

## Reihe 17

Biotechnik/  
Medizintechnik

Nr. 292

M.Sc. Dipl.-Ing. (FH) Sven-Oliver Borchert,  
Hamburg

## Entwicklung, Beobachtung und Steuerung integrierter, quasi-kontinuierlicher pharmazeutischer Produktions- prozesse mit Methoden der Multivariaten Datenverarbeitung



**Entwicklung, Beobachtung und Steuerung integrierter,  
quasi-kontinuierlicher pharmazeutischer Produktionsprozesse  
mit Methoden der Multivariaten Datenverarbeitung**

Von der Naturwissenschaftlichen Fakultät der  
Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades

Doktor der Naturwissenschaften

- Dr. rer. nat. -

genehmigte Dissertation

von

M.Sc. Dipl.-Ing. (FH) Sven-Oliver Borchert  
geboren am 26. Juli 1981 in Göttingen

2016

Referent: Prof. Dr. rer. nat. Thomas Scheper

Korreferent: Prof. Dr. rer. nat. Bernd Hitzmann

Tag der Promotion: 21.12.2015

# Fortschritt-Berichte VDI

Reihe 17

Biotechnik/  
Medizintechnik

M. Sc. Dipl.-Ing. (FH)  
Sven-Oliver Borchert,  
Hamburg

Nr. 292

Entwicklung, Beobachtung  
und Steuerung integrierter,  
quasi-kontinuierlicher  
pharmazeutischer Produktions-  
prozesse mit Methoden der  
Multivariaten Datenverarbeitung

VDI verlag

Borchert, Sven-Oliver

## **Entwicklung, Beobachtung und Steuerung integrierter, quasi-kontinuierlicher pharmazeutischer Produktionsprozesse mit Methoden der Multivariaten Datenverarbeitung**

Fortschr.-Ber. VDI Reihe 17 Nr. 292. Düsseldorf: VDI Verlag 2016.

174 Seiten, 104 Bilder, 33 Tabellen.

ISBN 978-3-18-329217-2, ISSN 0178-9600,

€ 62,00/VDI-Mitgliederpreis € 55,80.

**Für die Dokumentation:** Design of Experiments – Golden Batch Monitoring – Integrierter Bio-prozess – Malaria Vakzinkandidaten – Multivariate Datenverarbeitung – Process Analytical Technology – *Pichia pastoris* – Prozesssteuerung und -automatisierung – Quality by Design – Quasi-kontinuierliche Prozessstrategie

Die vorliegende Arbeit befasst sich mit Aspekten einer modernen Bioverfahrenstechnik am Beispiel von Prozessen zur Herstellung rekombinanter potentieller Malariaimpfstoffe. Dabei wurden zwei quasi-kontinuierliche Prozesse aus herkömmlichen *Batch-Unit Operationen* aufgebaut, in denen die Anwendung von *Process Analytical Technology* im Vordergrund steht. Das Hauptaugenmerk dieser Arbeit lag dabei auf einer Implementierung der *Multivariate Data Analysis* zum Monitoring und zur Evaluierung des zyklischen Prozessablaufes und seiner Reproduzierbarkeit. Im Bereich der *Principal Component Analysis* wurde die Methode der Prozessüberwachung mit dem *Golden Batch-Tunnel* angewendet. Mit dem *Golden Batch-Ansatz* wurden Methoden zur Prozessprädiktion implementiert und mit einer *Model Predictive Multivariate Control* auch zur Steuerung von realen Prozessen erprobt. Darüber hinaus wurde die MVDA zur Prädiktion von Medienkomponenten sowie deren zellspezifische Reaktionsraten aus klassischen Online-Messungen mit *Partial Least Squares Regression* eingesetzt.

### **Bibliographische Information der Deutschen Bibliothek**

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliographie; detaillierte bibliographische Daten sind im Internet unter <http://dnb.ddb.de> abrufbar.

### **Bibliographic information published by the Deutsche Bibliothek**

(German National Library)

The Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliographie (German National Bibliography); detailed bibliographic data is available via Internet at <http://dnb.ddb.de>.

© VDI Verlag GmbH · Düsseldorf 2016

Alle Rechte, auch das des auszugsweisen Nachdruckes, der auszugsweisen oder vollständigen Wiedergabe (Fotokopie, Mikrokopie), der Speicherung in Datenverarbeitungsanlagen, im Internet und das der Übersetzung, vorbehalten.

Als Manuskript gedruckt. Printed in Germany.

ISSN 0178-9600

ISBN 978-3-18-329217-2

## Vorwort

Die vorliegende Arbeit wurde in der Zeit von September 2011 bis September 2015 im Forschungs- und Transferzentrum für Bioprozess- und Analysetechnik der Hochschule für Angewandte Wissenschaften Hamburg unter der Leitung von Herrn Prof. Dr.-Ing. Reiner Luttmann angefertigt.

