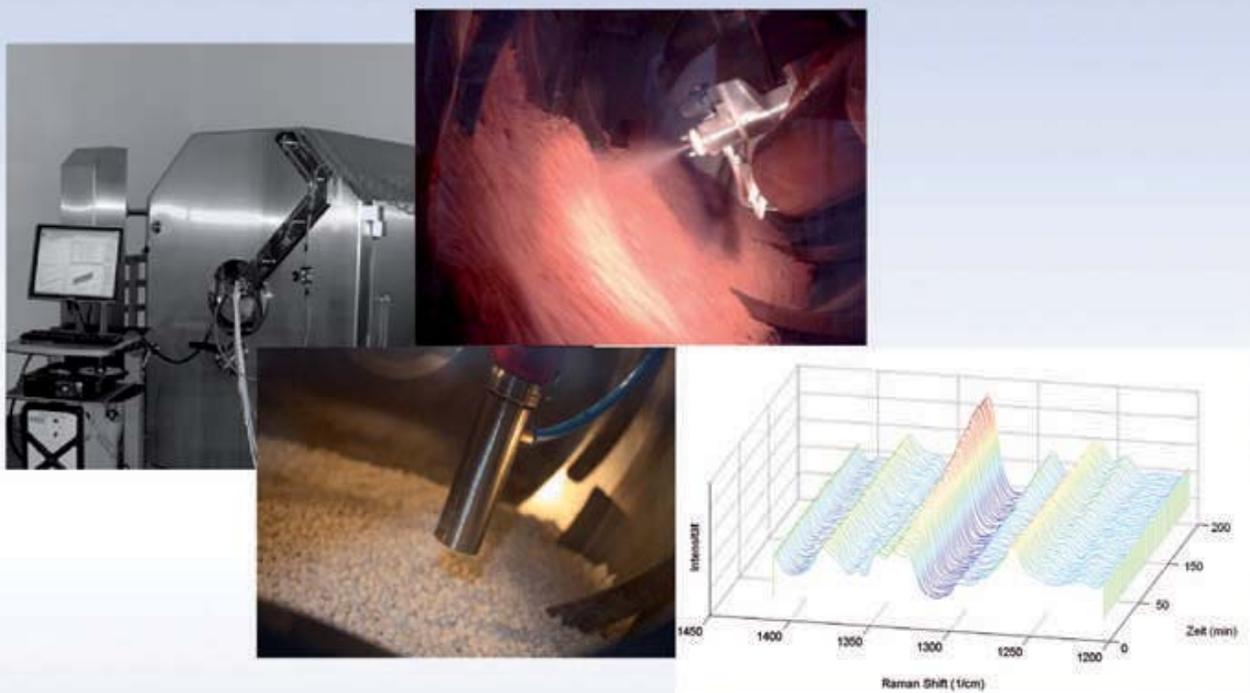


---

# Ramanspektroskopie als PAT-Methode beim Coating von Tabletten im Trommelcoater

---







# **Ramanspektroskopie als PAT-Methode beim Coating von Tabletten im Trommelcoater**

Inaugural-Dissertation

zur Erlangung des Doktorgrades  
der Mathematisch-Naturwissenschaftlichen Fakultät  
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

**Joshua Müller**

aus Darmstadt

Düsseldorf 2010

## **Bibliografische Information der Deutschen Nationalbibliothek**

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

1. Aufl. - Göttingen : Cuvillier, 2010

Zugl.: Düsseldorf, Univ. Diss., 2010

978-3-86955-406-8

aus dem Institut für Pharmazeutische Technologie und Biopharmazie  
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der  
Mathematisch-Naturwissenschaftlichen Fakultät der  
Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. Peter Kleinebudde

Koreferent: Prof. Dr. Jörg Breitzkreutz

Tag der mündlichen Prüfung: 01.07.2010

© CUVILLIER VERLAG, Göttingen 2010

Nonnenstieg 8, 37075 Göttingen

Telefon: 0551-54724-0

Telefax: 0551-54724-21

[www.cuvillier.de](http://www.cuvillier.de)

Alle Rechte vorbehalten. Ohne ausdrückliche Genehmigung des Verlages ist es nicht gestattet, das Buch oder Teile daraus auf fotomechanischem Weg (Fotokopie, Mikrokopie) zu vervielfältigen.

1. Auflage, 2010

Gedruckt auf säurefreiem Papier

978-3-86955-406-8

# Inhalt

Inhalt.....	III
1 Einleitung .....	1
1.1 Prozessanalytische Technologie in der pharmazeutischen Industrie .....	1
1.2 Instrumente der Prozessanalytischen Technologie.....	3
1.2.1 Einleitung .....	3
1.2.2 Risikoanalyse.....	3
1.2.3 Statistische Versuchsplanung .....	3
1.2.4 Prozessführung .....	3
1.2.5 In-/Online Messtechnik .....	4
1.2.6 Multivariate Datenanalyse.....	4
1.2.7 Statistische Prozesskontrolle .....	4
1.3 Multivariate Datenanalyse.....	5
1.3.1 Einleitung .....	5
1.3.2 Hauptkomponentenanalyse (PCA).....	6
1.3.3 Projektion auf latente Strukturen (PLS) .....	7
1.4 Ramanspektroskopie .....	9
1.4.1 Einleitung .....	9
1.4.2 Raman-Effekt .....	9
1.4.3 Anwendung der Ramanspektroskopie in der Pharmazie.....	12
1.5 Filmcoating.....	13
1.5.1 Einleitung .....	13
1.5.2 Prozessanalytische Technologie beim Coating .....	13
1.5.2.1 Einleitung .....	13
1.5.2.2 Nahinfrarotspektroskopie .....	14
1.5.2.3 Ramanspektroskopie .....	15
1.5.2.4 Terahertzspektroskopie .....	16
2 Zielsetzung der Arbeit .....	18
3 Ergebnisse und Diskussion.....	20
3.1 Methodvalidierung nach ICH – Guideline Q2 – Validation of Analytical Procedures .....	20
3.1.1 Einleitung und Zielsetzung.....	20
3.1.2 Referenzmethode UV-Spektroskopie.....	21
3.1.3 PAT-Methode Ramanspektroskopie und Modellerstellung .....	21
3.1.4 Validierungselemente und deren Überprüfung .....	26
3.1.4.1 Einleitung .....	26
3.1.4.2 Spezifität.....	27
3.1.4.3 Richtigkeit .....	31
3.1.4.4 Linearität .....	34
3.1.4.5 Nachweis- und Bestimmungsgrenze .....	36
3.1.4.6 Präzision .....	37
3.1.4.6.1 Einleitung .....	37

