

# Fortschritt-Berichte VDI

VDI

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 Schepers

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 Malariavakzine. 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 Prozesses 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 Analysentechnik 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 Bioingenieur. Gemeinsam haben wir das Potential dieser Arbeit erkannt und stetig auf das Ziel – eine erfolgreiche Promotion – hingearbeitet. Trotz der damit verbundenen Strapazen ist es uns dabei gelungen nicht den Spaß an der Schuftelei 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	Konsekutive Pre-Optimierung mit dem ÄKTA™purifier 100 .....	46
5.7	Insitu-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 Sartocon Slice MF-Modulen .....	70
6.5.3	Ultrafiltration mit Sartocon Slice UF-Modulen .....	71
6.5.4	Chromatographische Produktreinigung mit einem AEKTA™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 $q_{IX}$ -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 <sub>IK</sub>	:= UV-Absorption der Messung von Komponente I in Teilsystem K	[AU]
a <sub>i</sub>	:= Koeffizient des physikalischen DoE-Modells	[xx]
A <sub>P<sup>IM</sup></sub>	:= UV-Absorption des Zielproteins P1 in der Medienphase	[AU]
B <sub>jKi</sub>	:= relatives Flächenmaß einer Bande j in der SDS-PAGE i aus K	[ $\cdot$ ]
b <sub>i</sub>	:= Koeffizient des codierten DoE-Modells	[xx]
c <sub>IK</sub>	:= Konzentration von Komponente I in Teilsystem K	[g l <sup>-1</sup> ]
C <sub>IK</sub>	:= 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> ]
$F_{nI}$	:= Begasungsrate der Komponente I unter Normbedingungen	[l h <sup>-1</sup> ]
$g_{Pi/X}$	:= zellspezifischer Gehalt an Produkt P <sub>i</sub>	[gg <sup>-1</sup> ]
$IA_{P1Eout}$	:= Integral der UV-Absorption von Produkt P <sub>1</sub> 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	[g mol <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 <sub>K</sub>	:= 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 <sub>1</sub> 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 <sub>1</sub> im Teilsystem K	[·]
RQ	:= molarer Respirationsquotient	[mol mol <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	[·]
$y_{IJ}$	:= 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	[m Scm <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-Faktionensammler	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	:= Biofeuchtpause

### Komponenten

Ac	:= Säure (acid)	P1	:= Produkt 1 (Zielprodukt)
AF	:= Antischaum (anti foam)	P2	:= Produkt 2 (Alkoholoxidase)
AIR	:= Luft (air)	Ptot	:= Gesamtprotein
AI	:= 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	:= ultraviolet
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
$p_{\text{lof}}$	:= 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
$y_r$	:= 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 Malariavakzine.

Dabei wurden zwei quasi-kontinuierliche Prozesse aus herkömmlichen *Batch-Unit Operations* 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 ihren 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 Prozesses 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

## 1. Einleitung und Zielsetzung

### 1.1 Paradigmenwechsel in der pharmazeutischen Produktion

Steigende Kosten im Gesundheitswesen und bei der Entwicklung von pharmazeutischen Produkten stehen fallende Rückvergütungen bei den Großen Pharmaunternehmen (*big pharma*) durch ein entstehendes *Patent Cliff* sowie eine allgemeine wirtschaftliche Rezession gegenüber. Dies lässt die Anforderungen an die Effizienz bei der Entwicklung von Prozessen für die Herstellung neuer Medikamente ansteigen [Au, 2014]. Zusätzlich zu den wirtschaftlichen Aspekten lastet ein hoher regulatorischer Druck durch Behörden (insbesondere der *European Medicines Agency* (EMA) und der *US Food and Drug Administration* (FDA)) auf den Unternehmen [Aksu et al., 2012].

In der Vergangenheit wurden die regulatorischen Rahmenbedingungen von der Industrie kritisiert. Diese seien durch viele Vorgaben bei einer Implementierung von modernen Herstellungsmethoden, die eine stetige Prozessoptimierung und damit Effizienzsteigerung ermöglichen könnten, hinderlich.

Begründet liegt dies zum Teil in dem hohen bürokratischen Aufwand, der bei der Dokumentation von Änderungen erforderlich ist, selbst wenn die Modifikationen von den Unternehmen als unbedeutend eingestuft werden. Als Ergebnis dessen waren pharmazeutische Firmen bisher zögerlich bei der Umsetzung von Erneuerungen an bestehenden Herstellungsverfahren [Bondi und Drennen, 2011].

Um einer Stagnation bei der Umsetzung von Innovationen entgegenzuwirken und die Freigabeprozedur für neue Medikamente zu vereinfachen, wurden im Jahr 2004 von der FDA die Richtlinien *Pharmaceutical cGMPs<sup>4</sup> for the 21st Century: A Risk-Based Approach* [FDA, 2004a] und *PAT – A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance* [FDA, 2004b] veröffentlicht.

Beide Dokumente beschreiben einen *Quality by Design* (QbD) Ansatz, bei dem die Qualität nicht erst nach Beendigung eines biotechnologischen Produktionsprozesses durch aufwendige Analyseverfahren am Endprodukt festgestellt wird. Stattdessen soll diese nunmehr durch eine wissensbasierte Entwicklung des Prozesses bis hin zum routinemäßigen Produktionsablauf (Einsatz gleichbleibender Materialien und Prozessverläufe) gewährleistet werden.

Dies bedingt den vermehrten Einsatz von *Process Analytical Technology* (PAT), d.h. von systemtheoretischen und automatisierungstechnischen Methoden in allen Prozessschritten. Ein Leitsatz dieser Initiative lautet: *Quality cannot be tested into products; it should be built-in or should be by design* [FDA, 2004b].

Diese Richtlinien haben einen eindeutigen Paradigmenwechsel in der biotechnologischen Entwicklung und Produktion, weg von statisch-endproduktorientierten Qualitätsmerkmalen hin zu dynamisch-prozessorientiert definierten Qualitätskriterien, herbeigeführt [Trautmann, 2011]. Dies wurde durch die Bekanntgabe von drei Produktionsrichtlinien der *International Conference on Harmonisation* (ICH), *ICH Q8(R2)*, *Pharmaceutical Development* [ICH, 2009a], *ICH Q9, Quality Risk Management* [ICH, 2009b] und *ICH Q10, Pharmaceutical Quality Systems* [ICH, 2009c], noch verstärkt.

Die Unternehmen sollen dadurch ermutigt werden, innovative Technologien in der pharmazeutischen Entwicklung, Herstellung und Qualitätssicherung zu implementieren. Das erklärte

<sup>4</sup> current Good Manufacturing Practice

Ziel besteht darin, gut charakterisierte Herstellungsprozesse, die den Richtlinien eines risikobasierten Ansatzes folgen, weniger zu überprüfen als traditionelle Prozesse [Bondi und Drennen, 2011].

Das Resultat einer vollständigen *Compliance* mit den Prinzipien von PAT und QbD soll der sogenannte *Real Time Release* (RTR) des Produktes sein, bei dem nach Beendigung eines Herstellungsverfahrens keine zusätzlichen Daten zu Sicherheit und Wirksamkeit von den Behörden eingefordert werden müssen, da die Qualität des Produktes bereits während seiner Herstellung sichergestellt ist [Pestieau et al., 2014].

## 1.2 Ansätze einer integrierten kontinuierlichen pharmazeutischen Produktion

Aus heutiger Sicht läuft nach wie vor ein Großteil aller Verfahren im Bereich der Bioprozesstechnik – sowohl im Upstream als auch im Downstream – im Batchbetrieb. Um in Zukunft die Kosten zu reduzieren und die Produktionskapazität zu steigern, wurden kleine, kontinuierlich betriebene Prozesse mit einer Integration und Intensivierung der Prozessschritte als Gegenentwurf zu den aktuellen Batch-Prozessen vorgeschlagen [Konstantinov and Cooney, 2014].

Langer und Rader fassen die Vorteile, die sich aus einer kontinuierlichen Produktion ergeben könnten, wie folgt zusammen (frei übersetzt aus [Langer and Rader, 2014a]):

- **Reduzierte Kosten:** Kontinuierliches Arbeiten erlaubt die Verwendung von maßstäblich kleineren Geräten, z.B. Bioreaktoren mit geringem Volumen, die durch längere Betriebszeiten ebenso viel Produkt erzeugen, wie wesentlich größere Reaktoren im Fed-Batch-Betrieb. Kleinere Gerätschaften haben einen geringeren Anschaffungswert und verursachen einen geringeren *Footprint* – sie benötigen weniger Platz. Zusätzlich ergibt sich ein kostengünstiger Materialeinsatz, insbesondere wenn *Single-Use-Material* verwendet wird.
- **Gesteigerte Produktivität:** Da die meisten Anlagenteile kontinuierlich betrieben werden, gibt es weniger Bedarf an Lager- und Transferbehältern und keine Haltepunkte zwischen den Teilprozessen. Diese Prozesse tendieren eher dazu, reibungslos abzulaufen. Höhere Zeldichten können erreicht werden, was zu höherer Produktausbeute und Produktkonzentration führt. Außerdem wird die Anzahl der benötigten Mitarbeiter verringert und ihre Arbeit ist körperlich weniger belastend.
- **Verbesserte Qualität:** Biologische Moleküle werden in der Natur kontinuierlich hergestellt. Im Vergleich zu einer Batch-Kultur neigt eine kontinuierliche Kultur dazu, besser kontrollierbar und weniger intensiv und anstrengend zu sein – einschließlich geringerer Scherung und konstant gehaltenem Nährstoffgehalt im Medium. Die Produktvariabilität, z.B. durch Verlust der Lebensfähigkeit der Zellen in späteren Kulturstufen oder eine veränderte Glykosylierung, wird verringert und mit kontinuierlicher Prozessführung naturgemäß gleichbleibend und robust. Probleme, die mit proteolytischen oder anderen zeitlichen Degradationen im Bioreaktor verbunden sind, können minimiert werden. Bei auftretenden Problemen muss nur ein Teil, aber nicht die gesamte Charge verworfen werden.
- **Erhöhte Flexibilität:** Kontinuierliche Produktionen besitzen eine bessere Anpassungsfähigkeit und effizientere Raumnutzung, sie bieten ähnliche Vorteile wie die Verwendung von Single-Use-Geräten. Eine biotechnologische Produktion wird mobiler und Einrichtungen besser kopierbar. In Verbindung mit dem Trend zu modular aufgebauten Bioprozess-Systemen mit mehreren kleinen, kontinuierlichen Bioprozesslinien an unterschiedlichen Standorten weltweit, wird erwartet, dass dieser Ansatz zunehmend in der kommerziellen industriellen Fertigung Einzug erhält.

Zum jetzigen Zeitpunkt einer Prozessentwicklung scheint das Konzept eines vollständig kontinuierlich ablaufenden biotechnologischen Prozesses allerdings noch unvereinbar mit einer Vielzahl von State-of-the-Art Technologien und biologischen Anforderungen zu sein.

Im Upstream-Bereich sind Perfusions- und Chemostat-Kultivierungen für bestimmte Zelllinien und auch Mikroorganismen erprobt, die Downstream Operationen bilden aber ein erhebliches Bottle-Neck für einen kontinuierlichen Betrieb [Langer and Rader, 2014b]. Hier sind z.B. kontinuierliche Chromatographieverfahren wie *Simulated Moving Bed* (SMB) [Angarita et al., 2015] oder *Periodic Counter-Current* (PCC) Chromatographie [Mahajan et al., 2012] noch nicht bis zur Marktreife hin entwickelt.

Die geringen Ernteströme – ob zellfrei (Perfusion) oder zellbehaftet (Chemostat) – machen einen adäquaten Betrieb einer für den Batch ausgelegten Downstreamoperation impraktikabel. So hat zum Beispiel die Fa. *DSM Biologics* auf der *Integrated Continuous Biomanufacturing* Konferenz 2013 in Barcelona ihren Ansatz für einen integrierten kontinuierlichen Bioprozess *Continuous XD®* vorgestellt, bei dem eine im Chemostat betriebene Zellkultur zur Herstellung eines monoklonalen Antikörpers mit einer *Expanded Bed Adsorption* (EBA) Chromatographie gekoppelt ist. Der Erntestrom musste jedoch über mehrere Tage in einem Haltegefäß gesammelt werden, um anschließend ein ausreichendes Volumen für den Betrieb der EBA bereit zu stellen [Zijlstra, 2013].

Folglich könnte durch eine intelligente Vernetzung von kleineren Grundoperationen, die für den Batchbetrieb entwickelt und ausgelegt sind, eine Übergangslösung hin zu einem quasi-kontinuierlichen Betrieb geschaffen werden. Dieser Ansatz ist sowohl in einer KMU<sup>5</sup>-Umgebung als auch bei großen cGMP-Lohnherstellern mit geringen Investitionen umsetzbar.

### 1.3 Zielsetzung der Arbeit

Im Labor für Bioprozessautomatisierung der Hochschule für Angewandte Wissenschaften Hamburg wird schon seit geraumer Zeit unter der Leitung von Prof. Dr.-Ing. Reiner Luttmann ideenreich an der Entwicklung von Integrierten Bioprozessen geforscht. Dabei sind die Produkte von unterschiedlicher Natur. Je nach Anwendung kann der Upstream- und/oder Downstream-Sektor dabei aus mehreren Grundoperationen zusammengesetzt sein [Cornelissen et al., 2003; Martens et al., 2011; Luttmann et al., 2015].

Bild 1.1 zeigt zwei Ansätze für Integrierte Bioprozesse, die in Verbindung mit einer sequentiell ausgelegten Kultivierung in quasi-kontinuierlicher Fahrweise betrieben werden.

Der erste Prozess besteht im Upstream aus einem Kultivierungsschritt mit zyklischer Prozesswiederholung. Der integrierte Downstream startet mit einem Konditionierungsschritt, gefolgt von einer EBA-Chromatographie. Diese fasst die herkömmlichen Prozessschritte Klärung, Produkt-Capture und Produktkonzentrierung in einer Grundoperation zusammen, wobei hier eine *Immobilized Metalion Affinity Chromatography* (IMAC)-Methode eingesetzt wird. Die EBA eignet sich somit für ein *In Situ Product Removal* (ISPR) [Van Hecke et al., 2014]. Die Aufarbeitung endet mit einem Schritt zur Feinreinigung des Produktes mit Hilfe einer *Size Exclusion Chromatographie* (SEC).

Der zweite Prozess besteht im Upstream aus einer zweistufigen Kultivierung, die in ein zirkuläres *cell breeding* und eine straight forward *protein production* aufgeteilt ist, wobei jeder Schritt in einem eigenen Bioreaktor erfolgt. Der erste Schritt des vierstufigen Downstreams ist die *broth clarification* mit einem Tellerseparatoren. Auf diese folgen zwei Filtrationsschritte,

<sup>5</sup> Kleine und mittlere Unternehmen

eine *microfiltration* (MF) und eine *ultrafiltration* (UF), die jeweils im *cross-flow* (CF) betrieben werden. Am Ende dieses Prozesses steht mit der *purification* eine chromatographische Reinigung, ebenfalls mit einer IMAC.

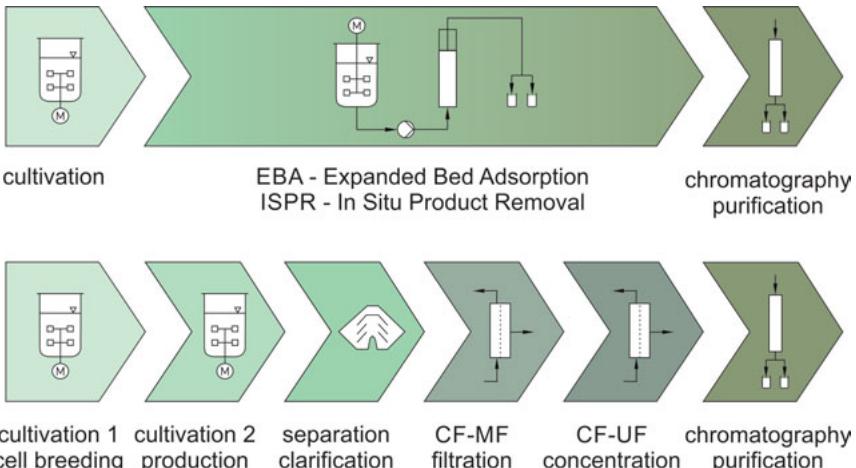


Bild 1.1: Zwei Ansätze für den Aufbau der Grundoperationen eines quasi-kontinuierlich betriebenen Integrierten Bioprozesses

Als pharmazeutisch relevante Produktionsprozesse wurden in diesen Anlagen die potentiellen Malariavakzine D1M1 und D1M1H [Faber et al., 2013] mit *Pichia pastoris* hergestellt. Dabei erfolgte im Umfeld der Integrierten Produktionsprozesse auch ein erheblicher Forschungsaufwand hinsichtlich der Prozesscharakterisierung und Prozessoptimierung.

Im Vorhinein sind z.B. in einer vollautomatischen Multibioreaktoranlage optimale Produktionsparameter über *Design of Experiments* (DoE) [Fricke, 2015] und in einer sequentiellen Einreaktorstrategie die reaktionskinetischen Zusammenhänge eines rekombinanten D1M1H-Prozesses, ebenfalls über DoE, ermittelt worden [Martens, 2014].

In dieser Arbeit werden die Anlagen um eine komplexe Datenerfassungs- und Datenverarbeitungssoftware ergänzt, um sequentiell/parallele integrierte Produktionsabläufe mit Hilfe der Multivariaten Datenanalyse (MVDA) zu begleiten.

In der Bioreaktor-EBA Anlage steht dabei die Optimierung der Downstream-Schritte mit einer mehrstufigen DoE-Strategie im Vordergrund.

In der integrierten, industriell-orientierten Produktionsanlage werden hingegen Methoden der *Principal Component Analysis* (PCA) zur Bewertung der Reproduzierbarkeit des zyklischen Prozesses über ein Offline- und Online-*Golden Batch Monitoring* bis hin zur *Model Predictive Multivariate Control* (MPMC) der ersten Hauptkomponente intensiv erprobt.

Die *Partial Least Squares Regression* (PLS) wird hingegen angewendet, um aus klassischen Online-Daten die schwer online messbaren Medienkomponenten Zelldichte, Glycerolkonzentration und sekretiertes Zielprotein über MVDA abzuschätzen.

Daraus lassen sich, ebenfalls über MVDA, die Verläufe der zellspezifischen Reaktionsraten und deren charakteristische Parameter offline und auch online ermitteln.

## 2. Zielprodukte, Hostsystem und Zielprodukt-Analytik

### 2.1 Potentielle Malariaimpfstoffe D1M1 und D1M1H

Etwa die Hälfte der Weltbevölkerung ist dem gesundheitlichen Risiko von Malaria-Infektionen, die durch Parasiten der Gattung *Plasmodium* verursacht werden, ausgesetzt. Der Erreger wird von Mensch zu Mensch durch die weibliche Anopheles-Mücke übertragen. Derzeit wird versucht, die Ausbreitung der Krankheit in den betroffenen Gebieten durch die Verteilung von Moskitonetzen und die Entwicklung neuer Insektizide zu minimieren. Ein wirk-samer Impfstoff könnte dabei helfen, einen entscheidenden Schritt gegen das Malariaproblem zu machen [<http://www.who.int>].

Trotz eines über Jahrzehnte andauernden intensiven Aufwands in der Forschung und Ent-wicklung ist noch kein Malariaimpfstoff kommerziell erhältlich. Lediglich für den vielverspre-chenden Impfstoffkandidaten RTS,S/AS01 der Fa. Glaxo Smith Kline befinden sich zur Zeit Unterlagen in einem Zulassungsverfahren bei der europäischen Arzneimittel-Agentur (EMA) [S Clinical Trials Partnership, 2015]. Ein Impfstoff allein kann allerdings niemals die Vielzahl der Erreger und Ihre natürlich vorkommenden Allele abdecken.

Am *Biomedical Primate Research Center* (BPRC) in den Niederlanden wurden ebenfalls Impfstoffkandidaten entwickelt, die den Erreger *Plasmodium falciparum*, der die besonders gefährliche Form *Malaria tropica* hervorruft, in seiner *merozoitischen Form (blood stage)* am Eindringen in die roten Blutkörperchen des Menschen hindern sollen. Dadurch wäre der Vermehrungszyklus des Parasiten (siehe: <http://www.cdc.gov>) an dieser Stelle unterbrochen.

Aus diesem Forschungsansatz sind verschiedene *Diversity Covering* (DiCo) Fusionsproteine [Remarque et al., 2008] aus genetisch veränderten Anteilen der Oberflächenantigene *apical membrane antigene 1* (AMA1) und *merozoite surface protein 1* (MSP1) [Tolle et al., 1995] entstanden.

Das Ziel dieser Vakzinkandidaten liegt darin, die stark polymorphe Struktur der Oberflächen-proteine mit einer statistischen Wahrscheinlichkeit von über 95 % abzubilden, so dass eine breite Immunantwort gegen die eindringenden Plasmodien durch entsprechende Antikörper entstehen kann.

Die modifizierten AMA1-Versionen werden mit D1, D2 und D3 bezeichnet, während die MSP1-Untereinheiten mit M1 und M2 gekennzeichnet sind [Faber et al., 2007].

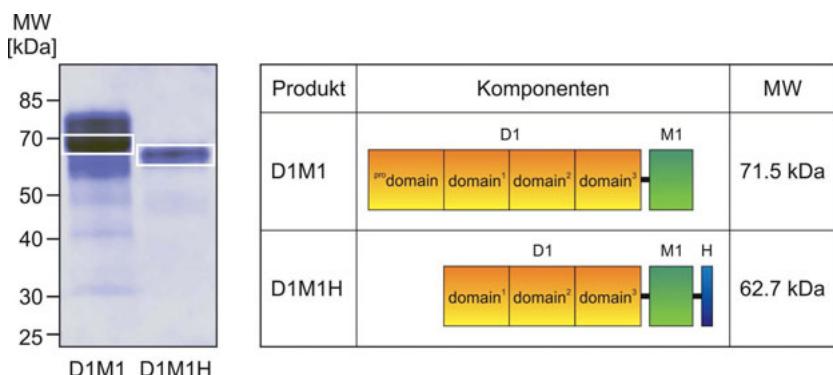


Bild 2.1: Charakteristika der eingesetzten Zielprodukte D1M1 und D1M1H

In dieser Arbeit wurden das 71,5 kDa schwere D1M1, bestehend aus einem AMA1 (D1) mit einer histidinreichen Prodomäne und einem MSP1<sub>19</sub> (M1), sowie das 62,7 kDa schwere D1M1H, bestehend aus der D1M1-Struktur ohne Prodomäne, aber dafür mit einem Histidin-Tag (H) für eine bessere Aufreinigung versehen, als Zielproteine verwendet [Faber et al., 2013]. Bild 2.1 veranschaulicht die charakteristischen Merkmale beider Produkte.

Für beide bietet sich eine Metalchelatchromatographie als Aufarbeitungsmethode an. Eine detaillierte Beschreibung dieser Proteine und ihrer physikochemischen Eigenschaften findet sich bei Martens [Martens, 2014].

## 2.2 Das Hostsystem *Pichia pastoris*

Bei dem eingesetzten Expressionssystem *Pichia pastoris* handelt es sich um eine fakultativ methylotrophe Hefe. Diese Zelle wird für die Herstellung einer Vielzahl von rekombinannten Proteinen erfolgreich eingesetzt [Cregg, 1999].

Neben Glycerol und Glucose ist *P. pastoris* auch in der Lage, mit Methanol als einziger Kohlenstoffquelle zu wachsen. Der erste Schritt in dem Methanolstoffwechsel ist die Oxidation von Methanol zu Formaldehyd. Die Reaktion wird durch das Enzym *Alkoholoxidase* (AOX) katalysiert, als Endprodukt entsteht dabei Wasserstoffperoxid. Aufgrund dessen hoher Toxizität findet dieser Vorgang in gesonderten Zellorganellen, den Peroxisomen, statt. Anschließend zerstört das Enzym *Katalase* das gebildete Wasserstoffperoxid zu Sauerstoff und Wasser [Ellis et al., 1985].

Es gibt zwei Genabschnitte in *P. pastoris*, die durch Methanol induziert werden und die genetische Information zur Herstellung der Enzyme AOX1 und AOX2 enthalten. Die Promotoren der beiden Genabschnitte sind unterschiedlich stark.

Der Promotor P<sub>AOX1</sub> ist der stärkere und für den überwiegenden Teil der Alkoholoxidase-Aktivität in der Zelle verantwortlich. Er unterliegt einer starken Repression durch Glycerol und wird über 1000-fach stärker induziert, wenn die Zellen mit Methanol als einziger Kohlenstoffquelle versorgt werden [Cereghino et al., 2002].

Dieser Induktionsmechanismus zusammen mit der starken Repression/Derepression wird zur Regulation der Zielproteinexpression genutzt. Dabei wird das *gene of interest* anstelle des AOX1-Genabschnitts in das Genom eingebracht. Die Transkription des Zielproteins wird dann bei einer Methanolzugabe über den Promotor P<sub>AOX1</sub> gesteuert.

Die Sequenz mit der genetischen Information des jeweiligen Fusionsproteins wurde mittels PCR amplifiziert und in die EcoR1-Kpn1 Schnittstelle des pPICZαA-Vektors von *Escherichia coli* kloniert. Dieser Vektor wurde in DH5α-Zellen eingebracht und vermehrt.

Die isolierten Plasmide wurden dann zur Transformation von *P. pastoris* KM71H verwendet. Der Vektor enthält außer der Information des Proteins noch ein Zeocin-Resistenzgen, das ermöglicht, die erfolgreich transfizierten Zellen zu selektieren [www.invitrogen.com].

Die Sekretion des Zielproteins in das Medium ermöglicht eine α-Faktor-Signalsequenz. Diese stammt aus *Saccharomyces cerevisiae* und wird an das N-terminale Ende des Proteins angehängt [Cereghino and Cregg, 2000]. Nach der Sekretion wird dieser Faktor wieder vom Protein abgespalten. Das Protein liegt dann in gelöstem Zustand außerhalb der Zellen vor.

Zusammenfassend sind die großen Vorteile für eine Expression von heterologen Proteinen mit *Pichia pastoris* noch einmal genannt [Cereghino and Cregg, 2000]:

- die Zellen sind leicht genetisch manipulierbar und besitzen einen Promotor P<sub>AOX1</sub>, der hervorragend für die kontrollierte Expression von Fremdgenen geeignet ist, sowie einen Promotor P<sub>AOX2</sub>, dessen Gen weiter für die Herstellung von AOX2 genutzt wird. Somit werden *Methanol Utilizing Slow* (MUT<sup>S</sup>) Stämme eingesetzt
- der einzellige Organismus ist leicht bis zu hohen Zelldichten (*high cell density culture*) mit einer hohen intrazellulären Proteinexpression und entsprechender Sekretion zu kultivieren
- die Zellen besitzen die Fähigkeit zu höheren eukaryotischen Proteinmodifikationen, wie Glykosilierung, Disulfid-Brückenbildung und proteolytische Prozessierung
- Für die nachfolgende Aufarbeitung bieten sich erhebliche Vorteile gegenüber zellintern angereicherten oder membrangebundenen Proteinen an, da die Zellen nicht aufgeschlossen werden müssen. Dadurch entfällt zum einen der Schritt des mechanischen oder chemischen Aufschlusses und zum anderen wird die weitere Reinigung nicht durch Zeldebris oder freigesetzte DNA-Fragmente erschwert [Potvin et al., 2012]
- Darüber hinaus eröffnet *Pichia pastoris* somit die Möglichkeit, die nachfolgenden zyklischen Strategien bei der Entwicklung quasi-kontinuierlicher Produktionsprozesse umzusetzen.

## 2.3 Offline-Analytik der Zielprodukte

### 2.3.1 Ermittlung der Konzentration des sekretierten Gesamtproteins

Die Analyse von Prozessen zur Herstellung rekombinanter Proteine steht und fällt mit der Produktanalytik.

Für diese erfolgte zunächst offline eine Bestimmung der Konzentration an sekretiertem Protein c<sub>PtotK</sub> nach der Bradford Methode [Bradford, 1976]. Dabei wurden alle Proben eines untersuchten Teilsystems K, wie z.B. die Medienphase M des Bioreaktors, das Eluat einer Chromatographie P oder der Aline-Analysenstrom A, mit einem kommerziellen Bradford Kit (BioRad Laboratories, München) vermessen.

### 2.3.2 Bestimmung der Produktreinheit und der Konzentration des Zielproduktes

Die Konzentration des Zielproduktes c<sub>P1Ki</sub> in einer Probe i aus dem Teilsystem K zu dem Zeitpunkt t<sub>i</sub> wurde über die Produktreinheit P<sub>P1Ki</sub> aus den zugehörigen Banden einer *Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis* (SDS-PAGE) mit anschließender Coomassie-Färbung ermittelt.

Bild 2.2 zeigt das Bandenmuster einer aufgetragenen Probe i aus einer Herstellung des Malariavakzins D1M1. Es zeigt eine hohe Verunreinigung. Das gefärbte Gel wurde mit einem Imager und dem zugehörigen Quantity One® Programm (BioRad Laboratories) digital aufgenommen und anschließend mit der TotalLab Quant® Bildanalysesoftware densitometrisch ausgewertet.

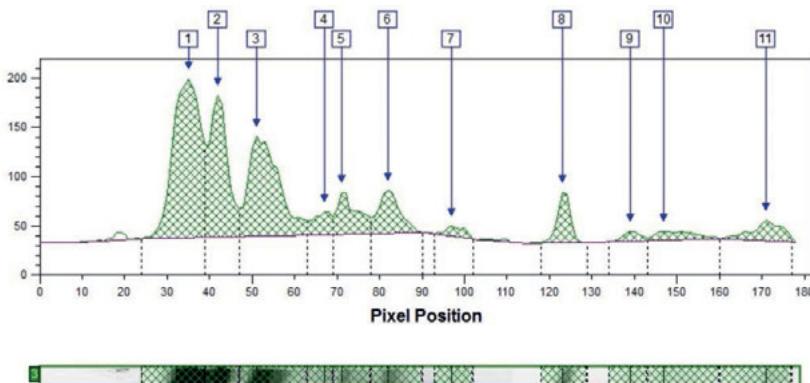


Bild 2.2: Analyse des Anteils spezifischer Proteinbanden am Gesamt muster einer Probe  $j$

Das Programm untersucht das Proteinpattern des Gels und detektiert dabei eine Anzahl  $b$  von Banden, die mit  $j = 1 \dots 11$  gekennzeichnet sind.

Die Bande 1 entspricht dem reinen Zielprodukt. Weitere Banden sind Produktbruchstücke oder andere Wirtsproteine und werden mit  $j \in [2, b]$  notiert.

Jeder Bande  $j$  wird entsprechend ihrer Höhe und Intensität ein Flächenmaß  $B_{jki}$  zugewiesen, das zur Probe  $i$  gehört.

Die Produktreinheit (*purity*)  $P_{P1Ki}$ ,

$$P_{P1K}(t_i) = P_{P1Ki} = \frac{B_{1Ki}}{\sum_{j=1}^b B_{jki}}, \quad (2.1)$$

mit

$P_{P1Ki}$  := Zielproduktreinheit in der Probe  $i$  aus Teilsystem  $K$  zum Zeitpunkt  $t_i$     [-]

$B_{jki}$  := relatives Flächenmaß einer Bande  $j$  in der zugehörigen SDS-PAGE  $i$     [-]

$B_{1Ki}$  := Produktbande in der SDS-PAGE  $i$     [-],

ergibt sich dabei aus dem Verhältnis der Zielproduktbande 1 zu allen Banden  $b$ .

Sie kann dann mit der Gesamtproteinkonzentration  $c_{\text{Plot}Ki}$  des Bradford-Assay zur Offline-Bestimmung der Konzentration des Zielproduktes  $c_{P1Ki}$ ,

$$c_{P1K}(t_i) = c_{P1Ki} = P_{P1Ki} \cdot c_{\text{Plot}Ki} \quad (2.2)$$

mit

$c_{\text{Plot}Ki}$  := sekretierte Konzentration an Gesamtprotein in Probe  $i$  aus  $K$     [ $\text{g l}^{-1}$ ],

verwendet werden.

Im weiteren Teil der Arbeit wird ebenfalls eine Atline-Bestimmung des Zielproduktes über eine *immobilized metal chelate affinity chromatography* (IMAC) HPLC-Methode eingesetzt und dort erläutert.

### 3. Grundoperationen eines integrierten Bioprozesses im Labormaßstab

#### 3.1 Prozesstechnische Zielsetzung

Im ersten Teil der Arbeit wird das in Kapitel 1 beschriebene Konzept einer Laboranlage zum *In situ Product Removal* (ISPR) umgesetzt.



Bild 3.1: Forschungsbioreaktor BIOSTAT ED5 mit Peripherie und EBA-Chromatographie

Bild 3.1 zeigt die dafür an der HAW Hamburg entwickelte Integrierte Pilotanlage. In der Bildmitte befindet sich der Bioreaktor BIOSTAT ED5 zur Proteinexpression mit seiner Versorgungseinheit und der *Frontend Control-Unit DCU* sowie seine erweiterte Vorlagenperipherie.

Rechts im Bild befinden sich eine HPLC zur Produktdetektion und die Prozessrechner zur Steuerung der Anlage über das Prozessleitsystem MFCS/win.

Auf der linken Seite befinden sich ein Erntetank mit Mess- und Steuereinheiten für die zellbehaftete Produktbrühe sowie eine eigenentwickelte *Expanded Bed Adsorption-Chromatographie* (EBA) zur Zellabtrennung und Produktvorreinigung.

Das zur Endreinigung des Produktes eingesetzte Chromatographie-System ÄKTA™purifier 100 steht rechts von den beiden Monitoren und wird später gezeigt.

In diesem Kapitel werden die bioverfahrenstechnischen Grundoperationen mit den zugehörigen gerätetechnischen Komponenten zum Aufbau eines integrierten Prozesses erläutert.

In dem nachfolgenden Kapitel 4 erfolgt dann die Beschreibung der entwickelten integrierten zyklischen Prozessführung und in Kapitel 5 die Optimierung des zweistufigen Aufreinigungsprozesses.

### 3.2 Die erweiterte Instrumentierung des Upstream-Prozesses

#### 3.2.1 Die eingesetzte Bioreaktoreinheit mit seiner MSR-Technik

Die erforderliche Bioreaktortechnik zur vollautomatischen Kultivierung von *Pichia pastoris* ist detailliert in mehreren Doktorarbeiten an der HAW Hamburg [Cornelissen, 2004; Martens, 2014; Ellert, 2015; Fricke, 2015; Lögering, 2015] beschrieben. Sie wird daher hier nur kurz vorgestellt.

Tabelle 3.1 zeigt die messtechnische Ausstattung des hier eingesetzten Bioreaktorsystems BIOSTAT® ED5 (B. Braun Biotech International, Melsungen). Geregelte Größen sind darin kursiv gekennzeichnet.

Tabelle 3.1: Ausstattung des Bioreaktorsystems BIOSTAT® ED5

Prozeßgröße	Bezeichnung	Ausstattung/Typ	Hersteller
Bioreaktor	BIOSTAT®	ED5	B. Braun Biotech
Maximales Arbeitsvolumen	V <sub>Lmax</sub>	5 l	
Minimales Arbeitsvolumen	V <sub>Lmin</sub>	3 l	
MSR-Einheit	DCU 2	Digital Control Unit 2	B. Braun Biotech
Ratio Höhe/Durchmesser	H:D	3:1	
Antischaumzugabe		Struktol J673	Schill & Seilacher
Laugezugabe		12,5 % Ammoniak	Carl Roth
Säurezugabe		2 M Phosphorsäure	Carl Roth
Rührer		3 Sechsstahl-Scheibenrührer	
Rührerdrehzahl	N <sub>St</sub>	Drehstrom, N <sub>Stmax</sub> = 1500 min <sup>-1</sup>	Georgii Kobold
Schaumdetektion	D <sub>F</sub>	Kapazitive Sonde (on/off)	B. Braun Biotech
Reaktionstemperatur	θ <sub>L</sub>	Pt100-Sonde, 0 - 150 °C	B. Braun Biotech
Doppelmanteltemperatur	θ <sub>DJ</sub>	Pt100-Sonde, 0 - 150 °C	B. Braun Biotech
pH-Wert	pH	405-DPAS-SC-K8S/120, pH 0 - 12	Mettler Toledo
Druck im Kopfraum	p <sub>G</sub>	Druckaufnehmer P13, 0 - 2 bar	Philips
Gelöstsauerstoffpegel	pO <sub>2</sub>	InPro ® 6800-Sonde, 0 - 100 %	Mettler Toledo
Trübung	E <sub>L</sub>	Trübungssonde AS16, 0 - 6 AU	optek-Danulat
Leitfähigkeit	σ <sub>L</sub>	CS60, 1 μS cm <sup>-1</sup> - 850 mS cm <sup>-1</sup>	optek-Danulat
Methanolkonzentration	Cs2M	Alcoline® Silicone Tubing Sonde	Biotech. Kempe
Begasungsrate Luft	F <sub>nAIR</sub>	Massflowcontroller, 0 - 10 l min <sup>-1</sup>	Bronkhorst
Begasungsrate O <sub>2</sub>	F <sub>nO2</sub>	Massflowcontroller, 0 - 10 l min <sup>-1</sup>	Bronkhorst
Begasungsrate N <sub>2</sub>	F <sub>nN2</sub>	Massflowcontroller, 0 - 10 l min <sup>-1</sup>	Bronkhorst
Abgasmolenbruch O <sub>2</sub>	x <sub>O2</sub>	Gas Sensor BPC-O <sub>2</sub> , 0 - 50 %	BlueSens
Abgasmolenbruch CO <sub>2</sub>	x <sub>CO2</sub>	Gas Sensor BPC-CO <sub>2</sub> , 0 - 25 %	BlueSens
Masse Reaktionsvolumen	m <sub>L</sub>	Waage 0 - 60 kg	Phillips
Masse Säurereservoir	m <sub>T1</sub>	Waage LP2200, 0 - 2,2 kg	Sartorius
Masse Laugereservoir	m <sub>T2</sub>	Waage LP6200, 0 - 6,2 kg	Sartorius
Antischaumreservoir	m <sub>AF</sub>	---	
Masse Glycerolreservoir	m <sub>R1</sub>	Waage LP6200, 0 - 6,2 kg	Sartorius
Masse Methanolreservoir	m <sub>R2</sub>	Waage QA7DCE-S, 0 - 7,5 kg	Sartorius
Masse Refreshreservoir	m <sub>R3</sub>	Waage QA7DCE-S, 0 - 7,5 kg	Sartorius
Zellfreier Analysenstrom	F <sub>A</sub>	Filtrationssonde ESIP-Lab 5442	Trace Analytics
Zielproduktkonzentration	C <sub>P1M</sub>	HPLC, IMAC-Methode	VWR International

In Bild 3.2 wurde die beschriebene Grundautomatisierung des Bioreaktors um eine erforderliche Infrastruktur von Waagen und Pumpen für automatisierbare zyklische Feed- und Harvest-Operationen erweitert.

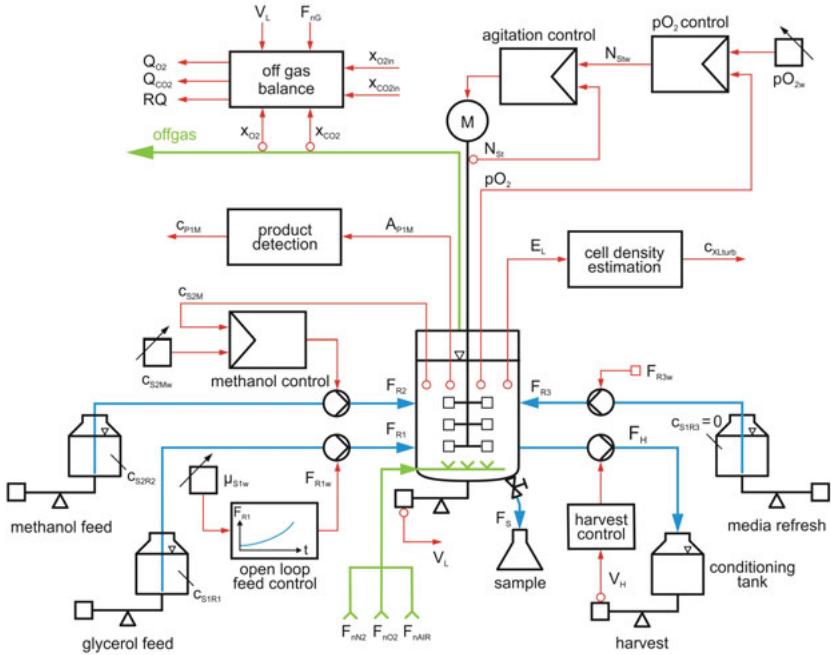


Bild 3.2: Erweiterte MSR-Technik für die Beobachtung einer Kultivierung von *Pichia pastoris*

Die Beobachtung des Reaktionsprozesses erfolgt über die Implementierung eines O<sub>2</sub>- und eines CO<sub>2</sub>-Abgassensors (BlueSens gas sensor GmbH, Herten), einer Leitfähigkeits- und einer Trübungssonde (beide optek-Danulat, Essen) sowie mit einer HPLC (VWR International, Darmstadt) zur Atline-Produktdetektion, die über eine *Eppendorf Sterilizable Inline Probe* (ESIP) eingebunden ist.

### 3.2.2 Bestimmung der Zeldichte über das Trübungssignal

Im BIOSTAT ED5 wird die Trübung  $E_L$  mit der Sonde AS16 sowie die Leitfähigkeit  $\sigma_L$  mit der Sonde CS16 (beide optek-Danulat, Essen) im Reaktionsraum (Flüssigphase L) inline gemessen.

In Bild 3.3 ist ein zyklischer Zellanzuchtprozess gezeigt, in dem der Reaktor alle 24 h teilweise abgeertet und dann wieder mit neuem Medium *refreshed* wurde. Das Volumen  $V_L$  bleibt dabei relativ konstant.

In den vier Zyklen wurden die Zellen jeweils über 7 h in einem Glycerol-Fed Batch angezüchtet und dann im Fed Batch auf Methanol zur Produktbildung induziert.

Das Trübungssignal  $E_L$  folgt zunächst dem erwarteten exponentiellen Wachstumsprozess, geht dann aber in eine Sättigung über.

Zur Abschätzung des Zelldichte  $c_{XL}$  muss daher  $c_{XL,turb}$ ,

$$c_{XL,turb}(t) = k_1 \cdot (e^{k_2 \cdot E_L(t)} - 1) \quad (3.1)$$

mit

$E_L$  := Signal der Trübungsmessung [AU]

$k_1$  := Zelldichteparameter [ $gl^{-1}$ ]

$k_2$  := Trübungsparameter [ $AU^{-1}$ ],

aus dem Inline-Trübungssignal  $E_L$  rekonstruiert werden.

Die Anpassung erfolgt über eine nichtlineare (exponentielle) Regression der aus der Biotrockenmasse (*cell dry weight*) ermittelten Zelldichte  $c_{XL,cdw}(t)$  mit zugeordneten Trübungsmessungen  $E_L(t)$ . Die genaue Vorgehensweise ist bei Kaiser [Kaiser, 2010] zu finden.

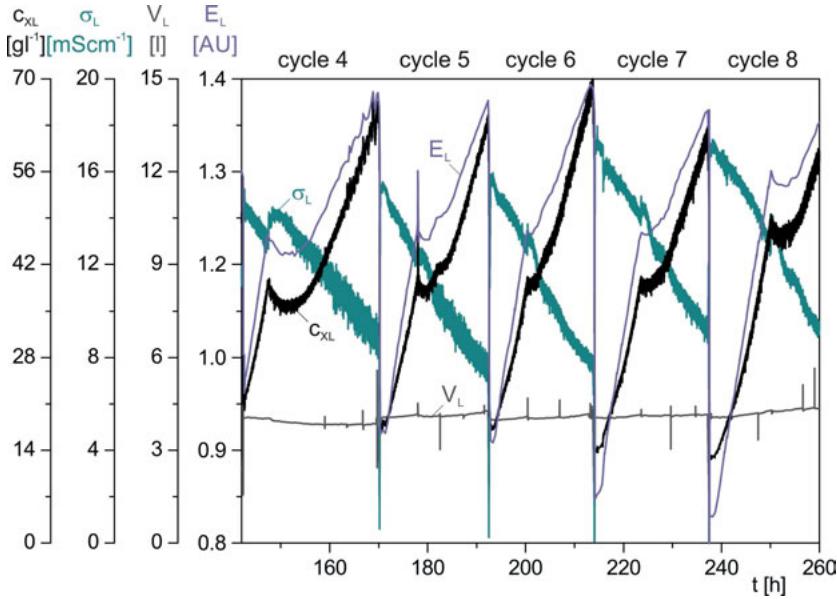


Bild 3.3: Verlauf von Leitfähigkeit  $\sigma_L$ , Volumen  $V_L$  sowie optischer Trübung  $E_L$  und daraus abgeschätzter Zelldichte  $c_{XL,turb}$

Es ist sinnvoll, für die Glycerol- und für die Methanolphase unterschiedliche Parametersätze für Gl. (3.1) im Prozessleitsystem zu hinterlegen. Die in dieser Arbeit gefundenen Parameter sind in Tabelle 3.2 zusammengefasst.