Herrn Prof. Dr.-Ing. Reiner Luttmann gilt an dieser Stelle ein besonderer und freundschaftlicher Dank für die Unterstützung auf dem Weg vom Studenten bis hin zum promovierten Biotechniker. Gemeinsam haben wir das Potential dieser Arbeit erkannt und stetig auf das Ziel – eine erfolgreiche Promotion – hingearbeitet. Trotz der damit verbunden Strapazen ist es uns dabei gelungen nicht den Spaß an der Schufterei aus den Augen zu verlieren.

Frau Prof. Dr. Gesine Cornelissen möchte ich für die unkomplizierte und konstruktive Zusammenarbeit danken.

Herrn Prof. Dr. Thomas Scheper<sup>1</sup>, Herrn Prof. Dr. Bernd Hitzmann<sup>2</sup>, Prof. Dr. Jürgen Alves<sup>3</sup> und Prof. Dr.-Ing. Karl-Heinz Bellgard<sup>1</sup> danke ich für die Übernahme der Referate.

Ich bedanke mich für das Max-Buchner-Forschungsstipendium der DECHEMA (Gesellschaft für Chemische Technik und Biotechnologie e.V.) mit der Förderkennziffer 2890.

Für die umfangreiche und gut durchdachte Laborinfrastruktur besonders im Bereich der Prozesstechnik und Automatisierung bedanke ich mich bei Ullrich Scheffler.

Meinem Freund Hans-Peter Bertelsen danke ich für die technische Unterstützung genauso, wie für die vielfältigen und überaus aufmunternden Gespräche und eine wunderbare gemeinsame Zeit!

Ein ebenso freudiger wie freundschaftlicher Gruß geht an meine beiden Wegbegleiter und Leidensgenossen Jens und David. Ein ausgewogenes Verhältnis zwischen beruflichen Anforderungen und privaten Bedürfnissen ist auch während der Promotion ein wichtiger Baustein des letztendlichen Erfolgs! Die beiden haben einiges dazugetan und damit einen nicht unerheblichen Teil zum Gelingen dieser Arbeit beigetragen.

Fabian, Ehsan, Tobi, Merle, Stephan, Eule und Flo danke ich für ihren unermüdlichen Einsatz sowie die im Rahmen von Master- und Diplomarbeiten eingebrachten Ergebnisse.

Für die tolle Arbeitsatmosphäre bedanke ich mich bei allen Mitarbeitern des Labors BPA. Vor allem Sanja, Jessica, Janett, Maria, Christian, Kai, Kristof, Jan-Patrick, Andree, Frithjof und Jan.

Hanna gilt mein Dank für Ihre Heiterkeit und unsere gemeinsame Zeit am Meer.

Meinen Eltern Waltraud und Herbert und meinen Geschwistern Thies und Björn danke ich für Ihre Liebe und Ihr Vertrauen während der gesamten Studien- und Promotionszeit.

---

<sup>1</sup> Institut für Technische Chemie, Leibniz Universität Hannover

<sup>2</sup> Institut für Lebensmittelwissenschaft und Biotechnologie, Universität Hohenheim

<sup>3</sup> Hannover Institut für Biophysikalische Chemie, Medizinische Hochschule





**Inhaltsverzeichnis**

<b>1. Einleitung und Zielsetzung .....</b>	<b>1</b>
1.1 Paradigmenwechsel in der pharmazeutischen Produktion .....	1
1.2 Ansätze einer integrierten kontinuierlichen pharmazeutischen Produktion .....	2
1.3 Zielsetzung der Arbeit .....	3
<b>2. Zielprodukte, Hostsystem und Zielprodukt-Analytik .....</b>	<b>5</b>
2.1 Potentielle Malariaimpfstoffe D1M1 und D1M1H .....	5
2.2 Das Hostsystem <i>Pichia pastoris</i> .....	6
2.3 Offline-Analytik der Zielprodukte .....	7
2.3.1 Ermittlung der Konzentration des sekretierten Gesamtproteins .....	7
2.3.2 Bestimmung der Produktreinheit und der Konzentration des Zielproduktes .....	7
<b>3. Grundoperationen eines integrierten Bioprozesses im Labormaßstab .....</b>	<b>9</b>
3.1 Prozesstechnische Zielsetzung .....	9
3.2 Die erweiterte Instrumentierung des Upstream-Prozesses .....	10
3.2.1 Die eingesetzte Bioreaktoreinheit mit seiner MSR-Technik .....	10
3.2.2 Bestimmung der Zelldichte über das Trübungssignal .....	11
3.2.3 O <sub>2</sub> - und CO <sub>2</sub> -Gasbilanzen zur Online-Prozessbeobachtung .....	13
3.3 Zellabtrennung und Proteinvorreinigung .....	15
3.3.1 Grundlagen der Expanded Bed Adsorptions-Chromatografie .....	15
3.3.2 Das Adsorbermaterial und die eingesetzte Säule .....	17
3.3.3 Aufbau und MSR-Technik der EBA-Anlage .....	18
3.3.4 Ablauf einer Aufarbeitungsprozedur .....	21
3.3.5 Automatisierung der Aufarbeitungsprozedur .....	24
3.4 Feinreinigung des Zielproduktes mit einem FPLC-System .....	25
<b>4. Integrierte Prozessführung mit einer EBA-Chromatografie .....</b>	<b>28</b>
4.1 Das Konzept des Integrierten EBA-Bioprozesses .....	28
4.2 Sequentielle Zellkultivierung und Proteinexpression .....	28
4.3 Die Erprobung der integrierten Produktionsweise .....	31
4.3.1 Prozessverlauf mit nicht optimalen Prozessparametern .....	31
4.3.2 Verbesserung der Produktqualität mit optimalen Kultivierungsparametern .....	32
<b>5. Implementierung von Vorgaben des Quality by Design .....</b>	<b>34</b>
5.1 Zielsetzungen von QbD in der Bioprozessentwicklung .....	34
5.2 Definition der Critical Quality Attributes im Downstream .....	35
5.2.1 Optimierungspotential und Gütekriterium des EBA Prozesses .....	35
5.2.2 Auswahl möglicher Critical Process Parameter .....	36
5.3 Die Idee des Design of Experiments .....	37

