – Run F5 - variables contributions for Q residuals statistic for time: a) between 2000 and 2600 s, b) between 5200 and 5600 s, c) 2000 and 2600s, d) 5200 and 5600s.

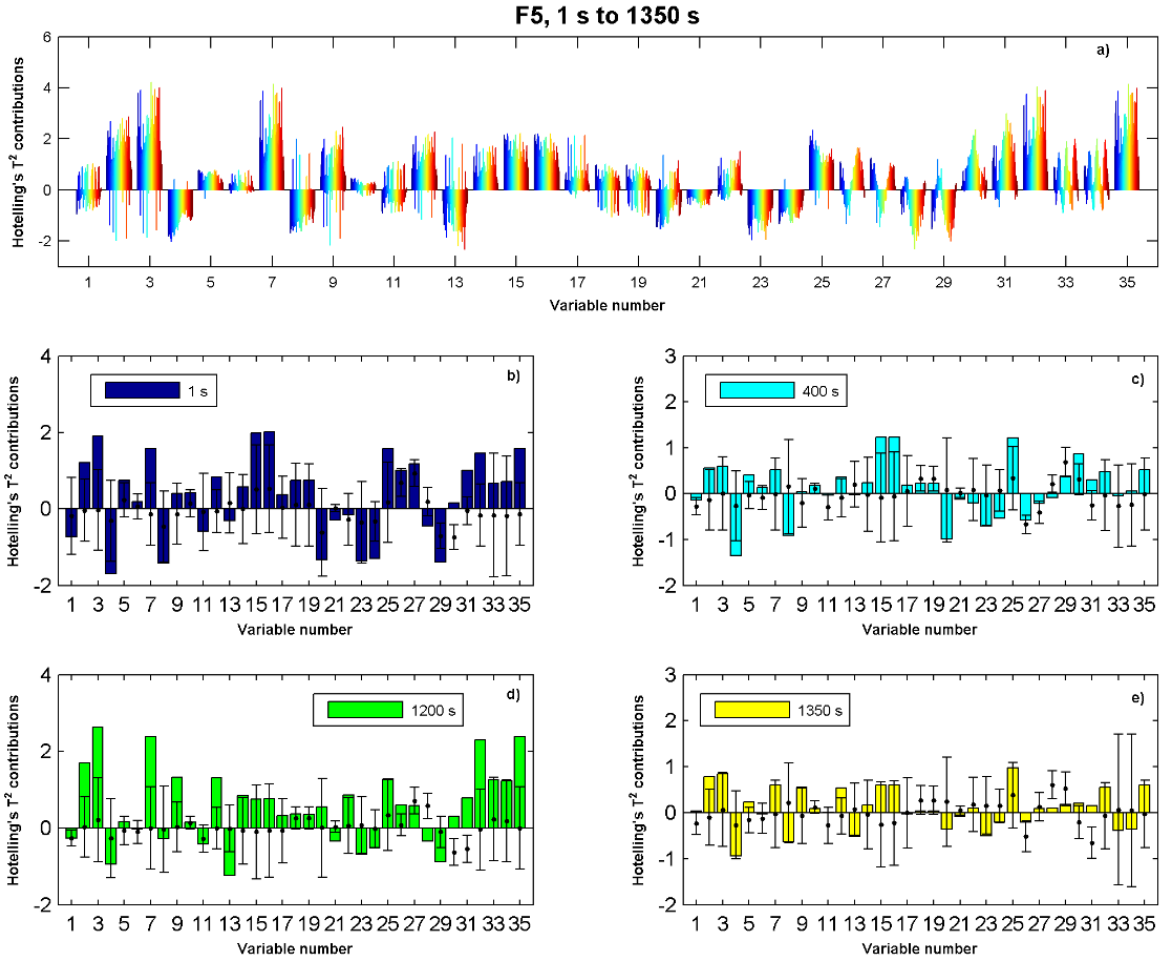


Figure 77 – Run F5 – variables contribution for Hotelling's T^2 statistic for time: a) between 1 s and 1350 s, b) 1 s, c) 400s, d) 1200s, e) 1350s.

Run F6 (powder mass flow)

For run F6 the powder dosing unit mass flow (V6) was changed between 25 to 21.7 kg/h. The Hotelling's T^2 (Figure 64) and Q residuals (Figure 65) are affected by the change in the setpoint. The plots of the variable contributions for this period reveals that multiple variables were contributing to this increase of the Hotelling's T^2 values (Figure 78). The contribution plots for the Q residuals over the time of the provoked changed (Figure 79) show that the variable responsible is, as expected, the powder dosing unit mass flow (V6).

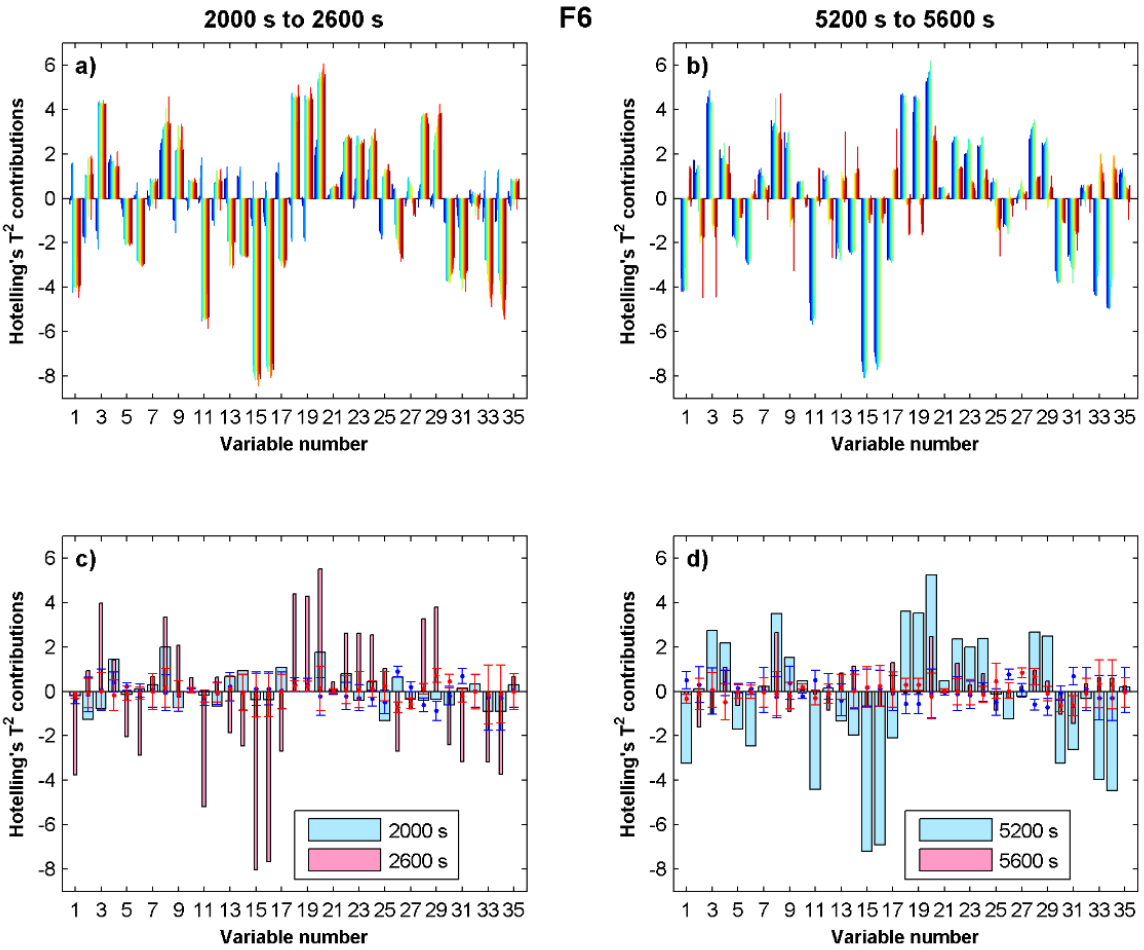


Figure 78. Run F6 - variables contributions for the Hotelling's T^2 statistic for time: a) between 2000 and 2600 s, b) between 5200 and 5600 s, c) 2000 and 2600 s, d) 5200 and 5600 s.