3.1.4.6.2	Wiederholungspräzision.....	38
3.1.4.6.3	Vergleichspräzision.....	39
3.1.4.7	Varianzhomogenität.....	42
3.1.4.8	Arbeitsbereich.....	42
3.1.5	Übertragung der entwickelten Methode auf die Anwendung für inline Messungen.....	43
3.1.6	Zusammenfassung.....	44
3.2	Modellentwicklung zur inline Verfolgung beim Wirkstoffcoating.....	46
3.2.1	Einleitung und Zielsetzung.....	46
3.2.2	Aufbau der Versuche.....	46
3.2.3	Modellerstellung.....	46
3.2.4	Placebotabletten.....	47
3.2.4.1	Einleitung.....	47
3.2.4.2	Referenzmethode UV-Spektroskopie.....	47
3.2.4.3	Referenzmethode Gravimetrie.....	50
3.2.4.4	Vergleich der Referenzmethoden UV-Spektroskopie und Gravimetrie.....	52
3.2.5	Diprophyllintabletten.....	54
3.2.5.1	Einleitung.....	54
3.2.5.2	Hohe Wirkstoffbeladung im Film.....	54
3.2.5.3	Niedrige Wirkstoffbeladung im Film.....	57
3.2.5.4	Wirkstofffreier Film.....	61
3.2.5.5	Zusammenfassung.....	62
3.3	Anwendung des Modells beim Scale up vom Laborcoater BFC 5 auf den BFC 50 Tripan.....	63
3.3.1	Einleitung und Zielsetzung.....	63
3.3.2	Anwendung des im Laborcoater BFC 5 erstellten Modells.....	63
3.3.3	Neue Modellerstellung.....	65
3.3.4	Zusammenfassung.....	66
3.4	Vergleich verschiedener Datenvorbehandlungen und unterschiedlicher chemometrischer Verfahren zur Modellbildung.....	67
3.4.1	Einleitung und Zielsetzung.....	67
3.4.2	Datenquelle.....	67
3.4.3	Auswahl des geeigneten spektralen Bereichs und Vergleich der Datenvorbehandlungen bei der Modellerstellung mit Hilfe der PLS.....	68
3.4.3.1	Einleitung.....	68
3.4.3.2	Partial Least Squares (PLS).....	68
3.4.3.2.1	Einleitung.....	68
3.4.3.2.2	Wellenzahlbereich 1200-1400 cm <sup>-1</sup> .....	69
3.4.3.2.3	Wellenzahlbereich 500-600 cm <sup>-1</sup> .....	70
3.4.3.2.4	Wellenzahlbereich 1550-1720 cm <sup>-1</sup> .....	71
3.4.3.2.5	Wellenzahlbereich 350-1800 cm <sup>-1</sup> .....	72
3.4.3.3	Multivariate Curve Resolution (MCR).....	73
3.4.3.4	Univariate Regression.....	74
3.4.3.5	Untersuchung der Vergleichspräzision von der univariaten Regression im Vergleich zu der PLS und MCR.....	75
3.4.4	Zusammenfassung.....	78
3.5	Robustheit des entwickelten Modells.....	80
3.5.1	Einleitung und Zielsetzung.....	80
3.5.2	Trommeldrehgeschwindigkeit.....	80
3.5.3	Sondenabstand.....	83
3.5.4	Zusammenfassung.....	87

3.6	Anwendung der Ramanspektroskopie zur Bestimmung der Coating Uniformity beim Wirkstoffcoating.....	88
3.6.1	Einleitung und Zielsetzung.....	88
3.6.2	Vorversuch mittels Drehscheibe .....	88
3.6.2.1	Versuchsdurchführung .....	88
3.6.2.2	Bestimmung der aufgetragenen Wirkstoffmenge mittels Ramanspektroskopie .....	90
3.6.2.3	Vergleich der anhand der Messungen erhaltenen Ergebnisse bei unterschiedlichen Anordnungen in Abhängigkeit von der Kreisgeschwindigkeit und der Messzeit.....	95
3.6.2.4	Mit Programm R berechnete theoretisch vorhergesagte Wirkstoffmenge ...	97
3.6.2.5	Vergleich der anhand Programm R berechneten Ergebnisse bei unterschiedlichen Anordnungen in Abhängigkeit von der Kreisgeschwindigkeit und der Messzeit.....	99
3.6.2.6	Zusammenfassung .....	100
3.7	Modellentwicklung zur inline Verfolgung beim Coating mit einem funktionellen Polymer .....	102
3.7.1	Einleitung und Zielsetzung.....	102
3.7.2	Modellerstellung.....	102
3.7.3	Referenzmethode Freisetzung .....	104
3.7.4	Referenzmethode Terahertzspektroskopie .....	107
3.7.5	Vergleich der Referenzmethoden Terahertzspektroskopie und Freisetzung..	112
3.7.6	Zusammenfassung .....	114
4	Zusammenfassung der Arbeit.....	115
5	Summary .....	117
6	Experimenteller Teil.....	119
6.1	Methodenvalidierung.....	119
6.1.1	Zusammensetzung der Coatingrezepturen und deren Herstellung.....	119
6.1.2	Eigenschaften der verwendeten Tabletten.....	119
6.1.3	Prozessparameter beim Coaten .....	119
6.1.4	Referenzmethode UV-Spektroskopie.....	119
6.1.5	PAT-Methode Ramanspektroskopie .....	120
6.1.5.1	P <sup>h</sup> AT-System .....	120
6.1.5.2	Modellerstellung.....	122
6.1.5.3	Durchführungsqualifizierung nach USP 33 <1120> .....	123
6.1.6	Validierungselemente und deren Überprüfung .....	124
6.1.6.1	Einleitung .....	124
6.1.6.2	Spezifität.....	124
6.1.6.3	Richtigkeit .....	125
6.1.6.4	Linearität .....	126
6.1.6.5	Nachweis- und Bestimmungsgrenze .....	128
6.1.6.6	Präzision .....	129
6.1.6.6.1	Einleitung .....	129
6.1.6.6.2	Wiederholungspräzision.....	130
6.1.6.6.3	Vergleichspräzision.....	130
6.1.6.7	Varianzhomogenität .....	131
6.1.6.8	Arbeitsbereich .....	132
6.1.7	Übertragung der entwickelten Methode auf die Anwendung für inline Messungen.....	132
6.2	Modellentwicklung zur inline Verfolgung beim Wirkstoffcoating.....	133
6.2.1	Zusammensetzung der Filmrezepturen und deren Herstellung.....	133