5.4	Lösungsansatz zum Optimierungsproblem der Aufarbeitung mit der EBA .....	42
5.5	Screening des mehrparametrischen Problems im ÄKTA™purifier 100 .....	43
5.6	Konsequente Pre-Optimierung mit dem ÄKTA™purifier 100 .....	46
5.7	In situ-Optimierung mit der EBA im integrierten Prozess .....	49
5.8	Erweiterung der integrierten Anlage um einen Feinreinigungsschritt .....	54
<b>6.</b>	<b>Aufbau und Erprobung eines quasi-kontinuierlichen Bioprozesses .....</b>	<b>57</b>
6.1	Entwicklung einer industrie-kompatiblen integrierten Produktionsanlage .....	57
6.2	Konzept eines sechsstufigen integrierten Prozesses .....	58
6.3	Die Upstream Unit-Operationen .....	60
6.3.1	Aufbau der zweistufigen Upstream-Anlage .....	60
6.3.2	Die zweistufige zyklische Produktionsstrategie .....	60
6.4	Eingesetzte Mess-, Steuer- und Regelungstechnik .....	63
6.4.1	Grundautomatisierung der Bioreaktoren .....	63
6.4.2	Messung des Sekundärsubstrates Methanol .....	63
6.4.3	Abgasmessung und Gasbilanzen .....	64
6.4.4	Einsatz einer NIR-Spektroskopie .....	64
6.4.5	Überwachung der Expression mit einer Atline-HPLC .....	66
6.5	Die Downstream Unit-Operationen .....	69
6.5.1	Zellabtrennung mit dem Separator SC1 .....	69
6.5.2	Mikrofiltration mit Sartoclon Slice MF-Modulen .....	70
6.5.3	Ultrafiltration mit Sartoclon Slice UF-Modulen .....	71
6.5.4	Chromatographische Produktreinigung mit einem ÄKTA™purifier 100 .....	73
6.6	Automatisierung des integrierten Prozesses .....	73
6.6.1	Übersicht der dezentralen Automatisierungskomponenten .....	73
6.6.2	Automatisierung des Upstream-Prozesses mit MFCS/win .....	75
6.6.3	Integration und Automatisierung der Gesamtanlage über PCS 7 .....	75
6.7	Die quasi-kontinuierliche Herstellung des Malaria-Vakzinkandidaten D1M1H .....	76
6.7.1	Verlauf eines integrierten Produktionszyklus .....	76
6.7.2	Quasi-kontinuierliche Produktion in sequentiell/paralleler Prozessführung .....	79
6.8	Erfassung und Vorverarbeitung von Prozessdaten zur MVDA .....	81
6.8.1	Erweiterung der Anlage um eine komplexe Datenverarbeitung .....	81
6.8.2	Erfassung und Datenverarbeitung mit PCS 7 und SIPAT .....	82
6.8.3	Das Offline-Modul SIMCA .....	83
6.8.4	Das Online-Modul SIMCA Q .....	83
6.8.5	Das Prädiktor- und Control-Modul SIMCA-online .....	84