Tabelle 3.2: Ermittelte Parametersätze zur Online-Estimierung der Zelldichte

	Glycerol-Batch	Glycerol-Fed Batch	Methanol-Fed Batch
$k_1 [gl^{-1}]$	6.0873	1.8172	2.6049
$k_2 [AU^{-1}]$	1.3629	2.3510	2.1134

### 3.2.3 O<sub>2</sub>- und CO<sub>2</sub>-Gasbilanzen zur Online-Prozessbeobachtung

Die O<sub>2</sub>- und CO<sub>2</sub>-Gasmassenbilanzen bilden ein verlässliches Maß für eine Online-Bewertung des volumetrischen respiratorischen Reaktionsverhaltens. Sie werden dabei auf das Reaktionsvolumen bezogen.

Das Volumen der Flüssigphase V<sub>L</sub>,

$$V_L(t) = \frac{m_L(t)}{\rho_L(t)} \quad (3.2)$$

mit

$m_L$  := Masse der Flüssigphase

[g]

$\rho_L$  := Dichte der Flüssigphase

[g l<sup>-1</sup>],

lässt sich über die Kalibrierung der Reaktorwaage auf die Masse der Flüssigphase m<sub>L</sub> und bei Abschätzung der Dichte ρ<sub>L</sub> der Flüssigphase (hier Annahme ρ<sub>L</sub> = 1004 g l<sup>-1</sup>) online berechnen.

Mit Kenntnis der Gesamtbegasungsrate F<sub>nG</sub>,

$$F_{nG}(t) = F_{nAIR}(t) + F_{nO2}(t) + F_{nN2}(t) \quad (3.3)$$

mit

F<sub>nl</sub> := auf Normbedingungen bezogene Begasungsrate des Gases I,

I = AIR, O<sub>2</sub>, N<sub>2</sub>, [l h<sup>-1</sup>],

berechnet aus den Einzelmessungen, lassen sich sowohl der O<sub>2</sub>-Stoffmengenanteil x<sub>OGIN</sub>,

$$x_{OGin}(t) = \frac{F_{nO2}(t) + x_{OAIR} \cdot F_{nAIR}(t)}{F_{nG}(t)} \quad (3.4)$$

als auch der CO<sub>2</sub>-Stoffmengenanteil x<sub>CGIN</sub>,

$$x_{CGin}(t) = \frac{F_{nCO2}(t) + x_{CAIR} \cdot F_{nAIR}(t)}{F_{nG}(t)} \quad (3.5)$$

mit

x<sub>OAIR</sub> := Stoffmengenanteil für O<sub>2</sub> in Luft (x<sub>OAIR</sub> = 0.2094)

x<sub>CAIR</sub> := Stoffmengenanteil für CO<sub>2</sub> in Luft (x<sub>CAIR</sub> = 0.00033),

beide am Eingang des Reaktors, leicht bestimmen.

Für nachfolgende Gasbilanzen werden bei Kenntnis von F<sub>nG</sub> und V<sub>L</sub> mit der jeweils maximalen Angebotsrate Q<sub>Imax</sub>,

$$Q_{Imax}(t) = \frac{F_{nG}(t) \cdot M_I}{V_{nM} \cdot V_L(t)} \quad (3.6)$$

mit

F<sub>nG</sub> := Gesamtbegasungsrate unter Normbedingungen [l h<sup>-1</sup>]

V<sub>L</sub> := Reaktionsvolumen [l]

M<sub>I</sub> := molare Masse der Komponente I (M<sub>O2</sub> = 32 g mol<sup>-1</sup>, M<sub>CO2</sub> = 44 g mol<sup>-1</sup>)

V<sub>nM</sub> := molares Normvolumen eines Gases (V<sub>nM</sub> = 22,414 l mol<sup>-1</sup>),

zwei volumenbezogene Verfahrenskenngrößen mit Q<sub>O2max</sub> und Q<sub>CO2max</sub> festgelegt.

Werden nun die Abgasmolenbrüche  $x_{O_2}$  und  $x_{CO_2}$ , beide im Abgas in Prozent gemessen, in den O<sub>2</sub>-Stoffmengenanteil  $x_{OG}$ ,

$$x_{OG}(t) = \frac{x_{O_2}(t)}{100\%}, \quad (3.7)$$

und den CO<sub>2</sub>-Stoffmengenanteil  $x_{CG}$ ,

$$x_{CG}(t) = \frac{x_{CO_2}(t)}{100\%}, \quad (3.8)$$

umgerechnet, so sind die Sauerstoffeintragsrate  $Q_{O_2}$ ,

$$Q_{O_2}(t) = Q_{O_{2max}}(t) \cdot \frac{x_{OGin}(t) \cdot (1-x_{CG}(t)) - x_{OG}(t) \cdot (1-x_{CGin}(t))}{1-x_{OG}(t)-x_{CG}(t)}, \quad (3.9)$$

und analog dazu die volumetrische Kohlendioxidaustragsrate  $Q_{CO_2}$ ,

$$Q_{CO_2}(t) = Q_{CO_{2max}}(t) \cdot \frac{x_{CG}(t) \cdot (1-x_{OGin}(t)) - x_{CGin}(t) \cdot (1-x_{OG}(t))}{1-x_{OG}(t)-x_{CG}(t)}, \quad (3.10)$$

sowie der molare Respirationsquotient RQ,

$$RQ(t) = \frac{Q_{CO_2}(t) \cdot M_{O_2}}{Q_{O_2}(t) \cdot M_{CO_2}}, \quad (3.11)$$

leicht online zu berechnen.

In Bild 3.4 ist eine Kultivierungsstrategie mit den Variablen der Sauerstoffversorgung in drei sich wiederholenden Prozesszyklen gezeigt. Der Prozess wurde mit konstanter Luftbegleitung sowie einem über die Drehzahl  $N_{St}$  geregelten pO<sub>2</sub> gefahren.

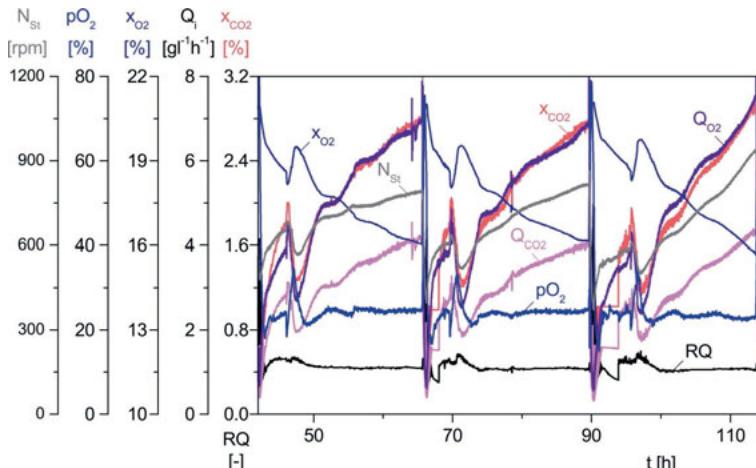


Bild 3.4: Variablen und Ergebnisse der Sauerstoffversorgung eines zyklischen Prozesses

Bild 3.3 und Bild 3.4 demonstrieren, dass man derartige Verläufe zur Bewertung der Reproduzierbarkeit von zyklischen Prozessen heranziehen kann. Diese erfolgt in einem späteren Kapitel.

### 3.3 Zellabtrennung und Proteinvorreinigung

#### 3.3.1 Grundlagen der Expanded Bed Adsorptions-Chromatografie

Bild 3.5 zeigt zwei unterschiedliche Methoden, wie ein extrazellulär vorliegendes Produkt aus einer Kulturbrühe aufgereinigt werden kann. Die Aufarbeitungsstrategien gliedern sich dabei in die Bereiche *capture* und *high-resolution techniques*, wobei für den *capture* zwischen zwei Alternativen zu wählen ist.

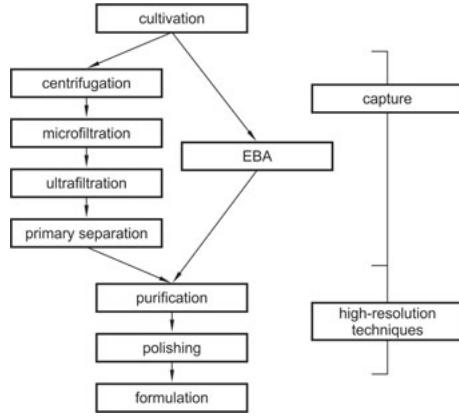


Bild 3.5: Aufarbeitung eines extrazellulären Proteins [Anspach et al., 1999]

Jeder Schritt einer Aufreinigungsstrategie bringt Proteinverluste mit sich, die sich in ihrer Folge multiplizieren. Ziel ist es daher, die Anzahl der Schritte auf ein Minimum zu reduzieren, um die Proteinverluste gering zu halten und Zeit und Kosten einzusparen [Anspach et al., 1999].

In dem linken Pfad folgt auf die *cultivation* eine *centrifugation*, um Zellen und produkthaltiges Medium voneinander zu trennen. Eine *microfiltration* dient zur Abtrennung von unlöslichen Bestandteilen, wie Zellbruchstücken, die bei der Zentrifugation noch nicht abgetrennt wurden, während eine *ultrafiltration* mit einem entsprechenden Ausschlussvolumen (*cut-off*) zum Pufferwechsel und zur Konzentrierung der Produktlösung herangezogen wird. Eine *primary separation* wird durch einen Adsorptionsprozess an einer Membran oder einem geeigneten Chromatographie-Träger realisiert.

Die *Expanded Bed Adsorption-Chromatographie* (EBA) ist ein integriertes Reinigungsverfahren. Sie ist im rechten Pfad gezeigt und fasst eine Zellabtrennung, eine Produktadsorption und eine Produktkonzentrierung in einer Operation *capture* zusammen. Hierfür werden sonst mehrere Einzelschritte und gegebenenfalls unterschiedliche Prozessanlagen benötigt.

Beide Pfade enden mit der Feinreinigung des Proteins, bestehend aus *purification* und *polishing*, die von der Produktkonfektionierung (*formulation*) gefolgt wird.

Wie die beschriebenen Prozessschritte des *capture* mit einer EBA-Operation zu realisieren sind, ist in Bild 3.6 schematisch dargestellt.

Die EBA-Säule besitzt durch einen beweglichen oberen Adapter ein variables Säulenvolume. Im Gegensatz zum gewöhnlichen Festbett befindet sich das Bett während der Probenaufgabe in einem expandierten Zustand. Die Expansion wird durch einen aufwärtigen Volumenstrom (*upflow*) durch die Säule erreicht.

Die Chromatographie erfolgt in der EBA an in der Flüssigkeit schwebenden *beads*. Der Kern dieser Beads muss eine höhere Dichte besitzen, als die Zellen und unlösliche Bestandteile im Medium. Er besteht deshalb aus Glas, Quarz oder Zirkon, damit die Beads nicht mit dem Volumenstrom aus der Säule gespült werden.

Der Kern ist von einer Agarose-Matrix umgeben. Diese trägt die funktionellen Gruppen zur Bindung des Produktes. Das Adsorbermaterial für die EBA besitzt eine definierte Verteilung der Partikelgröße. Die Durchmesser liegen zwischen 100 bis 300 µm und führen auf unterschiedliche Sedimentationsgeschwindigkeiten.

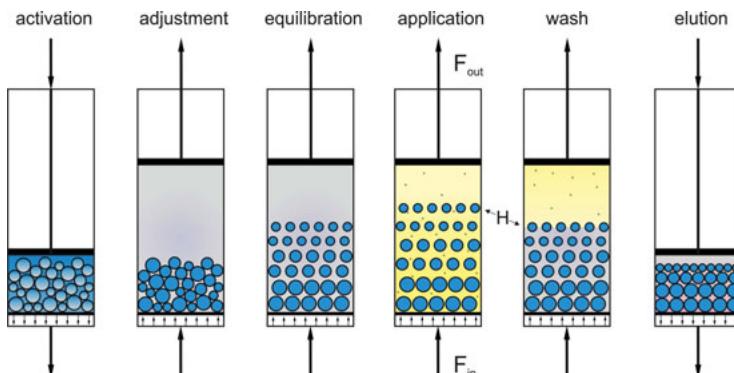


Bild 3.6: Schematischer Ablauf eines EBA-Zyklus

Für kleine Partikel sowie für solche in zähen Fluiden oder mit geringem Dichteunterschied  $\Delta\rho_E$ ,

$$\Delta\rho_E = \rho_B - \rho_F, \quad (3.12)$$

mit

$$\begin{aligned} \Delta\rho_E &:= \text{Dichteunterschied innerhalb der EBA} & [\text{kgm}^{-3}] \\ \rho_B &:= \text{Dichte der Beads B} & [\text{kgm}^{-3}] \\ \rho_F &:= \text{Dichte des Fluids F} & [\text{kgm}^{-3}], \end{aligned}$$

wie er in der EBA vorliegt, ergibt sich bei kleinen Reynolds-Zahlen ( $Re < 0,25$ ) eine Sedimentationsgeschwindigkeit  $v_s$ ,

$$v_s = \frac{\Delta\rho_E \cdot g \cdot d_B^2}{18 \cdot \eta_F}, \quad (3.13)$$

mit

$$\begin{aligned} v_s &:= \text{Sedimentationsgeschwindigkeit in der EBA} & [\text{ms}^{-1}] \\ g &:= \text{Gravitationsbeschleunigung} & [\text{ms}^{-2}] \\ d_B &:= \text{Durchmesser der Beads} & [\text{m}], \\ \eta_F &:= \text{Dynamische Viskosität des Fluids} & [\text{kgm}^{-1}\text{s}^{-1}], \end{aligned}$$

die mit der Stokes'schen Gleichung beschrieben ist [Peifer and Profos, 2008].

Für das verwendete Adsorbermaterial STREAMLINE Chelating ist die Dichteverteilung der Partikel, verglichen mit der Größenverteilung, uniform. Aus diesem Grund wird nur letztere betrachtet.

Die Sedimentationsgeschwindigkeit ändert sich mit dem Quadrat des Partikeldurchmessers. Beginnt das Bett zu expandieren, steigen kleinere Partikel somit schneller auf. Die Abstände zwischen den Partikeln am Boden der Säule sind geringer als im oberen Bereich und das Zwischenkornvolumen nimmt somit von unten nach oben hin zu.

Bild 3.6 zeigt die Verteilung der Partikel während der Expansion des Bettes. Der Volumenstrom  $F_{in}$  am Eintritt ist gleich  $F_{out}$  am Austritt und bleibt somit über die gesamte Säulenlänge konstant.

Mit der Zeit stellt sich ein Gleichgewicht zwischen Auftriebsgeschwindigkeit und Sinkgeschwindigkeit der Partikel ein. Diese befinden sich dann in der Schwebewelle und haben keine axiale (vertikale) Bewegung mehr. Ein stabiles expandiertes Bett stellt sich ein und es kommt zu keiner Vermischung der Partikel mehr.

Die Eigenschaften in diesem Zustand sind mit denen eines Festbettes vergleichbar. Es kommt nur zu einer geringen Horizontalbewegung und möglichst keiner Vertikalbewegung. Im Vergleich zu fluidisierten Betten, wie sie in der chemischen Reaktionstechnik verwendet werden, sind wesentlich geringere Rückvermischungen und Verwirbelungen zu beobachten.

Die Vergrößerung der Porosität des Packungsmaterials während der Expansion ermöglicht eine direkte Applikation von Biomasse, die ansonsten jede herkömmlich gepackte Säule verstopfen würde. Diese Besonderheit favorisiert ein EBA-System für die Entwicklung von Integrierten Bioprozessen, in denen das Zielprodukt extrazellulär anfällt.

### 3.3.2 Das Adsorbermaterial und die eingesetzte Säule

Für die EBA-Anlage wird eine STREAMLINE 50 Säule der Firma Pharmacia Biotech, heute GE-Healthcare, verwendet. Sie besteht aus einer durchsichtigen Borosilikatglasröhre mit einem Durchmesser von 5 cm und einer Höhe von 55 cm. Daraus ergibt sich ein maximales Volumen von 1080 cm<sup>3</sup>.

Die Säule ist mit einem beweglichen oberen Adapter ausgestattet. Durch diesen wird das Säulenvolumen in zwei variable Abschnitte getrennt. Im unteren Bereich befindet sich der Adsorber und damit der eigentliche Trennraum.

Um das Volumen des Trennbereichs zu ändern, muss im Kopfraum der Säule ein Über- oder Unterdruck erzeugt werden. Verwirklicht wird dies durch Pumpen einer Flüssigkeit in den Kopfraum hinein oder aus ihm heraus. Dadurch kann die Adapterhöhe je nach Prozessschritt wie Equilibrierung oder Elution angepasst werden.

Der untere Adapter ist so konstruiert, dass er eine gleichmäßige Anströmung der Adsorberpartikel über den gesamten Querschnitt der Säule ermöglicht. Der Fuß der Säule ist mit einer Libelle ausgestattet, die eine exakte Ausrichtung in vertikaler Richtung erlaubt. Hierzu sind die drei Standbeine der Anlage individuell höhenverstellbar. Eine schiefe Ausrichtung würde ein nicht gleichmäßig angeströmt Säulenbett zur Folge haben.

Beim Arbeiten mit der Säule darf der innere Säulendruck 1 bar nicht überschreiten. Gleichzeitig darf der Druck auf den oberen Adapter nicht über 3.5 bar liegen.

Das Säulenmaterial STREAMLINE Chelating wurde ebenfalls von der Firma GE Healthcare bezogen. Die Beads sind als IMAC-Medium, speziell für die EBA-Chromatographie, ausge-

legt. Sie sind chemisch sehr stabil und überstehen auch drastische Säuberungsprozeduren. Die Säule wurde mit 300 ml STREAMLINE Chelating bis zu einer Höhe von 15 cm gefüllt, um die vom Hersteller vorgeschlagene Mindesthöhe eines sedimentierten Bettens zu erfüllen.

In Tabelle 3.2 sind die technischen Spezifikationen der Beads zusammengefasst.

Tabelle 3.3: Technische Spezifikationen des verwendeten Adsorbermaterials von GE Healthcare

Composition	Macroporous cross-linked 6% agarose, containing a crystalline quartz core
Particle size	100 - 300 µm
Ligand	Imino diacetic acid (IDA)
Coupling chemistry	Epoxy
Mean particle density	~ 1.2 gml <sup>-1</sup>
Metal ion capacity	37 - 47 µmolml <sup>-1</sup>
pH stability (operational)	3 - 13

### 3.3.3 Aufbau und MSR-Technik der EBA-Anlage

Für die Entwicklung eines integrierten Produktionsprozesses wurde an der HAW Hamburg eine vollautomatisierte EBA-Anlage entwickelt. Diese ist im nachfolgenden Bild 3.7 gezeigt.



Bild 3.7: Entwickelte EBA-Anlage zur Einbindung in einen integrierten Prozess

Der Chromatographie ist der im Vordergrund stehende Konditionierungstank, der auf einer Industriewaage platziert ist, vorgeschaltet. Er besteht aus einem Chemap-Bioreaktor, der mit einer Edelstahlsäule mit Schaufenfenster erweitert ist, und von oben gerührt wird.

Dieser Tank ist mit einer MSR-Technik ausgerüstet und dient dem Produktionsreaktor als Erntegefäß. In diesem wird dann die geerntete Zellbrühe konditioniert. Hierbei werden die in Kapitel 5 ermittelten, kritischen Parameter der Produkt- und Zellbrühe eingestellt, um eine optimale EBA-Chromatographie zu ermöglichen.

Gewährleistet wird eine Messung des pH-Wertes  $pH_c$ , der Leitfähigkeit  $\sigma_c$  (beide Mettler Toledo, Gießen), der Temperatur  $\vartheta_c$  (B. Braun Biotech, Melsungen) und der Masse  $m_c$  (Sartorius, Göttingen).

Die EBA-STREAMLINE 50 ist in eine MSR-Technik eingebunden, die in das hintere Rack eingebaut ist. Die Verrohrung und die mess- und regelungstechnische Ausstattung ist dem RI-Fließbild in Bild 3.8 zu entnehmen.

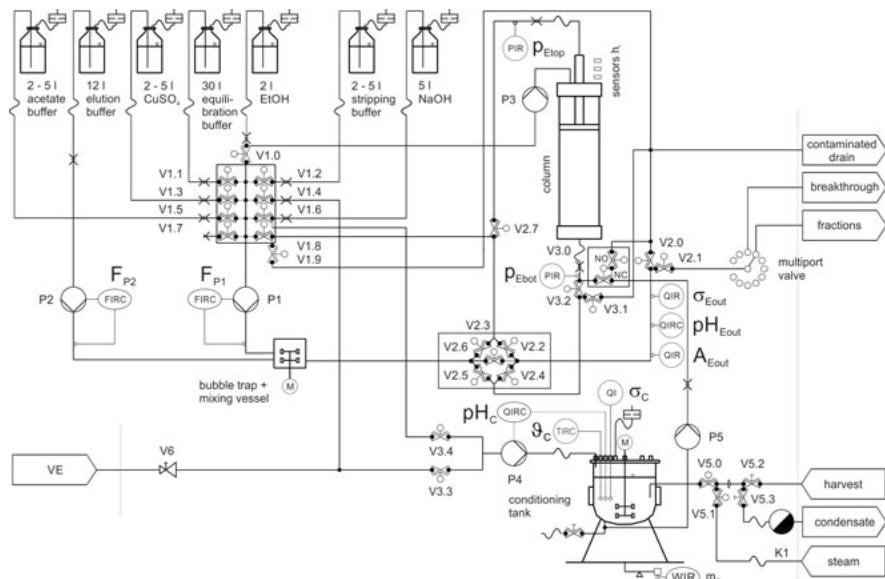


Bild 3.8: RI-Fließbild von der EBA und dem Konditionierungstank

Für den Betrieb der EBA-Anlage sind einige Kenngrößen wie Säulendruck oder Fließgeschwindigkeit unabdinglich, denn über diese ist ein einwandfreier und reproduzierbarer Betrieb möglich, ohne die Anlage zu schädigen.

Im Anschluss an die Pumpen  $P_1$  und  $P_2$  ist jeweils ein Flowmeter ( $F_{P1}$ ,  $F_{P2}$ ) X1000 (Krohne, Duisburg) integriert. Diese setzen eine Leitfähigkeit von mindestens  $0.02 \text{ mScm}^{-1}$  voraus. Somit ist eine Regelung von Fluiden wie Ethanol und VE-Wasser nicht möglich. Durch den Einsatz der Flowmeter zur Regelung der Pufferströme wird allerdings eine reproduzierbare Expansion des Bettens und der Gradientenbildung zur Elution gewährleistet.

Oberhalb der Säule sind drei Annäherungssensoren (Pepperl + Fuchs, Mannheim) in veränderlichen Höhen ( $h_1$ ,  $h_2$ ,  $h_3$ ) angebracht. Über diese wird die Position des beweglichen Stempels automatisch eingestellt. Die Sensoren reagieren, sobald ein Objekt mit einer höheren

Permeabilität als Luft in den Wirkungsbereich kommt. Hierzu ist am oberen Adapterkolben ein Aluminiumring befestigt.

Vor und hinter der Säule ist jeweils ein Drucktransmitter dTRANS p30 ( $p_{Ebot}$ ,  $p_{Etop}$ ) (Jumo, Fulda) integriert. Der Messbereich liegt bei 0 bis 1 bar. Mit Hilfe der Transmitter kann der maximale Belastungsdruck der Säule von 1 bar eingehalten werden.

Durch sich absetzende Biomasse am oberen Adapter kann es zu einem Druckanstieg kommen. In diesem Falle wird der Nutzer der Anlage bei Erreichen eines vorher definierten Drucklimits durch ein Warnsignal des Prozessleitsystems alarmiert, um die erforderlichen Korrekturen durchzuführen.

Nach Verlassen der Säule durchläuft der Volumenstrom eine UV-Messzelle, in der ein UV-Sensor AF 46 (optek-Danulat, Essen) integriert ist. Dieser verfügt über zwei Kanäle, die mit Filtern für die Wellenlängen von 254 nm und 280 nm besetzt sind.

Aktuell wird der größere optische Weg gewählt, um die relativ schwachen Signale  $A_{Eout}$  des Zielproteins besser analysieren zu können.

Im Anschluss an die UV-Messung ist eine weitere Messzelle (optek-Danulat), bestückt mit einer pH-Sonde (ABB, Mannheim) und einer Leitfähigkeitssonde CF60 (optek-Danulat), installiert. Die Messbereiche sind mit  $0.01 \mu\text{S/cm}^{-1}$  bis  $500 \text{ mS/cm}^{-1}$  für die Leitfähigkeit  $\sigma_{Eout}$  und mit 2 bis 12 für den pH-Wert  $pH_{Eout}$  angegeben.

Die Bauweise der Messzelle hat kaum Toträume und bietet ausgezeichnete Reinigungsmöglichkeiten.

In Tabelle 3.4 sind noch einmal die MSR-Größen von der EBA und dem Konditionierungstank zusammengefasst.

Tabelle 3.4: Messtechnische Ausstattung von EBA (E) und Konditionierungstank (C)

Prozessgröße	Bezeichnung	Ausstattung/Typ	Hersteller
Volumenstrom Equilibration	F <sub>B1</sub>	SC100A X1000	Altoflux
Volumenstrom Elution	F <sub>B2</sub>	SC100A X1000	Altoflux
pH-Wert Konditionierungstank	pH <sub>C</sub>	Sonde InPro 410, pH 2 - 12	Mettler Toledo
Leitfähigkeit Konditionierungstank	$\sigma_C$	SevenEasy, $0.01 \mu\text{S/cm}$ - $500 \text{ mS/cm}$	Mettler Toledo
Temperatur Konditionierungstank	$\vartheta_C$	Pt100-Sonde, $0$ - $150^\circ\text{C}$	B. Braun Biotech
Masse Konditionierungstank	m <sub>C</sub>	Waage Combics 30 CW1P, $0$ - $30 \text{ kg}$	Sartorius
pH-Wert EBA-Auslass	pH <sub>Eout</sub>	Adapter PF12 mit Sonde 405-DPAS	Mettler Toledo
Leitfähigkeit EBA-Auslass	$\sigma_{Eout}$	Sensor CF60, $1 \mu\text{S/cm}$ - $850 \text{ mS/cm}$	optek-Danulat
UV-Absorption EBA-Auslass	A <sub>Eout</sub>	UV Absorptionssensor AF46, 280 nm	optek-Danulat
Bodendruck der EBA-Säule	p <sub>Ebot</sub>	Umformer dTrans p30, $0$ - $1 \text{ bar}$	Jumo
Kopfdruck der EBA-Säule	p <sub>Etop</sub>	Umformer dTrans p30, $0$ - $1 \text{ bar}$	Jumo
Fließrate Pufferaufgabe	F <sub>P1</sub>	ISM444	Ismatec
Fließrate Elutionspuffer	F <sub>P2</sub>	FE 411	B. Braun Biotech
Fließrate Kolbenverstellung	F <sub>P3</sub>	520Du	Watson-Marlow
Fließrate Konditionierung	F <sub>P4</sub>	ISM597A	Ismatec
Fließrate Probenaufgabe	F <sub>P5</sub>	ISM444	Ismatec

### 3.3.4 Ablauf einer Aufarbeitungsprozedur

Vor Beginn der eigentlichen Aufarbeitungsprozedur wird das verwendete STREAMLINE Chelating Material mit Cu<sup>2+</sup>-Lösung im *downflow* aktiviert. Dabei befindet sich der bewegliche obere Adapter der Säule in einer Position direkt oberhalb des sedimentierten Bettes.

Bild 3.9 zeigt, wie sich das Material beim Durchströmen durch die Verbindung mit den Kupfer-ionen sukzessive blau einfärbt. Nachfolgend wird in der gleichen Strömungsrichtung mit NaAc-Lösung nachgespült, um schwach- sowie ungebundene Metallionen wieder aus der Säule heraus zu waschen.

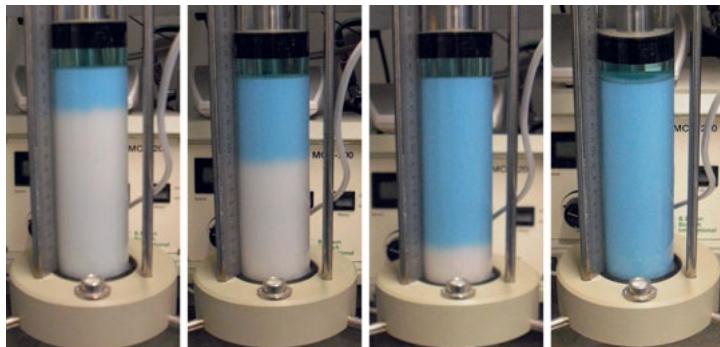


Bild 3.9: Aktivieren des Säulenmaterials mit einer Kupferlösung für IMAC

Bei der Equilibrierung wird die Säule im aufwärtigen Modus (*upflow*) von unten nach oben mit Laufpuffer durchströmt. Der Adapter der Säule bewegt sich dabei in die höchstmögliche Position. Nach etwa einer halben Stunde hat sich ein stabiler Zustand eingestellt. Die Partikel befinden sich, ihrer Größe nach verteilt, im expandierten Bett an verschiedenen Positionen in der Säule [Chang and Chase, 1996].

Die zellbehaftete Kulturbrühe wurde mit der Teilernte des Reaktors im Konditionierungstank C zwischengelagert.

Dort wird sie vor ihrer Aufgabe in die EBA auf die Bedingungen eingestellt, die eine möglichst optimale Bindung des Zielproteins an die Beads gewährleisten und den Einfluss der Zelldichte auf den Aufarbeitungsschritt minimiert.

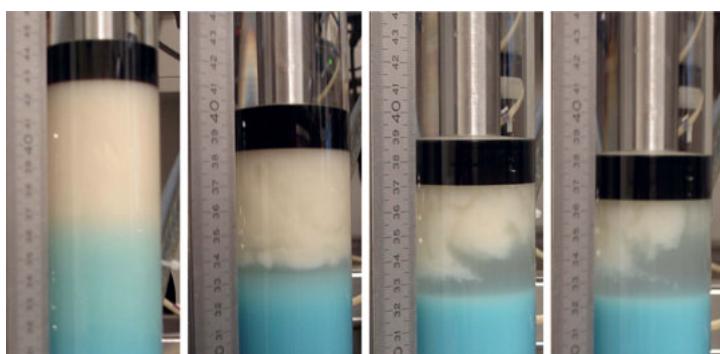


Bild 3.10: Zellaufgabe und anschließendes Herauswaschen aus der EBA-Säule im *upflow*

Die Variablen des Konditionierungstankes C, Zelldichte  $c_{XC}$ , Imidazolkonzentration  $C_{ImC}$ , Leitfähigkeit  $\sigma_C$ , pH-Wert  $pH_C$  und Tween20 Konzentration  $C_{T20C}$ , sind vor dem Chromatographieschritt einzustellen. Dies erfolgt zum Teil automatisiert.

Die Aufgabe der Kulturbrühe ist in Bild 3.10 gezeigt. Sie findet mit der gleichen Fließgeschwindigkeit statt, die zur Equilibrierung der Säule eingestellt war. Das Zielprotein bindet am Adsorbent, während Zellen und andere Verunreinigungen ungehindert das Bett passieren.

Bei der Aufgabe von biomassehaltigem Medium aus der Kultivierung ist eine Aggregation der Zellen zu vermeiden. Sobald die Aggregate durch ihre zunehmende Größe die Sedimentationsgeschwindigkeit der Adsorberpartikel überschreiten, wird eine Abtrennung unmöglich.

Bild 3.11 zeigt Untersuchungen des Durchbruchs einer EBA-Chromatographie. Dargestellt sind die Verläufe der Zelldichte am Austritt der Säule  $c_{Xeout}$ , der Absorptionsmessung  $A_{Eout}$  bei 280 nm und der Gesamtproteinkonzentration  $C_{PtotEout}$ .

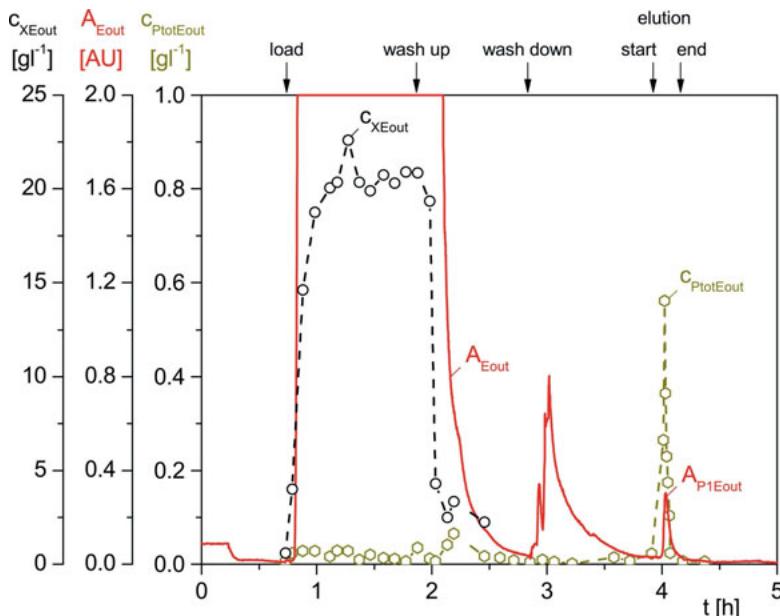


Bild 3.11: Untersuchungen eines Durchbruchs der EBA-Chromatographie

Für die Bestimmung von  $c_{Xeout}$  und  $C_{PtotEout}$  wurde der gesamte Durchbruch des Chromatographie-Schrittes ab dem Zeitpunkt der Zellaufgabe bei  $t = 0,8\text{ h}$  in Fraktionen aufgefangen. Die Zelldichte wurde über die Absorption bei 600 nm ( $OD_{600}$ ) gemessen, der Gesamtprotein gehalt der einzelnen Fraktionen über einen Bradford-Assay bestimmt.

Das Signal  $A_{Eout}$  des UV-Detektors am Austritt der Säule wurde online aufgezeichnet. Während der Zellaufgabe steigt es auf die obere Begrenzung von 2 AU an. Für die Zelldichte  $c_{Xeout}$  wurde im Durchbruch ca.  $20,3\text{ gl}^{-1}$  ermittelt. Diese Messung entspricht dem vorgelegten Wert im Konditionierungstank.

Die Zelldichte bleibt innerhalb der Probenaufgabe (ca. 1 h) am Ausgang relativ konstant. Das UV-Signal bleibt dort in der Übersteuerung.

Die Gesamtproteinkonzentration  $C_{\text{PtotEout}}$  verlief innerhalb dieser Zeitspanne sehr niedrig. Die gemessenen Werte bewegten sich um  $0,02 \text{ gl}^{-1}$ .

Nach Einleiten des ersten Waschschrittes im *upflow* Modus sank das UV-Signal stark ab. Die Zelldichte  $C_{\text{XLEout}}$  wurde zum letzten Mal bei  $t = 2,45 \text{ h}$  mit  $2,3 \text{ gl}^{-1}$  ermittelt. Aufgrund der nicht mehr beobachtbaren Trübung der Fraktionen erfolgte keine weitere Analyse mehr.

Bei  $t = 2,19 \text{ h}$  kommt es innerhalb von zwei Fraktionen zu einem kurzen Anstieg von  $C_{\text{PtotEout}}$  auf  $0,07 \text{ gl}^{-1}$ . Dieser könnte von schwach gebundenen Proteinen resultieren, die vom Imidazol im Waschpuffer von der Trägermatrix verdrängt wurden.

Während des zweiten Waschschritts im *downflow* Modus kommt es einem erneuten Anstieg auf  $0,8 \text{ AU}$  im UV-Signal. Dieser Anstieg wird von Zellen hervorgerufen, die sich während dieses Waschschrittes vom Adaptergitter lösen und abwärtig aus der Säule gespült werden. Dass es sich dabei nicht um Proteine handelt, wird dadurch bestätigt, dass es zu diesem Zeitpunkt zu keiner Erhöhung von  $C_{\text{PtotEout}}$  kommt.

Nach Wiedererreichen der Basislinie des UV-Signals wird in die Elution umgeschaltet. Hier werden die Proteine von ihrem konkurrierenden Substrat Imidazol im Elutionspuffer  $B_2$  vom Säulenmaterial verdrängt. Es kommt zu einem Anstieg des UV-Signals auf eine maximale Peakhöhe von  $0.3 \text{ AU}$ , die mit dem Produktsignal  $A_{\text{P1Eout}}$  gleich gesetzt wird. Parallel dazu steigt  $C_{\text{PtotEout}}$  auf ihren Maximalwert von  $0.56 \text{ gl}^{-1}$  an.

In Tabelle 3.5 sind die eingestellten Bedingungen, die beteiligten Volumenströme  $F_{\text{Bi}}$  und ihre Puffervorlagen  $B_i$  sowie deren Komponenten zusammengefasst.

Tabelle 3.5: Operationsschritte der EBA Chromatographie

Action	Direction	Flow	Value	Volume	Component
activation	downflow	$F_{\text{P1}}$	$1 \text{ lh}^{-1}$	2 CV	$0.1 \text{ M Cu}^{2+}$
rinse	downflow	$F_{\text{P1}}$	$2 \text{ lh}^{-1}$	2 CV	$50 \text{ mM NaAc}$
equilibration	upflow	$F_{\text{P1}}$	$4 \text{ lh}^{-1}$	10 CV	buffer $B_1$
sample load	upflow	$F_{\text{P5}}$	$3.5 \text{ lh}^{-1}$	5 l	cultivation broth
wash 1	upflow	$F_{\text{P1}}$	$4.5 \text{ lh}^{-1}$	10 CV	buffer $B_1$
wash 2	downflow	$F_{\text{P1}}$	$4.5 \text{ lh}^{-1}$	10 CV	buffer $B_1$
elution	downflow	$F_{\text{P2}}$	$5 \text{ lh}^{-1}$	4 CV	buffer $B_2$
Buffer $B_1$ : $0.5 \text{ M NaCl}$ , $20 \text{ mM NaH}_2\text{PO}_4\cdot2\text{H}_2\text{O}$ , $5 \text{ mM Imidazole}$ , pH 7.4					
Buffer $B_2$ : $0.5 \text{ M NaCl}$ , $20 \text{ mM NaH}_2\text{PO}_4\cdot2\text{H}_2\text{O}$ , $100 \text{ mM Imidazole}$ , pH 7.4					

Das Experiment hat gezeigt, dass sich quasi das gesamte Protein aus der Zellsuspension an das Säulenmaterial bindet. Es wird nicht während des Waschens von der Säule abgelöst, sondern erst während der Elution mit dem Elutionspuffer. Die Zellen dagegen werden während der Waschschritte vollständig aus der Säule herausgespült.

Während der Separation der Biomasse kommt die Anlage, vor allem die Adsorberpartikel, in Kontakt mit Zellen und anderen Verunreinigungen wie präzipitierte oder denaturierte Proteine oder DNA. Diese können zum Teil schwerlösliche Bindungen eingehen und zur Aggregation der Partikel führen. Diese stören dann die Bettstabilität und vermindern die Bindungskapazität.

Um eine lange Lebensdauer des Adsorbents zu gewährleisten, ist daher nach jedem Zyklus eine gründliche CIP-Reinigung der Anlage essentiell. Die Entnahme der Partikel aus der Säule wird dabei vermieden, um Verluste an Partikeln zu vermeiden und um Zeit zu sparen.

### 3.3.5 Automatisierung der Aufarbeiteungsprozedur

Um die Notwendigkeit für fehleranfällige manuelle Operationen während der EBA-Läufe zu minimieren, wurde ein vollautomatischer Ablauf der Prozedur angestrebt. Dazu wurde die komplette Downstream-Strategie aus dem vorangegangenen Abschnitt in S88 Rezept-Strukturen der Prozessleitsoftware MFCS/win programmiert. Eine Ansteuerung der EBA-Komponenten wurde dabei über einen CAN-Bus (Controller Area Network) realisiert.

Die *batch structure* ist in die Blöcke *start downstream*, *sterilization*, *harvest*, *product capture* und *end downstream* unterteilt. Jeder Block besteht aus einer *phase structure*. Die Batch-Struktur des Prozesses sowie die Phasen-Struktur des *product capture* sind in Bild 3.12 gezeigt.

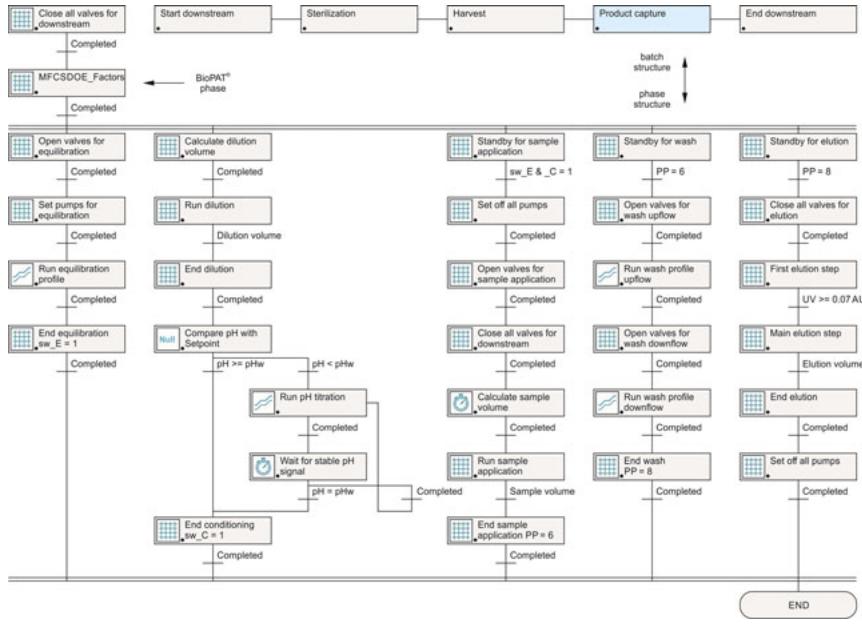


Bild 3.12: ISA S88 Rezeptstruktur der Aufarbeitung mit integrierter BioPAT® Phase

Die Phasen beinhalten dabei die eigentlichen Befehle der Prozessführung, z.B. die Vorgabe oder Änderung von Sollwerten, Änderungen von Reglerzuständen oder Parametern und das Aktivieren und Deaktivieren von On- und Offline-Formeln. Die Phasen können dabei auch parallel abgearbeitet werden. Einzelne Phasen sind über Transitionen verbunden, in denen Bedingungen für das Weiterschalten vorgegeben werden können.

Zudem wurde ein BioPAT® MFCS/win-Modul (Sartorius Stedim Systems, Göttingen) erfolgreich in die Rezeptstruktur implementiert. Das Modul ermöglicht es, Sollwerte, die von einem Versuchsplan für *Design of Experiments* vorgegeben sind, direkt auf die entsprechenden Regler aufzuschalten, sofern diese über das Rezept ansteuerbar sind. Dies wurde hier zum Beispiel für die pH-Regelung im Konditionierungstank realisiert.

Die Automatisierung der Anlage war eine Voraussetzung für die nachfolgende In situ-Optimierung der EBA in Kapitel 4. Hierdurch wurde ermöglicht, automatisch die geforderten Sollwerte der zu untersuchenden Prozessparameter (*factors*) einzustellen und die Versuche streng in der gleichen Weise auszuführen. Somit wurde sichergestellt, dass Variationen in den Prozessantworten (*responses*) nur auf die Variationen der vom Versuchsplan vorgegebenen Einstellungen, nicht aber auf Schwankungen in einer manuellen Prozesshandhabung zurückzuführen sind.

### 3.4 Feinreinigung des Zielproduktes mit einem FPLC-System

Für den abschließenden Feinreinigungsschritt und für eine nachfolgend beschriebene Pre-Optimierung der EBA-Chromatographie wurde das in Bild 3.13 gezeigte *fast protein liquid chromatography* (FPLC) System ÄKTA™purifier 100 (GE Healthcare, Schweden) eingesetzt.



Bild 3.13: ÄKTA™purifier 100 zur Feinreinigung des Zielproduktes

Die Anlage ist auf chromatographische Anwendungen im Hoch- und Niederdruckbereich ausgelegt. Alle eingesetzten Module sind miteinander verbunden und bedürfen nur einer zentralen Stromversorgung.

Der Turm besteht aus folgenden Modulen:

- Frac900 (Fraktionierer)
- Pump P-901 (Fließmittelpumpe)
- Pump P-960 (Probenpumpe)
- Monitor pH/C-900 (pH-/Leitfähigkeitsmessung)
- Monitor UV-900 (UV Absorptionsmessung)

Zur Programmierung des ÄKTA™purifier 100 wurde die Software Unicorn® 5.1 verwendet. Unicorn enthält alle benötigten Funktionen für die Steuerung und Überwachung des Systems. Es ist in die einzelnen Module *Manager*, *Method Editor*, *System Control* und *Evaluation* unterteilt.