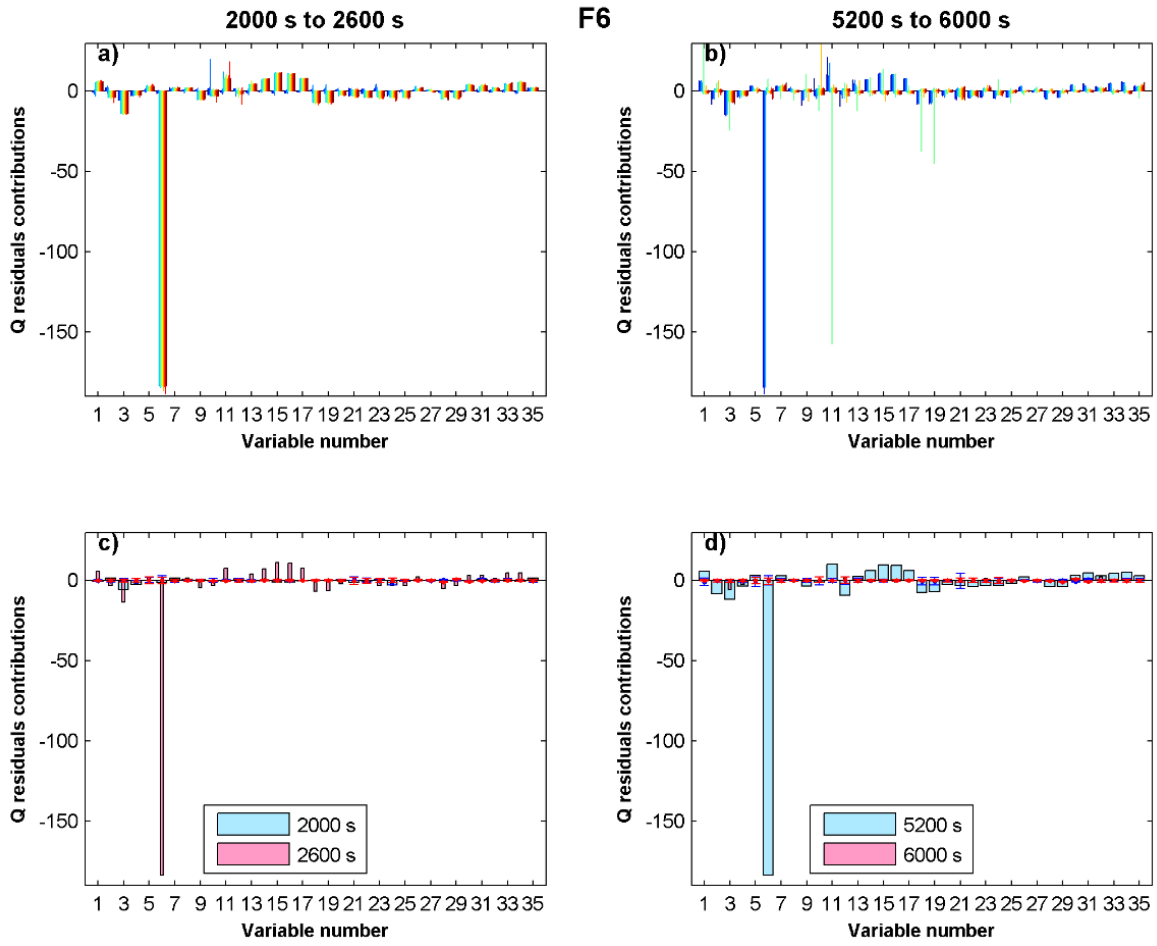


Figure 79. Run F6 - variables contributions for the Q residuals statistic for time: a) between 2000 and 2600 s, b) between 5200 and 6000 s, c) 2000 and 2600s, d) 5200 and 6000s.

Run F7 (granulator screws' speed)

The final setpoint changed was the granulator screws speed (V21) that was altered between 700 and 900 rpm. Again, both Hotelling's T^2 (Figure 64) and Q residuals (Figure 65) statistics were affected by the provoked change. Contribution plots show that several variables related with the granulator (V7, V32, V33 and V35) contribute to the change in the Hotelling's T^2 values as well as pressures at the outlet of the dryer before and after the HEPA filter (V15, and V16) (Figure 80). The variable that contribute to the change is the Q residuals values was the setpoint variable (V21) (Figure 81).

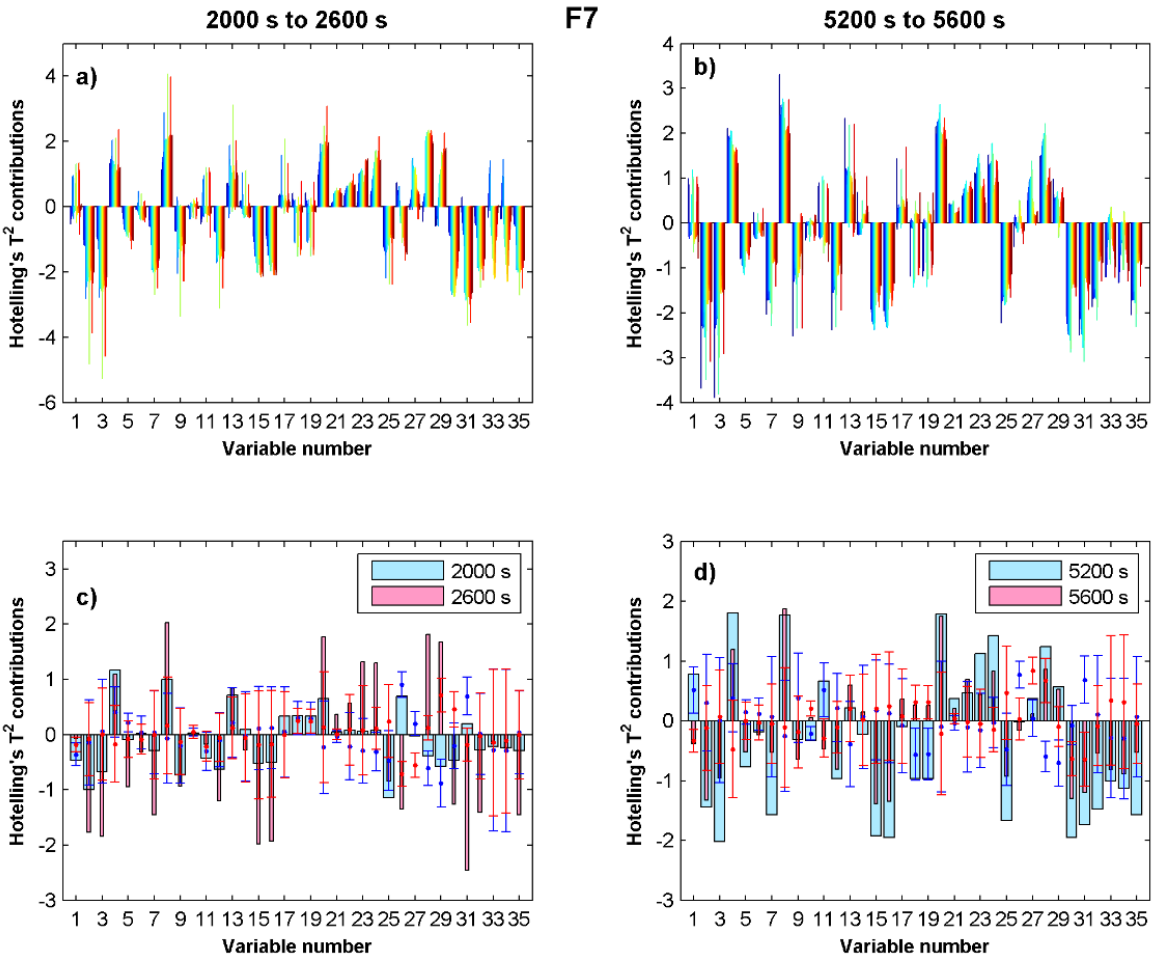


Figure 80 – Run F7 - variables contributions for the Hotelling's T^2 statistic for time: a) between 2000 and 2600 s, b) between 5200 and 5600 s, c) 2000 and 2600 s, d) 5200 and 5600 s.

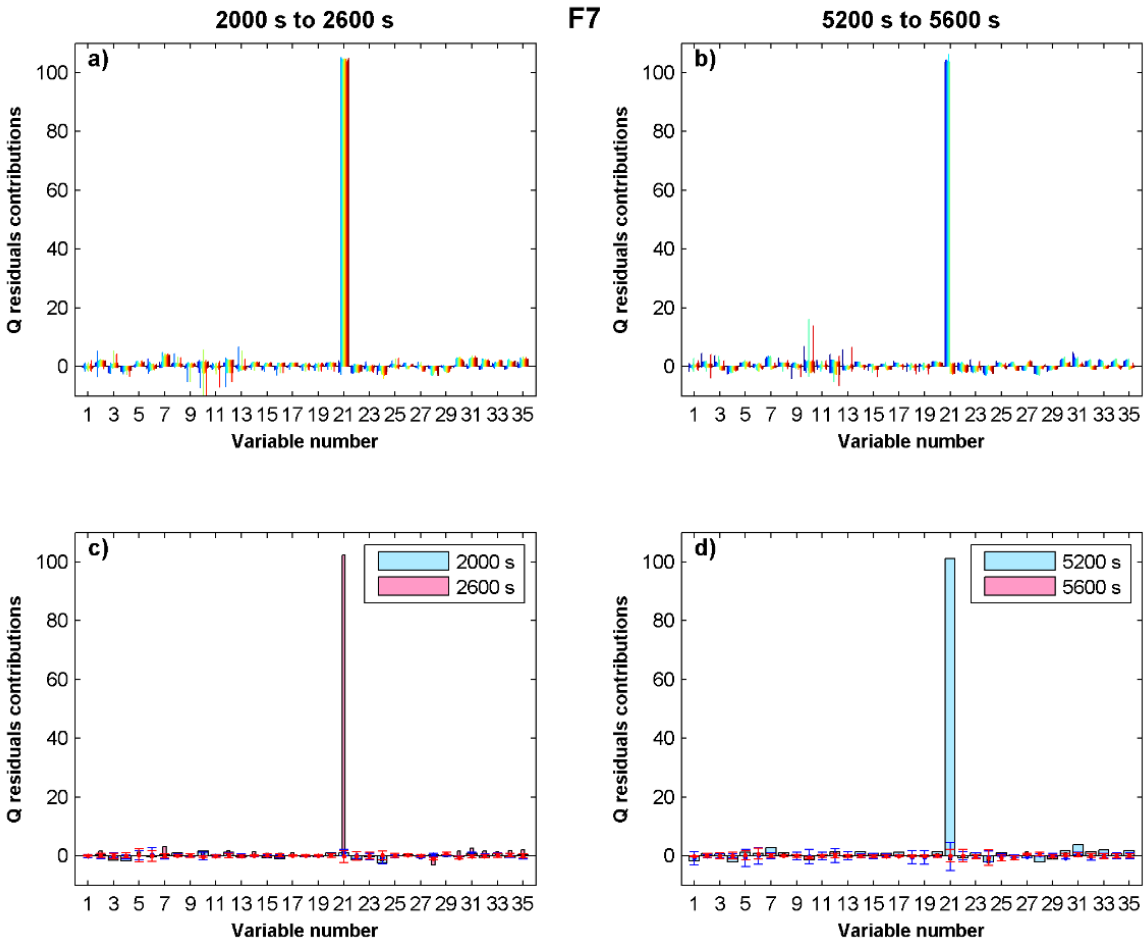


Figure 81 – Run F7 - variables contributions for the Q residuals statistic for time: a) between 2000 and 2600 s, b) between 5200 and 5600 s, c) 2000 and 2600 s, d) 5200 and 5600 s.