<b>7. Multivariate Datenanalyse zur Prozessbewertung.....</b>	<b>85</b>
7.1 Zielsetzung des MVDA-Einsatzes.....	85
7.2 MVDA als PAT-Tool zur Beherrschung großer Datenmengen .....	85
7.3 Vorbereitung der Prozessdaten zur Anwendung der MVDA.....	86
7.3.1 Aufbau der Prozessdatenmatrix D .....	86
7.3.2 Entfernen von Datenausreißern.....	86
7.3.3 Die skalierte Prozessdatenmatrix X.....	86
7.4 Eine kurze Einführung in PCA – Principal Component Analysis.....	88
7.5 Anwendung der MPCA zur Prozessbewertung.....	90
7.5.1 Aufbau einer dreidimensionalen Prozessdatenmatrix D.....	90
7.5.2 MPCA – Multiway Principal Component Analysis.....	91
7.6 Entwicklung des Golden Batch Modells.....	94
7.7 Anwendung der MPLS zur Prozessprädiktion .....	95
7.7.1 Hinführung auf das MVDA-Problem .....	95
7.7.2 MPLS – Multiway Partial Least Squares Regression .....	96
7.8 Modellvalidierung .....	100
<b>8. Golden Batch Diagnose, Monitoring und Control zyklischer Teilprozesse .....</b>	<b>106</b>
8.1 Forschungsbedarf einer MVDA basierten Prozessführung.....	106
8.2 Ermittlung des Golden Batch-Tunnels für den Upstream Prozess .....	107
8.3 Überwachung der Prozessreproduktion mit einem Golden Batch-Monitoring .....	108
8.4 Beobachtung, Prognose und Steuerung QbD-konformer Golden Batch-Verläufe.....	110
8.4.1 Erstellung QbD-konformer Golden Batch-Modelle .....	110
8.4.2 Prognose des Golden Batch-Verlaufes .....	112
8.4.3 Modellgestützte multivariate Steuerung der Prozessqualität.....	112
<b>9. Bestimmung zellspezifischer Reaktionsraten über MVDA .....</b>	<b>118</b>
9.1 Beteiligte Reaktionskomponenten .....	118
9.2 Herleitung von Massen- und Konzentrationsbilanzen.....	119
9.3 Herleitung der $q_{i/x}$ -Auswertungen des zyklischen Prozesses .....	123
9.4 Die Prädiktion der Zelldichte mit Hilfe eines PLS-Modells .....	126
9.4.1 Die Bereitstellung von Trainingsdaten .....	126
9.4.2 Prädiktion der Zelldichte .....	128
9.4.3 Prädiktion des sekretierten Zielproduktes.....	132
9.4.4 Prädiktion der Komponenten im cell breeding Prozess .....	133
9.4.5 PLS-Koeffizienten der entwickelten Prädiktionsmodelle.....	134
9.5 Ermittlung der zellspezifischen Reaktionsraten .....	136
9.5.1 Offline-Auswertung der drei Prozessphasen .....	136

9.5.2	Identifikation der maximalen zellspezifischen Reaktionsraten.....	137
9.5.3	Online-Bestimmung der zellspezifischen Raten .....	138
<b>10.</b>	<b>Zusammenfassung und Ausblick.....</b>	<b>141</b>
<b>11.</b>	<b>Anhang.....</b>	<b>143</b>
11.1	Kultivierungsmedien.....	143
11.2	Pufferlösungen für Filtration und Chromatographie .....	144
11.3	Offline-Messungen.....	145
11.3.1	Bestimmung der optischen Dichte, der Biofeucht- und Biotrockenmasse .....	145
11.3.2	Bestimmung des Gesamtproteingehalts (Bradford-Test) .....	146
11.3.3	SDS-PAGE zur Reinheitsbestimmung.....	146
<b>12.</b>	<b>Literaturverzeichnis .....</b>	<b>148</b>
12.1	Literatur.....	148
12.2	Veröffentlichungen des Autors .....	153

## Nomenklatur

### Allgemeine und prozessrelevante Abkürzungen

AMA1	:= Apical Membrane Antigen 1
AOX	:= Alkoholoxidase
BMBF	:= Bundesministerium für Bildung und Forschung
BPA	:= Bioprozessautomatisierung
BPRC	:= Biomedical Primate Research Center
BWF	:= Behörde für Wissenschaft und Forschung
D1M1	:= Fusionsprotein aus AMA1 und MSP1
DCU	:= Bioreaktor-Kontrolleinheit ( <i>digital control unit</i> )
DiCo	:= Diversity Covering
DSP	:= Downstream Processing
EBA	:= Expanded Bed Adsorption
EMA	:= European Medicines Agency
FDA	:= (United States) Food and Drug Administration
FPLC	:= Fast Protein Liquid Chromatography
HPLC	:= High Performance Liquid Chromatography
IMAC	:= Immobilized Metal Affinity Chromatography
IP	:= Integrated Processing
ISPR	:= In-Situ Product Removal
M&C	:= Measurement and Control
MSP1	:= Merozoite Surface Protein 1
Mut <sup>S</sup>	:= Methanol utilization slow
PAGE	:= PolyAcrylamide Gel Electrophoresis
<i>Pf</i>	:= <i>Plasmodium falciparum</i>
SCADA	:= Supervisory Control and Data Acquisition
SDS	:= Sodium Dodecyl Sulfate
SME	:= Small and Medium Enterprises
USP	:= Upstream Processing
WHO	:= World Health Organization

### Prozessgrößen

$A_{iK}$	:= UV-Absorption der Messung von Komponente I in Teilsystem K	[AU]
$a_i$	:= Koeffizient des physikalischen DoE-Modells	[xx]
$A_{P1M}$	:= UV-Absorption des Zielproteins P1 in der Medienphase	[AU]
$B_{jKi}$	:= relatives Flächenmaß einer Bande j in der SDS-PAGE i aus K	[-]
$b_i$	:= Koeffizient des codierten DoE-Modells	[xx]
$c_{iK}$	:= Konzentration von Komponente I in Teilsystem K	[g l <sup>-1</sup> ]
$C_{iK}$	:= molare Konzentration von Komponente I in Teilsystem K	[mol l <sup>-1</sup> ]