6.2.2	Eigenschaften der verwendeten Tabletten.....	134
6.2.3	Prozessparameter beim Coaten .....	135
6.2.4	Modellerstellung.....	135
6.2.4.1	Probenziehung .....	135
6.2.4.2	PAT–Methode Ramanspektroskopie.....	137
6.2.4.3	Referenzmethode UV-Spektroskopie.....	137
6.2.4.4	Referenzmethode Gravimetrie .....	138
6.2.5	Inline Messung .....	138
6.3	Anwendung des Modells beim Scale up vom Laborcoater BFC 5 auf den BFC 50	
Tripan	.....	139
6.3.1	Eigenschaften der verwendeten Tabletten.....	139
6.3.2	Zusammensetzung der Filmrezepturen und deren Herstellung.....	139
6.3.3	Prozessparameter beim Coaten .....	139
6.3.4	Modellerstellung.....	139
6.3.4.1	Probenziehung .....	139
6.3.4.2	PAT–Methode Ramanspektroskopie.....	140
6.3.4.3	Referenzmethode UV-Spektroskopie.....	140
6.3.5	Inline Messung .....	141
6.4	Vergleich verschiedener Datenvorbehandlungen und unterschiedlicher chemometrischer Verfahren zur Modellbildung .....	142
6.4.1	Datenvorbehandlungen.....	142
6.4.1.1	Allgemein .....	142
6.4.1.2	Multiplikative Streukorrektur (MSC).....	143
6.4.1.3	Standard Normal Variate (SNV) Transformation .....	143
6.4.1.4	1. und 2. Ableitung .....	143
6.4.2	Chemometrische Verfahren.....	144
6.4.2.1	Partial Least Squares Regression (PLS).....	144
6.4.2.2	Multivariate Curve Resolution (MCR).....	146
6.4.2.3	Univariate Regression .....	146
6.5	Robustheit des entwickelten Modells.....	148
6.5.1	Trommeldrehgeschwindigkeit.....	148
6.5.2	Sondenabstand.....	148
6.6	Anwendung der Ramanspektroskopie zur Bestimmung der Coating Uniformity beim Wirkstoffcoating.....	150
6.6.1	Herstellung der zu mischenden Chargen.....	150
6.6.1.1	Zusammensetzung der Filmrezeptur und deren Herstellung.....	150
6.6.1.2	Eigenschaften der verwendeten Tabletten.....	150
6.6.1.3	Prozessparameter beim Coaten .....	150
6.6.2	Versuchsaufbau und Durchführung.....	150
6.6.3	Mit dem Programm R berechnete theoretisch vorhergesagte Wirkstoffmenge ... .....	154
6.6.3.1	Einleitung .....	154
6.6.3.2	Programm R .....	154
6.6.3.3	Statistische Auswertung .....	154
6.7	Modellentwicklung zur inline Verfolgung beim Coating mit einem funktionellem Polymer .....	156
6.7.1	Zusammensetzung der Filmrezepturen und deren Herstellung.....	156
6.7.2	Eigenschaften der verwendeten Tabletten.....	157
6.7.3	Prozessparameter beim Coaten .....	157
6.7.4	Modellerstellung.....	158
6.7.4.1	Probenziehung .....	158

6.7.4.2	PAT – Methode Ramanspektroskopie.....	159
6.7.4.3	Referenzmethode Freisetzung .....	159
6.7.4.4	Referenzmethode Terahertz-Spektroskopie .....	160
6.7.4.4.1	Einleitung .....	160
6.7.4.4.2	Terahertz Pulsed Imaging.....	160
6.7.4.4.3	Elektrisches Feld des Terahertzsignals im Zeitbereich .....	162
6.7.5	Inline Messung .....	163
7	Anhang .....	164
8	Bezugsquellennachweis.....	175
9	Literaturverzeichnis.....	176

# Abkürzungsverzeichnis

a	additiver systematischer Fehler (Achsenabschnitt)
$\Delta\alpha$	Polarisationsänderung
$A_{\text{Tabl.}}$	Tablettenoberfläche
b	multiplikativer systematischer Fehler (Steigung)
BIAS	systematischer Fehler
CCD	Charge Coupled Devices
CTR	Centering
CU	Coating Uniformity
DIN	Deutsches Institut für Normung
DOE	Design of Experiments
$d_p$	Tablettendurchmesser
$DS^2$	Differenz der Abweichungsquadratsummen
E	Residuenmatrix
FDA	Food and Drug Administration
FMEA	Failure Mode and Effects Analysis
GMP	Good Manufacturing Practices
$h_k$	Kalottenhöhe
$I_0$	Intensität Anregungslaser
ICH	International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use
k	relative Ergebnisunsicherheit
KPIs	Key Performance Indicators
m	Wiederholungsmessung
MCR	Multivariate Curve Resolution
MDT	Mean Dissolution Time
MSC	Multiplicative Scatter Correction
MVDA	Multivariate Datenanalyse / Multivariate Data Analysis
n	Probenanzahl
N	Anzahl streuende Moleküle
v	Kreisgeschwindigkeit
p	Loadings der X-Daten
PAT	Prozessanalytische Technologie / Process Analytical Technology
PCA	Principle Component Analysis
PLS	Partial Least Squares
$P^T$	Hauptkomponentenmatrix
PW	Prüfwert
q	Loadings der Y-Daten
$Q^2$	Vorhersagegenauigkeit
$R^2$	Bestimmtheitsmaß
RMSEC	Root Mean Square Error of Calibration
RMSECV	Root Mean Square Error of Crossvalidation
RMSEP	Root Mean Square Error of Prediction
RSS	Residual Sum of Squares
RTR	Real Time Release
s	Steghöhe
Sa/Sb und Sc	Messsignal Lösung A/B und C
$S_L$	Standardabweichung der Messwerte der Leerprobe
SEC	Standard Error of Calibration
SEP	Standard Error of Prediction
SIMCA	Soft Independent Modeling of Class Analogy
SNV	Standard Normal Variate Transformation
SPC	statistische Prozesskontrolle / Statistical Process Control
$S_{xx}$	Quadratsumme von x
$S_y$	Reststandardabweichung
t	Scores der X-Daten
T	Scoresmatrix