5.5.4 Conclusions

The imposed disturbance in the dryer air flow (F1) was the most difficult root cause to diagnose based on the variables' contributions since it affected many pressures, temperatures, humidities and flows in both dryer and product control units. The imposed disturbance of the temperature of the granulator (F2) was more visibly captured in the Hotelling's T^2 chart. All other imposed disturbances from normal operation (temperature of the dryer air (F4), mass flow of the granulation liquid (F5), mass flow in the powder dosing unit (F6) and speed of the granulator screws (F7) could not be explained by the model and thus were more clearly captured in the Q residuals statistic, as expected. Several runs deviated from normal operation from the start of operation (even though they were all running at the NOC setpoints from the start), mainly due to differences in variables related to air flow, temperatures, humidities and pressures. In most cases, after the correction of the deviation, by setting the variable back to the setpoint, the process was returned closer

to the initial steady state of operation but, sometimes there are clear differences between the initial and final steady states. Other important process phenomena were identified such as the deposition of powder over the dryer filters over time and increases in the temperature of the granulator barrel, torque and power of the granulator drive (F1, F4 and F5). Not cleaning the system before a new run showed no significant impact (F3).

A MSPC strategy based on PCA was successfully developed as the model successfully captured the continuous manufacturing line's operating behavior. All deviations were usefully detected and contribution plots were insightful when assigning a possible root cause.

Chapter 6

Broader international context, relevance and future perspectives

6.1 Broader international context and relevance

In 2002 the FDA announced the PAT initiative, shortly followed up by the publication of a PAT guidance in September, 2004 [9]. Prior to publishing the guideline, the FDA conducted independent research and also worked together with companies such as Pfizer and Novartis, and academic institutions like Duquesne University Centre for Pharmaceutical Technology (DCPT), the Centre for Pharmaceutical Processing Research (CPPR) at Purdue University, and Engineering Research Centre for Structured Organic Composites (C–SOC) at Rutgers. This was done to ensure that science and technological concepts were incorporated into policy making, training and regulatory aspects in the implementation of PAT in the pharmaceutical industry. This initiated a change in the pharmaceutical industry in the direction of a risk and science-based approach and contributed in delivering the current Good Manufacturing Practices approach [26]. They became part of "a regulatory framework that will encourage the voluntary development and implementation of innovative approaches in pharmaceutical development, manufacturing, and quality assurance" [26]. P PAT is also a methodological framework sitting within a cadre of concepts that include, process understanding, QbD, risk-based regulatory approaches, CPPs and CQAs, design space, MSPC, process control, real-time release testing (RTRT) and continuous improvement. PAT concepts started being used to monitor and control processes, warranting process understanding, reducing the impact of raw materials and process variability in the final product quality.

Product quality started being built-in instead of tested. PAT is presently a hot topic principally in the pharmaceutical and fine chemical industries and it will remain as such in the coming years.

The PAT initiative has been since supported by the EMA [22] that set up a PAT team in November 2003 to support the activities in the European Union. In 2006, the EMA PAT team published a reflection on chemical, pharmaceutical and biological information to be included in PAT implementations. This document that highlighted specific points on manufacturing process development and control of critical steps, intermediates, excipients and drug product. Furthermore, risk management, DoE, data acquisition and establishing a design space based on a chemometric approach are also addressed, among others [228].

The Japanese Ministry of Health, Labor, and Welfare, the ICH guidelines [23, 24] and the STM International Technical Committee E55 have also been supportive of PAT.

The FDA PAT initiative itself has also grown since and now involves organizations and consortia from different parts of the globe that support the interpretation and delivery of PAT to real processes [229]:

- ASTM International Committee E55 (founded in 2003) – focuses on standardizing nomenclature, definition of terms, recommending practices, guides, test methods, specifications and performance standards for the manufacture of pharmaceutical products. Standards from E55 is published in the Annual Book of ASTM Standards, Volume 13.01.
- Center for Process Analytical Chemistry (CPAC, founded in 1984) – consortium formed at the University of Washington which includes of members from industry, U.S.-based laboratories and government agencies. It has been involved on the development of novel measurement approaches, endorsing research, communications and partnership in areas related to PAT and process control.
- Centre for Process Analysis and Control Technology (CPACT, founded in 1997) – currently includes seven universities and eighteen companies in the UK. Pursues at enabling discussion between chemical and process engineers, analytical chemists, chemometricians, control systems engineers, etc. from both academia and industry to produce solutions to common problems in process monitoring and control.
- International Society of Automation (ISA, founded in 1945) – aims at setting global standards for automation via certification, education, publication, developing standards and hosting conferences and exhibitions. Currently has

more than 4000 members, and has published over 150 standards for automation and control systems.

- International Society of Pharmaceutical Engineering (ISPE, founded in 1980) – is the world's largest not-for-profit society in pharmaceutical science and manufacturing. They educate and advance pharmaceutical industry and professionals. Has over 25000 members among experts, technologists, regulators, consultants and students in 90 countries.
- National Institute for Pharmaceutical Technology and Education (NIPTE) – not-for-profit organization formed buy ten academic members from North America. Together with the U.S. FDA funds research projects related to implementation of QbD and produces roadmaps for research and pharmaceutical technology education.
- The ICH of Technical Requirements for Registration of Pharmaceuticals for Human Use – collaboration between worldwide authorities and pharmaceutical industrial experts from Europe, Japan and the United States. Aims at international harmonization through discussions on scientific and technical aspects of new medicinal product registration.

In addition to the presented consortia, pharmaceutical companies worldwide have also independently allocated resources in the effort to the inclusion of the PAT concept in their daily activities [166, 230-232].

The value applying real-time monitoring strategies specifically in the commercial manufacture of drug substances has been a topic of great discussion within companies and also in various industry consortia. A superior process understanding leads to an improved manufacturing. While the savings will depend on the company a PAT program can result in major savings in four key areas: reduced cost of quality, reduced capital investment, reduced inventory, and increased speed to market [233].

While there is strong value in the use of PAT concepts in the pharmaceutical industry, cost of replacing traditional off-line analyses with real-time monitoring is less apparent, while the pharmaceutical industry still relies mainly on a batch manufacturing model with few opportunities for feedback control. The multipurpose manufacturing plants that produce limited batches of a given product make the cost-benefit proposition less captivating and suffering a greater risk from a technical implementation and regulatory acceptance perspective [234].

However, the pharmaceutical industry worldwide is now evolving from an industry that relies on batch processing to a fully-integrated continuous processing model [20-24]. PAT tools are an essential part of integrated continuous process. Without PAT to provide

process and product knowledge, RTRT would not be possible and this would significantly hinder the multiple advantages of producing continuously. Continuous manufacturing requires an in-depth process knowledge and also new technologies and equipment's plus a well-defined control strategy. To drive the implementation of the continuous manufacturing model and enable scientific discussion several industries and equipment manufacturers have come together in a global effort by forming several consortia [19]:

- Blue Sky vision project – a 65-million-dollar collaboration between Novartis and the Massachusetts Institute of Technology (MIT). It supports the development of a truly continuous manufacturing line from the primary manufacturing stage and as such, involves the development novel new synthesis approaches and technologies.
- Britest – a membership-based consortium including both pharmaceutical companies such as Pfizer, Hovione, Abbvie and Astra Zeneca and universities (University of Nottingham, Newcastle University, Purdue University, University of Limerick). Their main focus is to develop ground-breaking approaches to manufacturing and process design.
- Center of Structured Organic Particulate Systems (C-SOPS, 2006) – formed by academic partners (Rutgers University, Purdue University, New Jersey Institute of Technology and University of Puerto Rico) and funded by the National Science Foundation, about fourty industrial member companies and several equipment manufacturers. They focus mainly on three key areas: manufacturing science, composite synthesis and characterization and functionalization.
- European Consortium for Continuous Pharmaceutical Manufacturing – a partnership of the Research Centre Pharmaceutical Engineering (RCPE) (Graz, Austria) with several industrial partners such as AstraZeneca (UK), Automatic Pelletizing Systems, Bayer Health Care (Germany), GEA Pharma Systems (Belgium), Siemens (Austria), UCB Pharma (Belgium) and four academic partners (Graz University of Technology, University of Eastern Finland, Heinrich Heine University Düsseldorf, Ghent University). The main aim of this consortium focuses is the research, development and implementation of continuous manufacturing strategies for solid oral dosage forms.
- L. B. Böhle Technology Center – association between equipment manufacturers (Korsch, Gericke and Böhle), companies specialized in PAT equipment and control strategies (Kaiser Optical Systems, Kraemer

Elektronik and Sentronic) and academy (Heinrich Heine University Düsseldorf, RWTH Aachen University and RCPE Graz) focusing on the development and implementation of a modular production line for continuous manufacturing.

Continuous manufacturing is the future of the pharmaceutical industry which is advancing toward this simpler, faster, more efficient, less expensive, and more reliable production model. Moving from batch to continuous manufacturing improves final product quality, generates more flexibility in terms of product throughput, reduces times of processing and requires a lower capital investment.

6.2 Future perspectives

The work developed in this thesis open perspectives of future work on the use of PAT concepts and tools to supervise batch and continuous processes in the pharmaceutical industry.

Particle size is a paramount CQA in several pharmaceutical processes. In this dissertation, several in-process techniques were studied as potential PAT tools for particle size determination and compared to well-known, widespread, off-line reference methods. In this work, only nearly-spherical or spherical particles were analyzed. The effect of particle shape on the estimation of particle size distribution is of great interest and should be addressed in future. Ultimately, a future aim is to replace off-line techniques by these novel technologies during the process, for monitoring and/or control purposes with the objective of simultaneously improving product quality, and reducing manufacturing costs.

In-process monitoring of cocrystallization CQAs such as particle size and crystallinity can be key for ensuring that a cocrystal of desired quality is obtained. This dissertation includes the first comparison between PLS and OPLS for the monitoring of a batch co-crystallization process, using NIR spectroscopy. Future studies can include the implementation of the previously assessed on-line techniques, to obtain particle size information as well. This work opens up the possibility to further assess the potential of the still unexplored OPLS-based multivariate batch statistical process control to other cocrystallization processes as well as to other types of batch processes.

The ConsiGma™-25 continuous manufacturing line based on twin-screw granulation and fluid bed drying was described in this dissertation. An in-depth knowledge was acquired about the granulation by using in-process measurements made by univariate sensors implemented in the line. MSPC and BSPC concepts were successfully applied to this data. The next step is to link the MSPC model to raw material properties

and CQAs (e.g., particle size, bulk density, porosity, moisture content, etc.), establishing the link with product quality, and opening up the possibility of the implementation of control strategies to obtain product of desirable quality. Ultimately, the aim is to achieve a real-time release. Acceptable procedures for handling deviations including the removal of non-conforming material from the process stream must be investigated.