$D_{P1E}$	:= globale Desirability-Funktion der EBA-Optimierung	[-]
$d_{PP1F}$	:= Desirability-Funktion der Produktreinheit $P_{P1F}$ der EBA-Optimierung	[-]
$d_{RP1E}$	:= Desirability-Funktion der Wiederfindung $R_{P1E}$ der EBA-Optimierung	[-]
$E_K$	:= Trübung in Teilsystem K	[AU]
$F_K$	:= Volumenstrom aus/in Teilsystem K	[l·h <sup>-1</sup> ]
$F_{nl}$	:= Begasungsrate der Komponente I unter Normbedingungen	[l·h <sup>-1</sup> ]
$g_{Pi/X}$	:= zellspezifischer Gehalt an Produkt $P_i$	[gg <sup>-1</sup> ]
$IA_{P1Eout}$	:= Integral der UV-Absorption von Produkt $P_1$ im EBA-Auslass	[AUh]
$L_{PP1F}$	:= untere Grenze der Reinheit $P_{P1F}$	[-]
$L_{RP1E}$	:= untere Grenze der Wiederfindung $R_{P1E}$	[-]
$M_i$	:= Molmasse der Komponente I	[gmol <sup>-1</sup> ]
$m_{iK}$	:= Masse von Komponente I in Teilsystem K	[g]
$m_K$	:= Masse von Teilsystem K	[g]
$N_{St}$	:= Rührerdrehzahl	[min <sup>-1</sup> ]
$OD$	:= Optische Dichte	[AU]
$pH_K$	:= pH-Wert von Teilsystem K	[-]
$p_{Kx}$	:= Druck von Teilsystem K gemessen am Ort x	[bar]
$pO_2$	:= relativer Sauerstoffpartialdruck in der Flüssigphase	[%]
$P_{P1K}$	:= Reinheit des Produktes $P_1$ im Teilsystem K	[-]
$PRD_{int}$	:= Produktivität des Integrierten Prozesses	[gh <sup>-1</sup> ]
$PRD_{up}$	:= Produktivität des Upstream Prozesses	[gh <sup>-1</sup> ]
$p_{TMj}$	:= Transmembrandruck der Filtration j	[bar]
$Q_i$	:= volumetrische Ver- / Entsorgungsrate der Komponente I	[gl <sup>-1</sup> ·h <sup>-1</sup> ]
$q_{i/X}$	:= zellspezifische Reaktionsrate der Komponente I	[h <sup>-1</sup> ]
$R_{P1K}$	:= Wiederfindung des Produktes $P_1$ im Teilsystem K	[-]
$RQ$	:= molarer Respirationsquotient	[molmol <sup>-1</sup> ]
$S_K$	:= Streulichtmessung in Teilsystem K	[STU]
$t$	:= Zeit	[h]
$T_{int}$	:= Zeitintervall eines integrierten Produktionszyklus	[h]
$U_{PP1F}$	:= obere Grenze der Reinheit $P_{P1F}$	[-]
$U_{RP1E}$	:= obere Grenze der Wiederfindung $R_{P1E}$	[-]
$V_K$	:= Volumen des Teilsystems K	[l]
$V_{nM}$	:= Molvolumen unter Normbedingungen	[l·mol <sup>-1</sup> ]
$x_i$	:= Molenbruch der Komponente I in der Gasphase G	[%]
$x_{iG}$	:= Stoffmengenanteil der Komponente I in der Gasphase	[-]
$y_{i/J}$	:= Ausbeutekoeffizient der Komponente I aus Komponente J	[gg <sup>-1</sup> ]
$\alpha_{Z/X}$	:= Biofeucht-/Biotrockenmasseverhältnis	[-]
$\vartheta_K$	:= Temperatur in Teilsystem K	[°C]
$\mu$	:= zellspezifische Wachstumsrate	[h <sup>-1</sup> ]
$\rho_K$	:= Dichte des Teilsystems K	[gl <sup>-1</sup> ]
$\rho_Z$	:= Dichte feuchter Zellen	[gl <sup>-1</sup> ]
$\sigma_K$	:= Leitfähigkeit von Teilsystem K	[mScm <sup>-1</sup> ]