Im *Manager* finden dabei das Datei-Handling und die Verwaltung, z.B. Definition von System- und Anwenderprofilen usw., statt. Mit dem *Method Editor* werden Verfahren zur programmierten Steuerung des Chromatographie-Systems erstellt und bearbeitet. Die *System Control* wird benutzt, um das Trennverfahren durch eine Methode oder händisch zu steuern und den Prozess online zu überwachen. Mit dem *Evaluation* Modul können gespeicherte Ergebnisse analysiert und bewertet werden.

Wie dem Fließschema in Bild 3.14 zu entnehmen ist, können an diesem System maximal vier verschiedene Fließmittel eingesetzt werden. Das Pumpenmodul P-901 besitzt zwei separate Pumpen ( $P_{1A}$  und  $P_{1B}$ ), die je zwei Anschlüsse besitzen. Während eine Mischung zweier Fließmittel an einer Pumpe durch die Selektorenventile  $V_{1A}$  und  $V_{1B}$  ausgeschlossen ist, kann diese durch Anschluss an je eine Pumpe erreicht werden. Dabei ist der Volumenstrom mit 0,01% Schritten in einem Bereich von 0% bis 100% der Pumpe  $P_{1B}$  variabel, wobei die Pumpe  $P_{1A}$  komplementär zwischen 100% und 0% variiert wird.

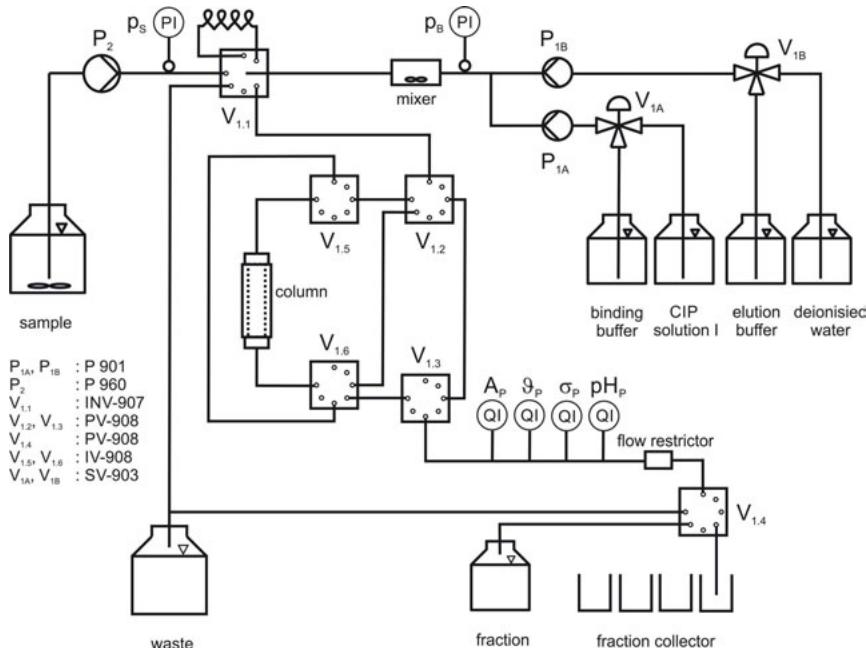


Bild 3.14: Fließschema der verwendeten Konfiguration des ÄKTA™ purifier 100

Die 2 ml Mischkammer sorgt für ein homogenes Gemisch der geförderten Puffer, so dass auch ein Stufen- und Gradientenbetrieb ermöglicht wird. Zusätzlich zu den Pumpen des Hauptmoduls ist eine weitere Probenpumpe  $P_2$  vorhanden, die durch Keramik-Pumpenköpfe und Zylinder für die Förderung von viskosen und partikelbehafteten Flüssigkeiten geeignet ist. Die Schaltungen der Ventile  $V_{1,2}$ ,  $V_{1,3}$ ,  $V_{1,5}$  und  $V_{1,6}$  ermöglichen nicht nur einen Gleichstrom-, sondern auch einen Gegenstrom- oder einen Bypassbetrieb.

Unabhängig von der verwendeten Einstellung wird der Flüssigkeitsstrom durch Detektoren für Leitfähigkeit  $\sigma_p$ , Temperatur  $\vartheta_p$ , UV-Absorption  $A_p$  und pH-Wert  $pH_p$  geleitet. Am abschließenden Ausgangsventil  $V_{1,4}$  ist ein Fraktionssammler angeschlossen, der Fraktionen in

einem Bereich von 0,1 ml bis 50 ml sammeln kann. Durch die Nutzung der verbleibenden Ausgänge an diesem Ventil können auch größere Fraktionen aufgefangen werden.

Im nächsten Kapitel 4 werden die vorgestellte EBA-Prozedur und die Feinreinigung mit der Proteinexpression im Bioreaktor zu einem sequentiellen Integrierten Bioprozess vereint. Im darauffolgenden Kapitel 5 werden dann die Operationsparameter der EBA optimiert, um einen QbD-konformen Produktionsprozess zu etablieren.

## 4. Integrierte Prozessführung mit einer EBA-Chromatografie

### 4.1 Das Konzept des Integrierten EBA-Biprozesses

In Bild 4.1 ist der Aufbau der integrierten Prozessanlage mit der Expanded Bed Adsorptions-Chromatografie schematisch gezeigt. Die Anlage gliedert sich prinzipiell in sechs Teilsysteme, die vornehmlich über fest verlegte Edelstahlrohre und autoklavierbare Silikonschläuche, mit Schnellschusskupplungen ankopplbar, verbunden sind.

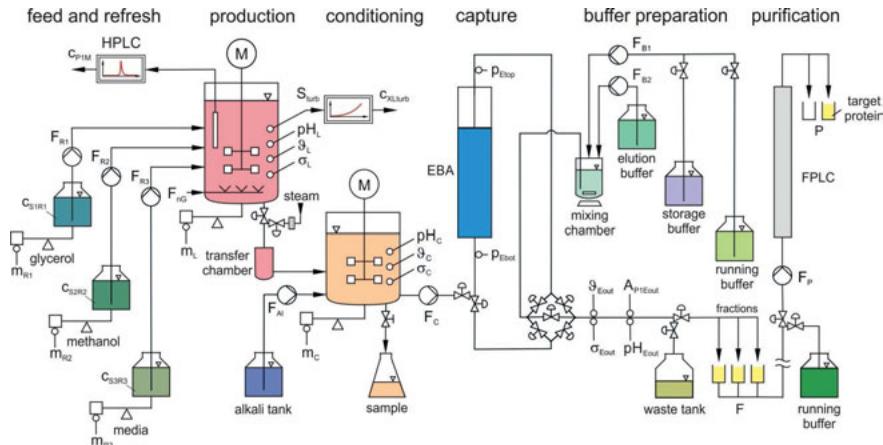


Bild 4.1: Schematische Darstellung der integrierten EBA-Prozessanlage

Die Teilsysteme bestehen aus einem 5 l BIOSTAT® ED Bioreaktor (*production*), einer Medien- und Substratversorgung für den Bioreaktor (*feed and refresh*), einem Konditionierungsgefäß (*conditioning*), einem EBA-Chromatographiesystem (*capture*), einer ÄKTA™purifier 100 FPLC-Anlage (*purification*), sowie einer Puffervorlage mit Mischgefäß (*buffer preparation*) für die Chromatographie.

### 4.2 Sequentielle Zellkultivierung und Proteinexpression

In einer bereits veröffentlichten Studie wurde eine vollautomatische Produktion von D1M1 entwickelt [Borchert et al., 2013]. Sie beschreibt eine sequentielle Produktionsstrategie, die eine Kombination von vor- und nachgelagerten Grundoperationen für das ISPR beinhaltet. Dazu wurden stabile Produktionsbedingungen etabliert, die einen reproduzierbaren Prozessverlauf ermöglichen und ein hohes Potential zur Optimierung des integrierten Prozesses bereitstellen.

Die entwickelte Kultivierungsstrategie setzt zu Beginn eines neuen Prozesses eine Batchphase und einen initialen Zyklus – bestehend aus einer Fed-Batch-Phase auf Glycerol zum schnellen Erreichen höherer Zeldichten sowie einer Produktionsphase, in der die Proteinexpression durch Methanol induziert wird – voraus (*startup*). Daran lässt sich eine beliebige Anzahl sich wiederholender Zyklen anschließen, die täglich Ausgangsmaterial für eine integrierte EBA-Chromatographie liefern. Bild 4.2 zeigt einen Startup-Zyklus mit sechs folgenden Produktions-Sequenzen.

Die Vorkulturen wurden in Schüttelkolben von 1 l mit je 170 ml Kultivierungsmedium aus einer aufgetauten 400 µl Kryokultur des Stammes KM71H-D1M1 angeimpft. Nach einer 12-stündigen Inkubationszeit bei 30 °C wurde daraus der Bioreaktor mit einer Anfangszelldichte  $c_{X0}$  von 0,5 g/l bei einem Ausgangsvolumen von 4 l inkuliert.

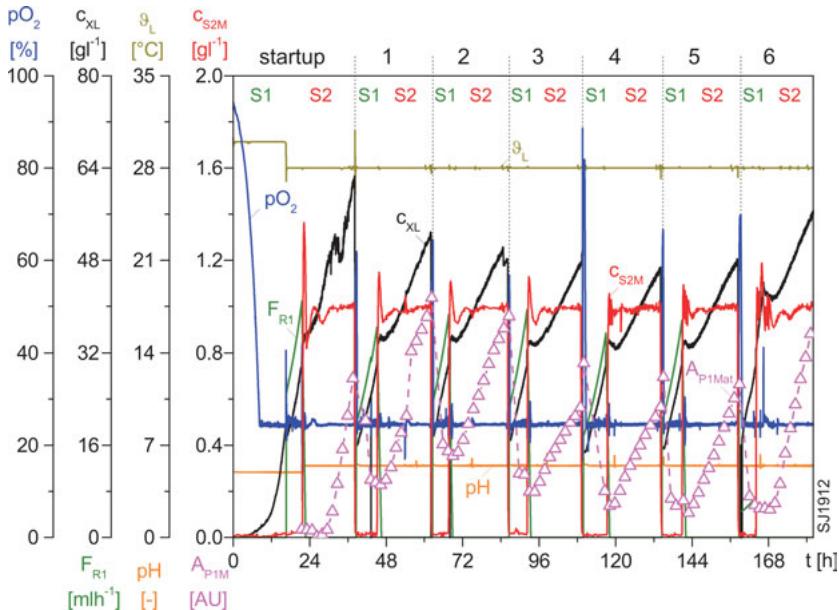


Bild 4.2: Sequentielle Kultivierung mit einem Startup-Zyklus und sechs nachfolgenden Produktions-Sequenzen

Das verwendete Batch-Medium wurde mit einer Glycerolkonzentration  $c_{S1L0}$  von 25 g/l so ausgelegt, dass die Zellen nach vollständigem Verbrauch des Primärsubstrates eine Zelldichte  $c_{X1L}$  von 20 g/l erreichen – woraus sich ein mittlerer Substratausbeutekoeffizient  $y_{X/S1}$  von 0,78 ableiten lässt. Der pH wird durch Titration mit 2 M  $H_3PO_4$  oder auf 12,5 %  $NH_4OH$  auf einen Sollwert von 5,0 geregelt. Der Reaktor wird mit 5,5 lmin<sup>-1</sup> Luft begast und ein Überdruck von 500 mbar im Kopfraum konstant gehalten. Nach Erreichen seines Sollwertes wird der relative Sauerstoffpartialdruck  $pO_2$  über eine Drehzahlkaskadierung konstant auf 20 % geregelt.

Am Ende der Glycerol Batch-Phase wird die Kultivierungstemperatur  $9_L$  von 30°C auf 28°C abgesenkt und der pH von 5.0 auf 5.5 angehoben.

Durch ein exponentielles Zufütterprofil für  $F_{R1}$  wird der Prozess im Glycerol Fed-Batch fortgeführt und dabei die zellspezifische Wachstumsrate  $q_{1/X}$  auf 0,1 h<sup>-1</sup> ( $\mu_{1w}$ ) gesteuert [Cornelissen et al., 2003].

Der glycerollimitierte Verlauf soll die Zellen bereits auf die nachfolgende Induktionsphase vorbereiten. Die Zeldichte  $c_{Xturb}$  wird aus dem Online-Trübungssignal  $E_L$  berechnet. Bei Erreichen von 30 g/l wird die Proteinexpression durch Zufüttern von Methanol als einzige Kohlenstoffquelle eingeleitet. Die Konzentration von Methanol im Medium  $c_{S2M}$  wird gemessen und auf 1 g/l geregelt. In beiden Phasen wird die UV-Absorption  $A_{P1M}$  des Zielproduktes über eine HPLC-Methode atline gemessen [Martens et al., 2011].

Der Startup-Zyklus endet nach 24 h mit einer Teilernte der Kultivierungsbrühe in den Konditionierungstank der EBA (siehe Bild 4.1). Die Ernte erfolgt über die Erntekammer portionsweise und ist so ausgelegt, dass eine finale Zellmasse  $m_{XLn}$  von 80 g für den folgenden Produktions-Zyklus im Reaktor zurückbleibt.

Nach Auffüllen des Reaktors mit frischem Medium auf ein Volumen  $V_{LO}$  von 4 l wird eine Anfangszelldichte  $c_{XL0}$  von 20  $gl^{-1}$  eingestellt. Dieses Vorgehen gewährleistet gleiche Anfangsbedingungen für jeden folgenden Produktionszyklus.

Der erste Zyklus beginnt wieder mit einem  $\mu$ -gesteuerten Glycerol-Fed Batch, diesmal über 7 h. Hier baut die Zelle die zellinternen Bestandteile des Methanolstoffwechsels wieder ab [Fricke et al., 2013], so dass zu Beginn der folgenden Induktionsphase AOX-arme Zellen für eine neue Produktion bereit stehen.

In der methanolinduzierten Produktionsphase beginnt nach einer Totzeit von ca. 2 h wieder die Sekretion des rekombinanten Zielproteins, beobachtbar am Signal  $A_{P1M}$  der Atline-HPLC. Die Produktion endet mit der bereits beschriebenen Teilernte mit nachfolgendem Refresh zur nächsten Zellanzucht.

Bild 4.3 zeigt aus einer späteren Projektphase drei aufeinanderfolgende Zyklen, aus denen sich bereits eine erhöhte Reproduzierbarkeit des Prozesses im Vergleich zur Kultivierung in Bild 4.2 ableiten lässt.

Während des letzten gezeigten Zyklus ist die Messung des Zielproduktes mit der HPLC gegen Ende des Prozesses ausgefallen, so dass ab diesem Zeitpunkt keine Atline-Messung mehr möglich war. Die Steigung des Signals lässt hier lediglich einen ähnlichen Verlauf wie in den anderen Zyklen vermuten.

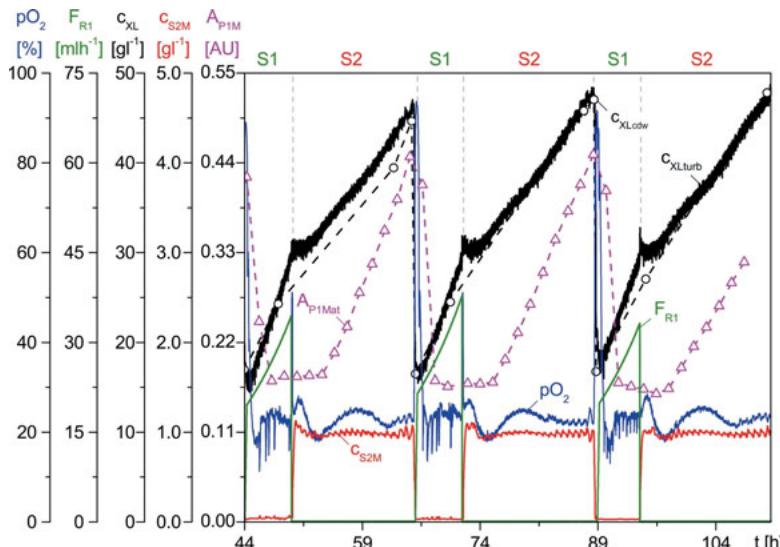


Bild 4.3: Drei Zyklen einer Kultivierung mit erhöhter Reproduzierbarkeit

Der zyklische Ablauf in Bild 4.3 ist über einen Zeitraum von 12 Tagen mehrmals erfolgreich durchgeführt worden. Er ermutigte, die Aufarbeitung mit der EBA in derartigen Prozessen mit ein zu beziehen. Diese Ergebnisse werden im nächsten Abschnitt erläutert.

## 4.3 Die Erprobung der integrierten Produktionsweise

### 4.3.1 Prozessverlauf mit nicht optimalen Prozessparametern

Im Verlauf einer ersten Erprobung der integrierten Produktionsweise wurden während der Herstellung des Vakzinkandidaten D1M1 die implementierte Automatisierungstechnik sowie alle verwendeten Bauteile auf ihre Funktionsweise hin untersucht. Außerdem stand die Performance des Upstream und des Downstream im Hinblick auf die Güte einer Herstellung des Zielproduktes im Fokus der Untersuchung.

Die zyklische Kultivierung erfolgte mit einem nicht optimierten Kultivierungsprotokoll [Martens & Borchert et al., 2011]. Für einen ersten Aufarbeitungsversuch mit der EBA wurde das in Abschnitt 3.3 erläuterte EBA-Protokoll verwendet.

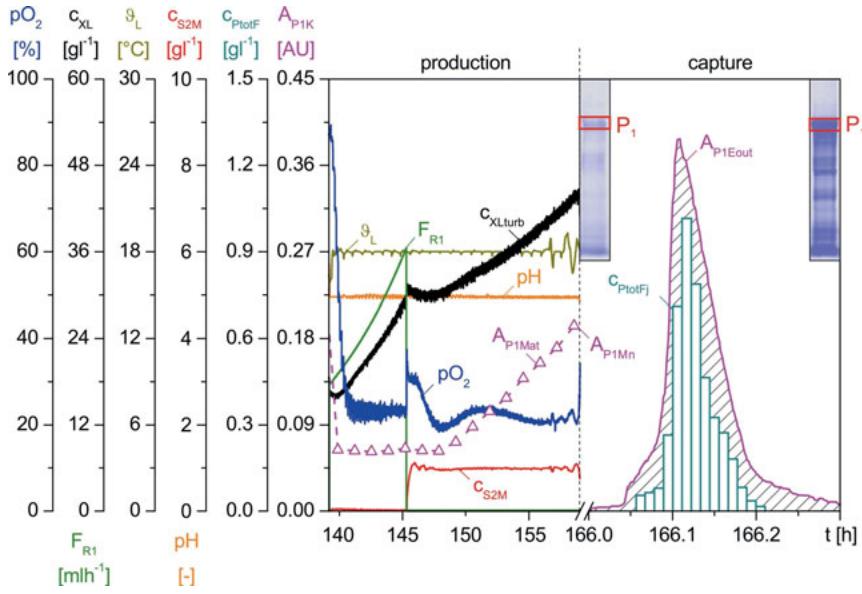


Bild 4.4: Ein Produktionszyklus von D1M1 mit nachfolgender EBA-Elution und SDS-PAGEs aus Proben am Ende der Produktion sowie der sechsten Fraktion des Captures

Bild 4.4 zeigt den Prozessverlauf der durchgeführten Kultivierung und den Elutionsverlauf des zugehörigen Aufreinigungsschrittes mit der EBA. Für den Upstream sind die Online-Variablen  $pO_2$ , pH,  $s_L$ ,  $c_{XLturb}$ ,  $F_{R1}$  und  $c_{S2M}$  dargestellt. Hinzu kommen im Upstream die mit der Atline-HPLC gemessene UV-Absorption  $A_{P1Mat}$  sowie im Downstream die im Outflow der EBA gemessene UV-Absorption  $A_{P1Eout}$  des Zielproduktes. Die Produktkonzentrationen  $c_{PlotFj}$  in den Fraktionen des EBA-Eluats wurden mit einem Bradford-Assay bestimmt.

An Hand der aufgezeichneten Prozessgrößen lässt sich kein ungünstiger Prozessverlauf feststellen. Der Kultivierungszyklus verläuft charakteristisch für den eingesetzten Stamm. In der Produktionsphase steigt die Produktkonzentration gemessen am Absorptionsignal  $A_{P1Mat}$  an. Eine SDS-PAGE-Analytik der Fermentationsendprobe weist jedoch lediglich eine schwache Bande in Höhe des Zielproduktes bei 73 kDa auf, enthält aber zusätzlich erhebliche Ver-

unreinigungen, die z.T. auf eine Degradation der Prodomäne zurückzuführen sind [Martens, 2014].

Der Elutionsverlauf ist symmetrisch. Hier ist weder ein ausgeprägtes *Fronting* noch *Tailing* zu erkennen. Während der Elution wurden insgesamt 14 Fraktionen aufgefangen und auf ihren Gesamtproteingehalt hin untersucht. Mit einem Anstieg der Absorption steigt auch der Gesamtproteingehalt in zeitlich dazugehörenden Fraktionen an. Die maximal angezeigte Absorption von  $A_{P1Eout}$  betrug 0.38 AU, die maximal gemessene Gesamtproteinconzentration  $C_{PtotF6}$  betrug 1.0  $gl^{-1}$  in Fraktion 6.

Zu der sechsten Fraktion wurde beispielhaft auch das Ergebnis der SDS-Analytik dargestellt. Es wirkt an dieser Stelle nur wie eine Verstärkung des Patterns der Kultivierungsprobe. Das Zielprodukt wurde durch den Aufarbeitungsschritt ebenso wie alle vorhandenen Verunreinigungen stark auf dem Säulenmaterial angereichert und erst mit der Elution wieder von diesem abgelöst. Für die Verunreinigungen sollte dies im Idealfall bereits während des Waschschrittes erfolgen.

Als ein Ergebnis der Erprobung konnte dennoch festgestellt werden, dass sowohl der vorgeschlagene Ablauf der zyklischen Kultivierung als auch die komplexe Aufarbeitungsstrategie mit der integrierten EBA erfolgreich und vollautomatisch abgelaufen sind. In der Produktion wurde das Zielprodukt in geringen Mengen gebildet und konnte im Überstand mit SDS-PAGE nachgewiesen werden – allerdings enthielt diese Ausgangsprobe für eine EBA-Chromatographie zu viele Verunreinigungen.

Das Zielprodukt wurde mit der EBA zwar von Zellen und Medienbestandteilen getrennt, dennoch war mit den gewählten Prozessparametern noch keine Trennung von Zielprotein und Verunreinigungen zu beobachten. Die geringe Qualität der Aufarbeitung lässt sich zum einen mit noch nicht optimierten Prozessbedingungen der EBA-Chromatographie begründen, zum andern liegt sie zu erheblichen Teilen aber an der vom Upstream bereitgestellten Ausgangsprobe, die von sehr schlechter Qualität war. Dies gilt es zunächst zu verbessern.

#### 4.3.2 Verbesserung der Produktqualität mit optimalen Kultivierungsparametern

Parallel zu dieser Arbeit wurde eine Optimierung der Kultivierungsbedingungen für die Herstellung des kleinsten Malariavakzin-Kandidaten D1 (53 kDa) mit dem gleichen *Pichia pastoris* Hostsystem KM71H durchgeführt [Fricke et al., 2011].

Die Gruppe um Fricke hat vollständig automatisierte mehrphasige Kultivierungsstrategien für mehrere rekombinante Malariavakzine etabliert. Für eine schnelle, zielorientierte Optimierung der biotechnologischen Produktion mittels DoE wurde dabei eine Multibioreaktoranlage entworfen und erfolgreich für die Optimierung eingesetzt [Fricke, 2015].

Die optimalen D1-Kultivierungsparameter wurden hier auch für die Herstellung von D1M1 angewendet und das Experiment in Bild 4.4 damit wiederholt.

Das Ergebnis ist in Bild 4.5 gezeigt. Die Zielproduktabsorption  $A_{P1M1}$  am Ende der Kultivierung ist doppelt so hoch wie im Ausgangsexperiment. Die zugehörige SDS-PAGE zeigt eine klare Produktbande mit sehr wenigen Verunreinigungen.

Der Elutionspeak in der EBA ist weitaus schlanker und bringt in den gepoolten Elutionsfraktionen eine Gesamtproteinconzentration  $C_{PtotF}$  von ca. 1,05  $gl^{-1}$ . Dies entspricht in etwa der Gesamtproteinconzentration der Elutionsfraktion, die in der nicht optimierten Kultivierung gewonnen wurde. Bei einer wesentlich verbesserten Reinheit  $P_{P1F}$  ist der Anteil an Zielpunkt P<sub>1</sub> aber wesentlich höher.

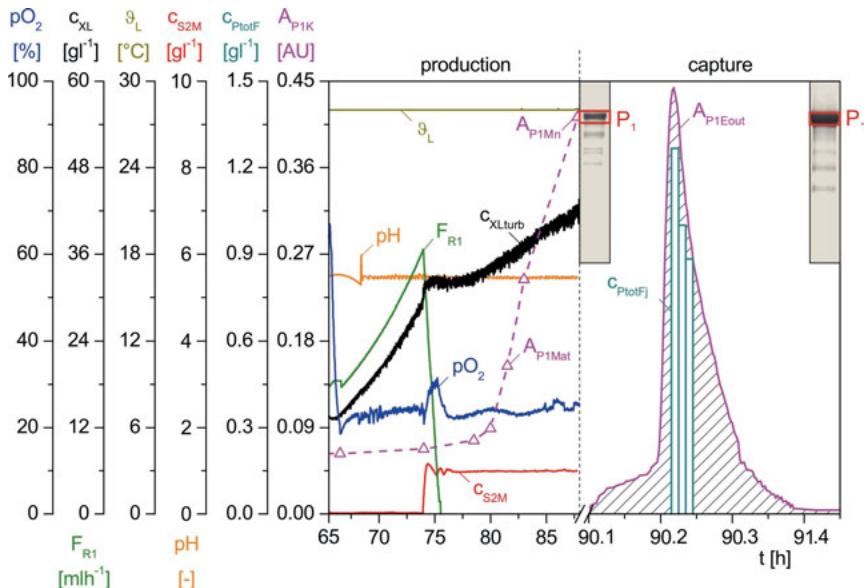


Bild 4.5: Ein optimierter Kultivierungszyklus mit EBA-Elution und SDS-PAGEs aus der Endprobe der Kultivierung sowie aus der ersten Fraktion des Captures

Zum Abschluss dieses Kapitels sind in Tabelle 4.1 noch einmal die beiden Experimente gegenübergestellt.

Tabelle 4.1: Vergleich der gewählten Prozessparameter und der finalen Produktmessung

	not optimized	optimized
$\vartheta_L$ [°C]	18.0	28.0
pH [-]	5.0	5.5
$c_{S2M}$ [gl <sup>-1</sup> ]	1.0	1.0
$A_{P1Mn}$ [AU]	0.2	0.4

Die SDS-PAGEs am Ende des optimierten Upstreams und im EBA-Eluat zeigen bereits eine hohe Reinheit des Produktes  $P_1$ . Dennoch ist jetzt nach der Maximierung des Upstream auch eine Optimierung der EBA-Operationsparameter hinsichtlich der Reinheit  $P_{1F}$  und der Wiederfindung  $R_{P1E}$  erforderlich. Diese wird in dem nachfolgenden Kapitel vorgenommen.

## 5. Implementierung von Vorgaben des Quality by Design

### 5.1 Zielsetzungen von QbD in der Bioprozessentwicklung

Der in Kapitel 4 entwickelte vollautomatische Ablauf eines Integrierten Prozesses soll im Weiteren den in Kapitel 1 vorgestellten Richtlinien der FDA, nach Vorgaben eines *Quality by Design*, entsprechen. Hierzu sind die von der ICH festgelegten Entwicklungsschritte in der Richtlinie Q8 *Pharmaceutical Development* in Bild 5.1 dargestellt und mit weiteren Zielsetzungen dieser Arbeit ergänzt.



Bild 5.1: Implementierung von QbD bei einer biotechnologischen Prozessentwicklung

An erster Stelle steht die Anfertigung eines *Quality Target Product Profile* (QTPP), das eine Zusammenfassung von Qualitätsmerkmalen unter Berücksichtigung von Sicherheit und Wirksamkeit des Produktes darstellt. Zu den Qualitätsmerkmalen gehören u.a. Darreichungsform, Pharmakokinetik und Sterilität [Rathore, 2009].

Diese Vorgaben wurden hinsichtlich der Produkte *potentielle Malariavakzine* von dem Projektpartner am *Biomedical Primate Research Center* (BPRC) der Niederlande in Rijswijk vorgenommen [Faber et al., 2007].

In den Arbeiten an der HAW Hamburg stand hingegen das QbD der zu entwickelnden Produktionsprozesse im Vordergrund. Das Wort kritisch bezieht sich daher im Folgenden auf die Reproduzierbarkeit des Produktionsprozesses im Hinblick auf die Reinheit (*purity*), und damit auf die immunologische Wirksamkeit der hergestellten Proteine, sowie auf die Produktivität (*productivity*) im Upstream und die Wiederfindung (*recovery*) im Downstream im Hinblick auf eine ökonomische Nutzung der entwickelten Strategien.

Die Bestimmung von *Critical Quality Attributes* (CQA) und *Critical Process Parameters* (CPP) stellt dabei den ersten Schritt bei der Implementierung der Produktionsstrategien dar. CQA sind physikalische, chemische oder biologische Prozessmerkmale, die in einem vordefinierten Bereich liegen müssen, um eine gewünschte Produktqualität hervorzubringen [Rathore und Winkle, 2009]. Ein Produkt- und ein Prozessmerkmal ist z.B. die Produkttreinheit.

Die CPP sind Prozessparameter, die einen wesentlichen Einfluss auf die CQA haben, wie z.B. Kultivierungstemperatur und pH-Wert. Zur Identifizierung des QTPP sowie der CQA und der CPP bieten sich Risikoanalysen, im speziellen eine *Failure Mode and Effects Analysis* (FMEA) nach ICH Q9 *Quality Risk Management*, an [Reifsnyder et al., 2011].

Um einen Zusammenhang zwischen den CQA (*responses*) und den CPP (*factors*) durch eine Prozessoptimierung zu finden, wird intensiv die Statistische Versuchsplanung, *Design of Experiments* (DoE), angewendet.

DoE ist in die drei Kategorien *Screening*, *Optimization* und *Robustness Testing* unterteilt. In einem *Screening* wird untersucht, welche der CPP signifikanten Einfluss auf die CQA haben. In einer nachfolgenden *Optimization* werden Prozessparameter etabliert, die einen mathematischen Zusammenhang mit den CQA hervorbringen und einen optimalen Prozessverlauf garantieren.

Das *Robustness Testing* untersucht abschließend, ob bei geringen Änderungen der optimalen Werte für die CPP das vordefinierte CQA und damit auch das QTPP erhalten bleibt [Menezes, 2011].

Der *Design Space* stellt einen stabilen Produktionsbereich dar, dessen Etablierung ein entscheidender Schritt bei der Implementierung von QbD darstellt. Innerhalb des *Design Space* führen Änderungen der CPP nicht zu signifikanten Änderungen der CQA, wodurch gleichbleibend die gewünschte Qualität erzielt wird.

Solange ein Prozess innerhalb der Grenzen seines *Design Space* betrieben wird, gilt dies nicht als eine Änderung am Prozess. Initiiert eine Prozessänderung jedoch eine Bewegung aus dem *Design Space* hinaus, ist im Normalfall eine regulatorische Nachprüfung bzw. Neuzulassung durch die Behörden [ICH, 2009a] angesagt.

Diese regulatorische Flexibilität ermöglicht es, einen Prozess nicht nur unter gleichbleibenden Bedingungen reproduzierbar zu fahren, sondern ihn auch gegen Schwankungen im Ausgangsmaterial robust zu gestalten [Menezes, 2011]. Diese Schwankungen können z.B. durch einen Herstellerwechsel für Medienkomponenten, unterschiedliche Startzelldichten und/oder veränderte Handhabungen bei einem Personalwechsel bedingt sein.

Die Herstellung potentieller Malariavakzine mit der vorgestellten zyklischen Produktionsstrategie unter Maßgabe von QbD-Vorgaben wurde an der HAW Hamburg bereits von Fricke [Fricke et al., 2013; Fricke, 2015] intensiv für den Upstream untersucht.

Ein ähnlicher Ansatz für den Aufreinigungsprozess mit der EBA wird in diesem Kapitel vorgestellt, während die weiteren QbD-Vorgaben in Bild 5.1 hinsichtlich einer Implementierung von Monitoring und Control der Reproduzierbarkeit zyklischer Produktionsprozesse in einer industrie-äquivalenten Forschungsumgebung in Kapitel 7 und 8 bearbeitet werden.

## 5.2 Definition der Critical Quality Attributes im Downstream

### 5.2.1 Optimierungspotential und Gütekriterium des EBA Prozesses

Für die Auslegung und Bewertung des Aufarbeitungsschrittes mit der EBA wurden zunächst die Gütekriterien zur Bewertung der Prozessqualität definiert.

Hierbei bietet sich die bereits in Abschnitt 2.3.2 vorgestellte Produktreinheit (*purity*)  $P_{P1K}$ ,

$$P_{P1K} = \frac{B_{1Kn}}{\sum_{j=1}^b B_{jKn}} \quad (5.1)$$

mit

$P_{P1K}$  := Reinheit eines Aufarbeitungsschrittes K [·]

$B_{jKn}$  := Bande j in einer SDS-PAGE in einer Probe n am Ende von K [·]

$B_{1Kn}$  := Zielproduktbande von  $B_{jKn}$  [·]

an, die nur aus einer Probe am Ende ( $t = t_0$ ) eines jeden Aufarbeitungsschrittes bestimmt wird und diesen Teilprozess K kennzeichnet.

Für das zweite Gütekriterium, die Wiederfindung (*recovery*), benötigt man die Reinheit, das Volumen und die Gesamtproteinkonzentration jeweils von der Aufgabe aus dem vorherigen Teilprozess J und der Ernte des betrachteten Schrittes K.

Die Wiederfindung  $R_{P1K}$ ,

$$R_{P1K} = \frac{m_{P1Kout}}{m_{P1Kin}} = \frac{V_{Kout} \cdot P_{P1K} \cdot C_{PlotKn}}{V_{Kin} \cdot P_{P1J} \cdot C_{PlotJn}} \quad (5.2)$$

mit

$R_{P1K}$  := Wiederfindung des Zielproduktes P<sub>1</sub> in Aufarbeitungsschritt K [-]

$V_{Kin}$  := aufgegebenes Volumen in Teilprozess K [l]

$V_{Kout}$  := geerntetes Volumen in Teilprozess K [l]

$m_{P1Kin}$  := Zielproduktmasse in der Aufgabe bei K [g]

$m_{P1Kout}$  := Zielproduktmasse in der Ernte von K [g]

$P_{P1K}$  := Reinheit am Ende von Teilprozess K [-]

$P_{P1J}$  := Reinheit am Ende des vorherigen Teilprozesses J [-]

$C_{Plotln}$  := Gesamtproteinkonzentration am Ende von Teilprozess I, I = J, K [g l<sup>-1</sup>],

wird über die Division der Produktmenge aus den nachfolgenden Teilschritten J und K bestimmt.

Häufig wird das Produkt aber in mehreren (f) Elutionsfraktionen k aufgefangen und gepoolt. Untersucht man jede Einzelfraktion F<sub>k</sub>, so ist die Masse an aufgefangenem Produkt m<sub>P1Kout</sub>,

$$m_{P1Kout} = \sum_{k=1}^f V_{Fk} \cdot P_{P1Fk} \cdot C_{PlotFk} \quad (5.3)$$

mit

$V_{Fk}$  := Erntevolumen der Elutionsfraktion F<sub>k</sub> [l]

$P_{P1Fk}$  := Reinheit in Fraktion F<sub>k</sub> [-]

$C_{PlotFk}$  := Gesamtproteinkonzentration in Fraktion F<sub>k</sub> [g l<sup>-1</sup>]

aus den Einzelmassen zu berechnen.

Es empfiehlt sich, die Reinheit für die Aufgabe und die Ernte in einer gemeinsamen SDS-PAGE gleichzeitig zu bestimmen.

## 5.2.2 Auswahl möglicher Critical Process Parameter

Um für die beiden CQA, *purity* P<sub>P1K</sub> und *recovery* R<sub>P1K</sub>, die kritischen Prozessparameter CPP zu identifizieren, wurde die Ishikawa-Methode angewendet, bei der zunächst alle möglichen Einflussfaktoren auf die CQA in einem Diagramm erfasst werden [Ishikawa, 1981].

Das für die Aufarbeitung mit der EBA erarbeitete Diagramm ist in Bild 5.2 dargestellt und ähnelt einer Fischgrätenstruktur, woraus sich die geläufige Bezeichnung *Fishbone-Diagram* herleiten lässt. Es fasst Prozessparameter, verwendete Rohmaterialien und Analysemethoden, die einen Einfluss auf die Qualität der Aufarbeitung haben könnten, zusammen.

Auf Grundlage dieser Zusammenstellung wurden durch eine individuell ausgelegte Fehlermöglichkeits- und Fehlereinflussanalyse (FMEA) nach ICH Q9 fünf Verfahrensparameter als kritisch für den Aufarbeitungsprozess angesehen. Diese sind, auch in Hinblick auf die erforderliche Optimierung über *Design of Experiments* (DoE), in nachfolgender Tabelle 5.1 zusammengestellt.

Aus diesem Parametersatz sind nun diejenigen zu identifizieren, die einen starken Einfluss auf die beiden Gütekriterien (*responses* y<sub>i</sub>), die Reinheit des EBA-Prozesses P<sub>P1F</sub> (y<sub>1</sub>) und die Wiederfindung im EBA-Prozess R<sub>P1E</sub> (y<sub>2</sub>), haben.

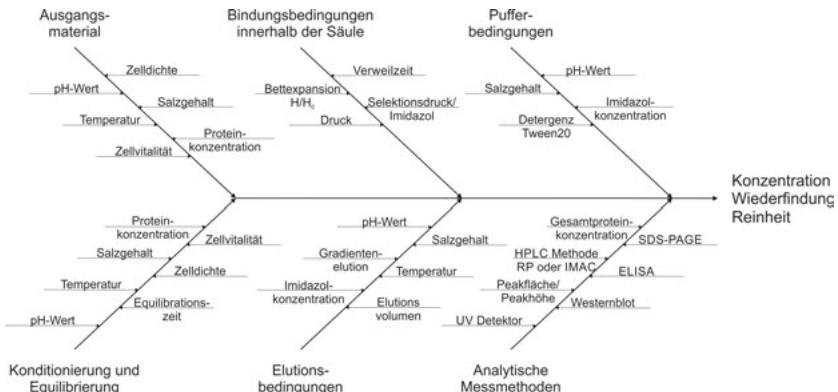


Bild 5.2: Ishikawa-Diagramm möglicher Einflussfaktoren auf die Gütekriterien

Tabelle 5.1: Potentielle kritische Betriebsparameter der EBA-Chromatographie

Variable	Bezeichnung	Einheit	DoE-Faktor	Suchbereich
pH-Wert im Konditionierungstank C	pH <sub>C</sub>	-	X <sub>1</sub>	5.5 - 8.5
Leitfähigkeit in C	$\sigma_C$	mScm <sup>-1</sup>	X <sub>2</sub>	15 - 65
Imidazol-Konzentration im Laufpuffer B1	C <sub>ImB1</sub>	mM	X <sub>3</sub>	0 - 20
Tween®20-Konzentration in C	C <sub>T20C</sub>	µl <sup>-1</sup>	X <sub>4</sub>	0 - 20
EBA-Flux durch die Säule	J <sub>E</sub>	cmh <sup>-1</sup>	X <sub>5</sub>	110 - 530

Diese Operationsgrößen werden mit einem geeigneten DoE untersucht und dann ggf. als *Critical Process Parameter* (CPP) identifiziert. Dieser Vorgang erfolgt mit dem DoE-Schritt Screening. Die nachfolgende Optimierung und Robustheits-Test der CPP bringen dann das beschreibende Modell und den *Design Space*.

### 5.3 Die Idee des Design of Experiments

*Design of Experiments* beruht auf der Vorgabe eines orthogonalen Suchraums, der durch die zu untersuchenden Parameterbereiche aufgespannt wird. Zur Anschauung wird das fünfdimensionale Suchproblem auf drei Parameter beschränkt.

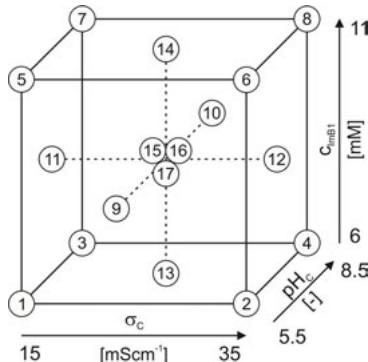


Bild 5.3: Region eines CCF-Designs für eine Pre-Optimierung der EBA-Konditionen

Bild 5.3 zeigt den Suchraum für  $o = 3$  Parameter,  $pH_C$  ( $X_1$ ),  $\sigma_C$  ( $X_2$ ) und  $C_{ImB1}$  ( $X_3$ ), die im physikalischen Modell mit  $X_i$  bezeichnet werden. In diesem als *Central Composite Face-centered* (CCF) bezeichneten Design sind alle 8 Ecken und die 6 Stirnseiten mittig besetzt sowie im Zentrum des Quaders drei *centerpoints* (CP) vorgegeben.

Für diesen Raum wird die Struktur des Modells für die Berechnung einer der beiden ( $r = 2$ ) *responses*  $y_r$ ,

$$y_r = a_{0r} + \sum_{i=1}^n a_{ir} \cdot X_i + \sum_{1 \leq i < j} a_{ijr} \cdot X_i \cdot X_j + \sum_{i=1}^n a_{iir} \cdot X_i^2 + e_r \quad (5.4)$$

mit

$y_r$  := Gütekriterium  $r$  des Prozesses

$X_i$  := physikalischer Prozessparameter

$n$  := Dimension des Parametersuchraums

$a_{0r}$  := Offset des Modells für  $y_r$

$a_{ir}$  := Koeffizienten des linearen Modellanteils

$a_{ijr}$  := Koeffizienten des Interaktions-Modellanteils

$a_{iir}$  := Koeffizienten des quadratischen Modellanteils

$e_r$  := Residuum des Modells bei der Berechnung von  $y_r$ ,

mit einem Offset sowie je drei linearen, bilinearen und quadratischen Termen vordefiniert.

Hierin sind mit so wenig Experimenten wie möglich, die unbekannten Koeffizienten  $a_{ir}$  zu finden und aus diesem Modell dann die optimalen Parameter  $X_{i, \text{opt}}$  für ein maximales Gütekriterium  $y_{\text{max}}$  zu identifizieren.

Allerdings muss der Suchraum so normiert werden, dass alle Parameter  $X_i$  mit gleicher Wichtigkeit in die Suche eingehen. Daher werden normierte Faktoren  $x_i$ ,

$$x_i = \frac{2 \cdot (X_i - X_{i, \text{CP}})}{X_{i, \text{high}} - X_{i, \text{low}}} \quad (5.5)$$

mit

$x_i$  := normierter Wert (Faktor) des physikalischen Parameters  $X_i$

$X_{i, \text{low}}$  := untere Einstellung von Parameter  $X_i$

$X_{i, \text{high}}$  := obere Einstellung von Parameter  $X_i$

$X_{i, \text{CP}}$  := Centerpoint Einstellung von Parameter  $X_i$ ,

eingeführt, die im CCF-Design nur Werte von -1, 0 oder +1 annehmen.

Im Centerpoint  $X_{i, \text{CP}}$ ,

$$X_{i, \text{CP}} = \frac{X_{i, \text{high}} + X_{i, \text{low}}}{2} \quad (5.6)$$

liegt der zugeordnete Faktor  $x_i$  bei 0.

Das neue Faktor/Response-Modell wird nun so umstrukturiert,

$$1 \cdot b_{0r} + \sum_{i=1}^n x_i \cdot b_{ir} + \sum_{1 \leq i < j} x_i \cdot x_j \cdot b_{ijr} + \sum_{i=1}^n x_i^2 \cdot b_{iir} = y_r \quad (5.7)$$

dass eine lineare Gleichung mit  $o$  unbekannten Koeffizienten  $b_{ir}$  entsteht. Bei  $n = 3$  Faktoren sind in diesem Modell  $o = 10$  Koeffizienten unbekannt.

Die werden im vorliegenden Design mit  $n = 14$  Experimenten gesucht, die mit  $k = 1$  bis 14 bezeichnet werden. Dabei werden die CP-Experimente nicht berücksichtigt.

Aus Gl. (5.7) entsteht mit  $n$ -fachem Einsatz der zugeordneten Faktorkombinationen des Designs in Bild 5.3 somit ein überbestimmtes Gleichungssystem,

$$F \cdot k_r = v_r \quad (5.8)$$

mit

$F := (n \times o)$  Design-Matrix der Faktoren und Faktorkombinationen in Gl. (5.7)

$k_r := (o \times 1)$  Spaltenvektor der unbekannten Koeffizienten  $b_{ir}$  in Gl. (5.7)

$v_r := (n \times 1)$  Spaltenvektor der  $n$  experimentellen Ergebnisse von  $y_{rk}$ ,

das über eine *Multiple Linear Regression* (MLR),

$$k_r = [F^T \cdot F]^{-1} \cdot F^T \cdot v_r, \quad (5.9)$$

mit Hilfe der pseudoinversen Matrix von  $F$  gelöst wird.