To make the adoption of PAT and QbD a reality, and turn continuous manufacturing into the industry standard, there are still aspects that need to be overcome in the future. The concern over delays on regulatory approval or misuse of data from PAT leads companies to use PAT methods in parallel with conventional methods for process insight, and control. This leads to duplication and inefficiency, spreading a sentiment of mistrust toward PAT and QbD. A systematic approach to education and professional development is required in order to make scientists, engineers, academics, regulators and suppliers familiar with the concepts and science supporting the PAT/QbD paradigm. In the latest years, several dedicated QbD/PAT programs have been introduced in Europe, to fill the gap of the undergraduate courses which lack specific knowledge regarding pharmaceutical manufacture and product design. Today, several universities across Europe provide study programs of different lengths where QbD and PAT are taught [235]. As the pharmaceutical industry and regulatory agencies gain more experience with PAT, QbD and continuous manufacturing, several regulatory aspects will need to adjusted to minimize the gap between principles and reality. To accomplish this, better communication channels between scientists, regulators and manufacturers is necessary. In the future, investment in PAT equipment and innovative sensors development is also of great importance.

Chapter 7

General conclusions

PAT tools for determining particle size in-line (Spatial Filtering Velocimetry, Focused Beam Reflectance Measurements, Photometric Stereo Imaging and the Eyecon™ technology) were compared with well-known and widespread off-line methodologies (sieve analysis and laser diffraction). Significant dissimilarities were found and explained according to previous knowledge about the assessed techniques and used instrumentation. Presenting the materials in an appropriate concentration was found important for some of the techniques (e.g. SFV, PSI); an accurate sampling technique was also one of the impacting factor on the obtained results (e.g. SFV, PSI, Eyecon™).

PAT tools based on the application of multivariate data analysis to data coming from multivariate (NIR spectroscopy) and univariate sensors (pressures, temperatures, humidity, etc.) were utilized to extract knowledge from batch and continuous processes. Moreover, the use of multivariate statistical process control concepts allowed the detection and diagnosis of disturbances imposed to the processes. OPLS and PLS-based BSPC methods were found suitable for the supervision of a co-crystallization process, monitored on-line, with near infrared spectroscopy. OPLS resulted in a better Hoteling's T^2 statistic calculation that permitted an improved detection of process disturbances. The developed approaches were not perfect since in the Hotelling's T^2 control charts some false positives and negatives were found. An in-depth knowledge was obtained about a granulation and drying process, performed on a ConsiGma™-25 continuous manufacturing line. Three manufacturing phases were identified: a startup-phased, a steady state phase and a shutdown phased. Several process phenomena such as the deposition of fines on the dryer filters and the material accumulation over the granulator barrel walls during startup could be extracted form the univariate data. Latent variable

models such as PCA, PLS and their extensions were used for the purpose of developing MSPC or BSPC strategies for monitoring the different units. Furthermore, a PCA-based MSPC approach for a continuous manufacturing line allowed to correctly detect and diagnose disturbances in the process.

Overall, this thesis presents several tools, which are meant to facilitate the introduction of PAT in the pharmaceutical industry for batch process monitoring, and as an essential part of a fully-integrated continuous manufacturing process.

References

[1] E. Petrova, Innovation in the pharmaceutical industry: the process of drug discovery and development, in: M. Ding, J. Eliashberg, S. Stremersch (Eds.) Innovation and Marketing in the Pharmaceutical Industry: Emerging Practices, Research, and Policies, Springer New York, New York, NY, 2014, pp. 19-81.
[2] A. Behr, V.A. Brehme, C.L.J. Ewers, H. Grön, T. Kimmel, S. Küppers, I. Symietz, New developments in chemical engineering for the production of drug substances, Engineering in Life Sciences, 4 (2004) 15-24.
[3] N.N. Malik, Key issues in the pharmaceutical industry: consequences on R&D, Expert Opinion on Drug Discovery, 4 (2009) 15-19.
[4] J. Karamehic, O. Ridic, G. Ridic, T. Jukic, J. Coric, D. Subasic, M. Panjeta, A. Saban, L. Zunic, I. Masic, Financial aspects and the future of the pharmaceutical industry in the United States of America, Materia Socio-Medica, 25 (2013) 286-290.
[5] M. Teżyk, B. Milanowski, A. Ernst, J. Lulek, Recent progress in continuous and semi-continuous processing of solid oral dosage forms: a review, Drug Development and Industrial Pharmacy, 42 (2016) 1195-1214.
[6] K. Plumb, Continuous processing in the pharmaceutical industry: changing the mind set, Chemical Engineering Research and Design, 83 (2005) 730-738.
[7] T. Melton, The benefits of lean manufacturing, Chemical Engineering Research and Design, 83 (2005) 662-673.
[8] Y. Zhang, J. An, Modeling and monitoring of multimode transition process based on reconstruction, Information Sciences, 279 (2014) 176-185.
[9] FDA, Guidance for industry PAT - a frameword for innovative pharmaceutical development, manufacturing and quality assurance, Process Analytical Technology Initiative, (2004).
[10] International Conference on Harmonization (ICH) of technical requirements for registration of pharmaceuticals for human use, Topic Q8(R2): Pharmaceutical Development, Geneva, 2009.
[11] International Conference on Harmonization (ICH) of technical requirements for registration of pharmaceuticals for human use, Topic Q9: Quality Risk Management, Geneva, 2009.
[12] International Conference on Harmonization (ICH) of technical requirements for registration of pharmaceuticals for human use, Topic Q10: Pharmaceutical Quality Systems, Geneva, 2009.
[13] S.L. Lee, T.F. O'Connor, X. Yang, C.N. Cruz, S. Chatterjee, R.D. Madurawe, C.M.V. Moore, L.X. Yu, J. Woodcock, Modernizing pharmaceutical manufacturing: from batch to continuous production, Journal of Pharmaceutical Innovation, 10 (2015) 191-199.
[14] S. Chatterjee, FDA perspective on continuous manufacturing, IFPAC Annual Meeting, Baltimore, MD, 2012.
[15] G. Allison, Y.T. Cain, C. Cooney, T. Garcia, T.G. Bizjak, O. Holte, N. Jagota, B. Komas, E. Korakianiti, D. Kourti, R. Madurawe, E. Morefield, F. Montgomery, M. Nasr, W. Randolph, J.-L. Robert, D. Rudd, D. Zezza, Regulatory and quality considerations for continuous manufacturing May 20–21, 2014 Continuous Manufacturing Symposium, Journal of Pharmaceutical Sciences, 104 (2015) 803-812.
[16] S. Byrn, M. Futran, H. Thomas, E. Jayjock, N. Maron, R.F. Meyer, A.S. Myerson, M.P. Thien, B.L. Trout, Achieving continuous manufacturing for final dosage formation: challenges and how to meet them. May 20-21, 2014 Continuous Manufacturing Symposium, Journal of Pharmaceutical Sciences, 104 792-802.
[17] M.J. Mollan Jr, M. Lodaya, Continuous processing in pharmaceutical manufacturing, Pharmaceutical Manufacturing Magazine, (2004) 1-9.

[18] K. Nepveux, F. Pavlou, S. Whitfield, K. Schoeters, A. Weiler, Continuous processing: is the pharma industry finally coming round to the idea, Pharmaceutical Technology Europe, 22 (2010).
[19] A. Pellek, P. Van Arnum, Continuous processing: moving with or against the manufacturing flow, Pharmaceutical Technology, 9 (2008) 52-58.
[20] J.D. Rockoff, Drug making breaks away from its old ways, The Wall Street Journal, 8 (2015).
[21] S. Massey, (2016), Making the switch: continuous manufacturing vs. batch processing of pharmaceuticals, Massey, Sarah; Retrieved 08/01/2017, from http://xtalks.com/Continuous-And-Batch-Manufacturing-Pharmaceuticals.ashx
[22] J. Butschli, (2013), AstraZeneca invests in continuous wet granulation unit; Retrieved 14/01/2017, from http://www.healthcarepackaging.com/trends-and-issues/processing-packaging-integration/astrazeneca-invests-continuous-wet-granulation
[23] K. Langhauser, Janssen's Historic FDA Approval. The FDA has approved -- for the first time in history -- a manufacturer's production method change from "batch" to continuous manufacturing, 2017; Retrieved 08/01/2017, from http://www.pharmamanufacturing.com/articles/2016/janssens-historic-fda-approval/
[24] A.M. Thayer, (2016), Continuous drug production advances; Retrieved 17/01/17, from http://cen.acs.org/articles/94/i11/Continuous-drug-production-advances.html?type=paidArticleContent
[25] L. Abboud, S. Heinsley, (2003), New prescription for drug makers: update the plants after years of neglect, industry focuses on Manufacturing; FDA acts as a catalyst, The Wall Street Journal. Retrieved 09/01/17, from https://www.wsj.com/articles/SB10625358403931000
[26] U. Food, D. Administration, Pharmaceutical CGMPs for the 21st Century—a risk-based approach, Final report. Rockville, MD, (2004).
[27] FDA, Guidance for industry: PAT — a framework for innovative pharmaceutical, (2004).
[28] K. Theodora, Quality by design in the pharmaceutical industry: process modelling, monitoring and control using latent variable methods, IFAC Proceedings Volumes, 42 (2009) 36-41.
[29] N. Cohen, PAT: process analytical technology an FDA and industry effort, ISPE Midwest Chapter Meeting, 2005.
[30] E.M. Agency, Quality by design Retrieved 19/12/2016, from http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/document_listing/document_listing_000162.jsp
[31] A.S. Rathore, R. Bhambure, V. Ghare, Process analytical technology (PAT) for biopharmaceutical products, Analytical and Bioanalytical Chemistry, 398 (2010) 137-154.
[32] J. Rantanen, J. Khinast, The future of pharmaceutical manufacturing sciences, Journal of Pharmaceutical Sciences, 104 (2015) 3612-3638.
[33] J.F. MacGregor, T. Kourti, Statistical process control of multivariate processes, Control Engineering Practice, 3 (1995) 403-414.
[34] J.V. Kresta, J.F. Macgregor, T.E. Marlin, Multivariate statistical monitoring of process operating performance, The Canadian Journal of Chemical Engineering, 69 (1991) 35-47.
[35] T. Kourti, J. Lee, J.F.J.F. MacGregor, Experiences with industrial applications of projection methods for multivariate statistical process control, Computers & Chemical Engineering, 20 (1996) S745-S750.
[36] T. Kourti, Process analytical technology beyond real-time analyzers: the role of multivariate analysis, Critical Reviews in Analytical Chemistry, 36 (2006) 257-278.
[37] T. Kourti, J.F.J.F. MacGregor, Process analysis, monitoring and diagnosis, using multivariate projection methods, Chemometrics and Intelligent Laboratory Systems, 28 (1995) 3-21.