## Prozessrelevante Indizes

### Teilsysteme

B1	:= Laufpuffer	perj	:= Permeat der Filtration j
B2	:= Elutionspuffer	pro	:= Produktionsbereich im USP
C	:= Chromatographiesystem	pur	:= Feinreinigung (purification)
C	:= Konditionierungstank	R11	:= Glycerolvorlage an Reaktor 1
DJp	:= Doppelmantel an Reaktor p	R2p	:= Methanolvorlage an Reaktor p
E	:= EBA-System	R3p	:= Refreshvorlage an Reaktor p
F	:= EBA-Fraktionensammler	retj	:= Retentat der Filtration j
feedj	:= Zulauf der Filtration j	S	:= Probe (sample)
Gp	:= Gasphase von Reaktor p	sep	:= Separator Ablauf
Hp	:= Ernte von Reaktor p	T1p	:= Säurevorlage an Reaktor p
int	:= Integrierter Anlagenbereich	T2p	:= Basevorlage an Reaktor p
Lp	:= Flüssigphase in Reaktor p	TMj	:= Transmembran in Filtration j
mf	:= Mikrofiltrationsbereich	uf	:= Ultrafiltrationsbereich
Mp	:= Medienphase in Reaktor p	up	:= Upstreambereich
N	:= Klarphase	Wj	:= Wastetank j
P	:= Produkttank der Filtration	X	:= Biotrockenmasse
P	:= AEKTA™purifier 100	Z	:= Biofeuchtpphase

### Komponenten

Ac	:= Säure (acid)	P1	:= Produkt 1 (Zielprodukt)
AF	:= Antischaum (anti foam)	P2	:= Produkt 2 (Alkoholoxidase)
AIR	:= Luft (air)	Ptot	:= Gesamtprotein
Al	:= Lauge (alkali)	S1	:= Substrat 1 (Glycerol)
C, CO2	:= Kohlenstoffdioxid	S2	:= Substrat 2 (Methanol)
Im	:= Imidazol	T20	:= Tween®20
O, O2	:= Sauerstoff		

### Zustände, Orte

^	:= Schätzwert	min	:= Minimalwert
—	:= Mittelwert	n	:= Finaler Messwert
0	:= Anfangsbedingung	n	:= Gas-Normbedingungen
at	:= at-line	opt	:= optimal
bottom	:= unterer Säuleneinlass /-auslass	out	:= Ausgang (output)
cdw	:= Biotrockenmasse	pre	:= Prädiktion
crit	:= kritischer Wert	rel	:= relativ
est	:= abgeschätzt	SP	:= Sollwert
gr	:= Wachstumsanteil (growth)	top	:= oberer Säuleneinlass /-auslass
high	:= oberer Faktorwert	tot	:= total / gesamt
in	:= Eingang (input)	trans	:= Transfer
low	:= unterer Faktorwert	turb	:= Trübung (turbidity)
m	:= Erhaltungsstoffwechsel	uv	:= ultraviolett
max	:= Maximalwert	w	:= Sollwert

## QbD- und MVDA-relevante Bezeichnungen

BLM	:= Batch Level Model
BOBYQA	:= Bound Optimization By Quadratic Approximation
CCC	:= Central Composite Circumscribed
CCF	:= Central Composite Face-centered
cGMP	:= current Good Manufacturing Practice
CM	:= Control Model
CPP	:= Critical Process Parameters
CQA	:= Critical Quality Attributes
DModX	:= Distance to Model X
DoE	:= Design of Experiments
F	:= F-ratio (F-distribution)
FMEA	:= Failure Mode and Effects Analysis
GB	:= Golden Batch
IBR	:= Imputation by Regression
ICH	:= International Conference on Harmonisation
J	:= Gütefunktional (objective function/cost function/performance index)
MLR	:= Multiple Linear Regression
MM	:= Monitoring Model
MPCA	:= Multiway Principal Component Analysis
MPLS	:= Multiway Partial Least Squares (regression)
MPMC	:= Model Predictive Multivariate Control
MV	:= Model Validity
MVDA	:= MultiVariate Data Analysis
NIPALS	:= Non-Linear Iterative Partial Least Squares
OLM	:= Observation Level Model
PAT	:= Process Analytical Technology
PCA	:= Principal Component Analysis
PCR	:= Principal Component Regression
$p_{lof}$	:= p-Wert für das Lack of Fit
PLS	:= Partial Least Squares (regression)
PRESS	:= Predictive Residual Sum of Squares
PS	:= Prediction Set
$Q^2$	:= Goodness of prediction
QbD	:= Quality by Design
$R^2$	:= Goodness of fit
RMSE	:= Root Mean Squared Error
RP	:= ReProducibility
SPE	:= Standard Prediction Error
SS	:= Sum of Squares
$T^2$	:= Hotellings $T^2$ -value
TPP	:= Target Product Profile
TS	:= Training Set
VS	:= Validation Set
$x_i$	:= DoE-Faktoren
$y_i$	:= DoE-Responses
$\theta$	:= Wichtungsfaktor (weight factor)