Die Elemente des Vektors  $k_r$ ,

$$k_r = [b_{0r}, b_{1r}, b_{2r}, b_{3r}, b_{12r}, b_{13r}, b_{23r}, b_{11r}, b_{22r}, b_{33r}]^T, \quad (5.10)$$

beinhalten dann die unbekannten Modellkoeffizienten  $b_{ir}$ .

Auf eine Erläuterung der statistischen Hintergründe dieser Methode wird hier verzichtet. Sie ist bei Erikson [Eriksson et al., 2008] näher beschrieben und wurde von Ellert [Ellert, 2015] an einer Kultivierung zur Maximierung des löslichen Anteils eines zellinternen rekombinannten Pharmaproductes in *E. coli* detailliert dargestellt.

Im Bereich der Herstellung rekombinanter Malariavakzine ist von Fricke [Fricke, 2015] mittels vollautomatischer Optimierungen in einer Multibioreaktoranlage auch der QbD-Gedanke mit eingebbracht.

Hier soll lediglich ein Schnittmuster der weiteren Vorgehensweise erläutert werden.

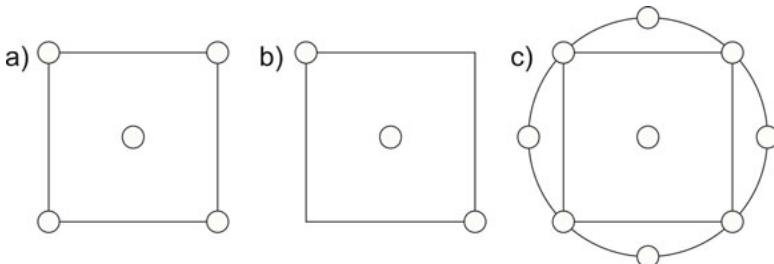


Bild 5.4: Versuchspläne für zwei Faktoren in (a) vollfaktorieller und (b) tealfaktorieller Ausführung sowie als (c) Central Composite Circumscribed (CCC) Design

Die Suche nach einem geeigneten Modell beginnt mit dem Screening, in dem in Gl. (5.7) die quadratischen Terme nicht berücksichtigt werden. Je nach Anzahl der untersuchten Parameter werden unterschiedliche Designs verwendet, von denen die wichtigsten in Bild 5.4 gezeigt sind.

Das vollfaktorielle Design (Bild 5.4a) wird bei relativ wenigen Parametern im Screening genutzt. Bei einer größeren Anzahl ( $n > 3$ ) genügt aufgrund der Orthogonalität des Suchraums auch häufig ein teilstatisches Design (Bild 5.4b).

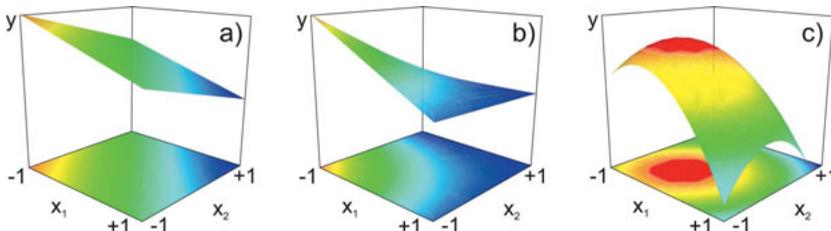


Bild 5.5: Response Surface Plots eines Linearen, Interaktions und Quadratischen Modells mit zwei Faktoren

Mögliche Antworten (*responses*) des Prozesses sind in Bild 5.5 mit dem *Response Surface Plot* und seiner Projektion in die Parameterebene, dem *Contour Plot*, gezeigt.

Beinhaltet das Screening einen nicht relevanten Parameter, so erhält man mit zwei linearen Modellanteilen eine ansteigende oder abfallende Ebene als Response (Bild 5.5a). Bei einem zusätzlichen Interaktionsterm entsteht eine gekrümmte Fläche im Raum (Bild 5.5b). Diese Modelle weisen ggf. in Richtung eines Optimums.

Bei der Suche nach dem Optimum wird der Suchraum verkleinert, das Modell mit den quadratischen Termen wieder komplexer gestaltet und die Anzahl der Experimente erhöht.

Aus diesem Ansatz entsteht ggf. das in Bild 5.5c gezeigte glockenförmige Prozessverhalten, aus dem ein Maximum zu identifizieren ist.

Um dies zu finden, ist das in Bild 5.4c gezeigte *Central Composite Circumscribed* (CCC) Design empfehlenswert.

Nach einer erfolgreichen Modellbildung, die z.B. mit Hilfe der hier genutzten Software MOD-DE (Umetrics, Schweden) vorgenommen wird, ist das Modell über statistische Gütemaße zu bewerten.

Die Güte der Anpassung des Modells (*goodness of fit*)  $R^2$ ,

$$R^2_r = \frac{\sum_{k=1}^n (\hat{y}_{rk} - \bar{y}_r)^2}{\sum_{k=1}^n (y_{rk} - \bar{y}_r)^2} = \frac{SQE_r}{SQT_r} \quad (5.11)$$

mit

$n$  := Anzahl der Gesamtversuche

$y_{rk}$  := gemessene k-te Response  $y_r$

$\hat{y}_{rk}$  := Abschätzung von  $y_{rk}$  über das gefundene Modell

$\bar{y}_r$  := Mittelwert der  $n$  Responses  $y_{rk}$ ,

ergibt sich über die Division der *sum of squares explained* ( $SQE_r$ ) durch die *sum of squares total* ( $SQT_r$ ).

Die Güte der Vorhersagbarkeit des Modells (*goodness of prediction*)  $Q^2$ ,

$$Q^2_r = 1 - \frac{\sum_{k=1}^n (\tilde{y}_{rk} - y_{rk})^2}{SQT_r} \quad (5.12)$$

mit

$\tilde{y}_{rk}$  := Abschätzung von  $y_{rk}$  bei einer Kreuzvalidierung,

ergibt sich aus seiner *Leave one out*-Kreuzvalidierung durch Modellierung der Response  $\tilde{y}_{rk}$  ohne  $y_{rk}$ .

Die Modelvalidität (*model validity*)  $MV_r$ ,

$$MV_r = 1 + 0.57647 \cdot \log(p_{lofr}) \quad (5.13)$$

mit

$p_{lofr}$  := Wahrscheinlichkeitswert aus dem *Lack of fit*-Test,

beschreibt die generelle Eignung des gewählten Modells für die Beschreibung der experimentellen Daten.

Die letzte Kenngröße für die Modellbildung, die Reproduzierbarkeit (*reproducibility*)  $RP_r$ ,

$$RP_r = 1 - \frac{\sum_{i=1}^m (y_{rCPi} - \bar{y}_{rCP})^2}{SQT_r} \cdot \frac{n-1}{m-1} \quad (5.14)$$

mit

$y_{rCPi}$  := gemessene Response  $y_i$  des  $i$ -ten Centerpoint Experiment

$\bar{y}_{rCP}$  := Mittelwert aller Responses der Centerpoint Experimente

$n$  := Anzahl der Gesamtversuche

$m$  := Anzahl der Centerpoint Versuche,

wird über die Centerpoint Experimente und die  $SQT_r$  bestimmt.

Tabelle 5.2 zeigt in welchen Bereichen die Gütemaße für ein exzellentes, brauchbares und unbrauchbares Modell liegen sollten.

Tabelle 5.2: Klassifizierung der DoE-Gütemaße [Fricke, 2015]

Gütemaß	exzellent	brauchbar	unbrauchbar
$R^2$	> 0.9	> 0.5	< 0.5
$Q^2$	> 0.9	> 0.5	< 0.5
$R^2 - Q^2$	< 0.1	< 0.2 – 0.3	> 0.3
$MV$	> 0.25	> 0.25	< 0.25
$RP$	> 0.5	> 0.5	< 0.5

Da diese Werte nur als Anhaltspunkte für eine Bewertung angesehen werden können, sollte jedes Modell individuell auf seine Plausibilität hin untersucht werden.

## 5.4 Lösungsansatz zum Optimierungsproblem der Aufarbeitung mit der EBA

Das zu lösende EBA-Optimierungsproblem ist in Bild 5.6 noch einmal aufgezeigt. Hier stehen zwei Prozessantworten (*responses*  $y_i$ ) fünf normierten Prozessparametern (*factors*  $x_i$ ) gegenüber.

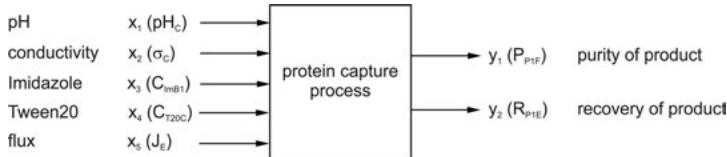


Bild 5.6: Optimierungsproblem der EBA-Chromatographie

Würde man hier ein vollfaktorielles Design für die fünf Faktoren anwenden, so wären 36 Experimente erforderlich. Bei einer In situ-Optimierung im integrierten Prozess würde das 36+1 Tage dauern und ist somit nicht mehr durchführbar.

Daher wurde ein anderer Weg beschritten, dessen Ergebnis in Bild 5.7 vorab gezeigt ist. Zunächst erfolgte in einem teilstatischen Design ein Screening mit den 5 Parametern in einem ÄKTA™purifier 100 (Systembezeichnung P).

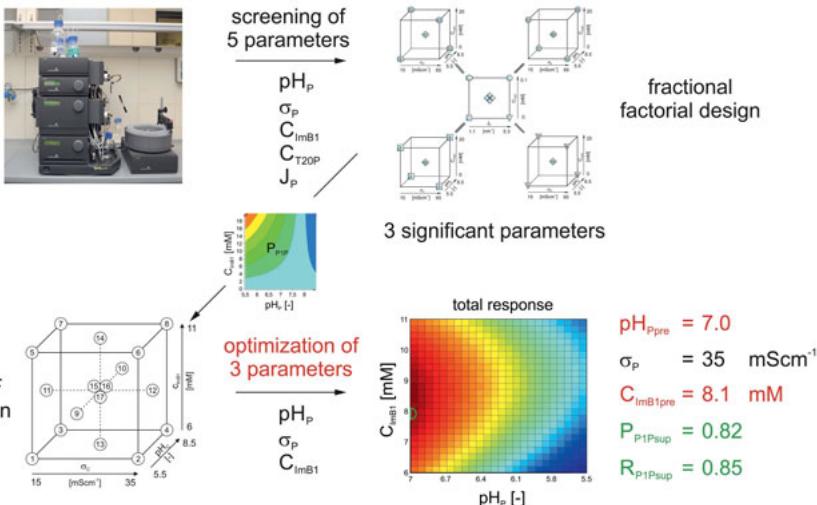


Bild 5.7: Voroptimierung des EBA-Prozesses mit dem ÄKTA™purifier 100

Hieraus ergaben sich mit  $\text{pH}_P$ ,  $\sigma_P$  und  $C_{\text{ImB}1}$  drei signifikante Parameter, mit denen nachfolgend eine Pre-Optimization, ebenfalls im ÄKTA™purifier 100, vorgenommen wurde. Dies erfolgte sowohl für die *purity*  $P_{\text{P1P}}$  als auch für die *recovery*  $R_{\text{P1P}}$ . Aus den Ergebnissen der Einzeloptimierungen wurde dann ein Kompromiss eines gemeinsamen Optimums mit den beiden *Critical Process Parameter*  $\text{pH}_{\text{P1P}}$  und  $C_{\text{ImB1P1P}}$  ermittelt.

Danach erfolgte, um das Maximum der Voroptimierung herum, mit einem CCC Design eine In situ-Optimierung der EBA innerhalb des Integrierten Prozessablaufs. Der gesamte Lösungsweg wird im Weiteren schrittweise erläutert.

## 5.5 Screening des mehrparametrischen Problems im ÄKTA™purifier 100

Bild 5.8 zeigt die *design region* eines  $2^{5-1}$  teilst faktoriellen Screeningplans. Jeder Faktor wird jeweils auf seiner hohen und seiner niedrigen Stufe sowie im Centerpoint untersucht. Das Design sieht dafür 20 Experimente mit unterschiedlichen Faktoreinstellungen vor, in denen 4 Centerpoint Experimente enthalten sind.

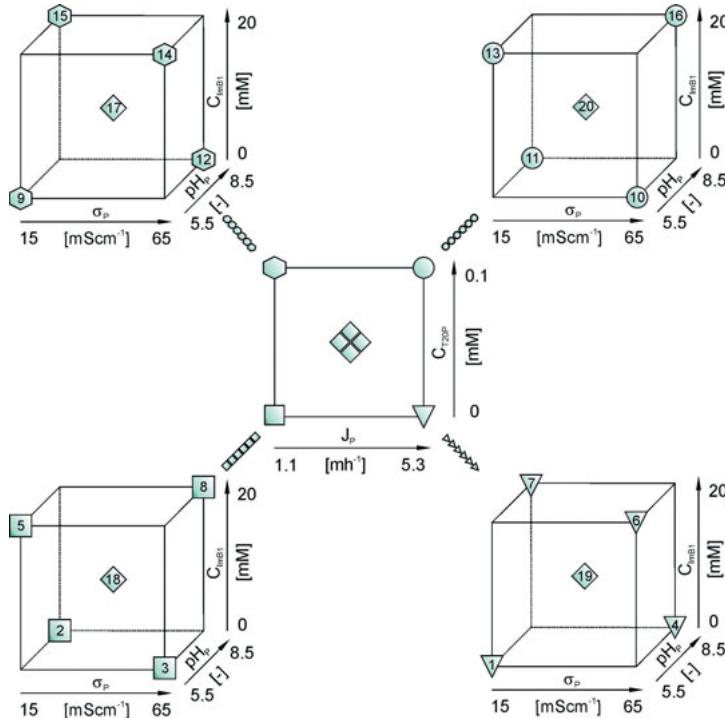


Bild 5.8: Design-Region des  $5^{2-1}$  teilst faktoriellen Screenings

Aus den Symbolen des inneren Quadrates lässt sich die Einstellung der Faktoren  $J_p$  und  $C_{T20P}$  ablesen. Der jeweils zugewandte Würfel zeigt die Werte der drei verbliebenen Faktoren  $pH_p$ ,  $\sigma_p$  und  $C_{ImB1}$  in dem jeweiligen numerisch gekennzeichneten Experiment, das auch in Tabelle 5.3 zu finden ist.

Alle Versuche des Screenings wurden mit 1 ml HisTrap™ FF Crude Säulen (GE Healthcare, München) durchgeführt. Das Säulenmaterial für IMAC wurde mit 2 CV 0,1 M  $\text{Cu}^{2+}$ -Lösung, gefolgt von einem Waschschritt mit 2 CV 50 mM NaAc-Lösung zum Herausspülen ungebundener Kupferionen, aktiviert.

Die Equilibrierung, die Probenaufgabe, das Waschen, die Elution und die Regenerierung der Säule wurden mit einer aufwärtigen Durchströmung (*upflow*) der Säule durchgeführt.

Der Puffer  $B_1$ , war für jedes Experiment individuell anzusetzen, um  $pH_p$ ,  $\sigma_p$  und  $C_{ImB1}$  entsprechend dem Versuchsplan anzupassen. Ebenso waren die Proben entsprechend einzustellen.

Tabelle 5.3: Faktoreinstellung, codierte Werte (in Klammern), und die Response für jedes Experiment k des  $2^{5-1}$  Screening Design

Run k	pH <sub>Pk</sub> [-]	$\sigma_{Pk}$ [mScm <sup>-1</sup> ]	C <sub>ImB1k</sub> [mM]	C <sub>T20PK</sub> [μl <sup>-1</sup> ]	J <sub>Pk</sub> [cmh <sup>-1</sup> ]	P <sub>P1PK</sub> [-]	R <sub>P1PK</sub> [-]
1	5.5 (-1)	15 (-1)	0 (-1)	0 (-1)	530 (+1)	0.053	0.072
2	8.5 (+1)	15 (-1)	0 (-1)	0 (-1)	110 (-1)	0.164	0.233
3	5.5 (-1)	65 (+1)	0 (-1)	0 (-1)	110 (-1)	0.048	0.056
4	8.5 (+1)	65 (+1)	0 (-1)	0 (-1)	530 (+1)	0.069	0.079
5	5.5 (-1)	15 (-1)	20 (+1)	0 (-1)	110 (-1)	0.657	0.645
6	8.5 (+1)	15 (-1)	20 (+1)	0 (-1)	530 (+1)	0.000	0.000
7	5.5 (-1)	65 (+1)	20 (+1)	0 (-1)	530 (+1)	0.518	0.520
8	8.5 (+1)	65 (+1)	20 (+1)	0 (-1)	110 (-1)	0.000	0.000
9	5.5 (-1)	15 (-1)	0 (-1)	20 (+1)	110 (-1)	0.079	0.099
10	8.5 (+1)	15 (-1)	0 (-1)	20 (+1)	530 (+1)	0.297	0.392
11	5.5 (-1)	65 (+1)	0 (-1)	20 (+1)	530 (+1)	0.305	0.362
12	8.5 (+1)	65 (+1)	0 (-1)	20 (+1)	110 (-1)	0.035	0.041
13	5.5 (-1)	15 (-1)	20 (+1)	20 (+1)	530 (+1)	1.000	0.332
14	8.5 (+1)	15 (-1)	20 (+1)	20 (+1)	110 (-1)	0.000	0.000
15	5.5 (-1)	65 (+1)	20 (+1)	20 (+1)	110 (-1)	0.792	0.697
16	8.5 (+1)	65 (+1)	20 (+1)	20 (+1)	530 (+1)	0.000	0.000
17	7.0 (0)	40 (0)	10 (0)	10 (0)	320 (0)	0.947	0.500
18	7.0 (0)	40 (0)	10 (0)	10 (0)	320 (0)	1.000	0.470
19	7.0 (0)	40 (0)	10 (0)	10 (0)	320 (0)	0.882	0.488
20	7.0 (0)	40 (0)	10 (0)	10 (0)	320 (0)	0.965	0.782

Aliquots des Kulturüberstands wurden dafür aufgetaut und Leitfähigkeit und pH-Wert durch Titration mit gesättigter NaCl-Lösung bzw. 2 M NaOH eingestellt sowie Tween 20 ohne Berücksichtigung der Volumenzunahme während dieser Einstellung zugegeben.

Der Puffer B<sub>2</sub> für die Elution enthielt in jedem Experiment 0,5 M NaCl, 20 mM NaH<sub>2</sub>PO<sub>4</sub> · 2H<sub>2</sub>O und 0,1 M Imidazol bei pH 7,4.

Danach war die Säule mit 5 CV Puffer B<sub>1</sub> bei einer konstanten Strömungsgeschwindigkeit F<sub>P</sub> von 1 mlmin<sup>-1</sup> zu equilibrieren. Parallel dazu wurde ein *sample flush* durchgeführt, um alle Schlauchleitungen vom Probenreservoir bis zum Injektionsventil mit der aktuellen Probe zu füllen. Anschließend daran wurden 20 CV der Probe auf die Säule aufgegeben. Die Strömungsgeschwindigkeit F<sub>P</sub> wurde dabei so gewählt, dass sich der erforderliche Flux J<sub>P</sub> einstellte.

Nach einer Probeaufgabe wurde die Säule mit 7 CV von Puffer B<sub>1</sub> gewaschen und das gebundene Protein im Anschluss in 5 CV Puffer B<sub>2</sub> eluiert. Während der Elution wurden Fraktionen mit einem Volumen von 1 ml aufgefangen, von denen die Fraktionen 3 bis 5 für die Prozessbewertung gepoolt wurden.

Um das System auf Startbedingungen zurückzusetzen, wurde die Säule dann mit 5 CV Puffer B<sub>1</sub> regeneriert. Das Waschen, die Elution und die Regenerierung verliefen wiederum bei einer Strömungsgeschwindigkeit F<sub>P</sub> von 1 mlmin<sup>-1</sup>.

In Bild 5.9 sind ausgewählte Experimente des Screening dargestellt. Bild 5.9a und Bild 5.9b beinhalten die ÄKTA-Chromatogramme von zwei Eckversuchen. Beide verdeutlichen den Einfluss der gewählten Faktoren auf das Reinigungsverhalten.

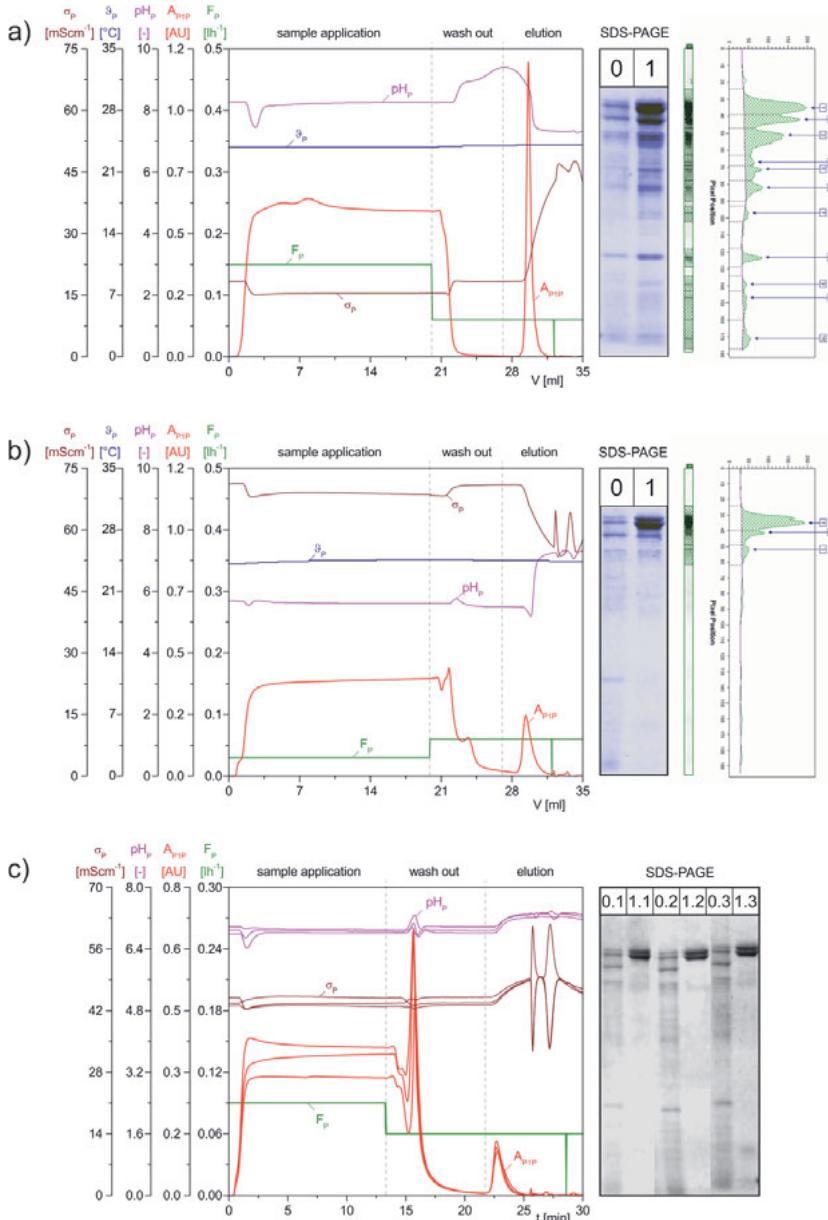


Bild 5.9: ÄKTA-Chromatogramme und zugehörige SDS-PAGE Analysen ausgewählter Screening Experimente, (0) Aufgabe, (1) Elution Fraktionen 3 bis 5

Gezeigt sind Leitfähigkeit  $\sigma_P$ , Temperatur  $\vartheta_P$ , pH-Wert  $pH_P$ , UV-Absorption  $A_{\text{P}1P}$ , und Fließgeschwindigkeit  $F_P$ , jeweils für Lauf 10 (Bild 5.9a) und Lauf 15 (Bild 5.9b). Zusätzlich finden sich eine SDS-PAGE der aufgegebenen Probe (0) sowie der gepoolten Fraktionen (1) und das resultierende densitometrische Diagramm der Produktbanden, jeweils rechts im Bild.

Bild 5.9a zeigt einen hohen, sehr definierten UV-Peak im Verlauf der Elution, der ein gutes Resultat suggeriert. Die SDS-PAGE widerlegt allerdings diese Vermutung, da hier die schlechte Qualität der Aufreinigung deutlich wird. Das Eluat enthält zwar das Zielprodukt, aber mit sehr vielen Verunreinigungen.

Im Gegensatz dazu, ist der UV-Peak der Elution in Bild 5.9b viel niedriger. Die Qualität des Eluats ist aber weitaus besser, da die SDS-PAGE nur wenige Verunreinigungen neben dem Zielprodukt erkennen lässt. Somit wurde verdeutlicht, dass eine Auswertung des Chromatographieschrittes anhand des UV-Signals allein nicht zielführend sein kann, sondern auch eine erweiterte Analyse durch eine SDS-PAGE notwendig ist.

Um eine Zufälligkeit der Ergebnisse auszuschließen, sind in Bild 5.9c drei Centerpoint-Läufe, 18 bis 20, gezeigt, die alle mit derselben mittleren Faktoreinstellung abgelaufen sind. Die Experimente weisen sowohl in ihren Verläufen als auch in der SDS-PAGE eine hohe Reproduzierbarkeit bei gleichbleibender Qualität auf. Zusätzlich kann diese in Bezug auf Reinheit und Recovery als sehr hoch eingestuft werden. Das weist bereits auf eine für den Prozess günstige Faktoreinstellung hin. Insgesamt wird dieser visuelle Eindruck auch durch die in Tabelle 5.3 enthaltenen Ergebnisse von  $P_{P1P}$  und  $R_{P1P}$  unterstützt.

Bei der Modelluntersuchung unter Zuhilfenahme einer *Multiple Linear Regression* (MLR) stellte sich der Einfluss der Faktoren  $J_P$  und  $C_{T20P}$  auf die Qualitätsmerkmale als nicht signifikant heraus.

Daher wurde die nachfolgende *Pre-Optimization* des EBA-Prozesses nur mit den drei signifikanten Faktoren  $p_{H_P}$ ,  $\sigma_P$ , und  $C_{imb1}$ , ebenfalls mit der FPLC, durchgeführt.

## 5.6 Konsekutive Pre-Optimierung mit dem ÄKTA™purifier 100

Bild 5.10 zeigt die Design-Region eines  $2^3$  vollfaktoriellen Versuchsplans für die Pre-Optimierung der EBA-Konditionen mit dem ÄKTA™purifier 100. Der Versuchsplan sieht insgesamt 11 Experimente vor, von denen 8 mit unterschiedlichen Faktorkombinationen und 3 auf *center point*-Niveau durchgeführt wurden.

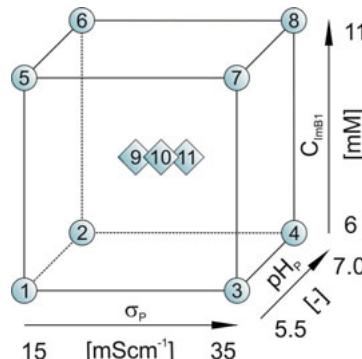


Bild 5.10: Design Region eines  $2^3$  vollfaktoriellen Designs für die Pre-Optimierung

In Tabelle 5.4 sind die Faktoreinstellungen inklusive codierter Werte und die Ergebnisse dieser Experimente aufgeführt.

Tabelle 5.4: Faktoreinstellung, codierte Werte (in Klammern) und Responses für jedes Experiment k des 2<sup>3</sup> Pre-Optimierungs Design

Run k	pH <sub>Pk</sub> [-]	σ <sub>Pk</sub> [mScm <sup>-1</sup> ]	C <sub>ImB1k</sub> [mM]	P <sub>P1Pk</sub> [-]	R <sub>P1Pk</sub> [-]
1	5.5 (-1)	15 (-1)	6 (-1)	0.299	0.201
2	7.0 (+1)	15 (-1)	6 (-1)	0.749	0.611
3	5.5 (-1)	35 (+1)	6 (-1)	0.401	0.298
4	7.0 (+1)	35 (+1)	6 (-1)	0.901	0.426
5	5.5 (-1)	15 (-1)	11 (+1)	0.374	0.205
6	7.0 (+1)	15 (-1)	11 (+1)	0.777	0.729
7	5.5 (-1)	35 (+1)	11 (+1)	0.467	0.377
8	7.0 (+1)	35 (+1)	11 (+1)	0.864	0.567
9	6.3 (0)	25 (0)	8.5 (0)	0.478	0.448
10	6.3 (0)	25 (0)	8.5 (0)	0.616	0.712
11	6.3 (0)	25 (0)	8.5 (0)	0.656	0.672

Für die Modellbildung wurde eine MLR mit einem Konfidenzintervall von 90% angewendet, um die Modellkoeffizienten zu identifizieren und auf ihre Signifikanz hin zu untersuchen.

Bei der Analyse des Modells wurde der Einfluss der Leitfähigkeit σ<sub>P</sub> als weniger signifikant eingestuft, woraufhin diese in späteren Experimenten in der EBA auf ihrem hohen Niveau von 35 mScm<sup>-1</sup> festgehalten wurde.

Somit ergeben sich mit pH<sub>P</sub> und C<sub>ImB1</sub> zwei *Critical Process Parameter* für die Aufreinigung.

Die resultierenden Modelle für die Produktreinheit P<sub>P1P</sub>,

$$P_{P1P} = a_{01} + a_{11} \cdot pH_P + a_{31} \cdot C_{ImB1} \quad (5.15)$$

mit den Koeffizienten

$$a_{01} = -1,397, a_{11} = 0,292 \text{ und } a_{31} = 0,022 \text{ mM}^{-1},$$

und für die Recovery R<sub>P1P</sub>,

$$R_{P1P} = a_{02} + a_{12} \cdot pH_P + a_{32} \cdot C_{ImB1} + a_{132} \cdot pH_P \cdot C_{ImB1} + a_{332} \cdot C_{ImB1}^2 \quad (5.16)$$

mit den Koeffizienten

$$a_{02} = -5,827, a_{12} = 0,558, a_{32} = 0,974 \text{ mM}^{-1}, a_{132} = -0,041 \text{ mM}^{-1} \text{ und } a_{332} = -0,042 \text{ mM}^{-2},$$

sind beide valide für den untersuchten Bereich von pH<sub>P</sub> ∈ [5,5, 7,0] und C<sub>ImB1</sub> ∈ [6, 11] mM.

Die statistischen Kenngrößen der aufgestellten Modelle sind Tabelle 5.5 zu entnehmen. Sie bestätigen die sehr hohe Güte der Modelle und der durchgeföhrten Experimente.

Tabelle 5.5: Modellstatistik für die Pre-Optimierung der EBA-Konditionen

		$P_{P1P}$	$R_{P1P}$
Goodness of fit	$R^2$	0.98	0.94
Goodness of prediction	$Q^2$	0.96	0.78
Difference of fit and prediction	$R^2 - Q^2$	0.02	0.14
model validity	MV	0.84	0.69
reproducibility	RP	0.98	0.98

Bild 5.11 zeigt mit den *Response Surface Plots* und den auf die Grundfläche projizierten *Contour Plots* eine grafische Darstellung der Optimierungsergebnisse.

Darin ist die Problematik der Optimierung einer Aufarbeitung zu erkennen. Der 95 %-Bereich für  $P_{P1P}$  und  $R_{P1P}$  ist jeweils dunkelrot gekennzeichnet. Beide Qualitätsmerkmale erreichen mit  $pH_P = 7$  die höchsten Werte.

Allerdings liegt das Optimum für die *recovery* bei einer Imidazolkonzentration  $C_{ImB1}$  von ca. 8 mM, während die *purity* einen Wert höher als 11 mM anfordert.

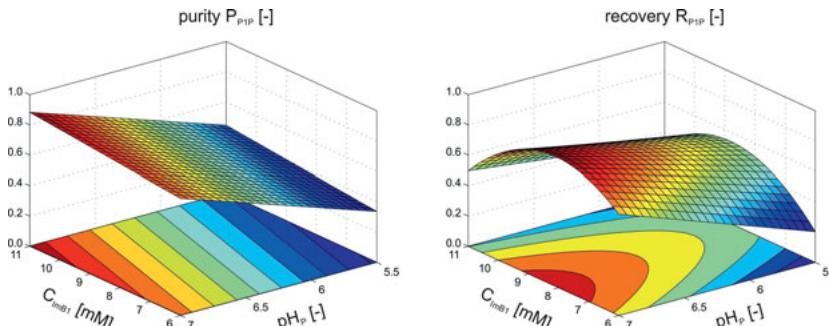


Bild 5.11: *Response Surface Plots* der Pre-Optimierung von Reinheit und Wiederfindung im ÄKTA™purifier 100

Somit ist für die Durchführung eines optimalen Prozesses ein Kompromiss zu finden. Der Weg dahin ist in Bild 5.12 gezeigt.

Im ersten Schritt werden beide Gütekriterien,

$$G_{nP1P} = \frac{G_{P1P} - G_{P1Pmin}}{G_{P1Pmax} - G_{P1Pmin}}, \quad G = P, R \quad (5.17)$$

mit

$G_{nP1P}$  := auf 0 bis 1 normiertes Gütekriterium  $G_{P1P}$

$G_{P1P}$  := Gütekriterium der Aufreinigung

$G_{P1Pmin}$  := unterer Grenzwert von  $G_{P1P}$

$G_{P1Pmax}$  := oberer Grenzwert von  $G_{P1P}$ ,

zwischen 0 und 1 normiert.

In Bild 5.12a erfolgt zunächst deren Überlappung und in Bild 5.12b eine Addition in eine Zusammengesetzte Response  $y_{nP1P}$ ,

$$y_{nP1P} = P_{nP1P} + R_{nP1P}, \quad (5.18)$$

aus dem ein Maximalwert und die zugehörigen suboptimalen Arbeitspunkte  $pH_{p\text{pre}}$  von 7,0 und  $C_{\text{ImB1pre}}$  von 8,1 zu finden sind.

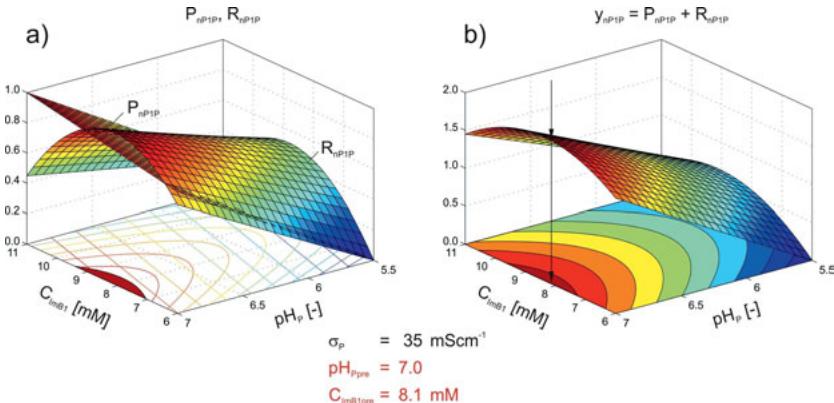


Bild 5.12: Ermittlung eines gewogenen Optimums für *purity*  $P_{P1P}$  und *recovery*  $R_{P1P}$

Hieraus ergeben sich mit den Ausgangsmodellen, Gln. (5.15) und (5.16), die zugehörigen suboptimalen Gütemaße  $P_{P1\text{sup}}$  mit 0.82 und  $R_{P1\text{Psup}}$  mit 0.85 mit dem ÄKTA™purifier 100.

Die zugehörigen Arbeitspunkte der Aufarbeitung bilden nun die Grundlage für die finale Optimierung des Integrierten EBA-Prozesses.

## 5.7 Insitu-Optimierung mit der EBA im integrierten Prozess

Eine finale Optimierung der Produktreinheit  $P_{P1F}$  und der Produktwiederfindung  $R_{P1E}$  erfolgte im Integrierten Prozess mit der EBA. Die Variation der Faktoren  $pH_c$  und  $C_{\text{ImB1}}$  wurde durch eine *Central Composite Circumscribed* (CCC) Design mit drei *center point*-Experimenten vorgegeben.

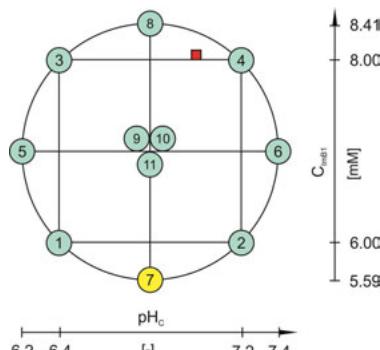


Bild 5.13: *Central Composite Circumscribed* (CCC) Design der EBA Insitu-Optimierung

Bild 5.13 zeigt grafisch den Versuchsplan, in den auch die vorgeschlagenen Parameter der Pre-Optimization mit  $pH_{Pre} = 7,0$  und  $C_{ImB1pre} = 8,1$  eingezeichnet sind. Das Design der Optimierung der EBA wurde mit  $pH_C \in [6,2, 7,4]$  und  $C_{ImB1} \in [5,59, 8,41]$  mM nicht zentriert über diesem suboptimalen Arbeitspunkt gelegt, da davon auszugehen war, dass bei einer zu hohen Imidazolkonzentration das Produkt bereits während des Waschschriffts von der Säule desorbiert wird. Dies beruht darauf, dass es beim Herauswaschen der Zellen zu einer wesentlich längeren Kontaktzeit des Waschpuffers mit dem Säulenmaterial kommt. Dies beinhaltet einen höheren Selektionsdruck durch das Imidazol.

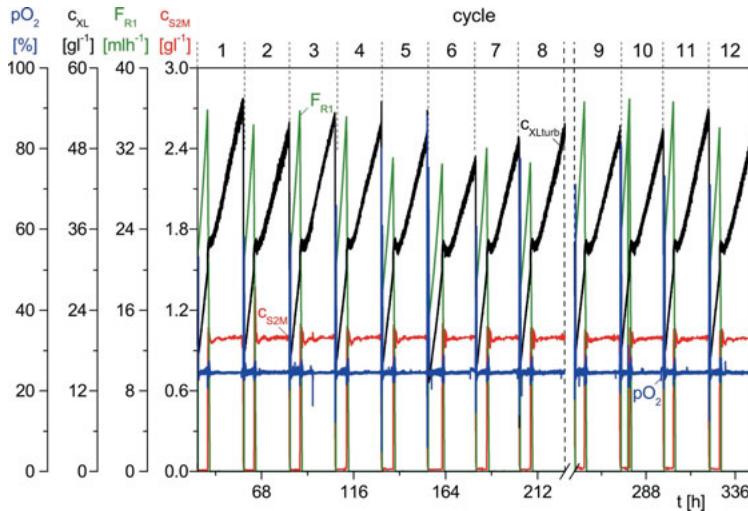


Bild 5.14: Verlauf eines sequentiellen Produktionsprozesses von D1M1 in 12 Zyklen.

Um hieraus ein Optimum für die EBA zu finden, waren 11 ganztägige Experimente notwendig. Eine sequentielle Kultivierung lieferte das Ausgangsmaterial für jeden Optimierungslauf. In Bild 5.14 sind die Verläufe der Prozessgrößen  $pO_2$ ,  $F_{R1}$ ,  $c_{S2M}$  und  $c_{XL,turb}$  in dieser Kultivierung, die eine hohe Reproduzierbarkeit der einzelnen Zyklen erkennen lässt, dargestellt.

Tabelle 5.6: EBA Parametereinstellungen, codierte Faktoren (in Klammern) und Responses des CCC-Designs für jeden Lauf k

Run k	$pH_{Ck}$ [-]	$C_{ImB1k}$ [mM]	$P_{P1Fk}$ [-]	$R_{P1Ek}$ [-]
1	6.40 (-1)	6.00 (-1)	0.725	0.318
2	7.01 (+1)	6.00 (-1)	0.920	0.204
3	6.45 (-1)	8.00 (+1)	0.790	0.144
4	7.17 (+1)	8.00 (+1)	0.855	0.226
5	6.23 (-1.41)	7.00 (0)	0.741	0.575
6	7.21 (+1.41)	7.00 (0)	0.848	0.329
7	6.81 (0)	5.59 (-1.41)	0.873	0.506
8	6.87 (0)	8.41 (+1.41)	0.957	0.196
9	6.84 (0)	7.00 (0)	0.860	0.525
10	6.82 (0)	7.00 (0)	0.804	0.402
11	6.87 (0)	7.00 (0)	0.822	0.415

Die EBA-Optimierungsbedingungen mit den tatsächlich eingestellten Prozessparametern und die Ergebnisse der Aufreinigung sind in Tabelle 5.6 aufgeführt

In Bild 5.15 ist das Ergebnis der beiden MLR-Modellbildungen gezeigt. Hier sind die *Coefficient Plots* mit skalierten und zentrierten Koeffizienten für beide Modelle (a), eine Zusammenfassung der gängigen Modellstatistik zur Bewertung der Modelle in *Summary of Fit Plots* (b) sowie die *Contour Plots* (c) grafisch dargestellt.

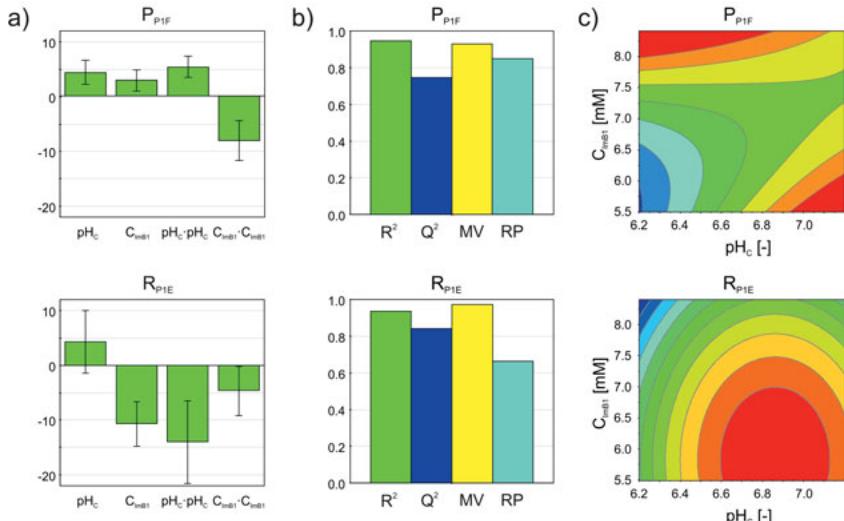


Bild 5.15: Ergebnisse der MLR für die Gütekriterien  $P_{P1F}$  und  $R_{P1E}$

Die damit verbundenen Modellgleichungen für die Produktreinheit  $P_{P1F}$ ,

$$P_{P1F} = a_{01} + a_{11} \cdot pH_c + a_{21} \cdot C_{ImB1} + a_{121} \cdot pH_c \cdot C_{ImB1} + a_{221} \cdot C_{ImB1}^2 \quad (5.19)$$

mit den Koeffizienten

$$a_{01} = -7,363, a_{11} = 1,555, a_{21} = 0,694 \text{ mM}^{-1}, a_{121} = -0,207 \text{ mM}^{-1} \text{ und } a_{221} = 0,053 \text{ mM}^{-2}$$

und für die Recovery  $R_{P1E}$ ,

$$R_{P1E} = a_{02} + a_{12} \cdot pH_c + a_{22} \cdot C_{ImB1} + a_{112} \cdot pH_c^2 + a_{222} \cdot C_{ImB1}^2 \quad (5.20)$$

mit den Koeffizienten

$$a_{02} = -42,106, a_{12} = 11,971, a_{22} = 0,530 \text{ mM}^{-1}, a_{112} = -0,872 \text{ und } a_{222} = -0,045 \text{ mM}^{-2}$$

sind in dem untersuchten Optimierungsraum valide. Beide Modelle wurden an die Daten mit Hilfe der MLR und einem Konfidenzintervall von 90% angepasst.

Die p-Werte für die Regression 0,003 ( $P_{P1F}$ ) und 0,012 ( $R_{P1E}$ ) untermauern die exzellenten Modelle der Aufarbeitungsgüte. Sie zeigen kein *lack of fit* mit  $p_{lof}$ -Werten von 0,74 bzw. 0,90.

Die Kenngrößen des Modells und die Bereiche mit hoher Aufarbeitungsqualität sind in Tabelle 5.7 zusammengefasst.

Tabelle 5.7: Zusammenfassung der Modellqualität

	R <sup>2</sup>	Q <sup>2</sup>	R <sup>2</sup> - Q <sup>2</sup>	MV	RP	pH <sub>opt</sub>	C <sub>ImB1opt</sub>	optimum
P <sub>P1F</sub>	0.98	0.96	0.02	0.84	0.98	6.64	8.36 mM	0.981
R <sub>P1E</sub>	0.94	0.78	0.14	0.69	0.98	6.86	5.83 mM	0.509

Am Ende einer Optimierung kann allerdings nur ein gemeinsamer Sollwert für die einzelnen Betriebsparameter in den Prozess übertragen werden. Bild 5.15c zeigt den Konflikt auf.