# 1 Einleitung

## 1.1 Prozessanalytische Technologie in der pharmazeutischen Industrie

Die Good Manufacturing Practices (GMP) beschreiben bewährte Vorgehensweisen in der Produktion und sind als empirische Standards in der pharmazeutischen Industrie etabliert. Dagegen wird in anderen Industriebereichen schon lange die Prozessanalytische Technologie (PAT) eingesetzt, wobei sie sich vor allem in der Chemie-, Erdöl-, Papier- [1,2] und in der Lebensmittelindustrie [3,4] etabliert hat. Die US-amerikanische Gesundheitsbehörde (Food and Drug Administration, FDA) erkannte die Notwendigkeit, die Effizienz in der Pharmaproduktion durch Einsatz von in anderen Industrien bewährten Technologien zu erhöhen und zugleich weiterhin die Qualität der produzierten Pharmazeutika sicherzustellen. Daraufhin veröffentlichte die FDA im Jahre 2004 die Guidance for Industry PAT-A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance [5]. Dadurch soll die Bereitschaft der Industrie gefördert werden, neue und innovative Methoden der Kontrolle und Überwachung in der pharmazeutischen Entwicklung und in der Qualitätssicherung mit einzubringen und zur Sicherung der Qualität nutzbringend anzuwenden. Die international anerkannte offizielle Definition des Begriffs „Prozessanalytische Technologie (PAT)“ lautet [5]:

"PAT is considered to be a system for designing, analyzing and controlling manufacturing through timely measurements (i.e., during processing) of critical quality and performance attributes of raw and in-process materials and processes, with the goal of ensuring final product quality. It is important to note that the term *analytical* in PAT is viewed broadly to include chemical, physical, microbiological, mathematical, and risk analysis conducted in an integrated manner. The goal of PAT is to enhance understanding and control the manufacturing process: *Quality cannot be tested into products; it should be built-in or should be by design.*"

PAT beinhaltet somit eine Ansammlung von Methoden, um Prozesse zu analysieren, zu verstehen und zu regeln, mit dem Ziel, die Qualität eines Arzneimittels zu gewährleisten bzw. zu verbessern. Ziel des Einsatzes von PAT ist letztendlich die Real-Time-Release (RTR), also die Chargenfreigabe im Sinne einer parametrischen Freigabe auf Basis der im laufenden Prozess erhobenen Daten, anstelle der separaten analytischen Freigabeprüfung am Fertigprodukt [5].

Ein Überblick über die Anwendung von PAT in der pharmazeutischen Industrie ist in der Literatur [6-8] aufgeführt. Hier kommt PAT beginnend mit der Kontrolle der

Wirkstoffsynthese und der Identifikation von Rohstoffen vor der Einwaage bis zur Wirkstoffgehaltsbestimmung bei der fertigen Formulierung zum Einsatz. In Tabelle 1-1 sind Beispiele von PAT-Anwendungen in der pharmazeutischen Industrie aufgeführt.

**Tabelle 1-1:** PAT–Anwendungen in der pharmazeutischen Industrie

<b>Grundoperation</b>	<b>Einsatz von PAT</b>	<b>Literatur</b>
➤ Wirkstoffsynthese	➤ Prozessüberwachung, Messung der Reaktionsgeschwindigkeitskonstante	➤ [9]
➤ Kristallisierung	➤ Kontrolle der Kristallisation, Analyse der polymorphen Zusammensetzung, Beobachtung von Veränderungen der polymorphen Zusammensetzung und anschließender Quantifizierung	➤ [10-13]
➤ Mahlprozesse	➤ Beobachtung von prozessinduzierten polymorphen Umwandlungen	➤ [14]
➤ Mischprozesse	➤ Prozessüberwachung und Endpunktbestimmung	➤ [15-20]
➤ Granulierung	➤ Prozessüberwachung und Beobachtung von prozessinduzierten polymorphen Umwandlungen	➤ [21-27]
➤ Trocknung	➤ Prozessüberwachung, Endpunktbestimmung, Beobachtung von prozessinduzierten polymorphen Umwandlungen	➤ [28-32]
➤ Tablettierung	➤ Quantifizierung Wirkstoffgehalt, Beobachtung von prozessinduzierten polymorphen Umwandlungen und anschließender Quantifizierung, Bestimmung der Tablettierhärte	➤ [6,33-38]
➤ Überziehen	➤ Prozessüberwachung, Filmdickenbestimmung, Bestimmung der Überzugsgleichmäßigkeit und der Auftragsmenge	➤ [39-51]
➤ Verpacken	➤ Identitätsprüfung	➤ [52]
➤ Haltbarkeit	➤ Stabilitätsüberwachung, Analyse von Abbauprodukten	➤ [53]

Für die Prozesskontrolle haben sich in der Prozessanalytischen Technologie vor allem die spektroskopischen Analysemethoden in Verbindung mit der multivariaten Datenanalyse etabliert. Vor allem die Nahinfrarotspektroskopie [17-20,25-28,45,46,54,55] findet breite Anwendung in der Prozessanalytischen Technologie und wird oft als Synonym für PAT verwendet. Aber inzwischen entwickelt sich die Ramanspektroskopie zur gleichwertigen

Alternative [7,9,15,16,21,29,37,39-42,48,51,53,56-69]. Weiterhin werden auch neue Technologien wie Terahertzspektroskopie [49,50,70-72] und Mikrowellensensoren [32] als PAT-Werkzeuge eingesetzt.