[38] A. Ferrer, Multivariate statistical process control based on principal component analysis (MSPC-PCA): some reflections and a case study in an autobody assembly process, Quality Engineering, 19 (2007) 311-325.
[39] S. Wold, K.I.M. Esbensen, P. Geladi, Principal component analysis, Chemometrics and Intelligent Laboratory Systems, 2 (1987) 37-52.
[40] G.H. Dunteman, Principal component analysis, 2002.
[41] I.T. Jolliffe, Principal component analysis, second edition, Encyclopedia of Statistics in Behavioral Science, 30 (2002) 487.
[42] A. Höskuldsson, A. Hoskuldsson, PLS regression methods, Journal of Chemometrics, 2 (1988) 211-228.
[43] S. Wold, E. Johansson, M. Cocchi, PSL - partial least-squares projections to latent structures, 3D QSAR Drug Design, 1993, pp. 523-550.
[44] S. Wold, M. Sjöström, L. Eriksson, PLS-regression: A basic tool of chemometrics, Chemometrics and Intelligent Laboratory Systems, 2001, pp. 109-130.
[45] P. Nomikos, J.F. MacGregor, Monitoring batch processes using multiway principal component analysis, AIChE Journal, 40 (1994) 1361-1375.
[46] P. Nomikos, J.F. MacGregor, Multi-way partial least squares in monitoring batch processes, Chemometrics and Intelligent Laboratory Systems, 30 (1995) 97-108.
[47] J.F.J.F. MacGregor, C. Jaeckle, C. Kiparissides, M. Koutoudi, Process monitoring and diagnosis by multiblock PLS methods, AIChE Journal, 40 (1994) 826-838.
[48] J.A. Westerhuis, T. Kourti, J.F. MacGregor, Analysis of multiblock and hierarchical PCA and PLS models, Journal of Chemometrics, 12 (1998) 301-321.
[49] T. Kourti, Multivariate dynamic data modeling for analysis and statistical process control of batch processes, start-ups and grade transitions, Journal of Chemometrics, 17 (2003) 93-109.
[50] J. Chen, K.C. Liu, On-line batch process monitoring using dynamic PCA and dynamic PLS models, Chemical Engineering Science, 57 (2002) 63-75.
[51] W. Ku, R.H. Storer, C. Georgakis, Disturbance detection and isolation by dynamic principal component analysis, Chemometrics and Intelligent Laboratory Systems, 1995, pp. 179-196.
[52] M. Mortsell, M. Gulliksson, An overview of some non-linear techniques in Chemometrics, Fibre Science and Communication Network Report Series, 2001.
[53] J. Trygg, S. Wold, Orthogonal projections to latent structures (O-PLS), Journal of Chemometrics, 16 (2002) 119-128.
[54] L. De Lathauwer, B. De Moor, J. Vandewalle, An introduction to independent component analysis, Journal of Chemometrics, 14 (2000) 123-149.
[55] a. Hyvärinen, E. Oja, Independent component analysis: algorithms and applications., Neural networks : the official journal of the International Neural Network Society, 13 (2000) 411-430.
[56] P. Comon, Independent component analysis, a new concept?, Signal Processing, 36 (1994) 287-314.
[57] A.K. Jain, K.M. Mohiuddin, Artificial neural networks: a tutorial, Computer, 29 (1996) 31-44.
[58] I.A. Basheer, M. Hajmeer, Artificial neural networks: fundamentals, computing, design, and application, Journal of Microbiological Methods, 43 (2000) 3-31.
[59] M.A. Hearst, S.T. Dumais, E. Osman, Support vector machines, IEEE Intelligent Systems and their Applications, 1 (1998) 1-8.
[60] R.G. Brereton, G.R. Lloyd, Support vector machines for classification and regression., Analyst, 135 (2010) 230-267.
[61] D. Meyer, Support vector machines, R News, 2 (2001) 23-26.
[62] S.T. Dumais, E. Osuna, J. Platt, Support vector machines, IEEE Intelligent Systems, 13 (1998) 18-28.
[63] F. Arteaga, A. Ferrer, Dealing with missing data in MSPC: Several methods, different interpretations, some examples, Journal of Chemometrics, 2002, pp. 408-418.

[64] F. Arteaga, A. Ferrer, Framework for regression-based missing data imputation methods in on-line MSPC, Journal of Chemometrics, 19 (2005) 439-447.
[65] P. Nomikos, W.B. Forum, Detection and diagnosis of abnormal batch operations based on multi-way principal component analysis World Batch Forum, Toronto, May 1996, ISA Transactions, 35 (1996) 259-266.
[66] P. Nomikos, J. MacGregor, Multivariate SPC charts for monitoring batch processes, Technometrics, (1995).
[67] R. Bro, PARAFAC. Tutorial and applications, Chemometrics and Intelligent Laboratory Systems, 1997, pp. 149-171.
[68] G. Tomasi, R. Bro, PARAFAC and missing values, Chemometrics and Intelligent Laboratory Systems, 75 (2005) 163-180.
[69] J.A. Westerhuis, T. Kourti, J.F. Macgregor, Comparing alternative approaches for multivariate statistical analysis of batch process data, Journal of Chemometrics, 13 (1999) 397-413.
[70] L.H. Chiang, R. Leardi, R.J. Pell, M.B. Seasholtz, Industrial experiences with multivariate statistical analysis of batch process data, Chemometrics and Intelligent Laboratory Systems, 81 (2006) 109-119.
[71] D.J. Louwerse, a.K. Smilde, Multivariate statistical process control of batch processes based on three-way models, Chemical Engineering Science, 55 (2000) 1225-1235.
[72] A. Simoglou, E.B. Martin, A.J. Morris, Multivariate statistical process control of an industrial fluidised-bed reactor, Control Engineering Practice, 8 (2000) 893-909.
[73] N. Souihi, A. Lindegren, L. Eriksson, J. Trygg, OPLS in batch monitoring - opens up new opportunities, Analytica Chimica Acta, 857 (2015) 28-38.
[74] T. Kourti, Abnormal situation detection and projection methods—industrial applications, Chemometrics and Intelligent Laboratory Systems, 76 (2005) 215-220.
[75] D. Neogi, C.E. Schlags, Multivariate statistical analysis of an emulsion batch process, Industrial & Engineering Chemistry Research, 37 (1998) 3971-3979.
[76] J.M.M. González-Martínez, a. Ferrer, J.a.A. Westerhuis, Real-time synchronization of batch trajectories for on-line multivariate statistical process control using Dynamic Time Warping, Chemometrics and Intelligent Laboratory Systems, 105 (2011) 195-206.
[77] A. Kassidas, J.F. MacGregor, P.A. Taylor, Synchronization of batch trajectories using dynamic time warping, AIChE Journal, 44 (1998) 864-875.
[78] C. Duchesne, J.F. MacGregor, Multivariate analysis and optimization of process variable trajectories for batch processes, Chemometrics and Intelligent Laboratory Systems, 51 (2000) 125-137.
[79] C. Duchesne, T. Kourti, J.F. MacGregor, Multivariate monitoring of startups, restarts and grade transitions using projection methods, Proceedings of the American Control Conference, 2003, pp. 5423-5426.
[80] T. Chen, On reducing false alarms in multivariate statistical process control, Chemical Engineering Research and Design, 88 (2010) 430-436.
[81] P. Miller, R.E. Swanson, C.E. Heckler, Contribution Plots: A Missing Link in Multivariate Quality Control, Applied Mathematics and Computer Science, 8 (1998) 775-792.
[82] J.A. Westerhuis, S.P. Gurden, A.K. Smilde, Generalized contribution plots in multivariate statistical process monitoring, Chemometrics and Intelligent Laboratory Systems, 51 (2000) 95-114.
[83] T. Kourti, Application of latent variable methods to process control and multivariate statistical process control in industry, International Journal of Adaptive Control and Signal Processing, 19 (2005) 213-246.
[84] B.M. Wise, Model maintenance: the unrecognized cost in PAT and QbD, EuroPACT, Eigenvector Research, Inc., Wenatchee, WA, USA, Barcelona (Spain), 2014.
[85] W. Chew, P. Sharratt, Trends in process analytical technology, Analytical Methods, 2010, pp. 1412.