Für die *recovery* R<sub>P1E</sub> ist ein klares Maximum bei pH<sub>C</sub> von 6,86 und C<sub>ImB1</sub> von 5,83 mM zu ermitteln. Die resultierende *recovery* R<sub>P1Emax</sub> liegt demnach bei 0,509.

Für die *purity* P<sub>P1F</sub> ergeben sich allerdings zwei diametrale Maximalbereiche im linken oberen und unteren rechten Eck. Das ermittelte Maximum liegt hier bei pH<sub>C</sub> von 6,64 und C<sub>ImB1</sub> von 8,36 mM am oberen Rand und ergibt eine *purity* P<sub>P1Fmax</sub> von 0,981.

Das heißt, hier muss wiederum ein Kompromiss gefunden werden, um im Mittel einen bestmöglichen Arbeitspunkt zu finden.

Um diesen suboptimalen Punkt zu finden, wurde ein *desirability function*-Ansatz nach Derringer und Suich [Derringer and Suich, 1980] verwendet.

Zur Berechnung dieser Funktionen werden P<sub>P1F</sub> und R<sub>P1E</sub> zunächst mit der unteren Grenze L<sub>x</sub> (*lower boundary*),

$$L_x = \begin{cases} 0, & X_{\min} < 0 \\ X_{\min}, & X_{\min} \geq 0 \end{cases}, \quad X = P_{P1F}, R_{P1E} \quad (5.21)$$

und der oberen Grenze U<sub>x</sub> (*upper boundary*),

$$U_x = \begin{cases} X_{\max}, & X_{\max} \leq 1 \\ 1, & X_{\max} > 1 \end{cases}, \quad X = P_{P1F}, R_{P1E} \quad (5.22)$$

zwischen 0 und 1 beschränkt.

Mit diesen Randbedingungen lassen sich für die *purity* des Produktes die *desirability function* d<sub>PP1F</sub>,

$$d_{PP1F} = \begin{cases} 0, & P_{P1F} < L_{P1F} \\ \frac{P_{P1F} - L_{P1F}}{U_{P1F} - L_{P1F}}, & L_{P1F} \leq P_{P1F} \leq U_{P1F}, \\ 1, & P_{P1F} > U_{P1F} \end{cases} \quad (5.23)$$

und für die *recovery* des Produktes die *desirability function* d<sub>RP1E</sub>,

$$d_{RP1E} = \begin{cases} 0, & R_{P1E} < L_{P1E} \\ \frac{R_{P1E} - L_{P1E}}{U_{P1E} - L_{P1E}}, & L_{P1E} \leq R_{P1E} \leq U_{P1E}, \\ 1, & R_{P1E} > U_{P1E} \end{cases} \quad (5.24)$$

leicht berechnen.

In Bild 5.16 sind die beiden *desirability functions* gezeigt. Die Beschreibung  $d_{PP1F}$  der *purity* weist eine konkave Form und die der *recovery*,  $d_{RP1E}$ , eine konvexe Form im Suchraum auf.

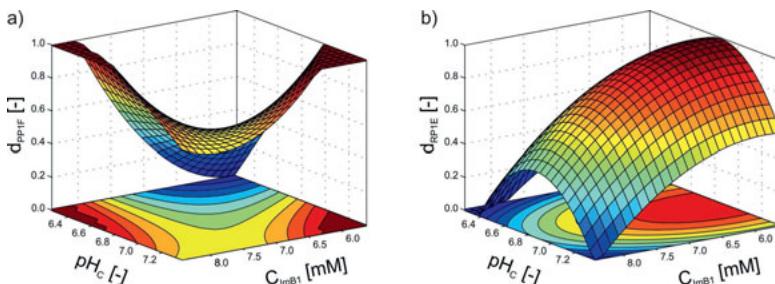


Bild 5.16: *Desirability functions*  $d_{PP1F}$  der *purity* (a) und  $d_{RP1E}$  der *recovery* (b)

Um ein gemeinsames Maximum zu finden, wird über das geometrische Mittel beider Funktionen, die *overall desirability function*  $D_{P1E}$ ,

$$D_{P1E} = \sqrt{d_{PP1F}(P_{P1F}) \cdot d_{RP1E}(R_{P1E})} \quad (5.25)$$

als ein gemeinsames Gütemaß für den EBA-Prozess definiert.

In Bild 5.17 ist mit dem berechneten  $D_{P1E}$  die geometrische Struktur der gemeinsamen Gütefunktion der Aufarbeitung mit der EBA gezeigt. Hieraus wird vorgeschlagen, den Prozess mit  $pH_{Copt}$  von 6,88 und  $C_{ImB1opt}$  von 5,6 mM zu fahren.

Mit diesen Einstellungen sind dann eine submaximale Reinheit  $P_{P1Fsub}$  von 0,87 und eine submaximale Wiederfindung  $R_{P1Esub}$  von 0,51 mit der EBA zu erzielen.

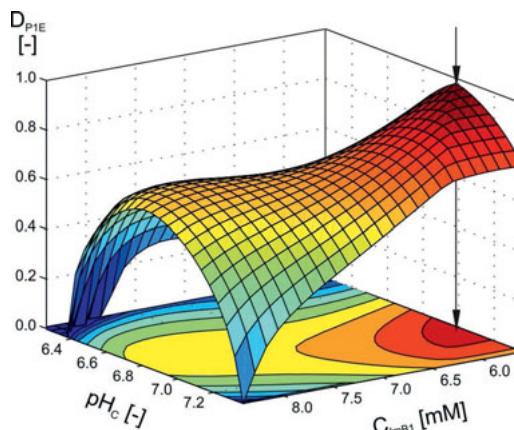


Bild 5.17: *Overall desirability function*  $D_{P1E}$  der Aufarbeitungsprozedur in der EBA

Mit Hilfe eines angemessenen Automatisierungskonzeptes und sauberem Regelverhalten lassen sich Varianzen sowohl zwischen Batches derselben Kampagne als auch zwischen Batches unterschiedlicher Kampagnen minimieren. Trotzdem können die tatsächlichen Werte von Regelgrößen und Einstellparametern von ihren Sollwerten abweichen, wie z.B. bei der pH-Regelung und/oder der Imidazol Konzentrationseinstellung.

Aus diesem Grunde ist es sinnvoll, nicht nur das exakt vorhergesagte Optimum seines Prozesses zu kennen, sondern vielmehr eine Region um dieses Optimum zu enthalten, innerhalb derer vorgegebene Ziele erreicht werden können.

Um solch eine Region zu exponieren, sind in Bild 5.18 drei *Sweet Spot Plots* gezeigt. Sie zeigen grafisch den Bereich in dem die tatsächlichen Werte der Faktoren liegen dürfen, damit beide (rot), eine (grün) oder keine Response (blau) der gesetzten Anforderung von 90 %, 95 % oder 98 % der Maximalwerte  $P_{P1Fmax}$  und  $R_{P1Emax}$  entsprechen.

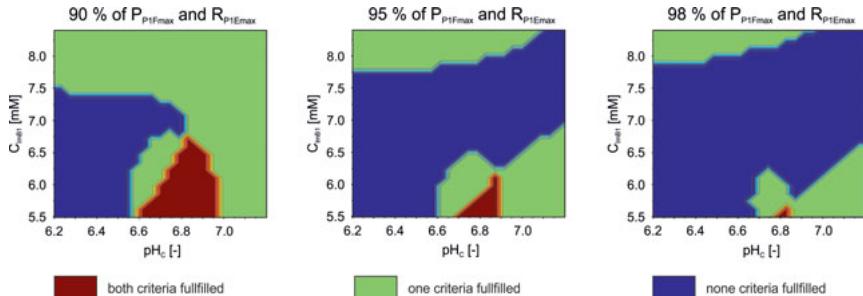


Bild 5.18: *Sweet spot-Plots* für  $P_{P1F}$  und  $R_{P1E}$  bei unterschiedlichen Randbedingungen

Hierin wird die Wahl der als optimal bezeichneten Operationsparameter  $pH_{Copt}$  und  $C_{ImB1opt}$  bestätigt.

## 5.8 Erweiterung der Integrierten Anlage um einen Feinreinigungsschritt

Die Optimierungsläufe des Integrierten EBA-Prozesses in Abschnitt 5.7 wurden auch um eine *Size Exclusion Chromatography* (SEC) zur Feinreinigung des Zielproduktes erweitert.

Auf eine Beschreibung dieser Methode wird hier verzichtet und lediglich ein Integrierter Zyklus sowie das Ergebnis von sechs Validierungzyklen vorgestellt.

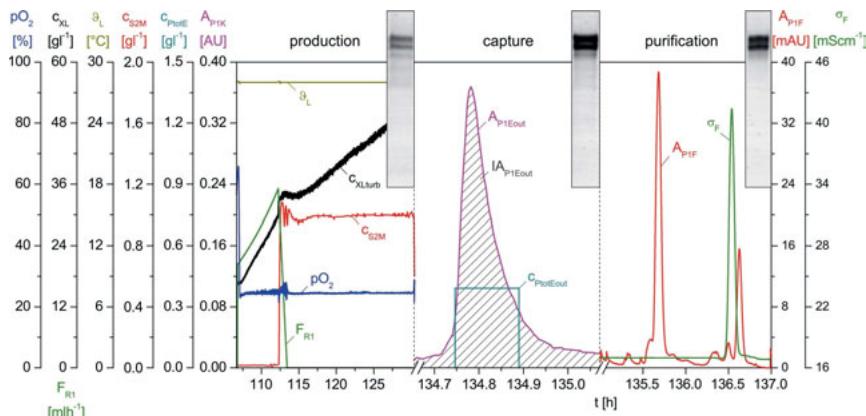


Bild 5.19: Optimaler Zyklus eines 3-stufigen Integrierten Bioprozesses

Das in Tabelle 5.6 und in Bild 5.13 gezeigte Experiment 7 (gelb) ist ein Sternpunkt des CCC-Design. Es befindet sich sehr nah an dem ermittelten Optimum und liegt innerhalb des Be-

reiches des *Sweet spot*-Plots, in dem *recovery* und *purity* 98 % ihres Maximalwertes erreichen.

Der Verlauf dieser Kultivierung, die EBA-Elution und die Feinreinigung des Eluats mit der SEC im ÄKTA™purifier 100 sind inklusive der SDS-PAGEs der jeweiligen Endprobe in Bild 5.19 gezeigt. Die zugehörigen Prozesseinstellungen und die Ergebnisse der Aufarbeitung sind in Tabelle 5.8 aufgeführt.

Tabelle 5.8: Einstellungen und Ergebnisse des optimierten Integrierten Zyklus in Bild 5.19

Variable	Value	Unit
$\vartheta_L$	28.0	°C
$pH_M$	5.5	-
$C_{S2M}$	1.0	g l <sup>-1</sup>
$J_E$	200.0	cm h <sup>-1</sup>
$C_{T20C}$	0.1	mM
$\sigma_C$	35.0	mScm <sup>-1</sup>
$pH_C$	6.8	-
$C_{ImB1}$	5.6	mM
$C_{P1F}$	0.3	g l <sup>-1</sup>
$m_{P1F}$	0.2	g
$P_{P1F}$	0.873	-
$R_{P1E}$	0.506	-
$P_{P1P}$	0.933	-
$R_{P1P}$	0.804	-

Um das erzielte Ergebnis in Tabelle 5.8 zu bestätigen, erfolgte ein Quasi-Robustheitstest. Dazu wurden sechs Zyklen des Integrierten Prozesses mit den optimalen Faktoreinstellungen durchgeführt und auf die Güte der Aufarbeitung hin untersucht. Die SDS-PAGE für die Prozessschritte *production*, *IMAC-capture* und *SEC-purification* sind in Bild 5.20 gezeigt. Aus den Messdaten der wiederholten Prozessschritte wurden die in Tabelle 5.9 dargestellten arithmetischen Mittel berechnet und zum Vergleich mit dem vorherigen Optimierungsergebnis herangezogen. Hier konnte eine hohe Übereinstimmung aller Werte festgestellt werden.

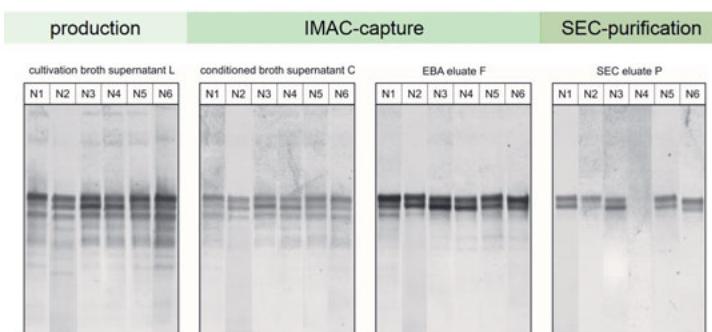


Bild 5.20: SDS-PAGEs des optimierten Integrierten Bioprozesses mit 6 Zyklen

Tabelle 5.9: Mittlere Ergebnisse der sechs Zyklen eines optimalen Integrierten EBA-Prozesses

Variable	Value	Unit
$\bar{c}_{PtotM}$	312	$mgl^{-1}$
$\bar{P}_{P1M}$	0.45	
$\bar{c}_{PtoIC}$	120	$mgl^{-1}$
$\bar{P}_{P1C}$	0.45	
$\bar{c}_{PtotF}$	240	$mgl^{-1}$
$\bar{P}_{P1F}$	0.72	
$\bar{R}_{P1E}$	0.64	
$\bar{c}_{PtotP}$	96	$mgl^{-1}$
$\bar{P}_{P1P}$	0.98	
$\bar{R}_{P1P}$	0.77	
$\bar{R}_{P1int}$	0.49	

Der in Tabelle 5.9 dargestellte Gesamtprozess zeigt bereits eine relativ hohe Reinheit  $P_{P1M}$  von 0,45 im Upstream. Die Reinheiten  $P_{P1F}$  von 0,72 nach der EBA und  $P_{P1P}$  von 0,98 nach der Chromatographie sind auch zufriedenstellend. Die Wiederfindung in der EBA (0,64) und im ÄKTA (0,77) führen auf einen Gesamtverlust von 51 %. Diesen gilt es noch zu minimieren.

Abschließend lässt sich zu diesem Kapitel festhalten, dass die Integration einer EBA-Chromatographie als primärer Aufarbeitungsschritt in einer sequentiellen Kultivierung des Impfstoffkandidaten D1M1 eine schnelle und reproduzierbare Insitu-Trennung des rekombinannten Proteins von Zellen, Wirtsproteinen und Medienkomponenten ermöglicht. Des Weiteren konnte durch eine SEC nochmals eine verbesserte Reinheit erzielt werden, was diesen Schritt als geeignete Feinreinigungsmethode bestätigt.

Ergebnisse aus Scale Down-Experimenten haben gezeigt, dass eine Vorabbestimmung der maßgeblichen Faktoren in einer Pre-Optimierung im FPLC-System die eigentliche Optimierung im präparativen EBA-Maßstab vereinfachen und beschleunigen kann. Die Software MODDE® hat sich dabei als eine ausgezeichnete Möglichkeit zur Planung und Auswertung von Versuchsreihen bewiesen.

## 6. Aufbau und Erprobung eines quasi-kontinuierlichen Bioprozesses

### 6.1 Entwicklung einer industrie-kompatiblen integrierten Produktionsanlage

In diesem Kapitel wird der Aufbau eines vollständig integrierten Bioprozesses beschrieben, der aus scale-down Komponenten industrieller Produktionsanlagen zusammengesetzt ist.

Ein Ziel dieser Forschungsarbeiten bestand darin, einen Großteil der im Bild 6.1 vorgegebenen Zielsetzungen der FDA – *United States Food and Drug Administration* für eine wissensbasierte Bioverfahrenstechnik umzusetzen. Diese sind unter dem Begriff PAT – *Process Analytical Technology* zusammengefasst [Bondi and Drennen, 2011].



Bild 6.1: Vorgaben der PAT-Initiative der FDA [Valero, 2013]

Neben der Entwicklung quasi-kontinuierlicher Prozessführungen auf der Basis konventioneller Batch-Operationen stand dabei insbesondere der Einsatz von MVDA – *Multivariate Data Analysis* zur Analyse des Bioreaktionsprozesses, dem Monitoring nicht direkt messbarer biotechnologischer Variablen und der Bewertung und Überwachung der Reproduktion zyklischer Abläufe im Vordergrund dieser Arbeit.

Hierzu wurden kommerziell erhältliche Softwarepakete eingesetzt, die im Weiteren noch erläutert werden.

Zunächst wurde allerdings in einem BMBF-Forschungsprojekt *Entwicklung vollautomatischer Integrierter Bioprozesse in einer industrie-kompatiblen Pilotanlage zur Herstellung rekombinanter Enzyme* im BMBF-Forschungscluster *Biokatalyse 2021* die in Bild 6.2 gezeigte Anlage für eine vollautomatische Herstellung einer rekombinanten Lipase CALB mit *Pichia pastoris* entwickelt und erprobt.

Diese Entwicklungsarbeiten wurden von Kai Lögering und Christian Müller vorgenommen und sind in [Lögering & Müller et al., 2011; Lögering, 2015] detailliert beschrieben. Die Anlage wurde dabei vollständig automatisiert, um die einzelnen Prozessschritte gekoppelt seriell abfahren zu können.

Zum Aufbau eines quasi-kontinuierlichen Produktionsprozesses wurden dabei die beiden Unit-Operationen im Upstream und der gesamte Downstream-Prozess, jeweils um einen Tag versetzt, auch parallel gefahren. Diese sequentiell/parallel genannte Fahrweise eröffnet ein hohes Potential an Prozessoptimierungen [Lögering & Müller et al., 2010].



Bild 6.2: Scale-down Forschungsanlage zur Herstellung rekombinanter Proteine

Die Nutzung dieser Anlage zur Herstellung von Malariavakzinen erforderte dann die Erweiterung der Downstream-Operationen um Schritte zur Reinigung der rekombinanten Produkte.

Der Aufbau des erweiterten integrierten Prozesses wird nachfolgend kurz beschrieben, während der eigentliche Pharmaprozess und die Lösung von QbD-Vorgaben der FDA hinsichtlich der Beobachtung und Reproduzierbarkeit sequentiell/paralleler Prozesse detailliert erläutert wird.

Hierzu war die komplexe Infrastruktur einer erweiterten Datenverarbeitung zu installieren. Deren Anwendung wird in den nachfolgenden Abschnitten beschrieben.

## 6.2 Konzept eines sechsstufigen integrierten Prozesses

Der Produktionsprozess ist aus den sechs Schritten, Zellanzucht (*cell breeding*), Protein-Expression (*protein production*), Klärung (*separation*), Querstrom-Mikrofiltration (*microfiltration*), Querstrom-Ultrafiltration (*ultrafiltration*) und einer chromatographischen Feinreinigung (*purification*), zusammengesetzt. Zellanzucht und Expression bilden den Upstreamprozess (USP) und nehmen jeweils genau 24 Stunden in Anspruch. Der gesamte 4-stufige Downstreamprozess (DSP) dauert hingegen nur 10 Stunden.

Bild 6.3 veranschaulicht den gesamten Produktionsprozess und die beteiligten Prozesseinheiten. Die Bioreaktoren, BIOSTAT® ED10 (B. Braun Biotech International GmbH, Melsungen) und BIOSTAT® C30 (Sartorius Stedim Systems, Guxhagen), werden jeweils durch eine digitale MSR-Einheit (DCU II / DCU III) geführt. Diese sind an das übergeordneten SCADA-System MFCS/win angeschlossenen (Sartorius Stedim Biotech, Guxhagen). Der USP wird über ISA-S88-Rezepte, die in MFSC/win hinterlegt sind, automatisiert.

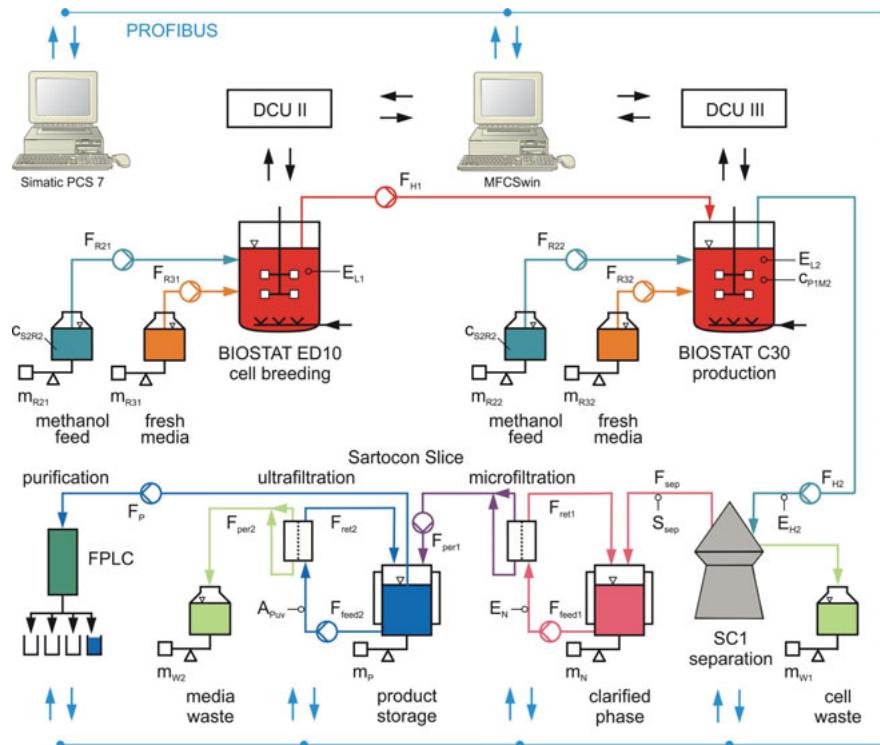


Bild 6.3: Schematische Darstellung des sechs-stufigen integrierten Produktionsprozesses

Der Separator SC1 (GEA Westfalia, Oelde) verfügt über eine eigene Steuereinheit SIMATIC S5 (Siemens, Karlsruhe) auf der alle Automatisierungsschritte, die diesen Prozessschritt betreffen, bereits hinterlegt und damit abrufbar sind.

Die Automatisierung der beiden Filtrationsschritte wird über Batchrezepte gesteuert, die auf der DCU III hinterlegt sind.

Das AKTApurifier 100 FPLC-System verfügt über eine eigene Steuerungssoftware UNICORN 6 (beides GE Healthcare, Schweden). Für eine Automatisierung des Feinreinigungsschrittes mit IMAC wurden ebenfalls Steuerezepte auf diesem System implementiert.

Eine Integration aller Prozess- und Automatisierungsschritte ist mit dem Prozessleitsystem SIMATIC PCS 7 (Siemens, Karlsruhe) realisiert [Lögering, 2015]. Durch einen Austausch von Prozessvariablen (*process phases*) können aus MFCS/win heraus Automatisierungsschritte, die in SIMATIC PCS 7 hinterlegt sind, angestoßen werden – und umgekehrt. Prozessdaten und Variablen werden dabei über eine OPC-Schnittstelle dem jeweils anderen System zur Verfügung gestellt.

Eine genaue Erläuterung der Datenverarbeitungsinfrastruktur erfolgt in Abschnitt 6.6.

## 6.3 Die Upstream Unit-Operationen

### 6.3.1 Aufbau der zweistufigen Upstream-Anlage

In Bild 6.4 wird der Aufbau und die Kopplung der beiden Upstreamschritte untereinander sowie mit dem ersten Downstreamschritt gezeigt.

Der *cell breeding* und der *protein production* Reaktor sind miteinander fest verrohrt. Die Ernterate  $F_{H1}$  wird nur über den Überdruck in Reaktor 1 durch eine getaktete Steuerung  $D_{H1}$  des Ernteventils bewirkt. Die Menge an transferierter Brühe wird über die Wägungen  $V_{L1}$  und  $V_{L2}$  gesteuert.

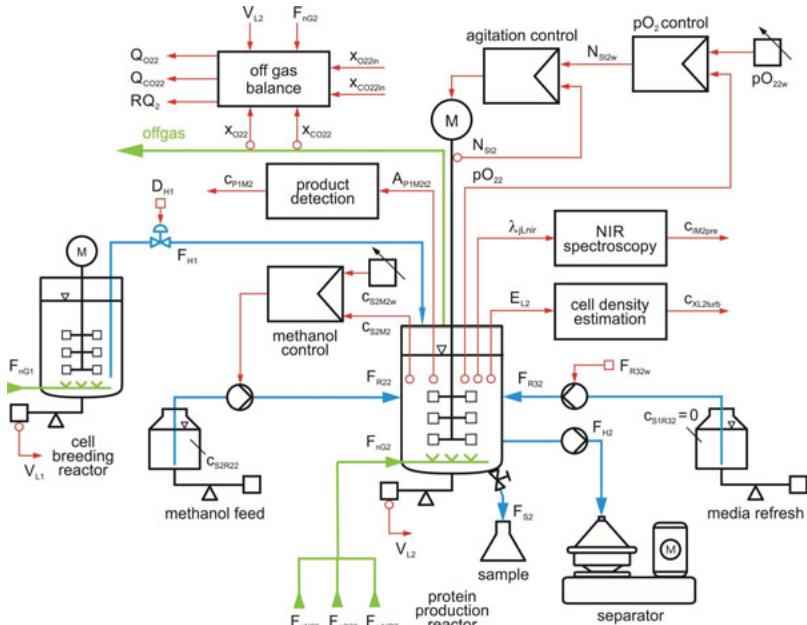


Bild 6.4: Erweiterte Instrumentierung des *protein production* Reaktors

Reaktor 2 und der Separator sind durch eine flexible Stahlleitung verbunden. Die Ernte erfolgt über die Eingangspumpe des Separators, die somit die Ernterate  $F_{H_2}$  steuert.

Beide Reaktoren haben die gleiche Grundausstattung wie der in Abschnitt 3.2 beschriebene BIORSTAT® ED5. Die Bestückung mit einer NIR-Inline Sonde und einer Atline-HPLC zur Produktmessung besteht nur im Produktionsreaktor. Die erweiterte Instrumentierung wird in Abschnitt 6.4 näher erläutert.

### 6.3.2 Die zweistufige zyklische Produktionsstrategie

Für eine Erläuterung der sequentiellen Produktionsstrategie in beiden Reaktoren ist der zweite Upstream-Zyklus einer längeren Produktionskampagne in Bild 6.5 dargestellt. Er teilt sich in eine wiederholte Zellanzucht und eine Protein-Produktion, die voneinander entkoppelt in zwei Bioreaktoren ablaufen.

In beiden Reaktoren werden die Zelldichten  $c_{XLp}$ , rekonstruiert aus den Inline-Trübungssignalen  $E_{lp}$  (optek-Danulat, Essen), und die Reaktionsvolumina  $V_{lp}$  online gemessen.

Bild 6.5 zeigt, wie der Zellanzuchtprozess mit einer etwa 10 stündigen exponentiellen Wachstumsphase beginnt. In dieser Phase wächst *Pichia pastoris* unlimitiert auf Glycerol. Dabei bauen die Zellen den im vorherigen Zyklus gebildeten zellinternen Gehalt an  $P_1$  und AOX wieder ab [Fricke et al., 2013].

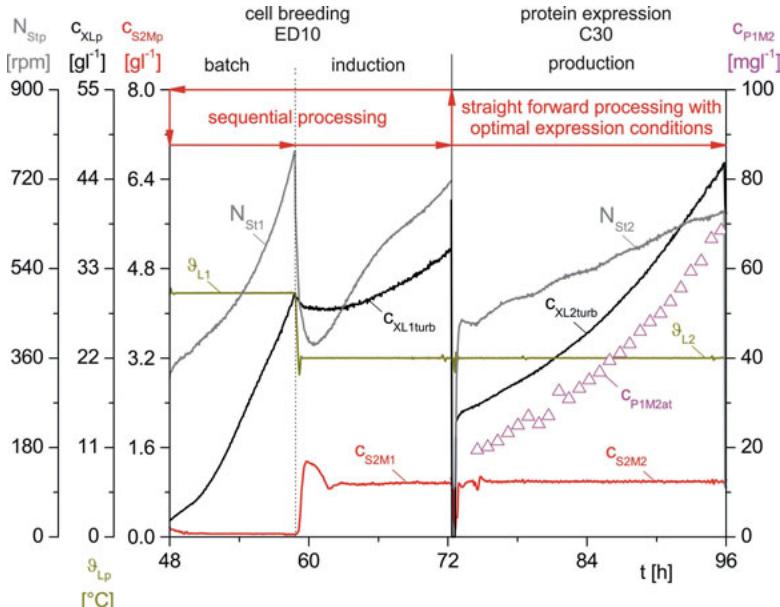


Bild 6.5: Ein Zyklus der automatisierten Produktion rekombinanter Proteine in einem integrierten Bioprozess

In der darauffolgenden 14-stündigen Pre-Induktionsphase, in der Methanol als einzige Kohlenstoffquelle kontrolliert zu gefüttert wird, stellen sich die Zellen erneut auf die bevorstehende Proteinexpression im Produktionsreaktor um.

Zu Beginn der Zellanzucht wird der Reaktor mit frischem Medium mit einer Glycerolkonzentration von  $43,6 \text{ gl}^{-1}$  versorgt. Bei einer Temperatur von  $28^\circ\text{C}$  wachsen die Zellen unlimitiert in einer Batchphase an, bis das vorgelegte Substrat vollständig verbraucht ist.

Anschließend wird der Prozess in einem Fed-Batch-Modus mit geregelter Methanolkonzentration gefahren, bis eine Zelldichte von etwa  $38 \text{ gl}^{-1}$  erreicht ist. In diesem Teil des Prozesses finden eine Produktion von zellinterner Alkoholoxidase (AOX) und parallel dazu die Induktion der Expression des Zielproteins D1M1H ( $P_1$ ) statt. Dadurch werden Zellen, die bereits an einen Methanolstoffwechsel gewöhnt sind, für den nachfolgenden Produktionsprozess in dem zweiten Bioreaktor bereitgestellt.

In dem Produktionsreaktor wird eine Zellmasse von 225 g benötigt, um eine Startzelldichte  $c_{XL20}$  von  $15 \text{ gl}^{-1}$  bei einem Startvolumen  $V_{L20}$  von 15 l zu erreichen.

Dies wird durch den Transfer eines abgestimmten Volumens von Zellbrühe  $V_{L1\text{trans}}$ ,

$$V_{L1\text{trans}} = \frac{c_{XL20} \cdot V_{L20}}{c_{XL1n}} \quad (6.1)$$

mit

$c_{XL1n}$  := Zelldichte am Ende des Zellanzuchzyklus in Reaktionsphase L<sub>1</sub> [gl<sup>-1</sup>]

$V_{L1\text{trans}}$  := Zellbrühevolumen das von Reaktor 1 zu Reaktor 2 überführt wird [l]

$c_{XL20}$  := Startzelldichte des folgenden Produktionszyklus in Reaktionsphase L<sub>2</sub> [gl<sup>-1</sup>]

$V_{L20}$  := Startvolumen des folgenden Produktionszyklus in Reaktionsphase L<sub>2</sub> [l],

von Reaktor 1 zu Reaktor 2 am Ende eines Zellanzuchzyklus realisiert.

Unmittelbar nach einem Transfer wird die Zellbrühe in Reaktor 2 mit frischem Medium, das keine Kohlenstoffquelle enthält, bis zu seinem festgelegten Anfangsvolumen  $V_{L20}$  aufgefüllt. Die Zellen werden dann für weitere 24 h über eine geregelte Methanolkonzentration  $c_{S2M2}$  im Medium M<sub>2</sub> mit Substrat versorgt. In dieser Phase produzieren sie das Zielprodukt P<sub>1</sub> bei einer optimalen Expressionstemperatur von 22 °C. Am Ende eines Produktionszyklus wird Reaktor 2 vollständig abgeerntet, wobei die geerntete Kulturbrühe in die Downstreamlinie des Prozesses gelangt.

Die verbliebene Zellbrühe in Reaktor 1 wird über ein Steigrohr bis auf ein Volumen  $V_{L1\text{rem}}$  von 700 ml abgeerntet. Dabei wird der Überschuss verworfen. Anschließend wird der Reaktor durch Befüllen mit glycerolhaltigem Batchmedium auf sein Startvolumen  $V_{L10}$  von 9 l für den darauffolgenden Zellanzuchzyklus vorbereitet.

Dieser Vorgang resultiert in einer Zelldichte  $c_{XL10}$ ,

$$c_{XL10} = \frac{c_{XL1n} \cdot V_{L1\text{rem}}}{V_{L10}} \quad (6.2)$$

mit

$c_{XL1n}$  := Zelldichte am Ende der vorherigen Zellanzucht in Reaktionsphase L<sub>1</sub> [gl<sup>-1</sup>]

$V_{L1\text{rem}}$  := Zellbrühevolumen das nach Transfer und Verwerfen in Reaktor 1 verbleibt [l]

$c_{XL10}$  := Startzelldichte des nächsten Zellanzuchzyklus in Reaktionsphase L<sub>1</sub> [gl<sup>-1</sup>]

$V_{L10}$  := Startvolumen des nächsten Zellanzuchzyklus in Reaktionsphase L<sub>1</sub> [l].

Mit dieser Strategie startet die sequentiell betriebene Zellanzucht in jedem folgenden Zyklus mit einer anfänglichen Zelldichte  $c_{XL10}$  von etwa 3 gl<sup>-1</sup>. Die Ernte der Bioreaktoren wird mit dampfsterilisierbaren, totraumarmen Verbindungen und Leitungen ermöglicht, um das Risiko einer Kontamination zu minimieren.

Die Proteinproduktion ist ein geradliniger Methanol Fed-Batch-Prozess [Lögering & Müller et al., 2010]. Hier wird durch eine vollautomatische Prozessstrategie mit einer Steuerung der Ausgangsbedingungen für jeden Zyklus, einhergehend mit einer exakten Einstellung der kritischen Prozessparameter, ein reproduzierbarer und QbD-konformer Prozessablauf garantiert.

Die Reproduzierbarkeit des Prozesses in Bild 6.6 stützt die These, dass die abwechselnde Zugabe von Glycerol und Methanol als einziger Kohlenstoffquelle und die damit verbundene wiederkehrende Umstellung im Stoffwechsel der Zellen dazu führt, diese über einen sehr langen Zeitraum vital und produktiv zu halten [Fricke et al., 2013].

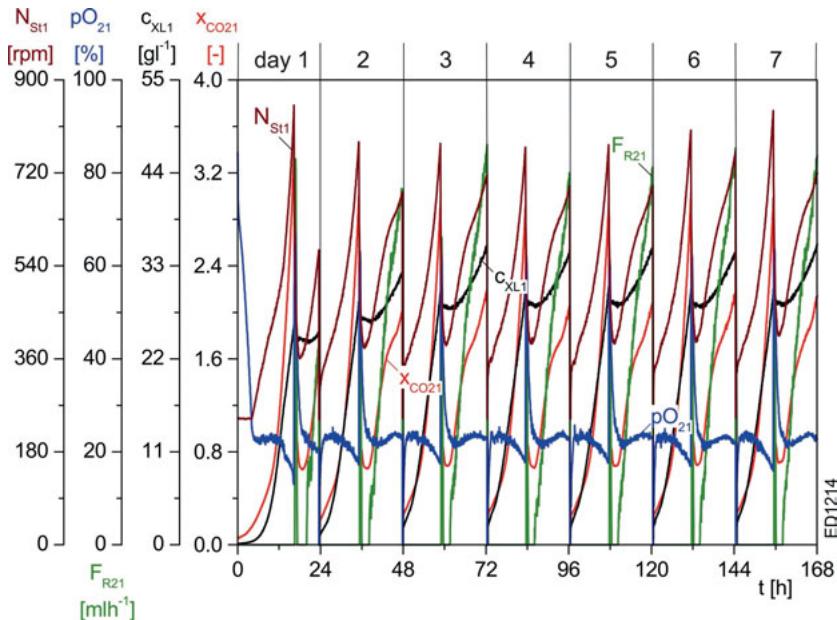


Bild 6.6: Zellanzuchtprozess aus einer Kampagne mit sechs aufeinanderfolgenden Zyklen

## 6.4 Eingesetzte Mess-, Steuer- und Regelungstechnik

### 6.4.1 Grundautomatisierung der Bioreaktoren

Die Grundautomatisierung wird für den BIOSTAT® ED10 mit einer DCU II und für den BIOSTAT® C30 mit einer DCU III vorgenommen. Die gemessenen Online-Variablen sind Tabelle 3.1 in Abschnitt 3.2 zu entnehmen. Geregelt werden die Inline-Größen Temperatur  $9_{LP}$ ,  $pH_P$  und  $pO_{2P}$  in der Flüssigphase sowie die Online-Größen Gasdruck  $p_{GP}$  im Kopfraum, Drehzahl  $N_{StP}$  und Begasungsrate  $F_{nAlR_P}$ .

Hervor zu heben ist, dass alle Vorlagen, die Reaktoren und alle Erntebehälter auf Waagen stehen, um die Prozesse vollständig bilanzieren zu können.

### 6.4.2 Messung des Sekundärsubstrates Methanol

Die Methanol-Konzentrationen  $c_{S2M_P}$  wurden unter Verwendung von Inline-Sonden (Biotechnologie Kempe, Kleinmachnow) über eine Diffusion von Methanol durch eine hydrophobe Membran in einen begasten Messraum (Tubingprinzip) mit einem Figarohalbleiter gemessen (<http://www.trace.de/methanol.html>) und über Zufütterraten  $F_{R2P}$  für eine Methanolzugabe geregelt.

Das Ergebnis einer Regelung der Substratkonzentration  $c_{S2M_2}$  im Produktionsreaktor C30 ist in Bild 6.7 dargestellt. Dadurch lassen sich über mehrere Zyklen stabile Expressionsbedingungen für die Mikroorganismen bereithalten, was eine Grundvoraussetzung für eine gleichbleibende Produktexpression darstellt.

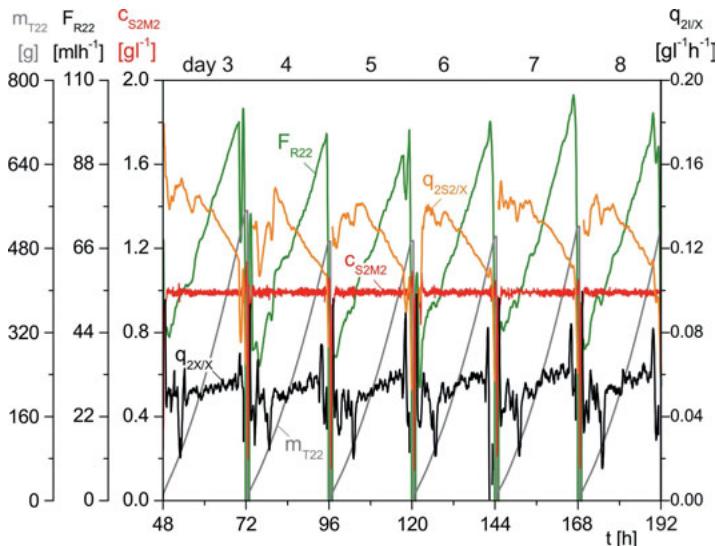


Bild 6.7: Methanolregelung im Produktionsreaktor über mehrere Zyklen

#### 6.4.3 Abgasmessung und Gasbilanzen

Für jeden Reaktor wurde eine Abgas-Messstation mit O<sub>2</sub>- und CO<sub>2</sub>-Sensoren (BlueSens, Herten) für eine Online-Berechnung der Sauerstoffversorgungsrate Q<sub>O2p</sub>, der Kohlendioxidbildungsrat Q<sub>CO2p</sub> und dem Respirations-Quotienten RQ<sub>p</sub> (s. Abschnitt 3.2) installiert.

#### 6.4.4 Einsatz einer NIR-Spektroskopie

Zur erweiterten Beobachtung des Prozesses wurde ein NIR-Spektrometer (Polytec, Deutschland) installiert. Der Messaufbau besteht aus einer Lichtquelle, einer mehrsträngigen Glasfaserleitung, einem Sondenkörper sowie einem Detektor.

Der Sondenkörper wird am Produktionsreaktor über einen 25 mm Seitenport im unteren Stutzenkranz angeschlossen. Er ist hohl und verfügt an seinem Ende über ein Fenster. Das Ende der Glasfaserleitung wird in den Sondenkörper eingeführt und endet ca. einen halben Millimeter vor dem Fenster. Das von der Lichtquelle ausgestrahlte Licht wird über den äußeren Faserring geführt und das reflektierte Licht wird im zentralen Strang zurück zum Detektor geleitet. Das Messprinzip beruht dabei auf einer diffusen Reflexion.

Es wird ein Spektrum im Bereich von 850 nm bis 1650 nm in einem Intervall von 2 nm aufgenommen. Durch Akkumulation direkt hintereinander gemessener Spektren wird der stochastische Fehler verringert. In Bild 6.8 ist der Verlauf der Spektren für einen Produktionszyklus dargestellt. Mit zunehmender Prozesszeit ist eine Abnahme der Absorption zu beobachten, was auf eine Zunahme der Reflektionen durch das Medium zurückzuführen ist.

Die Spektren eines Prozesses sind, ähnlich dem Fingerabdruck eines Menschen, charakteristisch für den Prozess und können daher auch ohne eine aufwendige Kalibrierung zu einer erweiterten Prozessbeobachtung genutzt werden. So können zum Beispiel über eine Hauptkomponentenanalyse eines Spektrums die relevanten Informationen extrahiert und u.a. für eine Bewertung der Reproduzierbarkeit herangezogen werden.

Darüber hinaus lassen sich über eine multivariate Kalibrierung der spektroskopischen Messungen zum Beispiel Zellmasse, Glycerol und sekretierte Proteine sowie der zellinterne AOX-Gehalt messen [Voss, J.-P. et al., 2014]. Hierzu sind dann allerdings intensive Offline-Analysen der betreffenden Verfahrensvariablen notwendig, die als Grundlage für eine solche Kalibrierung dienen.

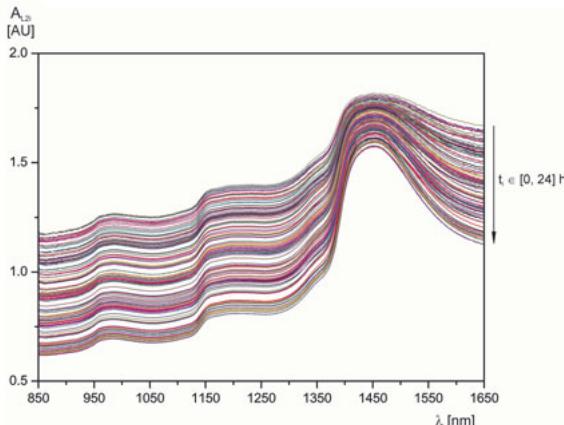


Bild 6.8: Verlauf der NIR-Spektren für einen Produktionszyklus

Bild 6.9 zeigt den NIR-Fingerprint einer Produktionskampagne mit sechs Zyklen. Gezeigt sind die klassischen Online-Messgrößen  $N_{S12}$ ,  $pO_{22}$  und  $c_{S2M2}$ , die online berechneten Variablen  $c_{XL2turb}$  und  $Q_{O22}$ , die Atline-Variable  $c_{P1M2}$  sowie die Absorption  $A_{L2}$  der NIR-Anregung  $\lambda_{nir}$ . Alle Variablen zeigen ähnliche Verläufe in den sich wiederholenden Zyklen.

Aus der gezeigten Darstellung lässt sich der erhöhte Informationsgehalt, den eine spektroskopische Messung für einen Bioprozess liefern kann, ablesen.

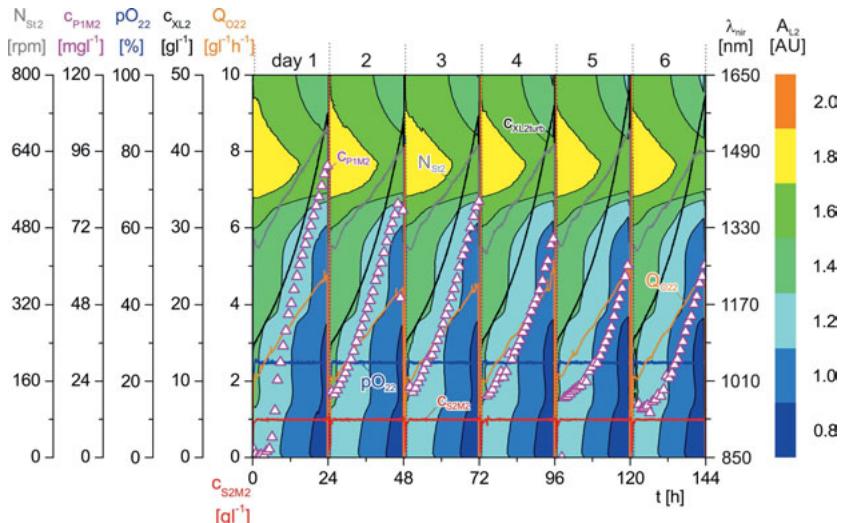


Bild 6.9: Verlauf von uni- und multivariaten Prozessvariablen in einer zyklischen Produktion

#### 6.4.5 Überwachung der Expression mit einer Atline-HPLC

Das Zielprotein D1M1H in dem Produktions-Bioreaktor wurde atline mit einer HPLC (VWR International, Darmstadt) nachgewiesen [Martens et al., 2011]. Da es N-terminal mit einem His-Tag fusioniert ist, wird eine Nickel-Chelatchromatographie eingesetzt. Der strukturelle Aufbau dieser Messmethode [Martens, 2014] ist in Bild 6.10 gezeigt.