## **1.2 Instrumente der Prozessanalytischen Technologie**

### **1.2.1 Einleitung**

Für die Einführung von PAT in der Prozessentwicklung und Produktion ist neben der Anwendung der notwendigen Methoden bzw. Werkzeuge der Prozessanalytischen Technologie vor allem auch die konsequente Umsetzung des Konzepts bei den Mitarbeitern notwendig. Im Folgenden soll eine Übersicht über die in der Prozessanalytischen Technologie zum Einsatz kommenden Werkzeuge gegeben werden [6,73].

### **1.2.2 Risikoanalyse**

Der erste Schritt für die Anwendung von PAT ist eine Risikobetrachtung mit dem Ziel, eine qualitative Bewertung des Einflusses einzelner Prozessschritte und der dortigen Einflussgrößen auf die Produkteigenschaften bzw. Endproduktqualität zu erhalten. Als Beispiele sind hier die Fehlermöglichkeits- und Einflussanalyse (Failure Mode and Effects Analysis, FMEA) und Ishikawa-Diagramme zu erwähnen.

### **1.2.3 Statistische Versuchsplanung**

Mit der Hilfe der statistischen Versuchsplanung (Design of Experiments, DOE) wird die Beziehung zwischen Änderungen der Prozessparameter und den kritischen Produkteigenschaften quantitativ erfasst. Dabei werden Versuche durchgeführt, bei denen die durch Risikoanalyse ermittelten Haupteinflussfaktoren um einen hypothetisch optimalen Betriebspunkt systematisch variiert werden. Darauf basierend wird ein mathematisches Modell erstellt, mit dem der hypothetische Betriebspunkt verbessert wird, und die Wirkung der Haupteinflussfaktoren und ihre Wechselwirkung quantitativ beschrieben werden.

### **1.2.4 Prozessführung**

Traditionell setzt der Qualitäts-Regelkreis beim Produkt an, dessen Qualität geprüft und das Ergebnis dann verwendet wird, die Produktion zu optimieren. Es erfolgt somit eine Optimierung der Produktion von Ansatz zu Ansatz. PAT stellt den Prozess in den Fokus mit dem Ziel, eine Optimierung noch während der laufenden Produktion zu erreichen. Dies erfordert den Einsatz von produktionsbegleitenden Messungen (in- und online Messungen, vgl. 1.2.5), wodurch kritische Prozessparameter und Qualitätsindikatoren (Key Performance

Indicators, KPIs) während der Produktion erfasst werden. Dabei entsprechen die KPIs den durch die Risikoanalyse und DOE ermittelten Haupteinflussfaktoren. Anhand dieser KPIs wird der Prozess durch inline Messungen verfolgt und zur Prozessführung genutzt, um bei auftretenden Störungen in den Prozess eingreifen zu können, damit das Endprodukt immer im zulässigen Bereich liegt (automatisierte Prozessführung).

### **1.2.5 In-/Online Messtechnik**

Für eine automatisierte Prozessführung eignet sich nicht die klassische Laboranalytik (offline) und in den meisten Fällen auch nicht die atline Messung aufgrund der nicht ausreichenden Messgeschwindigkeit. Um eine Optimierung während der laufenden Produktion zu erreichen (vgl. 1.2.4), müssen die KPIs in-/online erfasst werden, weshalb die in-/online Messtechnik ein wichtiges Element in der Prozessanalytischen Technologie darstellt. Bei den KPIs handelt es sich oft um stoffspezifische Parameter wie Konzentrationen in gasförmigen, flüssigen oder festen Systemen. Daher wird bezüglich der Messtechnik sowohl Schnelligkeit als auch Genauigkeit über den zu betrachtenden Konzentrationsbereich gefordert, wobei sich in den letzten Jahren die Spektroskopie für in-/online Anwendungen etabliert hat. Bei der Charakterisierung von Feststoffen hat sich vor allem die NIR-Spektroskopie als quantitative in-/online Methode hervorgehoben (vgl. 1.1).

### **1.2.6 Multivariate Datenanalyse**

Die zu überwachenden und steuernden KPIs sind oft nicht direkt messbar. Dabei handelt es sich um Kombinationen mehrerer Messgrößen, wodurch das zu betrachtende Problem multivariat ist. In diesem Zusammenhang kommen multivariate statistische Methoden wie die Partial Least Squares (PLS) zum Einsatz (vgl. 1.3.3), womit entweder im Rahmen der statistischen Versuchsplanung aus den experimentellen Daten Modelle abgeleitet oder aus Produktionsdaten Haupteinflussgrößen identifiziert und moduliert werden. Die multivariate Datenanalyse (MVDA) spielt zusammen mit der automatisierten Prozessführung und der in-/online Messtechnik eine wichtige Rolle in der Prozessanalytischen Technologie. Die MVDA filtert aus den Unmengen an Prozessdaten die KPIs heraus, die zur Prozessführung eingesetzt werden können.

### **1.2.7 Statistische Prozesskontrolle**

Bei der statistischen Prozesskontrolle (Statistical Process Control, SPC) werden die für den Prozess relevanten Haupteinflussfaktoren überwacht, und im Falle von statistisch signifikanten Prozessveränderungen kommt es zur Alarmierung. Die FDA-Initiative PAT

fordert die Prozessparameter (KPIs) zu überwachen, die einen identifizierten und quantifizierbaren Einfluss auf die Produktqualität haben. Dadurch liefert die statistische Prozesskontrolle einen leicht interpretierbaren Alarm, auf den mit vorhandenem Prozesswissen reagiert werden kann.

Durch konsequente Messung, Steuerung und Überwachung der relevanten Prozessparameter (KPIs) kann die Produktqualität im Prozess sichergestellt werden, womit dem Vorsatz der FDA „*Quality cannot be tested into products; it should be built-in or should be by design*“ entsprochen wird [5,74].