## MVDA-relevante Matrizen und Vektoren

- $D$  := ( $n \times m$ ) Original Online-Datenmatrix  
 $E$  := ( $n \times m$ ) Residuenmatrix von PCA/PLS Modellen für X  
 $F$  := ( $n \times v$ ) Residuenmatrix von PLS Modellen für Y  
 $G$  := ( $n \times v$ ) Residuenmatrix eines PLS-Modells zur Prädiktion von Y aus X  
 $I$  := ( $i \times i$ ) Einheitsmatrix  
 $M$  := ( $n \times v$ ) Original Offline-Datenmatrix  
 $P$  := ( $m \times r$ ) Loading-Matrix für X  
 $Q$  := ( $v \times r$ ) Loading-Matrix für Y  
 $T$  := ( $n \times r$ ) Score-Matrix für X  
 $U$  := ( $n \times r$ ) Score-Matrix für Y  
 $W$  := ( $m \times r$ ) PLS Weighting-Matrix für P und Q  
 $X$  := ( $n \times m$ ) autoskalierte Online-Prozessdatenmatrix  
 $Y$  := ( $n \times v$ ) erweiterte autoskalierte Offline-Datenmatrix  
 $d_j$  := ( $n \times 1$ ) Vektor einer Online-Messung mit n Messwerten (Observation)  
 $e_j$  := ( $n \times 1$ ) Residuen-Vektor von PCA/PLS Modellen für X  
 $f_h$  := ( $n \times 1$ ) Residuen-Vektor von PLS Modellen für Y (aus den X-Daten)  
 $g_h$  := ( $n \times 1$ ) Residuen-Vektor von PLS Modellen für Y (aus U-Scores)  
 $m_h$  := ( $s \times 1$ ) Vektor einer Offline-Messung mit s Analysenwerten  
 $p_{Ci}$  := ( $n \times 1$ ) Vektor der i-ten Hauptkomponente (principal component) von X  
 $p_i$  := ( $m \times 1$ ) Loading-Vektor aus Matrix P mit r Loadings  
 $q_i$  := ( $v \times 1$ ) Loading-Vektor aus Matrix Q mit r Loadings  
 $t_i$  := ( $n \times 1$ ) Score-Vektor aus Matrix T mit r Scores  
 $u_i$  := ( $n \times 1$ ) Score-Vektor aus Matrix U mit r Scores  
 $w_i$  := ( $m \times 1$ ) Weighting-Vektor mit r Gewichtungen  
 $x_j$  := ( $n \times 1$ ) autoskalierter Vektor von  $d_j$  mit n Messungen  
 $y_h$  := ( $n \times 1$ ) autoskalierter Vektor der erweiterten Offline-Datenmatrix Y  
 $\sigma$  := ( $n \times 1$ ) Vektor der Standardabweichungen

## MVDA-relevante Indizes

- $cp$  := Center Point  
 $crit$  := kritischer Wert  
 $gb$  := Golden Batch  
 $h$  := Laufindex der Variablen in Y  
 $i$  := Laufindex der Online-Observationen  
 $j$  := Laufindex der Online-Variablen  
 $k$  := Laufindex der Batches und der DoE-Versuche  
 $l$  := Laufindex der Hauptkomponenten  
 $lof$  := Lack of Fit  
 $m$  := Anzahl der Online-Variablen (Spalten in D/X)  
 $m$  := Anzahl der DoE-Center Point-Versuche  
 $MV$  := Manipulated Variable (Steuergrößen oder Sollwerte)  
 $n$  := Anzahl der Online-Observationen (Zeilen in D/X)  
 $n$  := Anzahl der DoE-Faktoren (Dimension des Suchraumes)  
 $o$  := Anzahl der Batches

pre	:= Wert aus einer PLS-Prädiktion
r	:= Anzahl der Hauptkomponenten (principal components) für X
rel	:= relativer Wert
SP	:= Sollwert im Gütefunktional J (BOBYQA)
sub	:= submaximaler Wert
v	:= Anzahl der Analysenvariablen bzw. Attribute in Y
$\alpha$	:= Signifikanzniveau

## Kurzfassung

Sven-Oliver Borchert

### Entwicklung, Beobachtung und Steuerung integrierter, quasi-kontinuierlicher pharmazeutischer Produktionsprozesse mit Methoden der Multivariaten Datenverarbeitung

Die vorliegende Arbeit befasst sich mit Aspekten einer modernen Bioverfahrenstechnik am Beispiel von Prozessen zur Herstellung rekombinanter potentieller Malariaavakzine.

Dabei wurden zwei quasi-kontinuierliche Prozesse aus herkömmlichen *Batch-Unit Operationen* zur Prozessintensivierung aufgebaut, in denen die Anwendung von PAT – *Process Analytical Technology* im Vordergrund steht. Die damit durchgeführten Prozessentwicklungen dienen insbesondere dazu, die von der FDA – *US Food and Drug Administration* vorgegebenen Randbedingungen eines QbD – *Quality by Design* zu erfüllen.

Im Bereich der Integrierten Herstellung rekombinanter Pharmaproteine wurde zunächst ein Laborprozess mit einer zyklischen Zellanzucht und nachfolgender Proteinexpression unter Einbindung einer Insitu EBA – *Expanded Bed Adsorptions* Chromatographie aufgebaut und in sequentiell/integrierter Prozessführung erprobt.