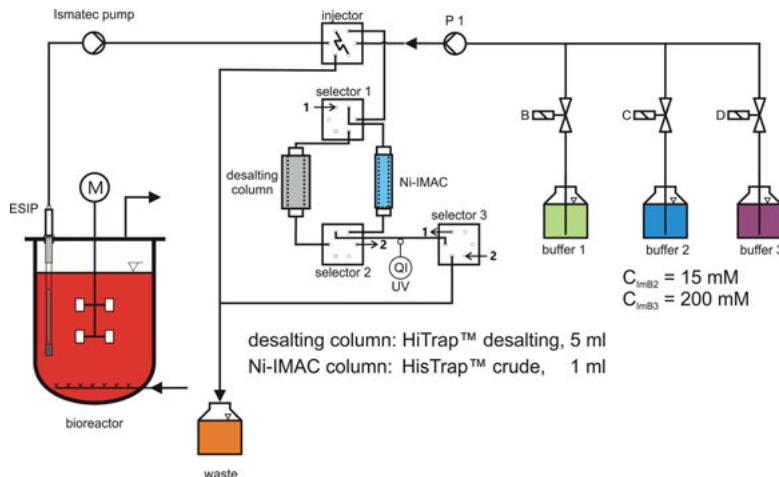


Bild 6.10: Prinzipieller Aufbau einer Atline-HPLC zur Proteinbestimmung mittels IMAC

Der kontinuierliche Probenstrom wird zellfrei über eine Inline-Filtrationssonde ESIP (*Eppendorf Sterilizable Inline Probe*) entnommen und über eine 4 ml Injektionsschleife zu der HPLC geleitet. Alle 52 min, was der Dauer eines Chromatographielaufes entspricht, wird der Probenstrom in das System injiziert. Nach der Injektion wird er über zwei 5 ml Hitrap Desalting-Säulen (GE Healthcare, Solingen) geleitet. Dies dient zur Abtrennung von Salzen, die bei der folgenden IMAC-Säule die Bindung der Proteine an die funktionellen Gruppen beeinträchtigen können.

Im Laufpuffer des Systems befindet sich kein Imidazol, was zu einer unspezifischen Bindung an das Säulenmaterial der Nickel-Chelatsäule führt. Nach einer Waschphase mit Laufpuffer über 5 Säulenvolumen (CV) wird eine zweistufige Elutionsphase eingeleitet. Zunächst wird ein erster Elutionspuffer über 7 CV mit 15 mM Imidazol verwendet, um unspezifisch gebundene Proteine zu entfernen.

Mit einem zweiten Elutionspuffer mit 200 mM Imidazol wird das über den His-Tag spezifisch gebundene Zielprotein eluiert. Dieser Elutionsschritt erfolgt über 10 CV. Das Eluat wird im Detektor bei 280 nm photometrisch vermessen. Über eine Integration der Peaksignale der unspezifisch gebundenen Proteine sowie des Zielproteins werden Massenkonzentrationen berechnet.

In Bild 6.11 sind die Verläufe von fünf Chromatogrammen der HPLC aus einem Produktionszyklus übereinandergelegt dargestellt. Mit zunehmender Produktionszeit steigt die Peakhöhe an, was den Zuwachs an Protein im Verlaufe des Prozesses indiziert.

Die Nickel-IMAC Säulen werden alle zwei Tage ausgetauscht, gereinigt und mit neuen Nickel-Ionen beladen.

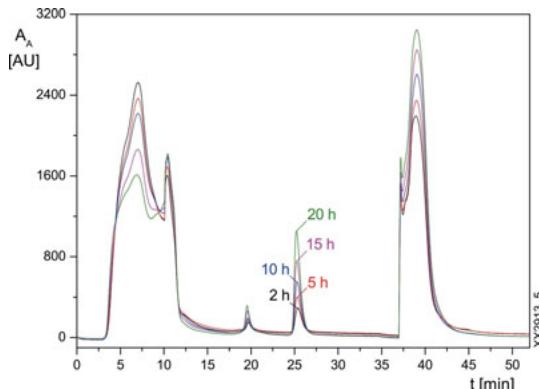


Bild 6.11: Übersicht von HPLC Chromatogrammen aus einem Produktionslauf

Zur Untersuchung des Effekts unterschiedlicher Säulen und der Filtration mit der ESIP auf die Güte der Messung wurde eine zusätzliche Offline-HPLC gleichwertig ausgestattet und mit denselben Standards kalibriert. Für eine anschließende Validierung der Atline-Messung konnten während einer 13-tägigen Fermentation insgesamt 68 Vergleichsproben entnommen werden, die jeweils mit der Atline-HPLC und der Offline-HPLC vermessen wurden.

Eine klassische Validierung sieht vor, zwei Datensets zu verwenden. Mit dem ersten (Trainingsset) wird ein Modell erstellt (Kalibriergerade) und mit dem zweiten (Validierungsset) wird dieses Modell validiert, indem die gemessenen Proben mit dem erstellten Modell vorhergesagt und die Summe der Fehlerquadrate berechnet wird.

Dies wurde realisiert, indem die 68 Datenpunkte der Offline-Messung der Größe nach sortiert und fortlaufend von 1 bis 68 durch nummeriert wurden. Daraufhin wurde jeder Messwert mit gerader Nummerierung dem Trainingsset zu geordnet, während der Rest dem Validierungsset entspricht.

Das Ergebnis der Anpassung ist in Bild 6.12 gezeigt.

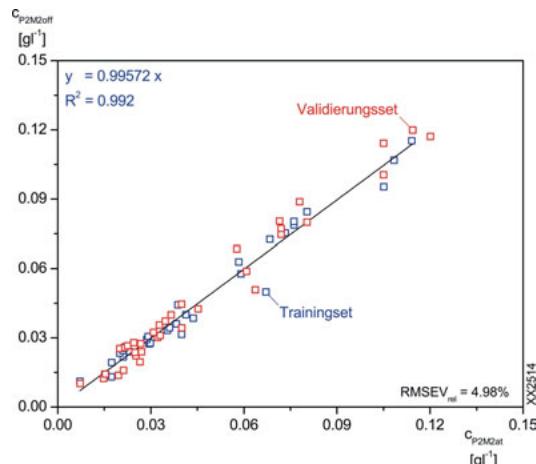


Bild 6.12: Offline-Validierung einer Atline-HPLC Messmethode zur Produktdetektion

Das Validierungsergebnis kann mit einem RMSEV – *root mean square error of validation*,

$$\text{RMSEV} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}, \quad (6.3)$$

mit

$y_i$  := Offline-Messwerte des Validierungsset

$\hat{y}_i$  := Schätzung der Offline-Werte aus den Atline-Werten

$n$  := Anzahl der Messwerte,

in einen RMSEV<sub>rel</sub> – *relative root mean square error of validation*,

$$\text{RMSEV}_{\text{rel}} = \frac{\text{RMSEV}}{y_{\max} - y_{\min}} \cdot 100\%, \quad (6.4)$$

mit

$y_{\max}$  := Maximalwert der Offline-Messwerte

$y_{\min}$  := Minimalwert der Offline-Messwerte,

transformiert und damit in einen Wert von 4,98 % zusammengefasst werden.

Durch Auftragung der Unterschiede von offline und atline gemessenen Proteinkonzentrationen gegen die Gesamtprozesszeit konnte festgestellt werden, dass sich der Filtrationseffekt erst nach 120 Stunden auf die Messung auswirkt. Die Validierungsergebnisse belegen jedoch, dass sich die erkennbaren Auswirkungen im tolerierbaren Rahmen bewegen.

Diese Erkenntnis wird durch die in Bild 6.13 gezeigte Kultivierung noch einmal bestätigt. Hier wurde unter variierenden Expressionsbedingungen bezüglich pH<sub>2</sub> und g<sub>L2</sub> die Atline-Produktdetektion (magenta) mit der Offline-Validierung (lila) verglichen.

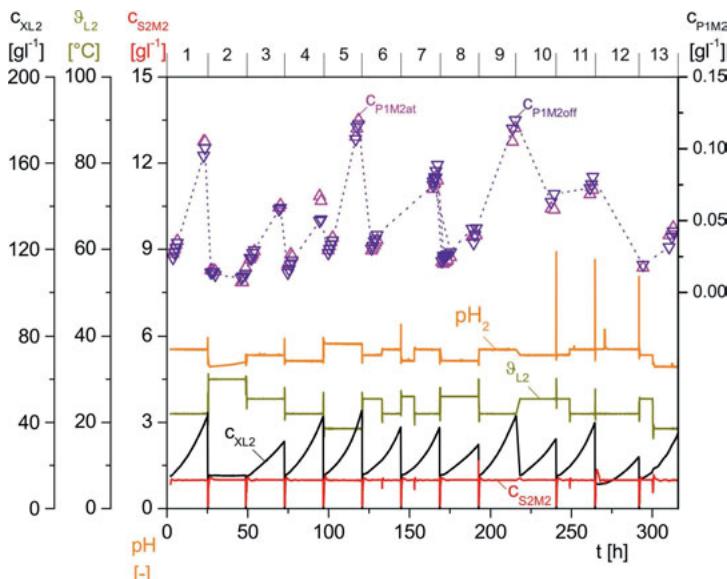


Bild 6.13: Validierung der Atline-Produktdetektion unter variierenden Bedingungen

## 6.5 Die Downstream Unit-Operationen

### 6.5.1 Zellabtrennung mit dem Separator SC1

Nach Beendigung einer 24 stündigen Produktionsphase im BIOSTAT® C30 erfolgt für die weitere Proteinaufreinigung zunächst eine Abtrennung der Zellen von der Klarphase, da das Zielprotein extrazellulär im Medium vorliegt.

Der in Bild 6.14 gezeigte selbstevakuierende Klärseparator SC1 wurde von der Fa. GEA Westphalia Separator entwickelt. Er ist über die Prozesskontrolleinheit SIMATIC S5 (Siemens, Karlsruhe) automatisiert und stellt den ersten Schritt der 4-stufigen Downstream-Linie des integrierten Prozesses dar.

Der Separator verfügt über eine äquivalente Klärfäche von  $1000 \text{ m}^2$  bei einem relativ kleinen Trommelvolumen von 1,1 l und einem Feststoffraum von 0,7 l.

Im Einlassstrom wird der Druck  $p_{H_2}$  (WIKA, Klingenberg/Main) gemessen und die Trübung  $E_{H_2}$  mit einem NIR-Durchflussmesser (730-970 nm) (optek-Danulat, Essen) bestimmt, während die Durchflussrate  $F_{H_2}$  (Bürkert, Ingelfingen) gemessen und mit einem Sollwert von  $40 \text{ l h}^{-1}$  durch Ansteuerung der Eingangspumpe geregelt wird.

Am Ausgang wird der Volumenstrom  $F_{sep}$  (Bürkert) ebenfalls gemessen und die Streuung  $S_{sep}$  mit einem hochpräzisen Streulichtmesser (400-1100 nm) (optek-Danulat) bestimmt, während der Ausgangsdruck  $p_{sep}$  (WIKA) gemessen und mit einem Sollwert von 3,4 bar durch Ansteuerung des Auslassventils (Bürkert) geregelt wird.

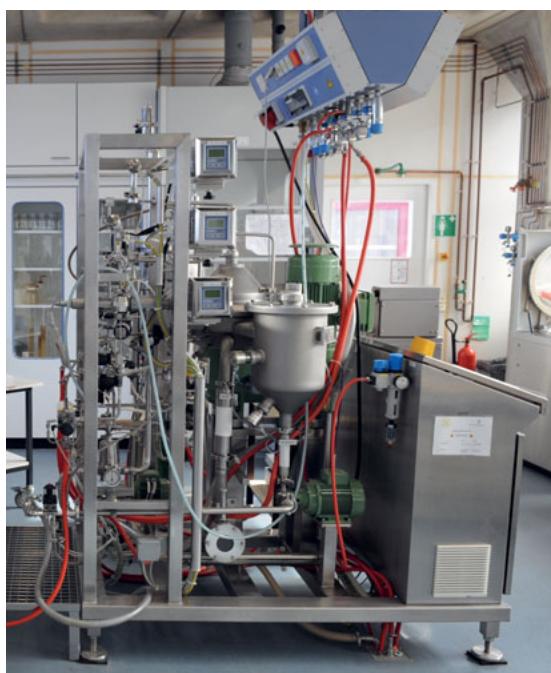


Bild 6.14: Klärseparator SC1 der Fa. GEA Westfalia Separator

Wie in Bild 6.15 gezeigt, wird der Klärprozess periodisch ausgeführt. Zuerst wird die Trommel vollständig gefüllt und die Zellbrühe im System rezirkuliert. Die Rezirkulation verläuft ohne weitere Zufütterung von Zellbrühe aus dem Reaktor bis ein Gleichgewichtszustand der Klärung erreicht ist, der an einem konstant niedrigen Streulicht  $S_{sep}$  im Ausgang ablesbar ist.

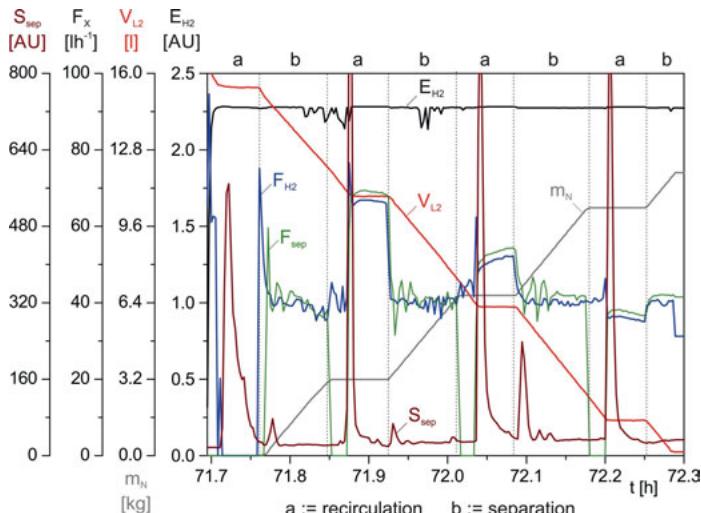


Bild 6.15: Ablauf einer Separation von Zellen aus dem Medium im Separator SC1

Auf die Rezirkulation (a) folgt eine kontinuierliche Klärung (b) der zugeführten Zellbrühe, erkennbar an der Masse  $m_N$  im Klarphasentank N, bis der Feststoffraum des Separators mit Zellen gefüllt ist. Die selbstevakuierende Trommel ermöglicht es, die Festphase bei voller Geschwindigkeit in den Abfalltank  $W_1$  zu entladen. Darauf folgend beginnt das Trennverfahren von vorne, bis der Produktionsreaktor vollständig geleert ist.

Der Klarphasentank N dient als Feed-Reservoir für eine nachfolgende Mikrofiltration zur Ablösung der Klärung.

### 6.5.2 Mikrofiltration mit Sartocon Slice MF-Modulen

Das SARTOCON Slice®-System zur Querstromfiltration wird mit 2 Modulen für eine Mikrofiltration mit einer Porengröße von  $0,2 \mu\text{m}$  und einer Gesamtfilterfläche von  $0,2 \text{ m}^2$  (Sartorius Stedim Biotech, Göttingen) ausgestattet. Der Betrieb wird durch die DCU III des Produktionsreaktors BIOSTAT® C30 gesteuert.

Bild 6.16 zeigt den Ablauf einer Mikrofiltration. Dabei wird die Masse der Klarphase  $m_N$  und die Masse  $m_P$  des im Produkttank aufgefangenen Permeats gewogen (Sartorius, Göttingen). Im Zulauf wird die Trübung  $E_N$  über NIR-Absorption (optek-Danulat) gemessen. Diese Messung spricht erst an, wenn der Klarphasentank fast leer ist.

Die Volumenströme  $F_{feed1}$ ,  $F_{ret1}$  und  $F_{per1}$  (alle Krohne, Duisburg) werden ebenso wie die Drücke  $p_{feed1}$ ,  $p_{ret1}$  und  $p_{per1}$  (alle WIKA) gemessen.

Der Eingangsdruck  $p_{feed1}$  wird mit einem Sollwert von 1,5 bar durch Steuerung des Eingangsstromes  $F_{feed1}$  und der Retentatdruck  $p_{ret1}$  mit einem Sollwert von 1,0 bar durch Manipulation des Retentatventils (GEMÜ) geregelt.

Auf der Permeatseite wird das Permeatventil (GEMÜ) vollständig geöffnet, so dass sich  $p_{per1}$  ungeregelt einstellt.

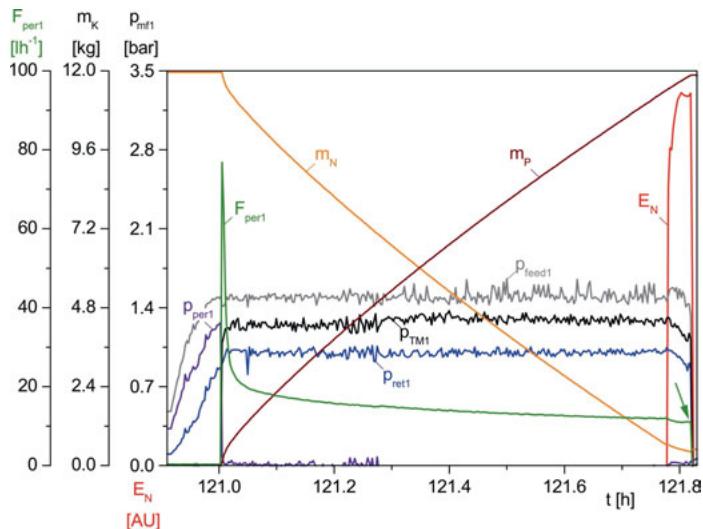


Bild 6.16: Ablauf einer Mikrofiltration zur abschließenden Reinigung der Klarphase

Das Retentat wird wieder in den Klarphasentank N zurückgeführt. Das Permeat wird im Produkttank P gesammelt, der ebenfalls auf einer Waage steht. Aus diesem Behälter wird die nachfolgende Ultrafiltration gespeist. Das Ende der Mikrofiltration ist durch eine scharfe Abnahme (Pfeil) der Fließrate des Permeats  $F_{per1}$  gekennzeichnet.

### 6.5.3 Ultrafiltration mit Sartocon Slice UF-Modulen

Für die Ultrafiltration wird ebenfalls ein Querstromfiltrationssystem Sartocon Slice<sup>®</sup> mit zwei Modulen bei einem Cutoff von 30 kDa Molekulargewicht und einer Gesamtfilterfläche von  $0,2 \text{ m}^2$  eingesetzt. In Bild 6.17 sind die MSR-Elemente der Ultra- und Diafiltration gezeigt.

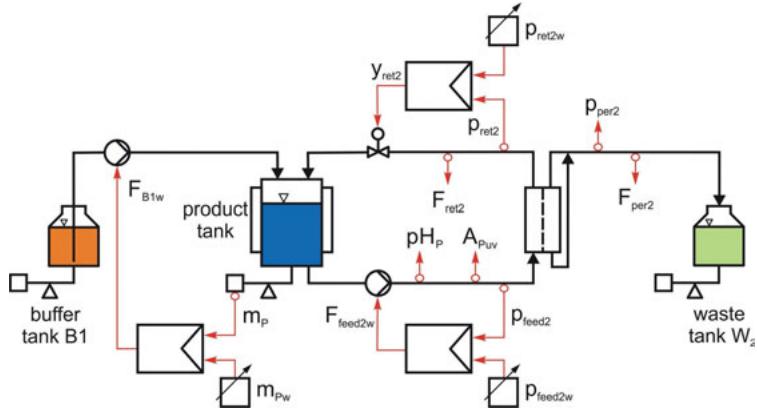


Bild 6.17: MSR-Elemente der Ultra- und Diafiltration

Im Zulauf des Filtrationsmoduls befindet sich, ebenfalls von optek-Danulat, eine pH-Sonde ( $\text{pH}_P$ ), eine Leitfähigkeitssonde ( $\sigma_P$ ) und ein Durchfluss-UV-Photometer ( $A_{\text{Puv}}$ ). Die Durchfluss- und Druckmessung sowie die Regelungstechnik entsprechen der Mikrofiltration.

Die Ultrafiltration läuft ohne Gegendruck von der Permeatseite. Gespeist wird diese Anlage aus dem Produkttank P, in den das Retentat auch wieder zurückgeführt wird. Mit der Entsorgung des produktfreien Permeats in den Wastetank  $W_2$  wird das Produkt  $P_1$  in Tank P aufkonzentriert.

Mit einer nachfolgenden Diafiltration wird das Produkt für die spätere Chromatographie umgepuffert. Dies erfolgt mit einer Füllstandsregelung am Tank P. Dabei wird das ausfließende Permeat durch einen Puffer  $B_1$  ersetzt.

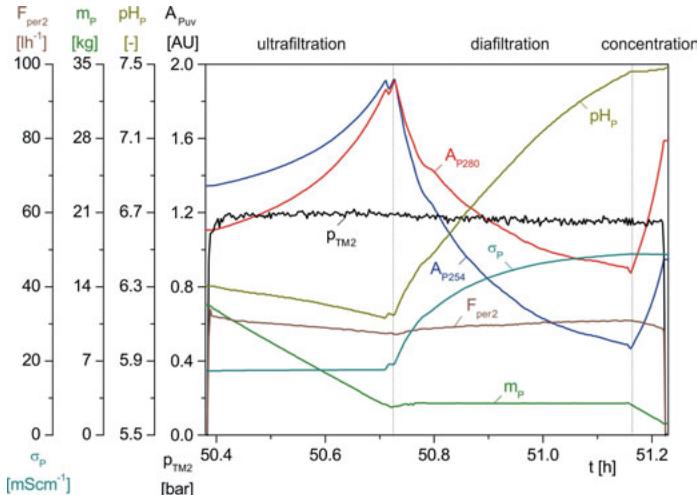


Bild 6.18: Ablauf einer Ultrafiltration zur abschließenden Klärung

In Bild 6.18 ist der Ablauf von Ultra- und Diafiltration gezeigt. Die Masse  $m_P$  im Produkttank P wird zunächst bis auf 3 kg abfiltriert. Dabei fällt der pH-Wert  $\text{pH}_P$  leicht ab, während die Konduktivität  $\sigma_P$  relativ konstant bleibt. In der nachfolgenden Diafiltration ab  $t = 50,72$  h erfolgt bei gleicher Masse  $m_P$  ein fünffacher Austausch mit dem Puffer  $B_1$ . Dadurch steigen  $\text{pH}_P$  und  $\sigma_P$  stetig an und nähern sich den Pufferwerten mit einer Verzögerung 1. Ordnung. In dem zweiten Ultrafiltrationsschritt ab  $t = 51,16$  h werden die zurück bleibenden Komponenten durch Verringerung der Brühe auf 1 kg weiter aufkonzentriert.

Die Aufkonzentrierung des Zielproduktes in den Ultrafiltrationsphasen ist in den Anstiegen der UV-Signale bei 254 nm ( $A_{\text{P}254}$ ) und bei 280 nm ( $A_{\text{P}280}$ ) zu erkennen.

Während der Umpufferung bleibt die Produktkonzentration  $c_{\text{P}1P}$  zwar konstant, die UV-Signale fallen aber exponentiell ab, d.h. in dieser Phase bestimmen die Pufferkomponenten den UV-Verlauf.

Am Ende von Mikro- bzw. Ultrafiltration werden die Filterkartuschen jeweils durch CIP – *cleaning in place* mit Lauge und Säure abwechselnd gereinigt und anschließend mit VE-Wasser gespült.

### 6.5.4 Chromatographische Produktreinigung mit einem AEKTA™purifier 100

Die Reinigung des Zielproduktes  $P_1$  erfolgt mit dem Chromatographiesystem C, einem ÄKTA™purifier 100 (GE Healthcare, Schweden), der durch die GE-Software UNICORN gesteuert wird. Das Produkt wird dabei über eine IMAC-Säule gereinigt, die zuvor mit Nickel-Ionen zu beladen ist [Paul et al., 2014]. Der Ablauf dieser Methode ist in Bild 6.19 gezeigt.

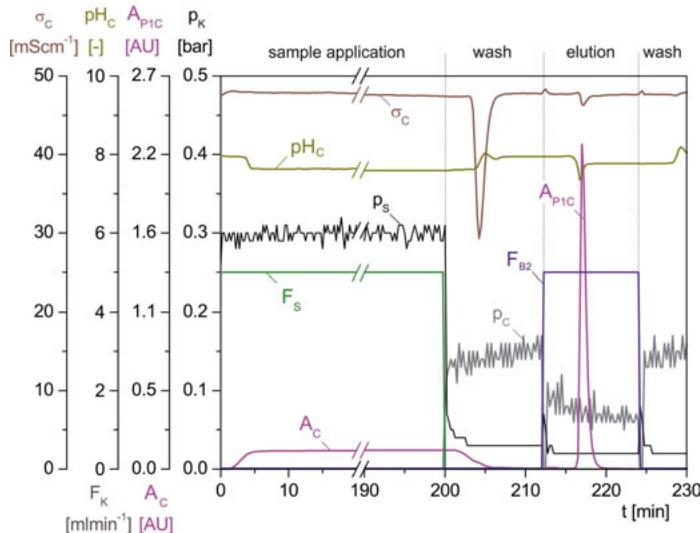


Bild 6.19: Ablauf der chromatographischen Produktreinigung mit einem AEKTA™purifier 100

Die Aufgabe der Produktlösung erfolgt mit einer Fließrate  $F_s$  von  $5 \text{ mlmin}^{-1}$ . Dabei wird der Probe  $15 \text{ mM}$  Imidazol zugesetzt, um Hauptverunreinigungen bereits im Durchbruch wiederzufinden.

Danach wird die Säule mit  $5$  Volumenmengen (CV) an Bindungspuffer gewaschen und anschließend mit  $5$  CV Elutionspuffer mit  $200 \text{ mM}$  Imidazol mit einer Fließrate  $F_{B2}$  von ebenfalls  $5 \text{ mlmin}^{-1}$  eluiert.

Das Zielprodukt ist in dieser Phase an dem starken Anstieg des UV-Signals  $A_c$  erkennbar. Die Fraktion des Produktpeaks  $A_{P1C}$  wird anschließend aufgefangen und enthält ein relativ hoch gereinigtes Zielprotein. Die Messung erfolgt bei  $280 \text{ nm}$ . Danach erfolgte ein weiterer Waschschnitt.

Im Weiteren wird nun die Zusammenführung der einzelnen Prozessschritte zu einer integrierten Anlage und deren Automatisierung beschrieben.

## 6.6 Automatisierung des Integrierten Prozesses

### 6.6.1 Übersicht der dezentralen Automatisierungskomponenten

Das Problem einer Integration von Anlagenteilen verschiedener Hersteller liegt in der unterschiedlichen Hardware und Software und den damit verbundenen Konzepten einer Automatisierung, Datenspeicherung und Datenverarbeitung.

In Bild 6.20 ist die historisch gewachsene Lösung einer hierarchischen Automatisierung der Anlage dargestellt. Die Scale-down Produktionsanlage beinhaltet vier Frontend-Automatisierungsstationen (blau). Der *cell breeding reactor* (M&C: ED10) wird von einer DCU – *Digital Control Unit II* grundautomatisiert, d.h. hierin sind die MSR-Einheiten des Reaktors enthalten.

Das gleiche gilt für den *protein production reactor* (M&C: C30), der über eine DCU III automatisiert wird. Daran sind allerdings auch beide *crossflow filtration units* Sartocon Slice (M&C: Xflow) angebunden, die über Batch-Rezepte in der DCU III geführt werden.

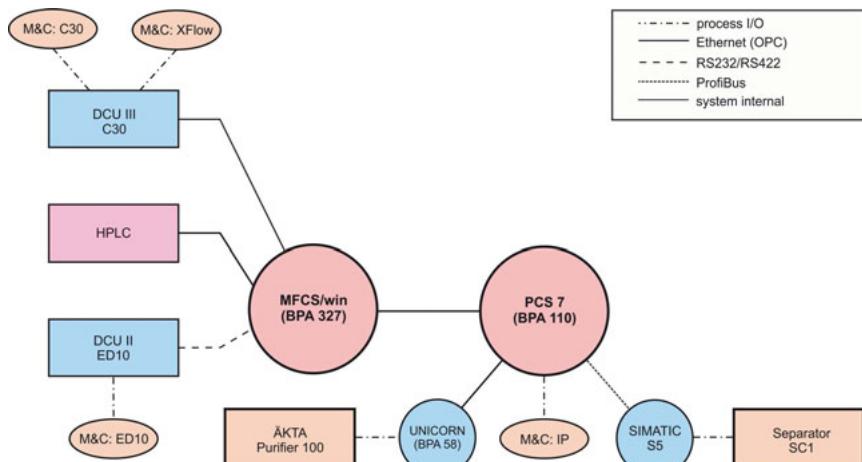


Bild 6.20: Automatisierung und Datenverarbeitung des integrierten Prozesses

Das SCADA – *Supervisory Control And Data Acquisition* System MFCS/win ist mit der DCU II über RS422 und mit der DCU III über Ethernet und OPC – *Open Process Control* verbunden und übernimmt die Datenspeicherung sowie die übergeordnete Steuerung des sequentiell/parallelen Upstream Prozesses.

MFCS/win läuft auf dem Rechner BPA327 und ist auch für die Prozessvisualisierung und die übergeordnete Regelung und Steuerung des USP zuständig. Die erforderlichen gekoppelten Kultivierungsabläufe wurden über S88 Rezepte nach ISA – *The International Society of Automation* in MFCS/win hinterlegt. Darüber hinaus werden dort auch die Ergebnisse der HPLC-Produktanalytik über OPC eingebunden.

Alle beschriebenen Automatisierungselemente der Bioreaktoren werden von der Fa. Sartorius Stedim Systems vertrieben.

Die *separation unit* SC1 (GEA Westfalia, Oelde) ist hingegen mit der Prozesssteuerung SIMATIC S5 (Siemens, München) automatisiert und ermöglicht eine Einbindung in die Siemens Prozesseleittechnik PCS 7 über das Profibus Interface.

Das *chromatography system* ÄKTA™purifier 100 erlaubt einen autonomen Betrieb über eine Automatisierung mit der Software UNICORN, die auf dem Rechner BPA58 installiert ist. Eine Einbindung in übergeordnete Strategien erfolgt ebenfalls über OPC via Ethernet.

### 6.6.2 Automatisierung des Upstream-Prozesses mit MFCS/win

Der Upstream Prozess wurde, auch historisch bedingt, allein über MFCS/win automatisiert.

Die Funktionalität dieser Software ist in 5 Module

- Datenerfassung
- Prozessvisualisierung
- Prozessdokumentation
- Prozessregelung mit PID-Modulen
- Prozessautomatisierung mit S88-Rezepten

zu unterteilen.

Die dafür zur Verfügung stehenden Softwarekomponenten sind in Bild 6.21 gezeigt.

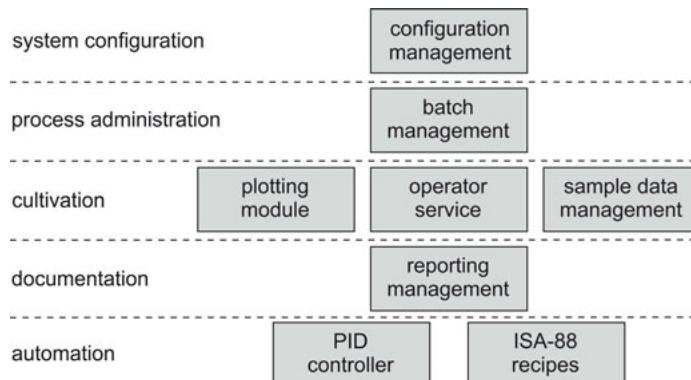


Bild 6.21: Modulare Struktur von MFCS/win [Kaiser, 2010]

### 6.6.3 Integration und Automatisierung der Gesamtanlage über PCS 7

Die übergeordnete Automatisierung des integrierten Gesamtprozesses erfolgte über das Prozessleitsystem PCS 7 (Siemens, München), dessen weitreichende Möglichkeiten allerdings nur in geringem Umfang genutzt wurden.

Dabei waren zunächst die einzelnen Unit-Operationen *cell breeding*, *protein production*, *separation*, *micro-filtration* und *ultra-filtration* über sterilisierbare Membranventile zu verbinden. Diese sind direkt über eine SIMATIC S7 Steuereinheit (M&C: IP) an das Hostsystem PCS 7 angebunden. Damit wurde mit PCS 7 die übergeordnete Steuerung der Prozessabfolge ermöglicht.

MFCS/win bildet, ebenso wie SIMATIC S5, die dafür erforderlichen Daten in PCS 7 ab. Dort wurden die entsprechenden Schrittfolgen sowie zugehörigen GUI – *graphical user interfaces* der Teilanlagen implementiert, so dass eine Prozessbedienung und Beobachtung, wie in Bild 6.22 gezeigt, erfolgen kann.

Da die Automatisierungsprobleme des sequentiell/parallelen Prozesses bereits in einer vorhergehenden Doktorarbeit [Lögering, 2015] gelöst und auch detailliert dokumentiert wurden, wird hier auf eine weitere Beschreibung verzichtet.

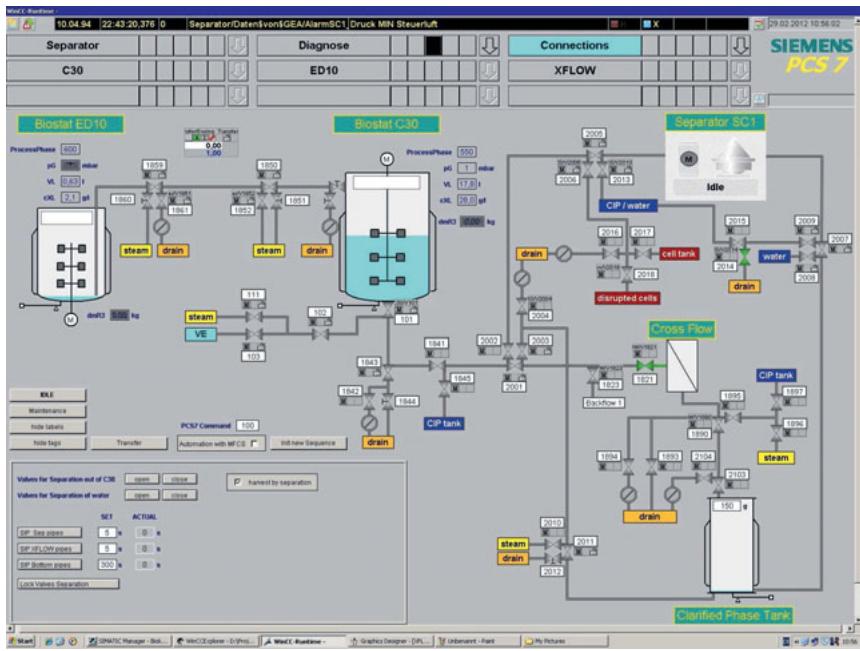


Bild 6.22: Bedien- und Beobachtungsebene des integrierten Prozesses [Lögering, 2015]

Der nachfolgende Abschnitt zeigt allerdings das Fortschreiten dieser Prozessentwicklung mit der Herstellung des rekombinanten Malaria-Vakzinkandidaten D1M1H. Daraus ergibt sich dann die Anforderung einer QbD-gerechten Bewertung der zyklischen Prozessführung.

Die dafür erforderliche Erweiterung der Anlage mit einer MVDA-gerechten Datenerfassung und die entsprechende Behandlung der Prozesszyklen mit MVDA-Tools bilden einen weiteren Schwerpunkt dieser Arbeit.

## 6.7 Die quasi-kontinuierliche Herstellung des Malaria-Vakzinkandidaten D1M1H

### 6.7.1 Verlauf eines integrierten Produktionszyklus

In Bild 6.23 wird ein vollautomatisierter, seriell integrierter Prozessablauf gezeigt. Die sechs Unit-Operationen *cell breeding*, *protein production*, *separation*, *micro-filtration*, *ultra-filtration* und *purification* werden durch ausgewählte, normierte Variablen präsentiert. Eine SDS-PAGE Analyse für den Produktions-, den Aufkonzentrierungs- und den Feinreinigungsschritt ist beigefügt.

Als Ergebnis endet der Upstream mit einer über die HPLC gemessenen Produktkonzentration  $C_{P1M2n}$  von  $72 \text{ mg l}^{-1}$  mit einer zugehörigen Reinheit  $P_{P1pro}$  von 0,25. Unter den gewählten Betriebsbedingungen wird der Flüssigkeitsverlust durch die Filtrationssonde durch eine Lautenttitration der pH-Regelung fast vollständig wieder ausgeglichen, so dass ein Zielvolumen  $V_{L2n}$  von 14,76 l und damit eine Produktmasse  $m_{P1M2n}$  von 960 mg erreicht wird.

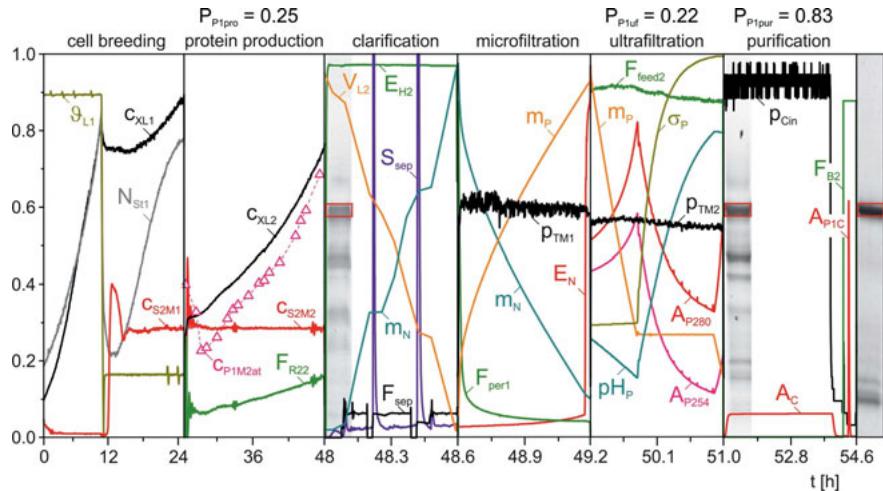


Bild 6.23: Verlauf einer vollautomatisierten integrierten Malariavakzin-Produktion

Tabelle 6.1: Skalierungsbereiche der Prozessvariablen in Bild 6.23

cell breeding (p=1) and protein production (p=2)		clarification			
$C_{XLp}$ :	cell density in reactor p	[0, 50] $\text{gl}^{-1}$	$E_{H2}$ :	turbidity clarifier inlet	[0, 2.2] AU
$g_{L1}$ :	cultivation temperature	[20, 30] °C	$S_{sep}$ :	scattering clarifier outlet	[0, 999] NTU
$N_{S11}$ :	stirrer speed	[250, 850] rpm	$F_{sep}$ :	clarifier outlet flow rate	[0, 660] $\text{lh}^{-1}$
$F_{R22}$ :	methanol feed rate	[0, 0.6] $\text{lh}^{-1}$	$V_{L2}$ :	volume in reactor 2	[0, 16] l
$C_{S2M2}$ :	methanol concentration	[0, 3.5] $\text{gl}^{-1}$	$m_N$ :	mass in supernatant tank N	[0, 13] g
$C_{P1M2}$ :	product concentration	[0, 0.1] $\text{gl}^{-1}$			
microfiltration		ultrafiltration			
$F_{per1}$ :	permeate flow rate	[0, 150] $\text{lh}^{-1}$	$F_{feed2}$ :	feed flow rate	[0, 70] $\text{lh}^{-1}$
$p_{TM1}$ :	transmembrane pressure	[0, 2.5] bar	$A_{P254}$ :	UV-absorption at 254 nm	[1, 3] AU
$E_N$ :	turbidity of feed	[0, 3.6] AU	$A_{P280}$ :	UV-absorption at 280 nm	[1, 2] AU
$m_N$ :	mass in supernatant tank N	[0, 13] kg	$p_{H_P}$ :	pH-value in product tank	[6, 9] -
$m_P$ :	mass in product tank P	[0, 15] kg	$\sigma_P$ :	conductivity in product tank	[0, 50] $\text{mScm}^{-1}$
			$p_{TM2}$ :	transmembrane pressure	[0, 2.5] bar
			$m_P$ :	mass in product tank	[0, 15] kg
purification		off-line analytics			
$A_C$ :	UV-absorption of feed	[0, 2.5] AU	$P_{P1pro}$ :	purity of production	0.25 -
$A_{P1C}$ :	UV-absorption of product	[0, 2.5] AU	$P_{P1uf}$ :	purity of ultrafiltration	0.22 -
$p_{Cin}$ :	inlet pressure	[0, 0.33] bar	$P_{P1pur}$ :	purity of purification	0.83 -
$F_{B2}$ :	flow rate of elution	[0, 1.5] $\text{lh}^{-1}$			

Die Produktivität des gesamten Upstreamverlaufes  $PRD_{up}$ ,

$$PRD_{up} = \frac{c_{P1M2n} \cdot V_{M2n}}{T_{up}} = \frac{c_{P1M2n} \cdot V_{L2n} \cdot (\rho_Z - \alpha_{Z/X} \cdot c_{XL2n})}{T_{up} \cdot \rho_Z} \quad (6.5)$$

mit

$PRD_{up}$  := Produktivität des gesamten Upstreamverlaufes  $[mgh^{-1}]$

$c_{P1M2n}$  := Zielproduktkonzentration in der Medienphase  $M_2$  am Ende eines Produktionszyklus bei  $t_{2n}$   $[mg l^{-1}]$

$V_{L2n}$  := Reaktionsvolumen in Reaktor 2 bei  $t_{2n}$   $[l]$

$c_{XL2n}$  := (Biotrocken-) Zelldichte im Reaktor 2 bei  $t_{2n}$   $[gl^{-1}]$

$\alpha_{Z/X}$  := Verhältnis von Biofeucht zu Biotrockenmasse ( $\alpha_{Z/X} = 3,54$ )

$\rho_Z$  := Dichte der feuchten Zellen ( $\rho_Z = 1.149 \cdot 10^3 \text{ gl}^{-1}$ )

$T_{up}$  := Zeitintervall des gesamten Kultivierungszyklus ( $T_{up} = 48 \text{ h}$ ),

wird auf den Kultivierungszyklus  $T_{up}$  bezogen.

In dem in Bild 6.23 gezeigten Upstream-Prozess wird eine  $PRD_{up}$  von  $20 \text{ mgh}^{-1}$  über 2 Tage erreicht.

Das Aufarbeitungsverhalten hinsichtlich Reinheit  $P_{P1K}$  und Wiederfindung  $R_{P1K}$  ist in Tabelle 6.2 zusammengefasst.

Tabelle 6.2: Resultierendes Zielproduktverhalten entlang der Aufreinigungskette

protein production	clarification	microfiltration
$c_{P1M2n} = 72 \text{ mg l}^{-1}$	$c_{P1Nn} = 67.5 \text{ mg l}^{-1}$	$c_{P1Pn} = 62 \text{ mg l}^{-1}$
$m_{P1M2n} = 960 \text{ mg}$	$m_{P1Nn} = 900 \text{ mg}$	$m_{P1Pn} = 740 \text{ mg}$
$P_{P1pro} = 0.25$	$R_{P1sep} = 0.938$	$R_{P1mf} = 0.822$
ultrafiltration	purification	integrated process
$c_{P1Pn} = 480 \text{ mg l}^{-1}$	$c_{P1Cn} = 21 \text{ gl}^{-1}$	$T_{int} = 54.6 \text{ h}$
$m_{P1Pn} = 480 \text{ mg}$	$m_{P1Cn} = 420 \text{ mg}$	$PRD_{int} = 8.6 \text{ mgh}^{-1}$
$R_{P1uf} = 0.649$	$R_{P1pur} = 0.875$	$R_{P1int} = 0.438$
$P_{P1uf} = 0.22$	$P_{P1pur} = 0.83$	$P_{P1int} = 0.83$

Die Proteinproduktion endet mit einer Zielproduktkonzentration  $c_{P1M2n}$  von  $72 \text{ mg l}^{-1}$  und einer Ernte von  $960 \text{ mg}$  in einem Volumen  $V_{L2n}$  von  $14,76 \text{ l}$ . Bei einer Zelldichte  $c_{XL2n}$  von  $31,4 \text{ gl}^{-1}$  ergibt sich ein Medienvolumen  $V_{M2n}$  von  $13,33 \text{ l}$ . Die Reinheit  $P_{P1pro}$  beträgt dabei 0.25.