### **1.3 Multivariate Datenanalyse**

#### **1.3.1 Einleitung**

Die multivariate Datenanalyse (MVDA) ist ein wichtiges Hilfsmittel bei der Analyse von großen Datenmengen und erarbeitet aus einer großen Zahl von Messwerten die für die gegebene Fragestellung relevanten Informationen heraus. Ziel der MVDA ist eine Datenreduktion, die durch Zusammenfassen von Variablen mit gleicher Information zu sogenannten Hauptkomponenten erfolgt. Dadurch wird jedes Objekt nur mit wenigen Hauptkomponenten beschrieben statt durch viele einzelne Variablen [75-77]. Durch das Finden der Hauptkomponenten trennt die MVDA die Variablen mit Informationsgehalt von denen ohne Informationsgehalt (Rauschen). Anschließend kann ein Modell erstellt werden, das abhängig von der Aufgabenstellung zur Klassifizierung (Hauptkomponentenanalyse, engl. Principle Component Analysis (PCA)) oder Regression (Projektion auf latente Strukturen, engl. Partial Least Squares Regression (PLS)) angewendet werden kann. Die erstellten Modelle der multivariaten Datenanalyse können zur qualitativen oder quantitativen Bestimmung unbekannter Proben verwendet werden, wobei man entweder die Klassenzugehörigkeit (PCA) oder Werte für die Zielgrößen (PLS) erhält.

Für die vorliegende Arbeit ist der Einsatz der MVDA notwendig, um mit Hilfe der Ramanspektroskopie den Coatingprozess anhand von inline Messungen zu verfolgen. Im Gegensatz zu der univariaten Regression werden bei der MVDA geeignete spektrale Bereiche zur Modellerstellung herangezogen (vgl. Kapitel 3.4). Dadurch gewinnt man mehr Informationen aus den erhaltenen Messdaten, die mit dem Coatingverlauf in Zusammenhang gebracht werden können. Weiterhin werden unter Verwendung der MVDA zur Modellerstellung die Methoden weniger empfindlich gegenüber äußeren Einflüssen (vgl. Kapitel 3.4.3.5 und Kapitel 3.5), wodurch man ein robusteres Modell erhält.

### 1.3.2 Hauptkomponentenanalyse (PCA)

Bei der Hauptkomponentenanalyse werden aus den ursprünglich gemessenen Ausgangsdaten (Variablen) die sog. Hauptkomponenten (latente Variablen) berechnet. Die Hauptkomponenten sind optimale Linearkombinationen, d.h. eine gewichtete Summe der ursprünglich gemessenen Variablen. Dadurch entstehen neue künstliche Variablen (Hauptkomponenten), die selbst nicht direkt messbar sind und mehr Information enthalten als jede einzelne Ausgangsvariable [75].

Zur Bestimmung der Hauptkomponenten sucht man die Richtung der maximalen Varianz in den Ausgangsdaten mit der Bedingung, dass die Hauptkomponenten orthogonal zueinander sind. Die auf diese Weise gefundenen Hauptkomponenten stellen ein neues Koordinatensystem dar, das die Ausgangsdaten besser beschreibt. Dabei erklärt die erste Hauptkomponente die größtmögliche Variation in den Daten und die zweite Hauptkomponente dementsprechend die zweitgrößte Variation. Mit ansteigender Anzahl an berechneten Hauptkomponenten sinkt der Anteil der erklärten Varianz der jeweiligen Hauptkomponente mit der Gefahr, dass die darauffolgende Hauptkomponente nur irrelevante Information (Rauschen) enthält. Die Form des Hauptkomponentenmodells kann man allgemein ausdrücken in:

$$X = TP^T + E \quad \text{Gl. 1-1}$$

Die Datenmatrix  $X$  (ursprünglich gemessene Daten) wird in die Scoresmatrix  $T$  und die Hauptkomponentenmatrix  $P$  zerlegt, mit denen die Ausgangsmatrix  $X$  reproduziert werden kann. Die Residuenmatrix  $E$  enthält die Differenz zwischen der originalen und der reproduzierten Datenmatrix (Abbildung 1-1).

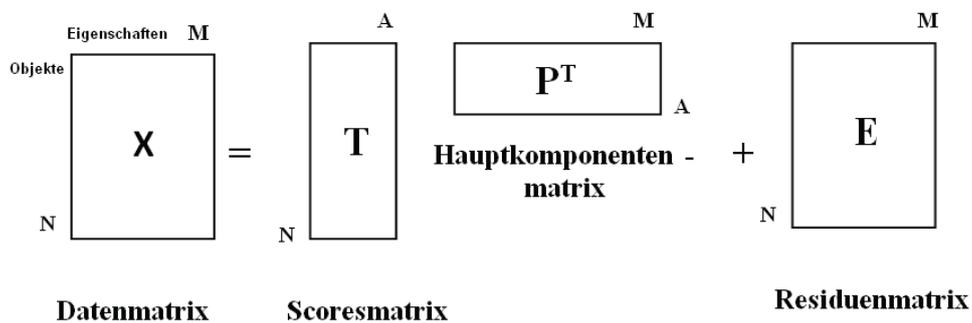


Abbildung 1-1: Matrizen der Hauptkomponentenanalyse [nach [75]]

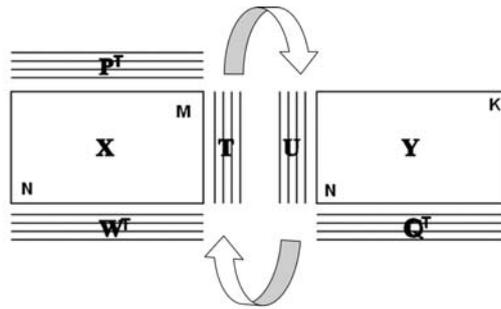
Zusammengesetzt ist die Datenmatrix  $X$  aus  $N$  Zeilen, in denen die Objekte (Proben) stehen, und  $M$  Spalten, die die Eigenschaften (Merkmale, Variablen) angeben. In der Hauptkomponentenmatrix  $P$  stellen die Spalten die berechneten Hauptkomponenten dar, dessen Elemente die Loadings der Hauptkomponenten sind. Da es neben des Heraushebens

der Information auch zur Datenreduktion kommt, werden normalerweise eine deutlich geringere Anzahl  $A$  an Hauptkomponenten berechnet, als es aufgrund der Anzahl  $M$  der Variablen möglich wäre ( $A \ll M$ ). Jede Variable muss vom ursprünglichen Koordinatensystem in das neue Hauptkomponentenkoordinatensystem überführt werden, und weil die Loadings die „Wegbeschreibung“ für diese Überführung sind, hat die Matrix  $P$  die gleiche Anzahl Zeilen  $M$  wie die Datenmatrix Spalten (Variable) hat.