[86] T. Rajalahti, O.M. Kvalheim, Multivariate data analysis in pharmaceutics: a tutorial review, International Journal of Pharmaceutics, 417 (2011) 280-290.
[87] A.L. Pomerantsev, O.Y. Rodionova, Process analytical technology: a critical view of the chemometricians, Journal of Chemometrics, 2012, pp. 299-310.
[88] S. Garcia-Munoz, D. Settell, Application of multivariate latent variable modeling to pilot-scale spray drying monitoring and fault detection: Monitoring with fundamental knowledge, Computers & Chemical Engineering, 33 (2009) 2106-2110.
[89] A. Burggraeve, T. Van Den Kerkhof, M. Hellings, J.P. Remon, C. Vervaet, T. De Beer, Batch statistical process control of a fluid bed granulation process using in-line spatial filter velocimetry and product temperature measurements, European Journal of Pharmaceutical Sciences, 42 (2011) 584-592.
[90] J.G. Rosas, M. Blanco, J.M. González, M. Alcalá, Quality by design approach of a pharmaceutical gel manufacturing process, part 1: determination of the design space, Journal of Pharmaceutical Sciences, 100 (2011) 4432-4441.
[91] R. Kona, H. Qu, R. Mattes, B. Jancsik, R.M. Fahmy, S.W. Hoag, Application of in-line near infrared spectroscopy and multivariate batch modeling for process monitoring in fluid bed granulation, International Journal of Pharmaceutics, 452 (2013) 63-72.
[92] V. Barla, R. Kumar, V. Nalluri, R. Gandhi, K. Venkatesh, A practical evaluation of qualitative and quantitative chemometric models for real-time monitoring of moisture content in a fluidised bed dryer using near infrared technology, Journal of Near Infrared Spectroscopy, 22 (2014) 221-228.
[93] M.C. Sarraguça, M. Paisana, J. Pinto, J.A. Lopes, Real-time monitoring of cocrystallization processes by solvent evaporation: a near infrared study, European Journal of Pharmaceutical Sciences, 90 (2016) 76-84.
[94] R.D. JiJi, M.H. Hammond, F.W. Williams, S.L. Rose-Pehrsson, Multivariate statistical process control for continuous monitoring of networked early warning fire detection (EWFD) systems, Sensors and Actuators B: Chemical, 93 (2003) 107-116.
[95] A. Nijhuis, S. de Jong, B.G.M. Vandeginste, Multivariate statistical process control in chromatography, Chemometrics and Intelligent Laboratory Systems, 38 (1997) 51-62.
[96] L. Zhu, R.G. Brereton, D.R. Thompson, P.L. Hopkins, R.E.a. Escott, On-line HPLC combined with multivariate statistical process control for the monitoring of reactions, Analytica Chimica Acta, 584 (2007) 370-378.
[97] M. Mattila, K. Saloheimo, K. Koskinen, Improving the robustness of particle size analysis by multivariate statistical process control, Particle & Particle Systems Characterization, 24 (2007) 173-183.
[98] K. Laursen, S.S. Frederiksen, C. Leuenhagen, R. Bro, Chemometric quality control of chromatographic purity, Journal of Chromatography A, 1217 (2010) 6503-6510.
[99] M. Fransson, A. Sparén, B. Lagerholm, L. Karlsson, On-line process control of liquid chromatography, Analytical Chemistry, 73 (2001) 1502-1508.
[100] S.M. de Lima, B.F.A. Silva, D.V. Pontes, C.F. Pereira, L. Stragevitch, M.F. Pimentel, In-line monitoring of the transesterification reactions for biodiesel production using NIR spectroscopy, Fuel, 115 (2014) 46-53.
[101] B. Lennox, G.a. Montague, H.G. Hiden, G. Kornfeld, P.R. Goulding, Process monitoring of an industrial fed-batch fermentation., Biotechnology and Bioengineering, 74 (2001) 125-135.
[102] S. Albert, R.D. Kinley, Multivariate statistical monitoring of batch processes: an industrial case study of fermentation supervision, Trends in Biotechnology, 19 (2001) 53-62.
[103] A.P. Ferreira, J.A. Lopes, J.C. Menezes, Study of the application of multiway multivariate techniques to model data from an industrial fermentation process, Analytica Chimica Acta, 595 (2007) 120-127.
[104] J.-M. Lee, C.K. Yoo, I.-B. Lee, Enhanced process monitoring of fed-batch penicillin cultivation using time-varying and multivariate statistical analysis., Journal of Biotechnology, 110 (2004) 119-136.

[105] B. Lennox, H. Hiden, G. Montague, Application of multivariate statistical process control to batch operations, Computers & Chemical Engineering, 24 (2000) 291-296.
[106] S. Yoon, J.F. MacGregor, Fault diagnosis with multivariate statistical models part I: using steady state fault signatures, Journal of Process Control, 11 (2001) 387-400.
[107] B.M. Wise, D.J. Veltkamp, B. Davis, Principal components analysis for monitoring the west valley liquid fed ceramic melter, Waste Management, (1988) 811-818.
[108] R.A. Potyrailo, Multivariate statistical process analysis systems and methods for the production of melt polycarbonate, Google Patents, 2003.
[109] T.E. Morud, Multivariate statistical process control; example from the chemical process industry, Journal of Chemometrics, 10 (1996) 669-675.
[110] C. Wikstrom, S. Rannar, E. Johansson, M. Sandberg, Multivariate process and quality monitoring applied to an electrolysis process Part I . Process supervision with multivariate control charts, (1998) 221-231.
[111] H. Huang, H. Qu, In-line monitoring of alcohol precipitation by near-infrared spectroscopy in conjunction with multivariate batch modeling, Analytica Chimica Acta, 707 (2011) 47-56.
[112] M. Clavaud, Y. Roggo, R. Von Daeniken, A. Liebler, J.-O. Schwabe, Chemometrics and in-line near infrared spectroscopic monitoring of a biopharmaceutical chinese hamster ovary cell culture: prediction of multiple cultivation variables, Talanta, 111 (2013) 28-38.
[113] Y.-J. Liu, S. André, L. Saint Cristau, S. Lagresle, Z. Hannas, É. Calvosa, O. Devos, L. Duponchel, Multivariate statistical process control (MSPC) using Raman spectroscopy for in-line culture cell monitoring considering time-varying batches synchronized with correlation optimized warping (COW), Analytica Chimica Acta, 952 (2017) 9-17.
[114] A. Simoglou, P. Georgieva, E.B. Martin, A.J. Morris, S. Feyo de Azevedo, On-line monitoring of a sugar crystallization process, Computers & Chemical Engineering, 29 (2005) 1411-1422.
[115] R. Dunia, G. Rochelle, T.F. Edgar, M. Nixon, Multivariate monitoring of a carbon dioxide removal process, Computers & Chemical Engineering, 60 (2014) 381-395.
[116] Z.L.Z. Lin, W.H.W. Huangang, X.W.X. Wenli, W.R.W. Rui, Z.H.Z. Haifeng, PCA based statistical process monitoring of grinding process, Control and Automation (ICCA), 2010 8th IEEE International Conference on, (2010).
[117] Y. Zhang, M.S. Dudzic, Online monitoring of steel casting processes using multivariate statistical technologies: From continuous to transitional operations, Journal of Process Control, 16 (2006) 819-829.
[118] I. Miletic, S. Quinn, M. Dudzic, V. Vaculik, M. Champagne, An industrial perspective on implementing on-line applications of multivariate statistics, Journal of Process Control, 14 (2004) 821-836.
[119] Y. Zhang, M.S. Dudzic, Industrial application of multivariate SPC to continuous caster start-up operations for breakout prevention, Control Engineering Practice, 14 (2006) 1357-1375.
[120] N.A. Abd Majid, M.P. Taylor, J.J.J. Chen, B.R. Young, Multivariate statistical monitoring of the aluminium smelting process, Computers & Chemical Engineering, 35 (2011) 2457-2468.
[121] B. Boonkhao, X.Z. Wang, Ultrasonic attenuation spectroscopy for multivariate statistical process control in nanomaterial processing, Particuology, 10 (2012) 196-202.
[122] A. Skoglund, A. Brundin, C.-F. Mandenius, Monitoring a paperboard machine using multivariate statistical process control, Chemometrics and Intelligent Laboratory Systems, 73 (2004) 3-6.
[123] A. AlGhazzawi, B. Lennox, Monitoring a complex refining process using multivariate statistics, Control Engineering Practice, 16 (2008) 294-307.
[124] K. Pöllänen, A. Häkkinen, S.-P. Reinikainen, J. Rantanen, M. Karjalainen, M. Louhi-Kultanen, L. Nyström, IR spectroscopy together with multivariate data analysis as a process analytical tool for in-line monitoring of crystallization process and solid-state

analysis of crystalline product, Journal of Pharmaceutical and Biomedical Analysis, 38 (2005) 275-284.
[125] M.C. Sarraguça, P.R.S. Ribeiro, A.O.D. Santos, J.A. Lopes, Batch statistical process monitoring approach to a cocrystallization process, Journal of Pharmaceutical Sciences, 104 (2015) 4099-4108.
[126] C. Bay, H. Kong, F. Wang, Stage-based multivariate statistical analysis for injection molding, (2000).
[127] X.-T. Doan, R. Srinivasan, Online monitoring of multi-phase batch processes using phase-based multivariate statistical process control, Computers & Chemical Engineering, 32 (2008) 230-243.
[128] X. Li, Multivariate statistical process monitoring of an industrial polypropylene catalyzer reactor with component analysis and kernel density estimation, 15 (2007) 524-532.
[129] Z. Xu, H. Zhang, H. Hou, L. Ren, Y. Sun, Q. Xu, Fault diagnosis method of feeding mechanism in printing machine based on multivariate statistical process control, in: Y. Ouyang, M. Xu, L. Yang, Y. Ouyang (Eds.) Advanced Graphic Communications, Packaging Technology and Materials, Springer Singapore, Singapore, 2016, pp. 659-665.
[130] S. García-Muñoz, T. Kourti, Troubleshooting of an industrial batch process using multivariate methods, Industrial & Engineering Chemistry Research, (2003) 3592-3601.
[131] F.S.L. Costa, R.H.P. Pedroza, D.L. Porto, M.V.P. Amorim, K.M.G. Lima, Multivariate control charts for simultaneous quality monitoring of isoniazid and rifampicin in a pharmaceutical formulation using a portable near infrared spectrometer, Journal of the Brazilian Chemical Society, 26 (2015) 64-73.
[132] S. Wold, N. Kettaneh, H. Fridén, A. Holmberg, Modelling and diagnostics of batch processes and analogous kinetic experiments, Chemometrics and Intelligent Laboratory Systems, 44 (1998) 331-340.
[133] K. Laursen, M.A. Rasmussen, R. Bro, Comprehensive control charting applied to chromatography, Chemometrics and Intelligent Laboratory Systems, 107 (2011) 215-225.
[134] A.R. Tôrres, S. Grangeiro Junior, W.D. Fragoso, Multivariate control charts for monitoring captopril stability, Microchemical Journal, 118 (2015) 259-265.
[135] L.E. Mujica, J. Vehí, M. Ruiz, M. Verleysen, W. Staszewski, K. Worden, Multivariate statistics process control for dimensionality reduction in structural assessment, Mechanical Systems and Signal Processing, 22 (2008) 155-171.
[136] C.V. Palau, F. Arregui, A. Ferrer, Using multivariate principal component analysis of injected water flows to detect anomalous behaviors in a water supply system - a case study, Water Science and Technology: Water Supply, 4 (2004) 169-182.
[137] Burst detection in water networks using principal component analysis, Journal of Water Resources Planning and Management, 138 (2012).
[138] D. Aguado, A. Ferrer, J. Ferrer, A. Seco, Multivariate SPC of a sequencing batch reactor for wastewater treatment, Chemometrics and Intelligent Laboratory Systems, 85 (2007) 82-93.
[139] D. Aguado, C. Rosen, Multivariate statistical monitoring of continuous wastewater treatment plants, Engineering Applications of Artificial Intelligence, 21 (2008) 1080-1091.
[140] M.W. Lee, S.H. Hong, H. Choi, J.-H. Kim, D.S. Lee, J.M. Park, Real-time remote monitoring of small-scaled biological wastewater treatment plants by a multivariate statistical process control and neural network-based software sensors, Process Biochemistry, 43 (2008) 1107-1113.
[141] R.K. Tomita, S.W. Park, O.a.Z. Sotomayor, Analysis of activated sludge process using multivariate statistical tools—a PCA approach, Chemical Engineering Journal, 90 (2002) 283-290.
[142] G. Tavares, Z. Zsigraiova, V. Semiao, M.d.G. Carvalho, Monitoring, fault detection and operation prediction of MSW incinerators using multivariate statistical methods, Waste Management, 31 (2011) 1635-1644.