Die Klärung mit dem Separator beinhaltet einen Produktverlust von 6.2 %. Dabei ist die Konzentration  $c_{P1Nn}$  im *supernatant tank N* gegenüber  $M_2$  leicht abgesenkt. Die vorgereinigte Produktdmenge  $m_{P1Nn}$  mit  $900 \text{ mg}$  führt auf eine Wiederfindung  $R_{P1sep}$  von 0.938.

Nach der Mikrofiltration ist die Konzentration  $c_{P1Pn}$  im Produkttank P auf  $62 \text{ mg l}^{-1}$  abgesunken. Die Wiederfindung  $R_{P1mf}$  beträgt 0.822 bei einer Masse  $m_{P1Pn}$  von  $740 \text{ mg}$ . Der Produktverlust beträgt in diesem Schritt 17.8 %.

In der Ultrafiltration wird das Produkt mit  $c_{P1Pn}$  von  $480 \text{ mg l}^{-1}$  in 1 l aufkonzentriert. Die Masse von  $m_{P1Pn}$  von  $480 \text{ mg}$  ergibt eine Wiederfindung  $R_{P1uf}$  von 0,649. Das entspricht einem Verlust von 35,1 %.

Die finale Feinreinigung endet nach einer Elution von 20 ml mit einer Konzentration  $c_{P1Cn}$  von  $21 \text{ gl}^{-1}$  und einer Masse  $m_{P1Cn}$  von  $420 \text{ mg}$ . Die Wiederfindung  $R_{P1pur}$  beträgt somit 0,875, d.h. der letzte Schritt bringt einen Verlust von 12,5 %.

Der integrierte Gesamtprozess lässt sich bei einer Länge  $T_{int}$  von 54,6 h über die verlustbehaftete Produktivität  $PRD_{int}$ ,

$$PRD_{int} = \frac{m_{P1Cn}}{T_{int}}, \quad (6.6)$$

und die Gesamtwiederfindung  $R_{P1int}$ ,

$$R_{P1int} = \frac{m_{P1Cn}}{m_{P1M2n}} = R_{P1sep} \cdot R_{P1mf} \cdot R_{P1uf} \cdot R_{P1pur}, \quad (6.7)$$

charakterisieren.

Die Produktivität  $PRD_{int}$  des gesamten Prozesses betrug somit  $8.6 \text{ mgh}^{-1}$ , während eine Wiederfindung  $R_{P1int}$  von 0.438 vorlag. Die Prozesszielgröße  $R_{P1int}$  gilt es durch eine Optimierung der einzelnen Aufarbeitungsschritte stark zu verbessern. Diese Schritte werden in dieser Arbeit allerdings nicht weiter verfolgt.

Die Erhöhung der Produktivität  $PRD_{int}$  wird hingegen durch eine quasi-kontinuierliche Prozessführung im nächsten Abschnitt mehr als verdoppelt.

### 6.7.2 Quasi-kontinuierliche Produktion in sequentiell/paralleler Prozessführung

Der sequentiell/parallele Kultivierungsmodus in Abschnitt 6.3.2 eröffnet nun die Durchführung eines quasi-kontinuierlich betriebenen integrierten Gesamtprozess.

Eine derartige Kampagne über acht Tage ist in Bild 6.23 gezeigt.

Der Prozess startet in Reaktor 1 mit dem ersten *cell breeding* über 24 h. Dieser wird am zweiten Tag wiederholt und von der ersten *protein production* in Reaktor 2 begleitet.

Nach einer 24-stündigen Expression und Sekretion des Zielproduktes wird Reaktor 2 vollständig über den Teller-Separator entleert. In diesem werden die Zellen nahezu vollständig aus dem Medium entfernt, das in den Klarphasentank N überführt wird. Die Klarphase läuft dann über die Mikrofiltration bis das Produkt im Produkttank P von allen Zellbruchstücken gereinigt vorliegt.

Danach erfolgt eine Ultrafiltration, in der das Produkt zunächst aufkonzentriert und mit einem Diafiltrationsschritt ein Pufferaustausch ausgeführt wird. Eine weitere Ultrafiltration stellt der nachfolgenden chromatographischen Reinigung über den ÄKTA™purifier 100 bereits ein hochkonzentriertes Produkt zur Verfügung.

Somit ist die erste Produktionsphase nach ca. 2,5 Tagen beendet.

Parallel zum ersten Downstream verlaufen bereits die zweite Produktbildung und die dritte Zellanzucht. Somit wird ab dem dritten Tag täglich eine Ernte des dreitägigen Prozesses vorgenommen.

Es wird noch einmal darauf hingewiesen, dass die quasi-kontinuierliche Strategie mit der gleichen Technologie wie die klassischer Batchprozesse durchgeführt wurde.

Bei einem Vergleich der Prozessdauer würde die Produktion mit einer konsekutiven Strategie 22 Tage zuzüglich der Zeit zur Vorbereitung der sieben individuellen Batch-Kultivierungen erfordern.

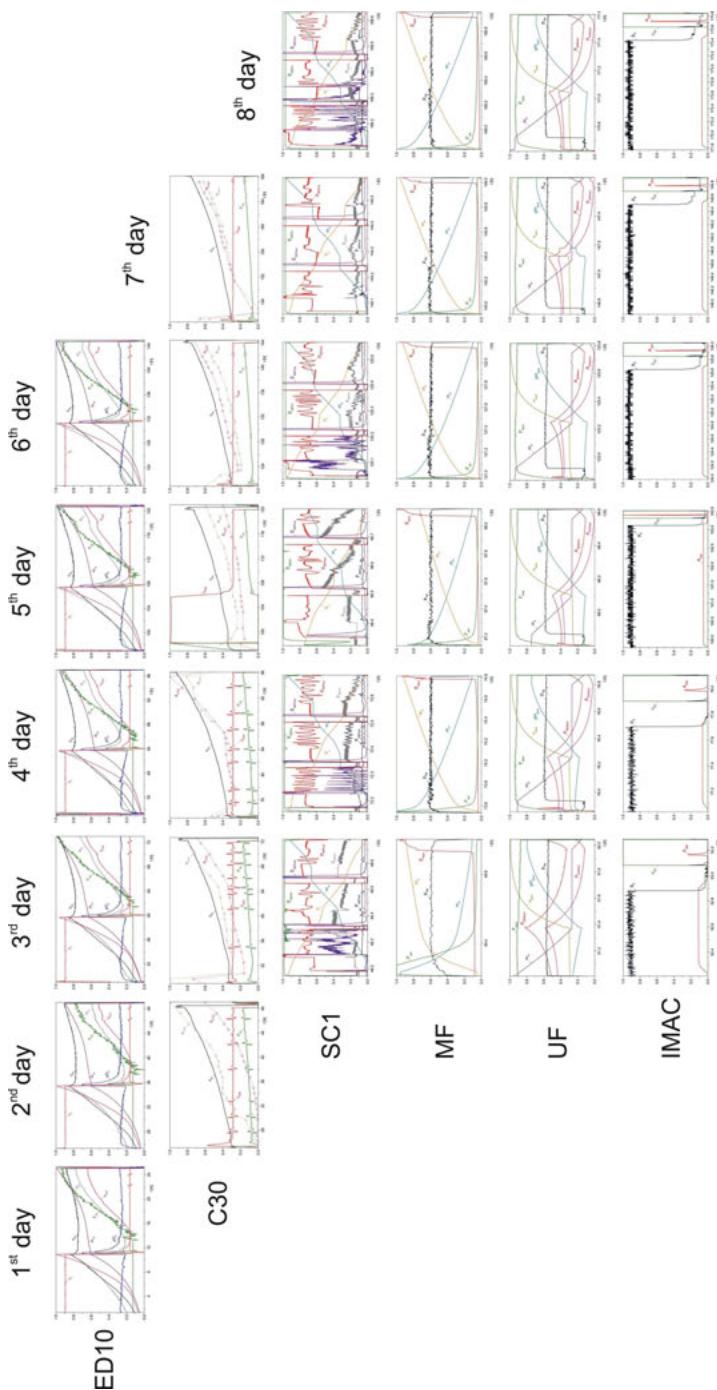


Bild 6.24: Quasi-kontinuierliche, integrierte, sequentiell/parallele Produktion von Malariavakzinen über 8 Tage

Mit der sequentiell/parallelen Prozessstrategie ist es möglich, 4,2 g gereinigtes Produkt in 12 Tagen herzustellen. Dies setzt allerdings eine stabile Produktivität in allen Zyklen voraus, die mit der in Bild 6.23 vorgestellten Einprozessstrategie vergleichbar ist.

Daher muss für jeden Prozesszyklus, wie auch für individuelle Batchprozesse, die Qualität des Prozessablaufs evaluiert werden, um eine QbD-gerechte Produktion zu demonstrieren.

Insbesondere der zyklische Zellanzuchtprozess und der jeweils nachfolgende Expressions-schritt müssen dafür näher untersucht werden.

Dies erfolgt im weiteren Teil der Arbeit über eine MVDA – *Multi Variate Data Analysis* Prozessbewertung. Die dafür erforderliche Ergänzung der Anlage mit einer MVDA-gerechten Datenverarbeitung wird im nächsten Abschnitt 6.8 erläutert, während die theoretischen Grundlagen und deren experimentelle Anwendungen ab Kapitel 7 zu finden sind.

## 6.8 Erfassung und Vorverarbeitung von Prozessdaten zur MVDA

### 6.8.1 Erweiterung der Anlage um eine komplexe Datenverarbeitung

Die Online-Anwendung der MVDA, z.B. zur Bearbeitung von Spektren der NIR-, Raman-, oder 2D-Fluoreszenz-Spektrometrie oder auch zur multivariaten Analyse von Prozessdaten scheitert häufig an der heterogenen Struktur von Prozessdaten aus unterschiedlichen Quellen und der fehlenden Einbindung von kommerziellen MVDA-Softwarepaketen in die Prozessleittechnik [Schneider, 2014].

Diese Probleme wurden in diesem Projekt durch Einbindung adäquater Software mit einem industriell geförderten Forschungsprojekt gelöst.

Dabei wurden die Softwarepakete SIPAT (Siemens, Karlsruhe) und SIMCA (Umetrics, Schweden) zur Lösung der Probleme eingesetzt.

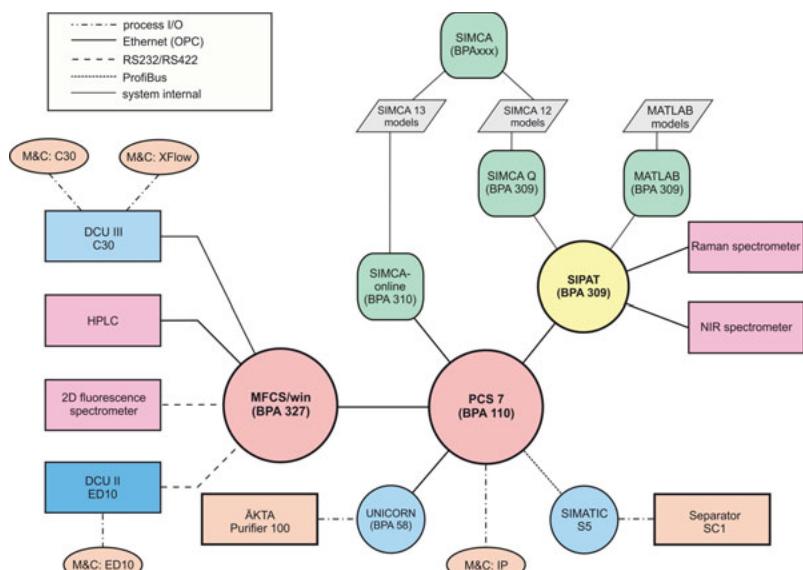


Bild 6.25: Erweiterung der Anlage mit kommerziellen Datenverarbeitungstools

In Bild 6.25 ist die daraus entwickelte Struktur einer erweiterten Datenerfassung und Datenverarbeitung entstanden. Die darin befindlichen Hauptkomponenten werden im Folgenden näher erläutert.

### 6.8.2 Erfassung und Datenverarbeitung mit PCS 7 und SIPAT

Die Fa. Siemens hat mit SIPAT eine Software entwickelt, die Datenquellen unterschiedlicher Art (analog, digital, seriell, spektral) in eine MVDA-verträgliche Art transformiert und ablegt.

Die Eigenschaften von SIPAT sind wie folgt charakterisierbar:

- Erfassung von Prozessdaten unterschiedlicher Art über entsprechende Treiber
- Weiterverarbeitung und Speicherung dieser Daten in eine matrizengerechte Form durch, wenn nötig, Interpolation und Ergänzung
- Triggern von modellbasierten Berechnungen auf der Basis von SIPAT-Datenerfassung mit MVDA-Tools (SIMCA) oder MATLAB-Skripten
- Rückführung von Prozesssteuerdaten zu eingebundenen Prozessleitsystemen

Der SIPAT-Rechner BPA309 ist über Ethernet (OPC) direkt mit dem PCS 7-Rechner BPA110 verbunden. Folglich stehen die PCS 7-Daten aus verschiedenen *collector stations* und, über eine OPC-Verbindung zu BPA327, auch die Daten von MFCS/win, der 2D-Fluoreszenz-Spektrometrie sowie der HPLC zur Produktmessung für eine Weiterverarbeitung mit SIPAT zur Verfügung.

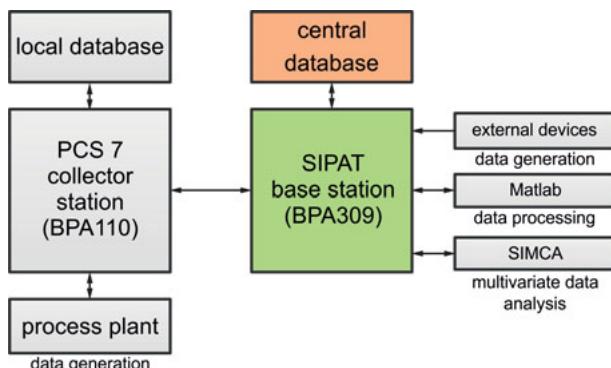


Bild 6.26: Prinzipielle Struktur von SIPAT

Darüber hinaus enthält SIPAT bereits Treiber für das Raman-Spektrometer (Kaiser Optics, Frankreich) und das NIR-Spektrometer (Polytec, Waldbronn).

Die Einspeisung von Daten zum Prozessmonitoring und zur Prozesssteuerung über MVDA-Modelle aus SIMCA sowie über MATLAB-Skripte erfolgt über die OPC-Verbindungen mit Rechner BPA309.

Die Struktur von SIPAT kann in zwei Komponenten, die *base station* und die *collector station* unterteilt werden, wie in Bild 6.26 dargestellt.

Die *base station* ist das Schlüsselement der Software und ist auf dem Computer BPA309 installiert. Sie verfügt über eine GUI, genannt SIPAT Client, und eine zentrale SQL-Datenbank zur Speicherung aller gesammelten und erzeugten Daten. Die *collector stations* sind auf externen Geräten installiert. Sie sammeln Prozessdaten und speichern diese zu-

nächst auf einer lokalen SQL-Datenbank. Die Daten werden an die *base station* dann nur auf Anfrage gesendet. Damit kann die Netzwerklast und das Risiko, Daten aufgrund von Netzwerkausfällen zu verlieren, reduziert werden. Im Fall eines Netzwerkfehlers stehen somit alle fehlenden Daten noch auf der *collector station* zur Verfügung, sobald das Netzwerk wieder hergestellt ist.

Um während eines Prozesses Daten zu sammeln und zu verarbeiten, werden SIPAT-Methoden innerhalb des zugehörigen Konfigurationsmoduls des SIPAT-Client erstellt. Eine Methode beinhaltet Informationen darüber, wo und wie bestimmte Arten von Daten abzurufen und welche Modelle, inklusive deren Input- und Output-Variablen, für deren weitere Berechnungen anzuwenden sind. Außerdem ist definiert, welche Daten und welche Variablen zurück zu dem OPC-Server geschrieben werden.

Die wichtigsten Komponenten bei einer Methodenkonfiguration sind die Kollektor- und Berechnungsmodule. Kollektoren werden ebenfalls im SIPAT-Client erstellt und dann auf den *collector stations* ausgeführt. Sie enthalten Informationen, wie Daten eines bestimmten Typs zu sammeln sind.

Modelle werden im Berechnungsmodul konfiguriert. Bevor die Modelle in eine Methode integriert werden können, müssen sie innerhalb des SIPAT-Client erzeugt werden. Um ein Modell zu erzeugen, muss zunächst der Modelltyp (entweder SIMCA oder MATLAB) definiert und die entsprechenden SIMCA- oder MATLAB-Dateien angegeben werden.

Daten, die von SIPAT in Echtzeit berechnet werden, können an das Prozessleitsystem, wo sie als Regler-Sollwerte oder Parameter dienen, ausgegeben werden. Zudem können die aus SIPAT extrahierten Daten in MFCS/win visualisiert und damit für eine Verbesserung der Prozessüberwachung gewonnen werden.

Im Falle des beschriebenen Upstream-Prozesses erfasst die SIPAT Methode die Daten alle drei Minuten. Die Variablen von jedem Sammelereignis werden in die entsprechende Datenmatrix (D) geleitet und anhand ihres Zeitstempels zugeordnet. Die Daten können als Textdatei oder im Excel-Format exportiert werden, wodurch sie direkt mit MATLAB und SIMCA Offline-Anwendungen kompatibel sind.

### 6.8.3 Das Offline-Modul SIMCA

Die Software-Lösung SIMCA 13 (Umetrics, Schweden) wird für die Berechnungen von multivariaten Prozessmodellen verwendet. Modelle werden als USP-Dateien (.usp), die von SIMCA-online für Online-Vorhersagen und Prozesssteuerungen eingesetzt werden, gespeichert. Offline berechnete Modelle können in eine SIPAT Methode für Online-Prognosen eingebettet werden. Jedoch kann die verfügbare Version von SIPAT (4.0.0.3) nur SIMCA-Modelle, die in einer älteren Version von SIMCA (SIMCA P<sup>+</sup>) entwickelt wurden, handhaben.

### 6.8.4 Das Online-Modul SIMCA Q

SIMCA Q ist eine Version von SIMCA, die über keine grafische Benutzeroberfläche verfügt. Sie kann als eine mathematische Maschine verstanden werden, die von SIPAT benutzt wird, um die mit SIMCA P<sup>+</sup> erstellten Modelle online zu berechnen. SIMCA Q kann auf verschiedenen Softwarelösungen, wie MATLAB, Microsoft Excel und LabVIEW, implementiert werden.

### 6.8.5 Das Prädiktor- und Control-Modul SIMCA-online

SIMCA-online ist eine Software-Lösung für die Echtzeitberechnung von multivariaten Batch-Modellen und für die Vorhersage zukünftiger Variablenverläufe, Chargen-Endbedingungen sowie angepassten Regler-Sollwerten für eine MPMC – *Model Predictive Multivariate Control* von Prozessqualitätsattributen. Die verwendete Software-Umgebung von SIMCA-online ist in Bild 6.27 gezeigt.

Der SIMCA-online-Server (SS) führt die Berechnung von multivariaten Modellen, die als SIMCA USP-Dateien integriert sind, aus. Er regelt zudem die Datenübertragung zwischen dem SIMCA-online-Client (GUI von SIMCA-online), den externen Datenquellen und der SIMCA-Online-Datenbank. Die Kommunikation zwischen der SS und anderen Datenquellen wird über sogenannte SimApi DLLs – *dynamic link libraries* (SimApis) ermöglicht.

Eine SimApi enthält Informationen dazu, wie man auf Daten aus einer bestimmten externen Datenquelle zugreifen kann, und in einigen Fällen auch, wie man auf diese schreiben kann. Für die meisten Arten von Datenquellen gibt es bereits eine SimApi, die als Download auf der Homepage der Fa. Umetrics erhältlich ist. Eine SimApi muss dann auf dem Host-Computer der SS installiert werden. Die SimApis werden innerhalb der Serveroptionen konfiguriert. Der SS ruft Daten aus drei verschiedenen Quellen (OPC-Server, SQL-Datenbank, DBMaker) ab.

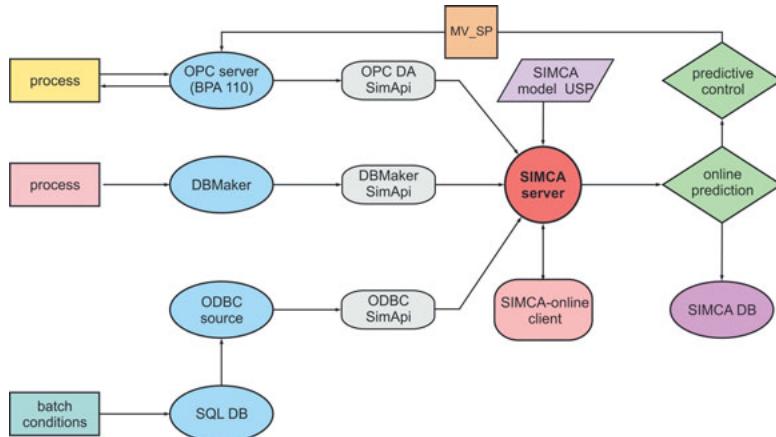


Bild 6.27: Verwendete Software-Umgebung von SIMCA-online

Die Batch-Identifikationsnummer, der Start- und Endzeitpunkt von jeder Charge sowie zeit-unabhängige Variablen, wie anfängliche Zelldichte oder Substratkonzentration, werden als *batch conditions* zusammengefasst. Diese Batch-Bedingungen werden in einer SQL-Datenbank vorab definiert. Um der SS diese Daten zur Verfügung zu stellen, müssen sie in einer ODBC – *Open Database Connectivity* Datenquelle, die von der ODBC SimApi gelesen werden kann, abgelegt sein.

Aktuelle Prozessdaten werden von dem OPC-Server, der auf dem externen Rechner BPA110 implementiert ist, unter Verwendung der entsprechenden OPC DA – *Data Access* SimApi, abgerufen. Diese SimApi erlaubt es SIMCA-online, zuvor berechnete Regler-Sollwerte der *manipulated variables* auf den OPC-Server zurück zu schreiben. Daten einer Prädiktion von Variablen werden in der SIMCA-Online-Datenbank (SDB) gespeichert.

## 7. Multivariate Datenanalyse zur Prozessbewertung

### 7.1 Zielsetzung des MVDA-Einsatzes

Das Ziel der weiteren Untersuchungen liegt darin, mit mathematischen Methoden die Reproduzierbarkeit der entwickelten sequentiell/parallelen Prozessführung letztendlich online zu verifizieren und ggf. durch Prozesseingriffe zu steuern. Die Beherrschung von großen Datensätzen ist dabei das zu lösende Problem.

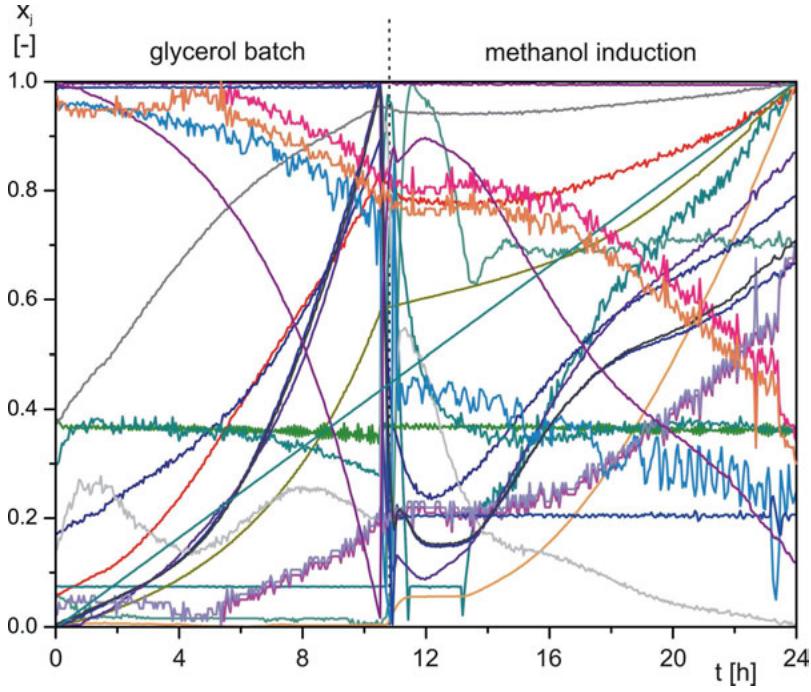


Bild 7.1: Normierte Darstellung eines Upstream-Zyklus

Dieses ist in Bild 7.1 näher erkennbar, in dem ein Zyklus des Upstream-Prozesses dargestellt ist. Hier wurden die Daten von 22 Online-Variablen, auf 0 bis 1 normiert, aufgetragen. Bei einer Abtastzeit von 3 min und einer Prozessdauer von 24 h ergeben sich daraus 480 Observationen und damit 10560 Messwerte (Objekte).

Aus dieser heterogenen Datenmenge sind vektorielle Kenngrößen oder auch zeitliche Trajektorien zu abstrahieren, die zur Prozessbewertung geeignet sind.

### 7.2 MVDA als PAT-Tool zur Beherrschung großer Datenmengen

Ein Schlüsselaspekt von PAT ist die Analyse von Prozessdaten und die Umwandlung in relevante Informationen. Mit Multivariater Datenanalyse (MVDA) kann durch eine Datenreduktion, Datenmodellierung und Vorhersage der Güte eines Prozesses zu diesem Aspekt beigetragen werden. MVDA ist daher ein wichtiges Werkzeug für die Implementierung von PAT in Produktionsprozessen.

Eine Reduktion von großen Datenmengen geschieht zum einen durch Extraktion prozessrelevanter Daten, aber auch durch eine Zusammenfassung von Daten mit gleichem bzw. ähnlichem Informationsgehalt. Zudem können durch eine simultane Betrachtung aller Daten oder von Datengruppen Informationen über nicht direkt messbare Prozesseigenschaften gewonnen werden. Dieser Gewinn an Information kann beispielsweise dafür Verwendung finden, ein multivariates Überwachungssystem einzurichten oder die Reproduzierbarkeit von Prozessen zu analysieren. Eine häufig angewandte Methode hierfür ist die *Principal Component Analysis* (PCA) [Yoon and MacGregor, 2004].

Ein weiteres wesentliches Nutzungsgebiet stellen multivariate Regressionsverfahren dar. Hierbei wird ein funktioneller Zusammenhang zwischen leicht bestimmbaren Messgrößen und schwierig messbaren Variablen, oft den CQAs, hergestellt. Hiermit kann im Idealfall durch das Prozessprofil auf diese Prozessgrößen geschlossen und so aufwendige analytische Verfahren unterstützt oder sogar auf diese verzichtet werden. Die am häufigsten verwendete Methode für diese Anwendung ist die *Partial Least Square Regression* (PLS).

Die in dieser Arbeit verwendete Software SIMCA [Umetrics, Schweden] ermöglicht die Umsetzung und Bewertung einer Vielzahl dieser multivariaten Anwendungen.

Im Weiteren wird beschrieben, wie unterschiedliche Methoden der MVDA eingesetzt werden können, um einen hohen Informationsgewinn, der nach Möglichkeit bereits online zur Verfügung steht, zu gewinnen.

## 7.3 Vorbereitung der Prozessdaten zur Anwendung der MVDA

### 7.3.1 Aufbau der Prozessdatenmatrix D

Aus den in Bild 7.1 gezeigten Online-Variablen wird in SIPAT die nicht normierte Datenmatrix D ( $n \times m$ ) mit  $j = 1, 2, \dots, m$  Spalten (Variablen) und  $i = 1, 2, \dots, n$  Zeilen (Observationen) mit äquidistanten Zeitintervallen über den gesamten Verlauf, ggf. durch Interpolation, aufgebaut.

### 7.3.2 Entfernen von Datenausreißern

Für eine Offline-Erstellung von Modellen wurden die verwendeten Daten batchweise vorbereitet, um Daten-Ausreißer zu entfernen.

Als Erstes wurden unerwünschte Daten (z.B. durch Systemfehler) beseitigt. Dazu wurden für jede Variable maximal und minimal zugelassene Werte definiert. Werte, die außerhalb dieser Grenzen lagen, wurden entfernt und durch ein leeres Matrixelement ersetzt. Beispielhaft wurde die untere Grenze für einen gemessenen pH-Wert mit 0 und die obere Grenze mit 14 festgelegt.

Anschließend wurde zur Identifikation von Ausreißern ein Grubbs-Test [Grubbs, 1950] mit einem Signifikanzniveau von 0,05 und einer Intervalllänge von 5 Abtastzeiten durchgeführt. Die als Ausreißer erkannten Datenpunkte wurden ebenfalls wie Duplikate, d.h. mehrere Messwerte zum selben Messzeitpunkt (z.B. von einer Atline-Messung), gelöscht und durch ein leeres Matrixelement ersetzt.

### 7.3.3 Die skalierte Prozessdatenmatrix X

Die ursprüngliche (unbearbeitete) Prozessdatenmatrix D eines einzelnen Batchs besteht – wie bereits beschrieben – aus  $n$  Zeilen (Beobachtungen) und  $m$  Spalten (Variablen). Daher

kann D als ein Datenraum, dessen Dimension gleich der Anzahl m der Variablen ist, angesehen werden.

Für weitere Erläuterungen wird eine Prozessdatenmatrix D mit n = 20 Beobachtungen und m = 3 Variablen betrachtet. Die Matrix D besteht aus den drei Datenvektoren  $d_1$ ,  $d_2$  und  $d_3$ .

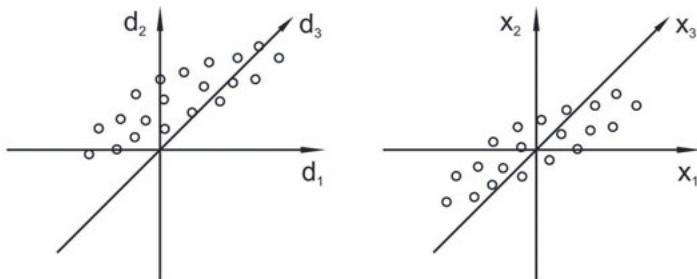


Bild 7.2: Koordinatensysteme von Rohdaten (Matrix D) und korrespondierenden, mittenzentrierten Daten (Matrix X), bestehend aus drei Variablen und 20 Objekten

Bild 7.2 (links) zeigt die Darstellung der Elemente von D im dreidimensionalen Raum. Diese werden für die Hauptkomponentenanalyse zweckmäßig autoskaliert, d.h. mittenzentriert und standardisiert, um eine skalierte Matrix X zu erhalten. Dadurch liegt der Datenschwerpunkt im Ursprung des Koordinatensystems und die Varianz ist für jede Variable eins. Bild 7.2 (rechts) zeigt die Verschiebung der Datenmatrix D durch Autoskalierung.

Dieses erweist sich als vorteilhaft in Fällen, in denen die Variablen aufgrund ihrer physikalischen Eigenschaften und Dimension unterschiedliche Größenordnungen aufweisen. Variablen höherer Größenordnung würden sonst aufgrund ihrer höheren absoluten Abweichungen (*sum of squares*) übermäßige Einflüsse auf ein PCA-Modell haben.

Um die Elemente  $x_{ij}$ ,

$$x_{ij} = \frac{d_{ij} - \bar{d}_j}{s_{dj}}, \quad (7.1)$$

der auto-skalierten Datenmatrix X zu erhalten, wird zunächst der Spaltenmittelwert  $\bar{d}_j$ ,

$$\bar{d}_j = \frac{1}{n} \cdot \sum_{i=1}^n d_{ij}, \quad (7.2)$$

und die Standardabweichung  $s_{dj}$ ,

$$s_{dj} = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^n (d_{ij} - \bar{d}_j)^2}, \quad (7.3)$$

mit

$d_{ij}$  := Element der Messdatenmatrix D

$\bar{d}_j$  := Mittelwert der Spalte j in der Messdatenmatrix D

n := Anzahl der Messwerte einer Variablen j,

bestimmt.

## 7.4 Eine kurze Einführung in PCA – Principal Component Analysis

Zur Erläuterung der weiteren Durchführungen wird eine kurze Einführung in die PCA (*Principal Component Analysis*) vorgenommen.

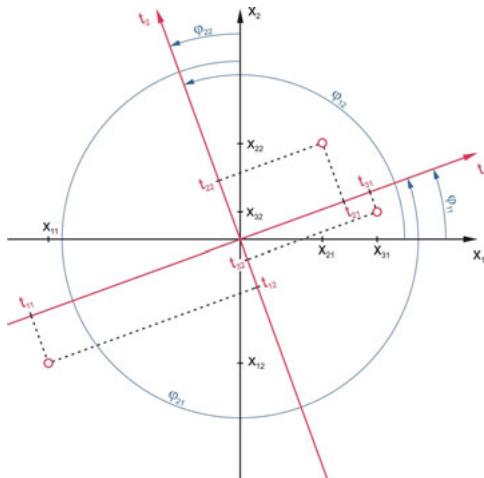


Bild 7.3: Schematische Darstellung der PCA

In Bild 7.3 sind zwei Messgrößen \$x\_1\$ und \$x\_2\$ (\$m = 2\$) aus drei Zeitpunkten (\$n = 3\$) aufgezeichnet. Das mit der skalierten Matrix \$X\$,

$$X = [x_1 \ x_2] = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{bmatrix}, \quad (7.4)$$

verbundene Achsenkreuz \$x\_1, x\_2\$ soll mit der Hauptkomponentenanalyse (PCA) durch die Transformationsvorschrift,

$$X = T \cdot P^T + E = [t_1 \ t_2] \cdot [p_1 \ p_2]^T + E \quad (7.5)$$

mit

\$T := (n \times r)\$ Score Matrix der PCA

\$P := (m \times r)\$ Loading Matrix der PCA

\$E := (n \times m)\$ Residuenmatrix,

so gedreht werden, dass die der Score Matrix \$T\$ zugeordneten latenten Variablen \$t\_1, t\_2\$ in ihrer ersten Hauptkomponente \$t\_1\$ den höchstmöglichen Informationsgehalt der Matrix \$X\$ beinhalten. Da \$T\$ in Bild 7.3 die gleiche Anzahl an Variablen wie \$X\$ besitzt (\$r = m\$), ist die Residuenmatrix \$E = 0\$. Somit beinhaltet die zweite Hauptkomponente \$t\_2\$ den Rest der Information.

Die Score Matrix \$T\$,

$$T = X \cdot P \cdot [P^T \cdot P]^{-1} = X \cdot P \cdot I, \quad (7.6)$$

ist bei Kenntnis der Loading Matrix \$P\$ der autoskalierten Prozessdatenmatrix \$X\$ fest zugeordnet.

Aus Bild 7.3 ist der Bezug zwischen den Variablen  $x_1$  und  $x_2$  und dem Score-Vektor  $t_i$ ,

$$t_i = \sum_{j=1}^2 x_j \cdot p_{ji} = x_1 \cdot p_{1i} + x_2 \cdot p_{2i} = x_1 \cdot \cos \varphi_{1i} + x_2 \cdot \cos \varphi_{2i}, \quad i = 1, 2, \quad (7.7)$$

leicht zu ermitteln. Jeder vektoriellen Variable  $x_j$  wird dabei nur eine Wichtung (Loading  $p_{ji}$ ) für die Berechnung des gesuchten Scores  $t_i$  zugeordnet. Diese ist über  $\cos \varphi_{ji}$  festgelegt.

Nach Studie der geometrischen Struktur in Bild 7.3 erkennt man, dass die Loadings in P,

$$\begin{aligned} p_{11} &= \cos \varphi_{11} \\ p_{21} &= \cos \varphi_{21} = \cos(1.5\pi + \varphi_{11}) = -\sin \varphi_{11} \\ p_{12} &= \cos \varphi_{12} = \cos(0.5\pi + \varphi_{11}) = -\sin \varphi_{11} \\ p_{22} &= \cos \varphi_{22} = \cos \varphi_{11}, \end{aligned} \quad (7.8)$$

allein über  $\cos \varphi_{11}$  bzw.  $\sin \varphi_{11}$  beschrieben werden können.

Die trigonometrische Verknüpfung aller Winkel führt dazu, dass die Inverse  $[P^T P]^{-1}$  zur Einheitsmatrix I wird. Dieser Zusammenhang gilt allgemein bei der PCA. Die unbekannten Winkel  $\varphi_{ji}$  bzw. deren Kosinus werden durch Minimierung der Summe der orthogonalen Abstände der Achse  $t_i$  zu den Wertepaaren ( $x_{ij}$ ) im m-dimensionalen Raum bestimmt. Dies ist ein nichtlineares Optimierungsproblem, das iterativ gelöst wird.

Die Identifizierung der unbekannten  $p_{ji}$  erfolgt mit dem in Tabelle 7.1 zusammengefassten NIPALS – Nonlinear Iterative Partial Least Squares Algorithmus [Wold, 1966], der mit dem Vektor der X-Matrix beginnt, der den höchsten Informationsgehalt enthält. Dieser lässt sich an der Varianz des Vektors messen.

Tabelle 7.1: Arbeitsschritte des NIPALS-Algorithmus [Schneider, 2007]

Schritt	Funktion	Beschreibung
1	$t_i := x_j$	Wählen einer beliebigen Spalte $x_j$ aus der Matrix X als initialen Score Vektor $t_i$
2	$p_i := \frac{X^T \cdot t_i}{t_i^T \cdot t_i}$	Projizieren der Matrix X auf den Unterraum $t_i$ , um den entsprechenden Loading Vektor $p_i$ zu erhalten
3	$p_i = \frac{p_i}{\ p_i\ }$	Normalisieren von $p_i$ auf die Länge 1, mit $\ p_i\  = \sqrt{\sum_{j=1}^m p_{ij}^2}$
4	$t_{i,old} := t_i$	Zwischenspeichern des aktuellen Scores $t_i$
5	$t_i := \frac{X \cdot p_i}{p_i^T \cdot p_i}$	Projizieren der Matrix X auf den Unterraum $p_i$ , um den entsprechenden neuen Score Vektor $t_i$ zu erhalten
6	$d := \ t_{i,old} - t_i\ $	Bestimmen einer Differenz d zwischen neuem und altem Scorevektor
7	$d < z$	Falls d größer als ein im Vorfeld definierter Schwellenwert z (z.B. $z = 10^{-6}$ ) ist, beginnt eine erneute Iteration des Algorithmus bei Schritt 2. Ansonsten fortfahren mit Schritt 8
8	$E := X - t_i \cdot p_i^T$	Entfernen der Information aus X, die bereits durch Hauptkomponente $t_i$ beschrieben wird
9	$X := E$	Um weitere Hauptkomponenten zu finden, startet der Algorithmus mit der neuen Datenmatrix und $i = i+1$ erneut bei Schritt 1

Da die autoskalierten Variablen in der Datenmatrix X alle die gleiche Varianz  $\sigma = 1$  besitzen, ist dieser Vektor ggf. aus der D-Matrix herauszufinden.

Mit den gefundenen Loadings ist für das Problem in Bild 7.3 die Score Matrix T,

$$T = X \cdot \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = X \cdot \begin{bmatrix} \cos \varphi_{11} & -\sin \varphi_{11} \\ \sin \varphi_{11} & \cos \varphi_{11} \end{bmatrix}, \quad (7.9)$$

dann einfach aus X berechenbar.

## 7.5 Anwendung der MPCA zur Prozessbewertung

### 7.5.1 Aufbau einer dreidimensionalen Prozessdatenmatrix D

Bei der Anwendung der MVDA zur Bewertung des sequentiell/parallelen Produktionsprozesses verschärft sich die Datenlage, da nunmehr eine Reihe von Prozesszyklen (o) simultan bearbeitet wird.

In Bild 7.4 ist die erweiterte Struktur der Prozessdatenmatrix D gezeigt.

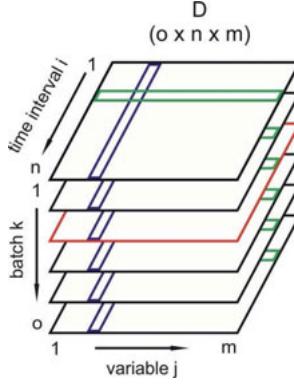


Bild 7.4: Dreidimensionale Struktur der Prozessdatenmatrix D

Jeder Zyklus bildet mit seiner  $n \times m$  Untermatrix  $X_k$  einen Batch. Ähnliche Daten existieren für die Anzahl  $k = 1, 2, \dots, o$  solcher Batches. Allerdings variiert aufgrund von unterschiedlichen Gesamtlängen der Teilprozesse die Anzahl  $n$  von Zeitintervallen pro Batch. Einige sind länger und andere etwas kürzer als der durchschnittliche Zyklus. Zu lange Abläufe werden daher durch Abschneiden der letzten Werte auf die mittlere Batchlänge angepasst. Zu kurze Zyklen werden mit den zuletzt gemessenen Werten bis hin zur Länge  $n$  aufgefüllt.

Alle Daten können dann in einer dreidimensionalen Datenmatrix D ( $o \times m \times n$ ) zusammengefasst werden. Eine derartige Matrix ist in Bild 7.4 dargestellt. Darin sind unterschiedliche Chargenläufe (k) entlang der vertikalen Seite und die Prozessgrößen (j) entlang der horizontalen Seite angeordnet. Die zeitliche Entwicklung (i) der Variablen (j) wird in Richtung der dritten Dimension beschrieben.

Jeder horizontale Schnitt (rot) durch dieses Array ist eine ( $n \times m$ ) Batch-Datenmatrix, die Zeitverläufe oder Trajektorien für alle Variablen eines Batches (k) enthält. Jede Stirnscheibe (grün) ist eine ( $o \times m$ ) Matrix, die die Werte aller Variablen von allen Batches zu einem gemeinsamen Zeitpunkt (i) enthält. Zusätzlich ist eine vertikale Scheibe (blau) eine ( $o \times n$ ) Mat-

rix, die den Zeitverlauf oder die Trajektorie einer bestimmten Variablen ( $j$ ) für alle Batches enthält.

### 7.5.2 MPCA – Multiway Principal Component Analysis

Im Unterschied zur PCA an einer zweidimensionalen Struktur wird bei der MPCA – *Multiway Principal Component Analysis* mit der dreidimensionalen Datenstruktur  $D$  gearbeitet, die zunächst entfaltet und dabei in verschiedene autoskalierte Matrizen  $X_A$ ,  $A = B, M$ , überführt werden kann.

In der Praxis haben sich zwei Methoden durchgesetzt, die in Bild 7.5 gezeigt sind.

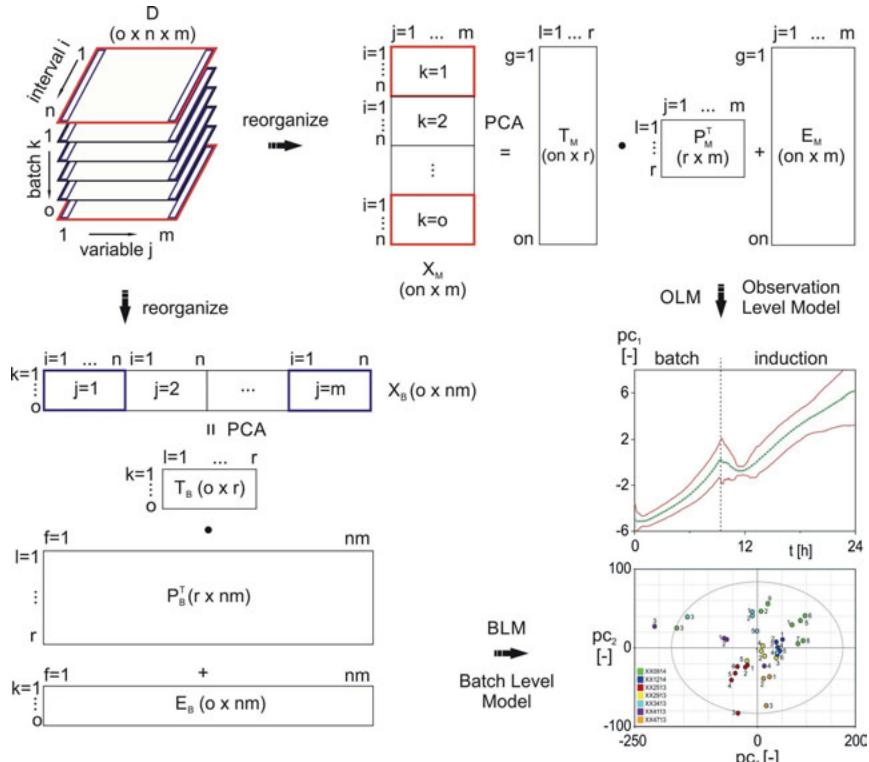


Bild 7.5: Entfaltung der dreidimensionalen Datenmatrix in zweidimensionale Datenmatrizen

Das Ziel dieser Operationen liegt darin, eine Prozessbewertung mit einer weitaus geringeren Anzahl ( $r$ ) an latenten Variablen als die zur Verfügung stehenden Online-Prozessgrößen ( $m$ ) zu entwickeln.

Für eine *Offline-Klassifizierung* schlagen Nomikos und MacGregor [Nomikos and MacGregor, 1994] vor,  $D$  so zu entfalten, dass nach der Autoskalierung eine Matrix  $X_B$  ( $o \times nm$ ) entsteht, in der jeder vertikale Schnitt ( $o \times n$ )  $m$ -fach wiederholt Seite an Seite angeordnet wird.