In der Matrix  $T$  stehen die Scores, die die Koordinatenwerte für die Objekte im durch die Hauptkomponenten neu erstellten Koordinatensystem darstellen. Jeweils für jedes Objekt und für jede berechnete Hauptkomponente steht in dieser Matrix  $T$  der Koordinatenwert bezüglich dieser Hauptkomponente. Daher hat die Matrix  $T$  genau so viele Zeilen  $N$  wie die Datenmatrix  $X$  Objekte hat und genauso viele Spalten  $A$  wie Hauptkomponenten berechnet wurden.

### 1.3.3 Projektion auf latente Strukturen (PLS)

Die PLS gehört zu den multivariaten Regressionsmethoden und kombiniert die Eigenschaften der Hauptkomponentenanalyse und der multiplen linearen Regression [75], um ein Regressionsmodell zwischen den Scores der X- und Y-Daten zu erstellen, wobei die X-Werte bei der multivariaten Regression häufig Spektren sind, und die Y-Werte die meist aufwendig zu bestimmenden Referenzwerte darstellen. Allgemein hat die Regressionsanalyse die Zielsetzung, einen funktionalen Zusammenhang zwischen den unabhängigen Variablen (X-Werte) und den davon abhängigen Variablen (Y-Werte) zu beschreiben und eine mathematische Formel zu erstellen, mit der man bei bekannten X-Variablen die zugehörigen Y-Werte vorhersagen kann. Das besondere an der PLS ist, dass die Hauptkomponenten der X- und Y-Daten gleichzeitig und voneinander abhängig berechnet werden. Es wird mit den X- und den Y-Daten eine PCA gemacht, wobei die X-Daten die PCA der Y-Daten beeinflusst und umgekehrt. Analog der PCA bilden die ursprünglich gemessenen Daten die Datenmatrix  $X$  mit den Dimensionen  $N$  (Objekte) \*  $M$  (Variable). Diesmal wird zu jedem Objekt eine Zielgröße  $y_i$  oder mehrere Zielgrößen  $y_{ik}$  gemessen, die den Vektor  $y$  (bei einer Zielgröße) bzw. die Matrix  $Y$  mit der Dimension  $N$  (Objekte) \*  $K$  (Anzahl der Zielgrößen) bilden (Abbildung 1-2).



**Abbildung 1-2:** Schematische Darstellung der PLS und der beteiligten Matrizen [nach [75]]

Bei der Ermittlung der Hauptkomponenten findet zwischen der X- und Y-Datenmatrix ein gegenseitiger Informationsaustausch statt. Aus den X-Daten werden über die PCA die Scores  $t$  und Loadings  $p$  berechnet. Als Zwischenschritt wird bei der PLS die W-Matrix berechnet, die die Loadings  $w$  (Gewichtsladungen) enthält und die Verbindung zwischen den X- und Y-Daten herstellt. Bei der PLS werden somit zwei verschiedene Arten von X-Loadings berechnet. Die  $p$ -Loadings entsprechen dem der PCA und drücken den Zusammenhang aus zwischen den X-Daten und deren  $t$ -Scores. Die  $w$ -Loadings sind die „effektiven Ladungen“ und geben die Beziehung zwischen den X- und Y-Daten wieder. Für das PLS-Regressionsmodell sind die  $p$ - und die  $w$ -Loadings wichtig und gehen bei der Berechnung der PLS-Regressionskoeffizienten  $b$  ein. Für die Y-Daten werden ebenfalls die Hauptkomponenten ermittelt und man erhält die Scores  $u$  und die Loadings  $q$ .

Wenn alle Hauptkomponenten berechnet wurden, stellen die Daten von X und Y, die nach Abzug der Information der letzten berechneten Hauptkomponente übrig bleiben, die jeweilige Restvarianz dar und werden als Residuenmatrix E bzw. F bezeichnet.

Man kann allgemein die Form der Hauptkomponentenmodelle von den X- und Y-Daten ausdrücken in:

$$X = TP^T + E \quad \text{und} \quad Y = UQ^T + F \quad \text{Gl. 1-2 und Gl. 1-3}$$

Im Anschluss werden nun die Regressionskoeffizienten  $b$  für die X-Variablen aus den Loadings  $w$ ,  $p$  und  $q$  berechnet. Im Falle von mehreren Zielgrößen  $y_{ik}$  ergibt sich eine Regressionskoeffizientenmatrix der Dimension  $N$  (Anzahl Objekte) \*  $K$  (Anzahl der Zielgrößen), wobei für jede Spalte  $K$  in den Y-Daten ein Regressionsvektor  $b_k$  berechnet wird. Wenn man diese Regressionskoeffizientenmatrix auf die Messwerte  $X_i$  anwendet, die für ein neues Objekt  $i$  ermittelt wurden, errechnet sich die Zielgröße  $y_k$  nach Gleichung 1-4:

$$y_{ik} = b_0 + x_i^T b_k \quad \text{Gl. 1-4}$$

Es gibt für jede Zielgröße  $y_k$  einen spezifischen Regressionskoeffizienten, aber es handelt sich um ein einziges Modell für alle X- und Y-Daten. Der Grund dafür ist, dass die beiden X-

Loadingsarten  $w$  und  $p$  aus allen X- und Y-Daten bestimmt werden und somit für alle Zielgrößen gleich sind. Die unterschiedlichen Regressionskoeffizienten für jede Zielgröße werden von der Matrix  $Q$  (Y-Loadings) verursacht, in der die individuellen Zielgrößen  $Y_k$  stecken.