[143] T.a. Lestander, C. Holmberg, L. Stenberg, R. Lehtonen, Towards multivariate statistical process control in the wood pellet industry, Biomass and Bioenergy, 45 (2012) 152-158.
[144] A. Burggraeve, T. Van Den Kerkhof, M. Hellings, J.P. Remon, C. Vervaet, T. De Beer, Evaluation of in-line spatial filter velocimetry as PAT monitoring tool for particle growth during fluid bed granulation, European Journal of Pharmaceutics and Biopharmaceutics, 76 (2010) 138-146.
[145] T. Narvanen, Particle size determination during fluid bed granulation - challenges and opportunities, European Journal of Pharmaceutical Sciences, 34 (2008) S12-S12.
[146] I. Masic, I. Ilic, R. Dreu, S. Ibric, J. Parojcic, Z. Duric, An investigation into the effect of formulation variables and process parameters on characteristics of granules obtained by in situ fluidized hot melt granulation, International Journal of Pharmaceutics, 423 (2012) 202-212.
[147] A. Burggraeve, N. Sandler, J. Heinamaki, H. Raikkonen, J.P. Remon, C. Vervaet, T. De Beer, J. Yliruusi, Real-time image-based investigation of spheronization and drying phenomena using different pellet formulations, European Journal of Pharmaceutical Sciences, 44 (2011) 635-642.
[148] A. Ali, N. Denis, A.R. Jose, Investigation of on-line optical particle characterization in reaction and cooling crystallization systems. Current state of the art, Measurement Science & Technology, 13 (2002) 349.
[149] Z.Q. Yu, R.B.H. Tan, P.S. Chow, Effects of operating conditions on agglomeration and habit of paracetamol crystals in anti-solvent crystallization, Journal of Crystal Growth, 279 (2005) 477-488.
[150] R.D. Braatz, Advanced control of crystallization processes, Annual Reviews in Control, 26 (2002) 87-99.
[151] A. Rawle, Basic principles of particle size analysis Malvern Isntruments Ltd Technical Paper, Worcestershire, UK.
[152] Z. Ma, H.G. Merkus, J.G.A.E. de Smet, C. Heffels, B. Scarlett, New developments in particle characterization by laser diffraction: size and shape, Powder technology, 111 (2000) 66-78.
[153] R.M. Jones, Particle size analysis by laser diffraction: ISO 13320, standard operating procedures, and Mie theory, American Laboratory, 35 (2003) 44-47.
[154] Mie Theory: The first 100 years, Inform, Malvern, 2010.
[155] Particle-size distribution estimation by analytical sieving, European Pharmacopoeia, 6 (2009).
[156] H. Heywood, Evaluation of powders, Journal of Pharmacy and Pharmacology, 15 (1963) T56-&.
[157] M.Z. Li, D. Wilkinson, K. Patchigolla, Determination of non-spherical particle size distribution from chord length measurements. Part 2: Experimental validation, Chemical Engineering Science, 60 (2005) 4992-5003.
[158] D. Petrak, Simultaneous measurement of particle size and particle velocity by the spatial filtering technique, Particle & Particle Systems Characterization, 19 (2002) 391-400.
[159] D. Petrak, H. Rauh, Optical probe for the in-line determination of particle shape, size, and velocity, Particulate Science and Technology, 24 (2006) 381-394.
[160] S. Schmidt-Lehr, H.U. Moritz, K.C. Jürgens, Online-control of the particle size during fluid-bed granulation/evaluation of a novel laser probe for a better control of particle size in fluid-bed granulation, Pharmazeutische Industrie, 69 (2007) 478-484.
[161] M. Fonteyne, J. Vercruysse, D.C. Diaz, D. Gildemyn, C. Vervaet, J.P. Remon, T. De Beer, Real-time assessment of critical quality attributes of a continuous granulation process, Pharmaceutical Development and Technology, 18 (2013) 85-97.
[162] D. Petrak, S. Dietrich, G. Eckardt, M. Köhler, In-line particle sizing for real-time process control by fibre-optical spatial filtering technique (SFT), Advanced Powder Technology, 22 (2011) 203-208.

[163] A.R. Heath, P.D. Fawell, P.A. Bahri, J.D. Swift, Estimating average particle size by focused beam reflectance measurement (FBRM), Particle & Particle Systems Characterization, 19 (2002) 84-95.
[164] P. Barrett, B. Glennon, characterizing the metastable zone width and solubility curve using Lasentec FBRM and PVM, Chemical Engineering Research and Design, 80 (2002) 799-805.
[165] E. Kougoulos, A.G. Jones, K.H. Jennings, M.W. Wood-Kaczmar, Use of focused beam reflectance measurement (FBRM) and process video imaging (PVI) in a modified mixed suspension mixed product removal (MSMPR) cooling crystallizer, Journal of Crystal Growth, 273 (2005) 529-534.
[166] F. Sistare, L. St. Pierre Berry, C.A. Mojica, Process analytical technology: an investment in process knowledge, Organic Process Research & Development, 9 (2005) 332-336.
[167] A. Tok, X. Goh, W. Ng, R. Tan, Monitoring Granulation Rate Processes Using Three PAT Tools in a Pilot-Scale Fluidized Bed, AAPS PharmSciTech, 9 (2008) 1083-1091.
[168] B. O'Sullivan, P. Barrett, G. Hsiao, A. Carr, B. Glennon, In situ monitoring of polymorphic transitions, Organic Process Research & Development, 7 (2003) 977-982.
[169] E. Kougoulos, A.G. Jones, M.W. Wood-Kaczmar, Modelling particle disruption of an organic fine chemical compound using Lasentec focussed beam reflectance monitoring (FBRM) in agitated suspensions, Powder technology, 155 (2005) 153-158.
[170] Y. Kim, J.R. Méndez del Río, R.W. Rousseau, Solubility and prediction of the heat of solution of sodium naproxen in aqueous solutions, Journal of Pharmaceutical Sciences, 94 (2005) 1941-1948.
[171] D. Greaves, J. Boxall, J. Mulligan, A. Montesi, J. Creek, E.D. Sloan, C.A. Koh, Measuring the particle size of a known distribution using the focused beam reflectance measurement technique, Chemical Engineering Science, 63 (2008) 5410-5419.
[172] A. Ruf, J. Worlitschek, M. Mazzotti, Modeling and experimental analysis of PSD measurements through FBRM, Particle & Particle Systems Characterization, 17 (2000) 167-179.
[173] P.A. Langston, Comparison of least-squares method and Bayes' theorem for deconvolution of mixture composition, Chemical Engineering Science, 57 (2002) 2371-2379.
[174] M.J.H. Simmons, P.A. Langston, A.S. Burbidge, Particle and droplet size analysis from chord distributions, Powder technology, 102 (1999) 75-83.
[175] E.J.W. Wynn, Relationship between particle-size and chord-length distributions in focused beam reflectance measurement: stability of direct inversion and weighting, Powder technology, 133 (2003) 125-133.
[176] P.A. Langston, A.S. Burbidge, T.F. Jones, M.J.H. Simmons, Particle and droplet size analysis from chord measurements using Bayes' theorem, Powder technology, 116 (2001) 33-42.
[177] M.Z. Li, D. Wilkinson, Determination of non-spherical particle size distribution from chord length measurements. Part 1: Theoretical analysis, Chemical Engineering Science, 60 (2005) 3251-3265.
[178] W. Liu, N.N. Clark, Relationships between distributions of chord lengths and distributions of bubble sizes including their statistical parameters, International Journal of Multiphase Flow, 21 (1995) 1073-1089.
[179] W.D. Liu, N.N. Clark, A.I. Karamavruc, Relationship between bubble size distributions and chord-length distribution in heterogeneously bubbling systems, Chemical Engineering Science, 53 (1998) 1267-1276.
[180] A. Tadayyon, S. Rohani, Determination of particle size distribution by Par-Tec (R) 100: modeling and experimental results, Particle & Particle Systems Characterization, 15 (1998) 127-135.