Der zweistufige EBA-Prozess wurde nach Maßgabe von QbD-Vorgaben entwickelt. Ausgehend von der Festlegung der CQA – *Critical Quality Attributes* und ihrer zugeordneten Operationsparameter wurden, über die Anwendung von DoE – *Design of Experiments*, die CPP – *Critical Process Parameter* mit ihren optimalen Prozessbedingungen und deren *Design Space* identifiziert.

In einem industriell orientierten Bereich wurde aus einer für den Batchbetrieb ausgelegten Scale-down Produktionsanlage, bestehend aus einem *cell breeding*, einer *protein production* sowie einem Downstream mit *cell clarification*, *micro-*, *ultra-* und *dia-crossflow filtration* mit abschließender *purification* ein integrierter Pharmaprozess aufgebaut.

Eine sequentielle, aber auch über drei Tage verteilte, parallele Prozessstrategie führte zu einer quasi-kontinuierlichen Fahrweise mit zyklischer Wiederholung eines 56-stündigen Prozesses, der über eine Reihe von Tagen alltäglich geerntet wurde.

Das Hauptaugenmerk dieser Arbeiten lag aber auf dem Gebiet der Anwendung der MVDA – *Multivariate Data Analysis* zum Monitoring und zur Evaluierung des zyklischen Prozessablaufes und seiner Reproduzierbarkeit.

Im Bereich der PCA – *Principal Component Analysis* wurde die Methode der Prozessüberwachung mit dem *Golden Batch*-Tunnel bis hin zur Online-Evaluierung intensiv angewendet.

Mit dem *Golden Batch*-Ansatz und dem Algorithmus IBR – *Imputation by Regression* wurden Methoden zur Prozessprädiktion implementiert und mit einem MPMC – *Model Predictive Multivariate Control* auch zur Steuerung der Qualität von realen Prozessen erprobt.

Darüber hinaus wurde die MVDA zur Prädiktion von Medienkomponenten wie Zellmasse, Glycerol und rekombinantem Produkt sowie deren zellspezifische Reaktionsraten aus klassischen Online-Messungen mit PLS – *Partial Least Squares* Regression eingesetzt.

Schlüsselwörter: Quasi-kontinuierlicher Prozess, *Pichia pastoris*, Malaria Vakzinkandidaten, parallel/sequentiell integrierte Prozessführung, Multivariate Datenverarbeitung, Golden Batch Monitoring, Modellgestützte multivariate Qualitätsregelung, QbD gerechte Prozessentwicklung

## Abstract

Sven-Oliver Borchert

### **Development, Monitoring and Control of Integrated, Quasi-continuous Pharmaceutical Production Processes with Methods of Multivariate Data Processing**

The present work is concerned with aspects of modern biochemical engineering using the example of production processes for potential recombinant malaria vaccines.

Here, two quasi-continuous processes were constructed out of conventional batch-to-batch unit operations for process intensification, where the application of PAT - *Process Analytical Technology* is at the forefront.

The so conducted process developments serve in particular to meet the constraints of a QbD – *Quality by Design*, laid down by the FDA – *US Food and Drug Administration*.

In the area of integrated processing of recombinant pharmaceutical proteins initially a laboratory process including a cyclic *cell breeding* with subsequent *protein expression* involving an Insitu EBA – *Expanded Bed Adsorption* Chromatography was tested with a sequential/integrated strategy.

The two-stage EBA process was developed in accordance with QbD guidelines. Based on determining the CQA – *Critical Quality Attributes* and their associated operation parameters, the CPP – *Critical Process Parameters*, with their optimal process conditions and *Design Space*, have been identified using DoE – *Design of Experiments*.

In a more industrially oriented field an integrated pharmaceutical process was developed from a scale-down production facility formerly designed for batch operation. The process includes a *cell breeding*, a *protein production* and a downstream line with *cell clarification*, *micro-*, *ultra-* and *dia-crossflow filtration* followed by a subsequent *purification*.

A sequential, but also spread over three days in parallel, process strategy resulted in a quasi-continuous procedure with cyclical repetition of the 56-hour process, which was harvested daily over a number of days.

However, the main focus of this work was in the field of MVDA - *Multivariate Data Analysis* applications for monitoring and evaluation of the cyclical process flow and its reproducibility.

In the scope of PCA – *Principal Component Analysis* the method of process-monitoring with a *Golden Batch* tunnel through to the online evaluation of the process evolution has been tested extensively.

With the *Golden Batch* approach and the algorithm IBR - *Imputation By Regression* methods for the prediction of future process conditions were implemented and even more a MPMC – *Model Predictive Multivariate Control* for controlling the quality of real processes was achieved.

Moreover, PLS - *Partial Least Squares* regression was used for the prediction of media components like cell mass, glycerol and recombinant product and their corresponding cell-specific reaction rates from conventional online measurements.

Keywords: Quasi-continuous processing, *Pichia pastoris*, Malaria-Vaccine candidates, parallel/sequential process strategy, Multivariate Data Analysis, Golden Batch monitoring, model predictive multivariate quality control, QbD-conform process development