## **1.4 Ramanspektroskopie**

### **1.4.1 Einleitung**

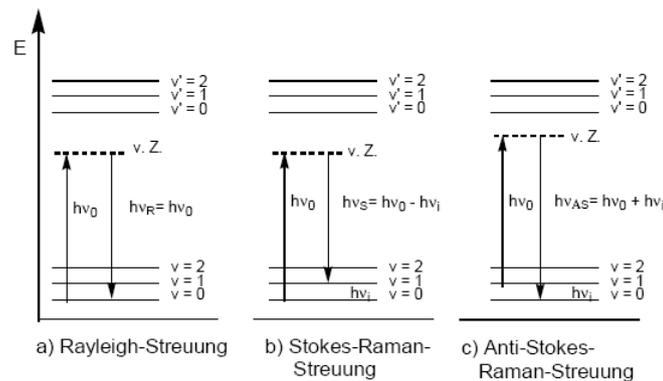
In der vorliegenden Arbeit wurde die Ramanspektroskopie verwendet, um den Coatingprozess zu verfolgen. Das besondere an der Ramanspektroskopie ist, dass sie die Vorteile der Nahinfrarot (NIR)- und Infrarotspektroskopie kombiniert. In Analogie zu der NIR-Spektroskopie können flüssige und feste Proben schnell, zerstörungsfrei und ohne Probenvorbereitung untersucht werden. Dadurch sind im Vergleich zu der Infrarotspektroskopie Messungen direkt im Prozess in Echtzeit (inline Messungen) möglich, das essentiell für eine geeignete PAT-Methode ist (vgl. Kapitel 1.2.5). Weiterhin ist im Gegensatz zum NIR-Spektrum das Ramanspektrum charakteristisch für das untersuchte Material. Damit ist wie bei der IR-Spektroskopie eine Identifizierung des Stoffes anhand des Spektrums möglich, ohne dass eine Auswertung mit Hilfe von Referenzspektren wie bei der NIR-Spektroskopie notwendig ist. Zusätzlich ist es möglich, mit Hilfe der Ramanspektroskopie durch inline Messungen verschiedene Modifikationen eines Arzneistoffes zu identifizieren bzw. zu quantifizieren, die sich während des Prozesses (Granulation, Extrusion, Coating) bilden können.

### **1.4.2 Raman-Effekt**

Beim Raman-Effekt handelt es sich um einen inelastischen Streuvorgang von Photonen an Molekülen oder Kristallen. Dieser Effekt wurde 1923 von G. A. Smekal [78] theoretisch vorausgesagt und 1928 vom indischen Physiker C. V. Raman experimentell mit gefiltertem und fokussiertem Sonnenlicht nachgewiesen [79].

Wird eine Probe mit monochromatischem Licht bestrahlt, geht der größte Teil des Lichts ungehindert durch (Transmission) oder wird absorbiert. Nur ein geringer Anteil des eingestrahnten Lichts wird gestreut. Neben dem unveränderten Anregungslicht (elastische Streuung, Rayleigh-Streuung) werden im Streulicht noch Spektrallinien anderer Frequenzen beobachtet (inelastische Streuung, Raman-Streuung), wodurch das Ramanspektrum eines Moleküls entsteht. Der auftretende Frequenzunterschied bei der Raman-Streuung ist

unabhängig von der Anregerfrequenz und ist charakteristisch für das zu streuende Molekül. Hierbei unterscheidet man zwischen der Stokes- und Antistokes-Raman-Streuung [57,80-82].



**Abbildung 1-3:** Energieniveaus bei der elastischen (a) und inelastischen Streuung (b-c) von Photonen an Molekülen [nach [80]]

Die Abbildung 1-3 verdeutlicht die bei der Rayleigh- und Raman-Streuung an Molekülen beteiligten Prozesse. Bei der Raman-Streuung wird ein virtueller Zwischenzustand (v. Z.), der unterhalb des ersten angeregten elektronischen Niveaus ( $v'$ ) liegt, angenommen. Das Molekül befindet sich am Anfang in der Regel im nicht schwingungsangeregten Grundzustand ( $v=0$ ). Bei der Rayleigh-Streuung (a) wechselwirkt das Molekül mit dem einfallenden Photon, nimmt im virtuellen Zwischenzustand keine Energie vom Photon auf und befindet sich anschließend wieder im nicht schwingungsangeregten Grundzustand. Das gestreute Licht hat somit die gleiche Frequenz wie das Anregerlicht. Dies ist der wahrscheinlichste Prozess und daher ist die Rayleigh-Linie im Raman-Spektrum sehr intensiv (Abbildung 1-4).

Bei der Stokes-Raman-Streuung (b) werden beim Streuprozess Schwingungen im Molekül angeregt. Die gestreute Strahlung ist um den Betrag der zur Schwingungsanregung benötigten Energie ärmer und weist somit eine niedrigere Frequenz als die einfallende Strahlung auf. Bei der Anti-Stokes-Raman-Streuung (c) befindet sich das Molekül vor dem Streuprozess in einem schwingungsangeregten Zustand (z.B.  $v=1$ ) und relaxiert nach der Wechselwirkung mit dem Photon in den nicht schwingungsangeregten Zustand ( $v=0$ ). Die Energie der Schwingung wird auf das Streulicht übertragen, welches dadurch eine höhere Energie und damit eine höhere Frequenz als die einfallende Streustrahlung aufweist. Nach der Boltzmann-Statistik ist der Grundzustand eines Moleküls bei einer niedrigen Temperatur höher besetzt als ein angeregter Zustand, was zu einem Intensitätsunterschied zwischen den Stokes- und Anti-Stokes-Linien zugunsten der Stokes-Linien bei Raumtemperatur führt. Die Abbildung 1-4 verdeutlicht, dass die Raman-Linien symmetrisch unter- und oberhalb der Rayleigh-Linie angeordnet sind. Die Intensität der jeweiligen Linien gibt die Wahrscheinlichkeit des Vorgangs wieder.