[181] P.A. Langston, T.F. Jones, Non-spherical 2-dimensional particle size analysis from chord measurements using Bayes' theorem, Particle & Particle Systems Characterization, 18 (2001) 12-21.
[182] H.H.J. Bloemen, M.G.M. De Kroon, Transformation of chord length distributions into particle size distributions using least squares techniques, Particulate Science and Technology, 23 (2005) 377-386.
[183] N. Sandler, Photometric imaging in particle size measurement and surface visualization, International Journal of Pharmaceutics, 417 (2011) 227-234.
[184] I. Soppela, S. Airaksinen, J. Hatara, H. Raikkonen, O. Antikainen, J. Yliruusi, N. Sandler, Rapid particle size measurement using 3D surface imaging, AAPS PharmSciTech, 12 (2011) 476-484.
[185] B. O'Sullivan, Introduction to FBRM technology, Mettler Toledo.
[186] United States Pharmacopeia and National Formulary (USP 28-NF), United States Pharmacopeial Convention, Rockville, MD, USA, 2007.
[187] M. Fonteyne, S. Soares, Y. Vercruysse, E. Peeters, A. Burggraeve, C. Vervaet, J.P. Remon, N. Sandler, T. De Beer, Prediction of quality attributes of continuously produced granules using complementary PAT tools, European Journal of Pharmaceutics and Biopharmaceutics, 2012.
[188] G. Crawley, Particle sizing online, Powder Metallurgy, 44 (2001) 304-306.
[189] G. Crawley, A. Malcolmson, Online particle sizing as a route to process optimization, Chemical Engineering (New York), 111 (2004) 37-41.
[190] M.-N. Pons, K. Milferstedt, E. Morgenroth, Modeling of chord length distributions, Chemical Engineering Science, 61 (2006) 3962-3973.
[191] E.J. Hukkanen, R.D. Braatz, Measurement of particle size distribution in suspension polymerization using in situ laser backscattering, Sensors and Actuators B: Chemical, 96 (2003) 451-459.
[192] J. Van Impe, G. Gins, An extensive reference dataset for fault detection and identification in batch processes, Chemometrics and Intelligent Laboratory Systems, 148 (2015) 20-31.
[193] T. Kourti, The process analytical technology initiative and multivariate process analysis, monitoring and control, Analytical and Bioanalytical Chemistry, 384 (2006) 1043-1048.
[194] R. Bro, PARAFAC. Tutorial and applications, Chemometrics and Intelligent Laboratory Systems, 38 (1997) 149-171.
[195] J.A. Lopes, J.C. Menezes, Industrial fermentation end-product modelling with multilinear PLS, Chemometrics and Intelligent Laboratory Systems, 68 (2003) 75-81.
[196] M.C. Sarraguça, T. De Beer, C. Vervaet, J.P. Remon, J.A. Lopes, A batch modelling approach to monitor a freeze-drying process using in-line Raman spectroscopy, Talanta, 83 (2010) 130-138.
[197] S. Wold, N. Kettaneh, H. Friden, A. Holmberg, Modelling and diagnostics of batch processes and analogous kinetic experiments, Chemometrics and Intelligent Laboratory Systems, 44 (1998) 331-340.
[198] J.M. Gonzalez-Martinez, A. Ferrer, J.A. Westerhuis, Real-time synchronization of batch trajectories for on-line multivariate statistical process control using Dynamic Time Warping, Chemometrics and Intelligent Laboratory Systems, 105 (2011) 195-206.
[199] G. Tomasi, F. van den Berg, C. Andersson, Correlation optimized warping and dynamic time warping as preprocessing methods for chromatographic data, Journal of Chemometrics, 18 (2004) 231-241.
[200] S. Wold, H. Antti, F. Lindgren, J. Ohman, Orthogonal signal correction of near-infrared spectra, Chemometrics and Intelligent Laboratory Systems, 44 (1998) 175-185.
[201] H. Stenlund, E. Johansson, J. Gottfries, J. Trygg, Unlocking interpretation in near infrared multivariate calibrations by orthogonal partial least squares, Analytical Chemistry, 81 (2009) 203-209.

[202] J. Boccard, D.N. Rutledge, A consensus orthogonal partial least squares discriminant analysis (OPLS-DA) strategy for multiblock Omics data fusion, Analytica Chimica Acta, 769 (2013) 30-39.
[203] M. Rantalainen, M. Bylesjo, O. Cloarec, J.K. Nicholson, E. Holmes, J. Trygg, Kernel-based orthogonal projections to latent structures (K-OPLS), Journal of Chemometrics, 21 (2007) 376-385.
[204] S. Wiklund, E. Johansson, L. Sjostrom, E.J. Mellerowicz, U. Edlund, J.P. Shockcor, J. Gottfries, T. Moritz, J. Trygg, Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models, Analytical Chemistry, 80 (2008) 115-122.
[205] M.C. Sarraguca, P.R.S. Ribeiro, A.O. Santos, M.C.D. Silva, J.A. Lopes, A PAT approach for the on-line monitoring of pharmaceutical co-crystals formation with near infrared spectroscopy, International Journal of Pharmaceutics, 471 (2014) 478-484.
[206] P. Sanphui, L. Rajput, Tuning solubility and stability of hydrochlorothiazide co-crystals, Acta Crystallographica Section B-Structural Science Crystal Engineering and Materials, 70 (2014) 81-90.
[207] S. Patel, A.M. Kaushal, A.K. Bansal, Compression physics in the formulation development of tablets, 23 (2006) 1-66.
[208] M.C. Gohel, P.D. Jogani, A review of co-processed directly compressible excipients, Journal of Pharmaceutical Sciences, 8 (2005) 76-93.
[209] C. Appelgren, Recent advances in granulation technology and equipment, Drug Development and Industrial Pharmacy, 11 (1985) 725-741.
[210] R. Perry, D. Green, J. Maloney, Perry's Chemical Engineers' Handbook, McGraw-Hill, New York, 1997.
[211] D.M. Parikh, Handbook of pharmaceutical granulation technology, CRC Press2016.
[212] S.M. Iveson, J.D. Litster, K. Hapgood, B.J. Ennis, Nucleation, growth and breakage phenomena in agitated wet granulation processes: a review, Powder technology, 117 (2001) 3-39.
[213] J.G. Osorio, A.U. Vanarase, R.J. Romañach, F.J. Muzzio, Continuous powder mixing, Pharmaceutical Blending and Mixing, (2015) 101-127.
[214] C. Vervaet, J. Remon, Continuous Granulation, Handbook of Pharmaceutical Granulation Technology, Third Edition, CRC Press, 2009, pp. 308-322.
[215] J. Kowalski, J.P. Lakshman, A.T. Serajuddin, W.-Q. Tong, Heated roller compaction process for making pharmaceutical compositions, Google Patents, 2007.
[216] S.S. Behzadi, S. Toegel, H. Viernstein, Innovations in coating technology, Recent patents on drug delivery & formulation, 2 (2008) 209-230.
[217] H. Leuenberger, New trends in the production of pharmaceutical granules: batch versus continuous processing, European Journal of Pharmaceutics and Biopharmaceutics, 52 (2001) 289-296.
[218] L. Chablani, M.K. Taylor, A. Mehrotra, P. Rameas, W.C. Stagner, Inline real-time near-infrared granule moisture measurements of a continuous granulation–drying–milling process, AAPS PharmSciTech, 12 (2011) 1050-1055.
[219] S. Wold, K. Esbensen, P. Geladi, Principal component analysis, Chemometrics and Intelligent Laboratory Systems, 2 (1987) 37-52.
[220] J.F. MacGregor, Using on-line process data to improve quality: challenges for statisticians, International Statistical Review, 65 (1997) 309-323.
[221] T. Kourti, Abnormal situation detection, three-way data and projection methods; robust data archiving and modeling for industrial applications, Annual Reviews in Control, 27 (2003) 131-139.
[222] I.T. Jolliffe, Principal component analysis, Second ed., Springer-Verlag, New York, 2002.
[223] J.F. Macgregor, C. Jaeckle, C. Kiparissides, M. Koutoudi, Process monitoring and diagnosis by multiblock PLS methods, AIChE Journal, 40 (1994) 826-838.

[224] T. Kourti, J.F. MacGregor, Multivariate SPC methods for process and product monitoring, Journal of Quality Technology, 28 (1996) 409-428.
[225] A.K. Conlin, E.B. Martin, A.J. Morris, Confidence limits for contribution plots, Journal of Chemometrics, 14 (2000) 725-736.
[226] H. Babamoradi, F. van den Berg, A. Rinnan, Confidence limits for contribution plots in multivariate statistical process control using bootstrap estimates, Analytica Chimica Acta, 908 (2016) 75-84.
[227] J. Vercruysse, U. Delaet, I. Van Assche, P. Cappuyns, F. Arata, G. Caporicci, T. De Beer, J.P. Remon, C. Vervaet, Stability and repeatability of a continuous twin screw granulation and drying system, European Journal of Pharmaceutics and Biopharmaceutics, 85 (2013) 1031-1038.
[228] E.M. Agency, Reflection Paper: Chemical, pharmaceutical and biological information to be included in dossiers when process analytical technology (PAT) is employed, (2006).
[229] W. Chew, P. Sharratt, Trends in process analytical technology, Analytical Methods, 2 (2010) 1412-1438.
[230] M. Krumme, Necessary change – perfecting process-analytical technology, WorldPharma, 2015.
[231] B. Scott, A. Wilcock, Process analytical technology in the pharmaceutical industry: a toolkit for continuous improvement, PDA journal of pharmaceutical science and technology, 60 (2006) 17-53.
[232] L.L. Simon, H. Pataki, G. Marosi, F. Meemken, K. Hungerbühler, A. Baiker, S. Tummala, B. Glennon, M. Kuentz, G. Steele, H.J.M. Kramer, J.W. Rydzak, Z. Chen, J. Morris, F. Kjell, R. Singh, R. Gani, K.V. Gernaey, M. Louhi-Kultanen, J. O'Reilly, N. Sandler, O. Antikainen, J. Yliruusi, P. Frohberg, J. Ulrich, R.D. Braatz, T. Leyssens, M. von Stosch, R. Oliveira, R.B.H. Tan, H. Wu, M. Khan, D. O'Grady, A. Pandey, R. Westra, E. Delle-Case, D. Pape, D. Angelosante, Y. Maret, O. Steiger, M. Lenner, K. Abbou-Oucherif, Z.K. Nagy, J.D. Litster, V.K. Kamaraju, M.-S. Chiu, Assessment of recent process analytical technology (PAT) trends: a multiauthor review, Organic Process Research & Development, 19 (2015) 3-62.
[233] R. Schneider, (2006), Achieving process understanding: the foundation of a strategic pat program; Retrieved 18/01/2017, from http://www.pharmamanufacturing.com/articles/2006/109/
[234] K. Gadamasetti, T. Braish, C. Mojica, F. Sistare, L. St. Pierre-Berry, Process analytical technology in the manufacture of bulk active pharmaceuticals—promise, practice, and challenges, Process Chemistry in the Pharmaceutical Industry, Volume 2: Challenges in an Ever Changing Climate, CRC Press2007, pp. 361-381.
[235] M. de Matas, T. De Beer, S. Folestad, J. Ketolainen, H. Lindén, J.A. Lopes, W. Oostra, M. Weimer, P. Öhrngren, J. Rantanen, Strategic framework for education and training in Quality by Design (QbD) and process analytical technology (PAT), European Journal of Pharmaceutical Sciences, 90 (2016) 2-7.