Aus dieser Anordnung wird das *Batch Level Model* (BLM),

$$X_B = T_B \cdot P_B^T + E_B \quad (7.10)$$

mit

$X_B$  := skalierte variablenorientierte ( $o \times nm$ ) Datenmatrix des BLM

$T_B$  := ( $o \times r$ ) Score-Matrix des BLM

$P_B$  := ( $r \times nm$ ) Loading-Matrix des BLM

$E_B$  := ( $o \times nm$ ) Residuen-Matrix des BLM,

generiert.

Die Differenz zwischen dem Messwert  $x_{Bkij}$  und dem mit Scores und Loadings geschätzten Wert  $\hat{x}_{Bkij}$  wird als Residuum  $e_{Bkij}$ ,

$$e_{Bkij} = x_{Bkij} - \hat{x}_{Bkij} = x_{Bkij} - t_{Bki} \cdot p_{Blij}^T, \quad (7.11)$$

ein Element der Residuen-Matrix  $E_B$ , bezeichnet.

Die Zahl der statistisch signifikanten latenten Variablen  $r$  wird durch eine Kreuzvalidierung vorgenommen, um einerseits ein *Underfitting* durch zu wenige Hauptkomponenten und andererseits ein *Overfitting* durch zu viele Hauptkomponenten zu vermeiden [Kessler, 2007].

Beschränkt man sich bei der Darstellung der BLM auf zwei Scores, lässt sich mit Hilfe eines in Bild 7.6 gezeigten *Score Scatter Plots* eine Analyse der beteiligten Zyklen durchführen. So liegen z.B. Batches, die dem mittleren Zyklus entsprechen, nahe am Koordinatenursprung, während Batches, die davon abweichen, weiter entfernt liegen. Starke Ausreißer lassen sich mit Hilfe der in Abschnitt 7.8 näher erläuterten Hotellings  $T^2$ -Statistik identifizieren.

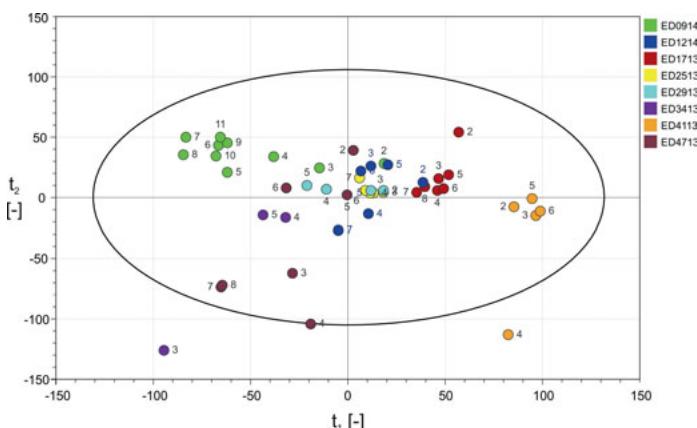


Bild 7.6: Score Scatter Plot eines Batch Level Models von acht Produktionskampagnen

In dem gezeigten Beispiel liegen die Batches ED3413\_3 und ED4113\_4 außerhalb der schwarzen Ellipse, die den Bereich kennzeichnet, der normale von abnormalen Zyklen trennt. Eine weitere Funktion dieses Plots ist die Aufdeckung von Mustern. Batches, die einer gleichen Kampagne angehören, d.h. zur selben Jahreszeit aus der gleichen Vorkultur durchgeführt wurden, sind dazu in einheitlicher Farbe markiert. Es lässt sich daraus erkennen, dass Batches einer gemeinsamen Kampagne oft im gleichen Quadranten des Plots liegen.

Diese Cluster entstehen zum Beispiel durch eine von den Jahreszeiten abhängige Umgebungstemperatur im Labor. Hohe Temperaturen, z.B. im Sommer, erfordern ein verstärktes Kühlen des Reaktors, wodurch die Variable  $g_{DJk}$  verändert in das Modell einfließt. Das betrifft dann jeweils alle Batches dieser Kampagne. Das BLM formt somit ein Allround-Bild eines Batches, indem es ihn als einen Datenpunkt zusammenfasst und im latenten Koordinatensystem abbildet [Kourt et al., 1995].

Für *Klassifizierungs-Modelle* in der Batch-Ebene enthalten vor allem die vorgestellten *Score Scatter Plots* wichtige Informationen und können daher für eine globale Bewertung von Batches eingesetzt werden.

Für ein *Online-Monitoring* ist es hingegen zweckmäßig, die Matrix D ( $o \times n \times m$ ) in der Art zu entfalten, dass eine Matrix  $X_M$  ( $o \times n \times m$ ) entsteht. Hierbei werden die Batches als horizontale Schnitte ( $n \times m$ ) untereinander angeordnet [Wold et al., 1998, 2009].

Ein aus dieser entfalteten Matrix hergeleitetes *Observation Level Model* (OLM),

$$X_M = T_M \cdot P_M^T + E_M \quad (7.12)$$

mit

$X_M$  := skalierte batchorientierte ( $o \times n \times m$ ) Datenmatrix des OLM

$T_M$  := ( $n \times r$ ) Score-Matrix des OLM

$P_M$  := ( $r \times m$ ) Loading-Matrix des OLM

$E_M$  := ( $o \times n \times m$ ) Residuen-Matrix des OLM,

beschreibt dabei mit den Scores  $t_i$  den zeitlichen Verlauf der Informationen, die in  $r$  orthogonalen Vektoren der Länge  $o \cdot n$  aufgeteilt sind. In Bild 7.7 ist ein derartiger Verlauf von drei Hauptkomponenten ( $i = 1, 2, 3$ ) in drei *cell breeding* Zyklen ( $k = 4, 5, 6$ ) aufgezeichnet.

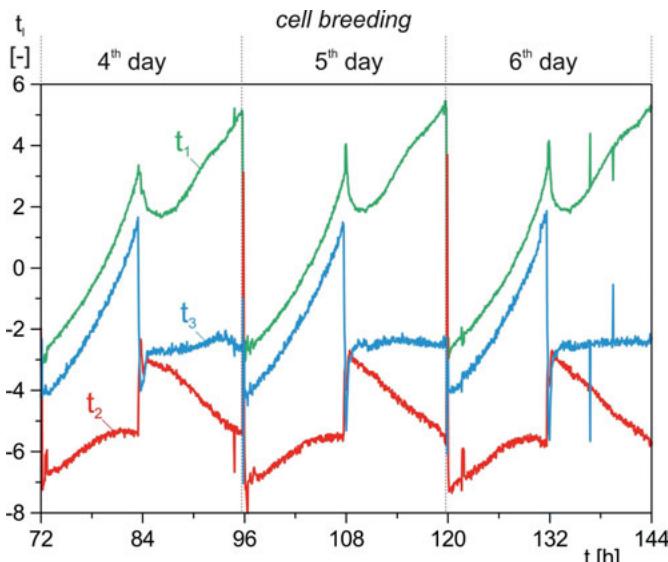


Bild 7.7: Online-Verlauf der drei Hauptkomponenten  $t_1$  (grün),  $t_2$  (rot) und  $t_3$  (blau) des *cell breeding* Prozesses über drei aufeinanderfolgende Zyklen

Im Gegensatz zum BLM kann das OLM für die Überwachung des Fortschritts von Chargenprozessen verwendet werden, indem der laufende Batch in Echtzeit gegen einen Referenzverlauf getestet wird.

Analysiert man die Scores  $t_i$  des Prozesses in Bild 7.7, so erkennt man, dass sie sich in den drei Prozesszyklen wiederholen. Das impliziert, die Scores jeweils übereinander zu legen und über die Mittelwerte  $\bar{t}_i$  die Reproduzierbarkeit zu bewerten.

Beschränkt man diese Methode auf die erste Hauptkomponente  $\bar{t}_1$  mit einer Varianzbreite von  $\pm 3\sigma_1$ , so wird diese Darstellung auch als *Golden Batch* oder auch *Golden Batch-Tunnel* bezeichnet.

Der Mittelwert  $\bar{t}_1$  wird im Weiteren  $pc_1$  (*principal component 1*) genannt. Er wird für die mittlere Zykluszeit von 24 h berechnet, indem statt einer PCA eine PLS auf eine künstliche Zeitachse  $y$  (*maturity*) durchgeführt wird.

## 7.6 Entwicklung des Golden Batch Modells

*"In monitoring, the concept is to capture in a model the acceptable "common cause" variations in the process and use this model to detect any deviations from such behavior"* [MacGregor and Cinar, 2012].

Ein *Golden Batch*-Modell soll nun den idealen Verlauf eines Produktionsprozesses in seiner ersten Hauptkomponente abbilden. Die Idee ist, dazu ein Modell auf der Grundlage von repräsentativen Batches zu erstellen, bei denen weder ungewöhnliche Abweichungen noch Bedienungsfehler am Prozess aufgetreten sind und die nachweislich eine hohe Prozess- und/oder Produktqualität geliefert haben.

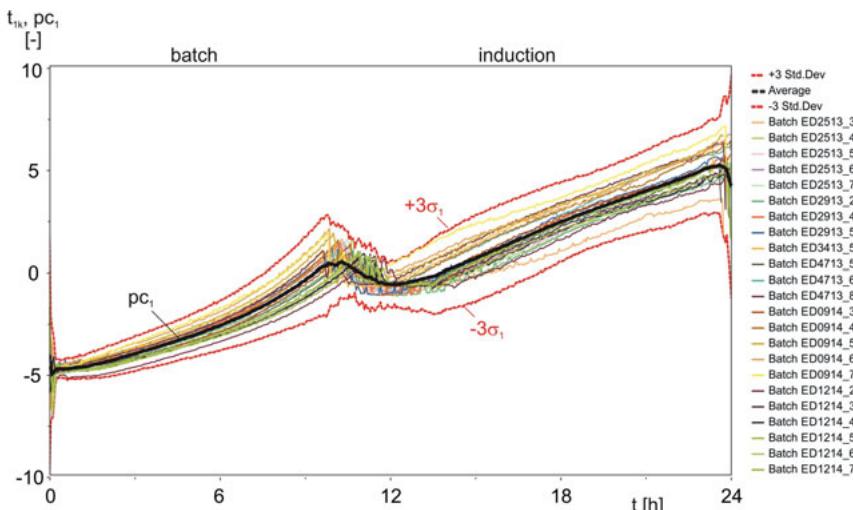


Bild 7.8: Verlauf der ersten Hauptkomponenten  $t_{ik}$  von 23 repräsentativen *cell breeding* Batches aus 6 Kampagnen zur Erstellung der *Golden Batch*-Trajektorie  $pc_1$  (sw) inklusive des *Golden Batch*-Tunnels  $\pm 3\sigma_1$  (rt)

Später soll dann für einen zukünftigen Prozess, der diesem Verlauf innerhalb vordefinierter Grenzen folgt, allein anhand dieser Trajektorie die Prozessqualität bewertet bzw. auch sichergestellt werden.

Bild 7.8 zeigt den zeitlichen Verlauf der ersten Hauptkomponente  $pc_1$  eines *Golden Batch*-Modells für den *cell breeding* Prozess. Die Grundlage des Modells lieferten  $o = 23$  repräsentative Batches.

Aus den Verläufen von  $t_{1k}$  für alle  $o$  Batches lässt sich damit der Mittelwert  $\bar{t}_1$ ,

$$\bar{t}_1 = \frac{1}{o} \cdot \sum_{k=1}^o t_{1k} = pc_1, \quad (7.13)$$

leicht errechnen, der den Verlauf der *Goldenen Batch*-Trajektorie beschreibt. Diese wird im Weiteren mit  $pc_1$  bezeichnet.

Zusätzlich lassen sich die Grenzen  $\pm 3\sigma_1$ ,

$$\pm 3\sigma_1 = \pm 3 \cdot \sqrt{\frac{1}{o-1} \sum_{k=1}^o (t_{1k} - \bar{t}_1)^2}, \quad (7.14)$$

bestimmen, die den *Golden Batch*-Tunnel bilden. So entsteht eine einfach beobachtbare Überwachung für zukünftige Prozesse.

Mit den bekannten Modellkoeffizienten (Loadings) ist die erste Hauptkomponente  $t_1$  künftiger Batches nach Gl. (7.9) berechenbar und kann ggf. online mit dem Referenzmodell, der *Golden Batch*-Trajektorie, verglichen werden.

Batches, deren Score  $t_1$  vollständig innerhalb des *Golden Batch*-Tunnels verläuft, wird eine hohe Prozessqualität zugeordnet, wohingegen Batches, in denen  $t_1$  den Tunnel verlässt, als fehlerhaft zu bezeichnen sind.

## 7.7 Anwendung der MPLS zur Prozessprädiktion

### 7.7.1 Hinführung auf das MVDA-Problem

Bei der MPCA gilt es, durch verschiedene Operationen mit der Prozessdatenmatrix D, einen vektoriellen oder zeitlichen Informationsgehalt zu erlangen. Dieser wird als Score bezeichnet.

In diesem Abschnitt wird beschrieben, wie man aus den Online-Daten (D) z.B. offline gemessene Analysenwerte (M) rekonstruieren kann. Diese Methode wird später dann auch online zur Bewertung des Reaktionsablaufes angewendet.

Hierbei gilt es, zunächst durch Projektion der Online-Datenmatrix D auf die Offline-Datenmatrix M ein Modell zu erstellen, das in der Lage ist, M mit D darzustellen.

Dies wird nach entsprechender Autoskalierung von D in X und M in Y mit Hilfe einer MPLS – *Multiway Partial Least Squares* Regression erreicht. Das zugehörige Modell muss allerdings mit verlässlichen Daten trainiert und verifiziert werden.

Tabelle 7.2 zeigt das Problem. Eine ( $n \times m$ ) Online-Datenmatrix D mit hoher Dichte (n) steht einer ( $s \times v$ ) Offline-Analysenmatrix M mit weitaus geringerer Messfrequenz (s) gegenüber. Um beide Matrizen auf die gleiche Anzahl von Messungen (n) zu bringen, gibt es zwei Wege. Die erste Möglichkeit besteht in der Ausdünnung der Zeilen von Matrix D, in dem nur Online-

Daten zu den Offline-Messzeitpunkten  $t_i$  verwendet werden. Dies birgt die Gefahr, dass Online- oder Offline-Messausreißer die Qualität der Vorhersage stark vermindern.

Tabelle 7.2: Gegenüberstellung von Prozessdatenmatrix D und Analysematrix M

$\backslash$	$N_{st}$ [rpm]	$F_{nG}$ [lh <sup>-1</sup> ]	...	$X_{O2}$ [%]	$C_{XL}$ [gl <sup>-1</sup> ]	$C_{S1M}$ [gl <sup>-1</sup> ]	...	$C_{P1M}$ [gl <sup>-1</sup> ]
$t_0$	$d_{11}$	$d_{12}$	...	$d_{1m}$	$m_{11}$	$m_{12}$	...	$m_{1v}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$t_{n-1}$	$d_{n1}$	$d_{n2}$	...	$d_{nm}$	$m_{s1}$	$m_{s2}$	...	$m_{sv}$

Der hier angewandte Weg liegt in der Vorabkonstruktion von ergänzten Offline-Datensätzen. Dieser Ansatz wird später Estimation genannt. Bei der Zelldichte ist z.B. der angepasste Wert  $c_{XL,turb}(t_i)$  als  $c_{XLest}(t_i)$  sofort zu übernehmen. Andere Offline-Verläufe werden ebenfalls durch nichtlineare Regressionen in mathematische Funktionen gezwängt.

Nach entsprechender Autoskalierung entsteht aus der erweiterten ( $n \times v$ ) Offline-Matrix  $M_{est}$  dann die ( $n \times v$ ) Matrix der Attribute Y. Diese gilt es mit X zu korrelieren.

### 7.7.2 MPLS – Multiway Partial Least Squares Regression

PLS ist ein Verfahren, das den mathematischen Bezug zwischen der Datenmatrix X und der Antwortmatrix Y herstellt. Beide Matrizen sind dabei aus mehreren Vektoren  $x_j$  ( $j = 1..m$ ) bzw.  $y_h$  ( $h = 1..v$ ) aufgebaut, wobei  $x_j$  nun den Verlauf eines Datenprozesseinganges und  $y_h$  den einer Prozessantwort repräsentiert.

Es gibt viele multivariate Regressionsmethoden, von denen PLS üblicherweise in der chemischen und biopharmazeutischen Industrie [Brereton, 2003; Kessler, 2007] eingesetzt wird. Neben PLS werden auch multiple lineare Regression (MLR) und Hauptkomponentenregression (PCA) für verschiedene Arten von Problemen angewendet.

MLR wird beispielsweise in DoE-Anwendungen benötigt. Dabei sind die X Matrizen orthogonal aufgebaut, so dass die Variablen (Faktoren) nicht miteinander korrelieren. Für MLR muss X umkehrbar sein, um die Modellkoeffizienten des DoE-Ansatzes zu berechnen.

Bei biotechnologischen Batchprozessen sind die Daten häufig mehr oder weniger stark korreliert, wodurch eine alleinige Anwendung von MLR nicht mehr sinnvoll ist. In solchen Situationen lässt sich dieses Problem durch PCA und die einhergehende Berechnung der Hauptkomponenten von X überwinden. Es wird dann die resultierende, orthogonale Ergebnismatrix T mit der Antwortmatrix Y durch MLR verknüpft.

Bei einer PLS werden die Matrizen X und Y gleichzeitig durch eine PCA modelliert, wobei sich beide Modelle untereinander beeinflussen. Dabei werden die Variablen  $x_j$ , die mit den Spaltenvektoren  $y_h$  korreliert sind, in den ersten Hauptkomponenten  $t_i$  abgebildet, auch wenn sie in einer von Y unabhängigen PCA einen geringen Einfluss auf diese Komponente hätten. Aus diesem Grund können sich die berechneten Hauptkomponenten  $t_i$  der gleichen X-Datenmatrix erheblich voneinander unterscheiden, je nachdem ob man eine PLS oder eine PCA anwendet.

PLS wird häufig verwendet, um aus spektroskopischen Messdaten die Konzentration einer oder mehrerer Komponenten in einer komplexen Lösung abzuschätzen.

In dieser Arbeit wurden dagegen routinemäßig erfasste Online-Variablen dazu verwendet, die Zelldichte sowie Substrat- und Produktkonzentrationen und deren Reaktionen im Reaktormedium abzuschätzen. Diese Anwendung wird in Kapitel 9 beschrieben.

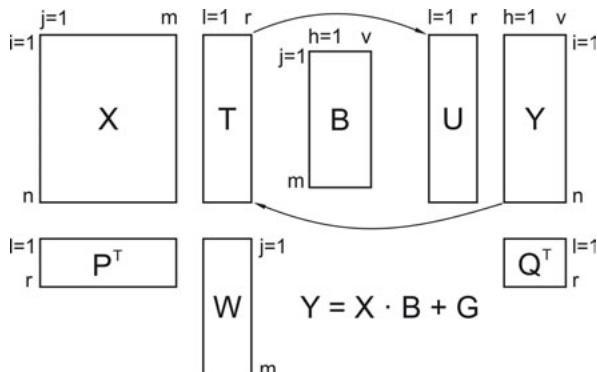


Bild 7.9: Korrelation der Datenmatrizen X und Y durch eine PLS Regression

Die Struktur von PLS ist in Bild 7.9 zusammengefasst. Y ist durch Autoskalierung aus der erweiterten Datenmatrix  $M_{est}$  entstanden. Die Matrizen X, T und P sind bereits von der PCA bekannt.

Da im Falle der PLS die Scores T dazu verwendet werden, X und Y gleichzeitig zu approximieren, unterscheiden sie sich von denen einer reinen PCA. Im Gegensatz zur PCR, wird die Antwortmatrix Y von einer separaten Score-Matrix U und einer Loadingsmatrix Q modelliert [Faber and Kowalski, 1997].

Das Ziel der PLS ist es, die Antwortvariablen in Y durch die Eingangsvariablen in X vorherzusagen. Dabei wird ein multivariater Ansatz verwendet, um ein Regressionsmodell für Y,

$$Y = X \cdot B + G \quad (7.15)$$

mit

$X := (n \times m)$  autoskalierte Prozessdatenmatrix

$Y := (n \times v)$  autoskalierte Matrix der bekannten Attribute

$B := (m \times v)$  Matrix der Regressionskoeffizienten

$G := (n \times v)$  Residuenmatrix der PLS-Anpassung,

zu erhalten.

Die dafür erforderliche Regressionskoeffizienten-Matrix B,

$$B = W \cdot (P^T \cdot W)^{-1} \cdot Q^T \quad (7.16)$$

mit

$B := (m \times v)$  Matrix der Regressionskoeffizienten zur Prädiktion von Y aus X

$W := (m \times r)$  Matrix zur Wichtung der Loadings

$P^T := (r \times m)$  transponierte Loadings der Matrix X

$Q^T := (r \times v)$  transponierte Loadings der Matrix Y,

wird aus den Loadings Q der Matrix Y und den gewichteten Loadings P der Matrix X berechnet.

Die Wichtung erfolgt dabei mit einer Matrix W. Die Elemente der drei unbekannten Matrizen Q, P und W gilt es durch geeignete PLS-Algorithmen zu finden. Hierzu gibt es zwei verschiedene Ansätze.

Der erste, PLS1 genannt, wird verwendet, um Attribute Y mit nur einer Antwort ( $v = 1$ ) zu modellieren. Der zweite, PLS2, ist in der Lage eine Matrix Y mit einer oder mehreren Antworten ( $v \geq 1$ ) zu modellieren. Dazu müssen die Spalten in Y allerdings miteinander korreliert sein.

In einem weiteren Lösungsansatz wird daher zunächst eine PCA für Y durchgeführt, um nicht korrelierte Variablen in Score-Klassen voneinander zu trennen. Im Anschluss kann dann für jede der erhaltenen Klassen eine separate PLS erfolgen.

Da PLS1 nur eine Sonderform von PLS2 darstellt, wird im Folgenden nur die letztere genauer beschrieben. Hierbei werden gleichzeitig zwei PCA Modelle verwendet und aneinander angepasst.

Die PLS lässt sich somit in eine PCA für X,

$$X = T \cdot P^T + E \quad (7.17)$$

mit

$T$  := (n×r) Scorematrix von X

$P$  := (m×r) Loadingmatrix von X

$E$  := (n×m) Residuenmatrix von X

und in eine für Y,

$$Y = U \cdot Q^T + F \quad (7.18)$$

mit

$U$  := (n×r) Scorematrix von Y

$Q$  := (v×r) Loadingmatrix von Y

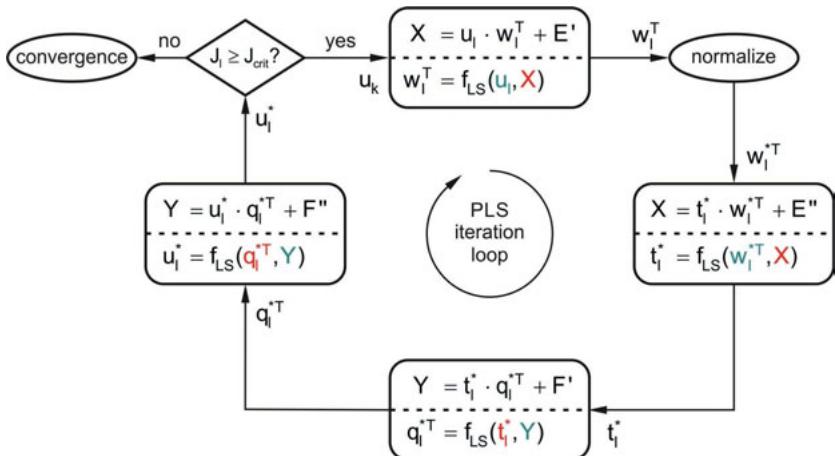
$F$  := (n×v) Residuenmatrix von Y,

aufspalten.

Die Informationen der beiden PCA-Modelle werden durch einen PLS-Algorithmus miteinander verbunden. Dabei verwendet dieser vektorielle Scores und Loadings von einem PCA-Modell, um beide jeweils für das andere Modell abzuschätzen bis dieser Vorgang konvergiert.

Es gibt vier Teilschritte in dieser Iteration, um die benötigten Modellvektoren zu berechnen. Jeder Vektor wird durch eine Least-Square-Lösung ( $f_{LS}$ ) berechnet, die aus einem lokalen Modell abgeleitet werden kann. Jede LS-Lösung ist dabei eine Funktion von X (oder Y) und einem zusätzlichen Vektor, der Informationen von der anderen Seite des PLS-Modells Y (oder X) bereitstellt.

Der an NIPALS angelehnte iterative PLS-Lösungsverlauf, die daran beteiligten Vektoren und Matrizen sowie die zugehörigen Gleichungen sind in Bild 7.10 dargestellt. Der genaue Ablauf ist bei Wold [Wold et al., 2001] beschrieben.



$X$  := ( $n \times m$ ) autoskalierte Prozessdatenmatrix

$\mathbf{Y} := (\mathbf{n} \times \mathbf{v})$  Matrix der gesuchten Attribute

$u_l$  := ( $n \times 1$ ) Scorevektor der l-ten Hauptkomponente von Y

$w^T$  : (1x m) gewichteter Loadingvektor der l-ten Hauptkomponente von X

$t_i$  :=  $(n \times 1)$ -Scorevektor der i-ten Hauptkomponente von  $X$

$c^T$  :=(1×v) gewichteter Loadingvektor der l-ten Hauptkomponente von Y

$E, E' := (\text{num})$  Residuenmatrizen eines lokalen Modells für  $X$

$E, E' := (\text{Rxxm})$  Residualmatrixen eines lokalen Modells für  $X$ ,  
 $E\bar{I}, E\bar{I}' := (\text{rxy})$  Residuumsmatrizen eines lokalen Modells für  $Y$ .

Bild 7.10: Iterationsschleife des PI-S Algorithmus (X: rot; Y: grün)

Nachdem die gewünschten Hauptkomponenten  $t$  identifiziert sind, stehen auch die Vektoren

$$p_i^T = \frac{t_i^T \cdot X}{t_i^T \cdot t}, \quad (7.19)$$

$$q_i^T = \frac{t_i^T \cdot Y}{t_i^T \cdot t}, \quad (7.20)$$

$$u_i = \frac{Y \cdot q_i}{q_i^T \cdot q} \quad \text{und} \quad (7.21)$$

$$w_i^T = \frac{u_i^T \cdot X}{\|u_i^T\|}, \quad (7.22)$$

der Matrizen  $P$ ,  $Q$  und  $W$  fest.

Mit diesen wird in Gl. (7.16) die Regressionsmatrix B berechnet. Diese ist dann in weiteren Experimenten zu nutzen, um mit Gl. (7.15) die Attribute Y und damit die nicht mehr zu messenden Offline-Größen M abzuschätzen.

Da das Berechnungsverfahren auf beiden Seiten (X und Y) des PLS-Modells basiert, ist PLS anderen Regressionsverfahren, besonders, wenn große Datenmengen zu handhaben sind, überlegen.

Trotzdem ist ein PLS-Modell auf die Informationen aus dem Referenzdatensatz X beschränkt. Dadurch kann es zu Ungenauigkeiten bei der Vorhersage kommen, wenn neue Daten einen anderen Ursprung haben als die Referenz-Daten.

Die Robustheit eines PLS-Modells kann dadurch verbessert werden, dass Daten eines Experimental Design mit einbezogen werden, die eine große Bandbreite von möglichen Störungen, die das untersuchte System beeinflussen könnten, abdecken.

In jedem Fall muss ein PLS-Modell bei jeder Änderung der experimentellen Voraussetzungen überprüft werden. Was ebenfalls für die PCA gilt. Dazu stehen viele Diagnosehilfsmittel zur Verfügung. Diejenigen, die in dieser Arbeit Anwendung gefunden haben, werden im folgenden Abschnitt 7.8 beschrieben.

## 7.8 Modellvalidierung

Ein großes Problem bei der Modellbildung ist die Identifizierung von Ausreißern und von Beobachtungen, die einen starken Einfluss auf das angestrebte Modell haben.

In Abschnitt 7.3 wurde bereits beschrieben, wie die Prozessdatenmatrix D von Ausreißern bereinigt wird. In diesem Abschnitt folgen weitere auf einer statistischen Auswertung basierende Tests, die später auch für die Prozessbeobachtung Anwendung finden werden.

Beobachtungen, die stark vom Modell abweichen, können mit Hilfe der Hotellings  $T^2$ -Statistik identifiziert werden [Hotelling, 1931].

Die Idee des  $T^2$ -Wertes,

$$T^2 = \sum_{l=1}^r \frac{(t_l - \bar{t}_l)^2}{\sigma_{ll}^2} \quad (7.23)$$

mit

$T^2$  := Hotellings  $T^2$ -Skalar über alle r Scores

$t_l$  := Scorevektor der Hauptkomponente l

$\bar{t}_l$  := Mittelwert des Scorevektors der Hauptkomponente l

$\sigma_{ll}^2$  := Varianz (*sum of squares*) aller Elemente des Scorevektors  $t_l$ ,

liegt darin, dass eine Beobachtung, die exakt dem Modell entspricht, genau auf den mittleren Scores  $\bar{t}_l$  des Modells verläuft und damit die Hotellings  $T^2$ -Werte gleich Null sind.

Daher wird die Wahrscheinlichkeit, dass eine Beobachtung nicht mit einem Kalibrierungsdatensatz (Trainingssatz) übereinstimmt, mit zunehmendem Hotellings  $T^2$ -Wert höher. Dieser lässt sich sowohl für das OLM als auch für das BLM berechnen.

Darüber hinaus werden auch Vertrauengrenzen (*confidence limits*) festgelegt.

Zum Beispiel sind die Hotellings  $T^2$ -Werte von 95 % aller Beobachtungen kleiner oder gleich einem kritischen Wert  $T_{\text{crit}, 95\%}^2$ ,

$$T_{\text{crit}, 95\%}^2 = \frac{r \cdot (z^2 - 1)}{z \cdot (z - r)} \cdot F_{(\alpha, r, z-r)}, \quad \alpha = 1 - 0.95 = 0.05 \quad (7.24)$$

mit

$T_{\text{crit}, 95\%}^2$  := 95 % Vertrauensbereich der Hotellings  $T^2$ -Werte

$\alpha$  := gewähltes Signifikanzniveau

$z$  := Anzahl der Beobachtungen (OLM:  $z = o \cdot n$ , BLM:  $z = o$ )

$r$  := Anzahl der verwendeten Hauptkomponenten

$F_{(\alpha, r, z-r)}$  := Kritischer Wert einer F-Verteilung mit einem Signifikanzniveau von  $\alpha$  sowie  $r$  und  $z - r$  Freiheitsgraden,

der von dem gewählten Signifikanzniveau  $\alpha$ , den erforderlichen  $r$  Hauptkomponenten und der Anzahl  $z$  der Beobachtungen geprägt wird. Beim BLM ist  $z$  die Anzahl der Batches  $o$  und beim OLM die Anzahl der  $o \cdot n$  Observationen.

Der kritische F-Wert wird aus dem F-Test [Billeter, 1972] abgeleitet und kann einer Tabelle, z.B. [Reinfeldt and Tränkle, 1976], entnommen werden.  $F$  ist dabei eine Funktion des gewählten Signifikanzniveaus ( $\alpha$ ) und den Freiheitsgraden der beiden Datensätze ( $r$ ,  $z - r$ ).

Eine weitere Möglichkeit mit der Hotellings  $T^2$ -Statistik starke Ausreißer zu detektieren, ist die Überprüfung, ob eine Beobachtung außerhalb des Vertrauensbereichs in einem *Score Scatter Plot* liegt.

Die Ellipse aus Bild 7.11a beschreibt einen Vertrauensbereich des BLM, bei dem 95 % aller Beobachtungen innerhalb der Ellipse liegen sollten.

An dieser Stelle wird deutlich, dass für sehr kleine Stichproben die Verwendung einer F-Verteilung sehr kritisch zu bewerten ist, da die Statistik für repräsentative Kohorten ausgelegt ist.

Ein kleiner Stichprobenumfang liegt allerdings häufig bei einem BLM vor, da nur eine limitierte Anzahl an Batches zur Verfügung steht.

Im Gegensatz dazu reicht die Anzahl an Beobachtungen in einem OLM für eine zuverlässige statistische Hotellings  $T^2$ -Untersuchung meistens aus.

Die beiden Radien  $r_{\text{ellipse}, l}$  der 95%-Ellipse für zwei beliebige Scores  $l$ ,

$$r_{\text{ellipse}, l} = \frac{2 \cdot (z^2 - 1)}{z \cdot (z - 2)} \cdot \sigma_{tl}^2 \cdot F_{(\alpha, 2, z-2)}, \quad \alpha = 1 - 0.95 = 0.05 \quad (7.25)$$

mit

$r_{\text{ellipse}, l}$  := Radius der Ellipse des Vertrauensintervalls in Richtung der Komponente  $l$

$F_{(\alpha, 2, z-2)}$  := Kritischer Wert einer F-Verteilung mit einem Signifikanzniveau von  $\alpha$  sowie 2 und  $z - 2$  Freiheitsgraden ( $z = o \cdot n$  für OLM,  $z = o$  für BLM)

$\sigma_{tl}^2$  := Varianz (*sum of squares*) aller Elemente des Scorevektors  $t_l$ ,

lassen sich dabei leicht errechnen.

Der Skalar des *standard prediction error* des OLM,

$$\text{SPE}_{xi} = \sigma_{\text{pred},xi}^2 = \sum_{j=1}^m (x_{ij} - \hat{x}_{ij})^2 = \sum_{j=1}^m e_{ij}^2 \quad (\text{OLM}) \quad (7.26)$$

summiert die Residuen zu einer Beobachtung  $i$  der  $(o \cdot n \times m)$  Modellmatrix  $X_M$  auf, während der Skalar des *standard prediction error* des BLM,

$$\text{SPE}_{xk} = \sigma_{\text{pred},xk}^2 = \sum_{t=1}^{n \cdot m} (x_{kf} - \hat{x}_{kf})^2 = \sum_{t=1}^{n \cdot m} e_{kf}^2 \quad (\text{BLM}) \quad (7.27)$$

die Residuen in einem Batch  $k$  der  $(o \times n \cdot m)$  Modellmatrix  $X_B$  berücksichtigt.

Teilt man den jeweiligen  $\text{SPE}_x$  durch die zur Verfügung stehenden  $z - r$  Freiheitsgrade und zieht daraus die Wurzel, so erhält man die *absolute distance to model*  $D\text{Mod}X_{\text{abs}}$ ,

$$D\text{Mod}X_{\text{abs}} = \sqrt{\frac{\text{SPE}_x}{z - r}}, \quad (7.28)$$

die ein Maß für den Abstand einer Beobachtung von ihrer Projektion in den reduzierten Raum der latenten Variablen ist. Um herauszufinden, ob eine Beobachtung dem Modell zugehörig ist, wird der  $D\text{Mod}X_{\text{abs}}$  normiert.

Dabei ergibt sich die *normalized distance to model*  $D\text{Mod}X_{\text{norm}}$ ,

$$D\text{Mod}X_{\text{norm}} = \sqrt{\frac{D\text{Mod}X_{\text{abs}}}{\sigma_{\text{resid}}^2}} = D\text{Mod}X, \quad (7.29)$$

aus der Division von  $D\text{Mod}X_{\text{abs}}$  durch die Gesamtvarianz aller Residuen des Trainingsdatensatzes  $\sigma_{\text{resid}}^2$ ,

$$\sigma_{\text{resid}}^2 = \begin{cases} \frac{\sum_{i=1}^n \sum_{j=1}^m e_{ij}^2}{(n - r - r_0) \cdot (m - r)}, & (\text{OLM}), \\ \frac{\sum_{k=1}^o \sum_{t=1}^{n \cdot m} e_{kf}^2}{(o - r - r_0) \cdot (n \cdot m - r)}, & (\text{BLM}), \end{cases} \quad (7.30)$$

mit

$r_0 := 1$  bei zentrierten Modellen, sonst 0 [Eriksson et al., 2006],

die wiederum für das OLM und das BLM unterschiedlich zu berechnen sind.

$D\text{Mod}X_{\text{norm}}$  wird im Weiteren mit  $D\text{Mod}X$  bezeichnet. Dieser besitzt bei ausreichend großer Stichprobe eine F-Verteilung und kann demnach mit einem kritischen  $D_{\text{crit},\alpha}$ , verglichen werden. Dieser Wert wird von der F-Verteilung und dem eingesetzten Signifikanzniveau abgeleitet [Eriksson et al., 2006].

Bild 7.11 und Bild 7.12 zeigen die beschriebenen Hilfsgrößen einer Hotellings  $T^2$ -Ellipse sowie den kritischen Hotellings  $T^2$ -Wert und den  $D\text{Mod}X$ -Wert an einem Beispiel für die Erstellung eines OLM sowie eines BLM für den *cell breeding*-Prozess.

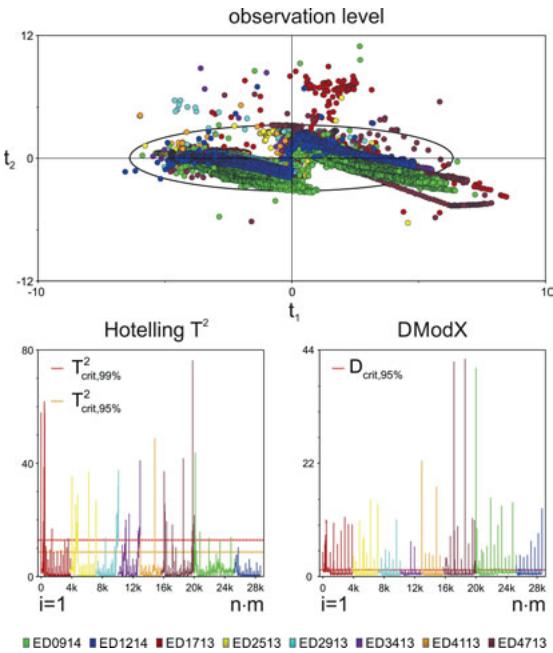


Bild 7.11: Modellvalidierung über statistische Hilfsgrößen für das OLM

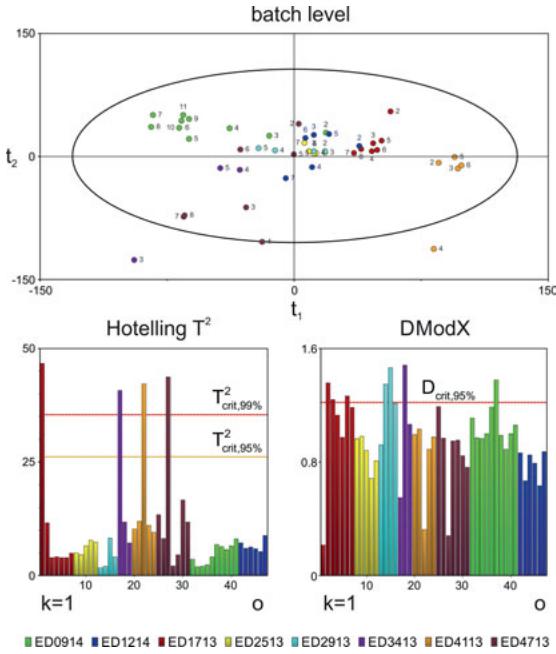


Bild 7.12: Modellvalidierung über statistische Hilfsgrößen für das BLM

Beide Modelle haben ein weitreichendes Optimierungspotential. Es ist eindeutig erkennbar, dass sowohl im OLM als auch im BLM Observationen weit außerhalb der kritischen Bereiche liegen. Durch eine Begutachtung dieser Observationen und ggf. deren Ausschluss aus der Modellbildung ließe sich in beiden Fällen noch erhebliche Verbesserung erzielen. Diese könnten anhand der im Folgenden beschriebenen Parameter gemessen werden.

Sobald ein PLS-Modell aufgestellt worden ist, sollte es validiert werden. Um die Performance eines Modells zu beschreiben werden zwei charakteristische Parameter verwendet. Ihre Berechnung erfolgt sowohl für das OLM als auch für das BLM, wird im Folgenden aber ausschließlich für das OLM vorgestellt, da die Berechnung für das BLM analog dazu verläuft.

Hierbei wird die Summe der quadrierten Residuen der X-Matrix  $RSS_X$ ,

$$RSS_X = \sum_{i=1}^n \sum_{j=1}^m (x_{ij} - \hat{x}_{ij})^2 = \sum_{i=1}^n \sum_{j=1}^m e_{ij}^2, \quad (7.31)$$

und der Y-Matrix  $RSS_Y$ ,

$$RSS_Y = \sum_{i=1}^n \sum_{h=1}^v (y_{ih} - \hat{y}_{ih})^2 = \sum_{i=1}^n \sum_{h=1}^v f_{ih}^2, \quad (7.32)$$

sowie die Gesamtvarianz der X-Matrix  $\sigma_X^2$ ,

$$\sigma_X^2 = \sum_{i=1}^n \sum_{j=1}^m (x_{ij} - \bar{x}_{ij})^2, \quad (7.33)$$

und der Y-Matrix  $\sigma_Y^2$ ,

$$\sigma_Y^2 = \sum_{i=1}^n \sum_{h=1}^v (y_{ih} - \bar{y}_{ih})^2, \quad (7.34)$$

benötigt, um daraus sowohl den Anteil der vom Modell beschriebenen *Varianz im X-Raum*  $R_{Xcum}^2$ ,

$$R_{Xcum}^2 = 1 - \frac{RSS_X}{\sigma_X^2}, \quad (7.35)$$

als auch den Anteil der vom Modell beschriebenen *Varianz im Y-Raum*  $R_{Ycum}^2$ ,

$$R_{Ycum}^2 = 1 - \frac{RSS_Y}{\sigma_Y^2}, \quad (7.36)$$

zu berechnen. Diese Parameter werden i. A. auch als *goodness of fit* bezeichnet.

Ein weiterer Parameter ist die *goodness of prediction*  $Q_{Ycum}^2$ ,

$$Q_{Ycum}^2 = 1 - \frac{\sum_{i=1}^n \sum_{h=1}^v (y_{ih} - \hat{y}_{ih,cv})^2}{\sigma_Y^2} = 1 - \frac{PRESS_Y}{\sigma_Y^2}, \quad (7.37)$$

mit

$\hat{y}_{ih,cv}$  := aus einer Kreuzvalidierung (cv) geschätzter Wert für  $y_{ih}$ ,

die auf der Grundlage einer Kreuzvalidierung bestimmt werden kann [Kessler, 2007].

Bei einer Kreuzvalidierung wird ein Datenpunkt nach dem anderen aus der Modellbildung ausgeschlossen und jeweils ein neues Modell konstruiert. Mit diesem Modell wird dann der Wert  $y_{ih}$  abgeschätzt. Der Schätzer wird als  $\hat{y}_{ih, cv}$  bezeichnet, wobei das  $cv$  für *cross validation* steht.

Ein weiterer bedeutender Wert zur Bewertung der Güte eines Modells ist der in Abschnitt 6.4 vorgestellte *root mean square error of validation* RMSEV,

$$\text{RMSEV} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}, \quad (7.38)$$

mit

$y_i$  := Werte des Trainingsset

$\hat{y}_i$  := Schätzwerte für die Daten des Trainingsset

$n$  := Anzahl der Messwerte,

der in einen *relative root mean square error of validation* RMSEV<sub>rel</sub>,

$$\text{RMSEV}_{\text{rel}} = \frac{\text{RMSEV}}{y_{\max} - y_{\min}} \cdot 100\%, \quad (7.39)$$

mit

$y_{\max}$  := maximaler Wert der Trainingsdaten

$y_{\min}$  := minimaler Wert der Trainingsdaten,

transformiert werden kann [Hyndman and Koehler, 2006].

Die in diesem Kapitel vorgestellten theoretischen Grundlagen werden nun im nachfolgenden Kapitel 8 zunächst zur Offline-Bewertung der zyklisch ablaufenden Produktionsprozesse eingesetzt. Darauf aufbauend erfolgt ein Online-Monitoring der Reproduzierbarkeit und damit der Qualität dieser Prozesse. Abschließend wird eine Steuerung des Golden Batches bei Verlassen des Vertrauensbereichs einer QbD-konform entwickelten Prozessführung vorgestellt.

## 8. Golden Batch Diagnose, Monitoring und Control zyklischer Teilprozesse

### 8.1 Forschungsbedarf einer MVDA basierten Prozessführung

Im Verlauf der Etablierung eines *Design Space* kann mit Hilfe einer Multivariaten Datenanalyse (MVDA) ein leicht zu interpretierendes Prozess-Monitoring aufgebaut werden. Prinzipiell teilt sich die MVDA dabei in die beiden Ebenen Batch Level (BL) und Observation Level (OL) auf [Wold et al., 1998].

Bei dem BL wird die Gesamtheit eines Batches als eine einzige vektorielle Information betrachtet. Die BL-Modelle werden i.d.R. offline für eine Klassifizierung von Prozessen in repräsentative und nicht-repräsentative Batches genutzt, um damit einen Einfluss der Anfangsbedingungen und der Prozessvariablen auf den Ausgang des Prozesses zu untersuchen.

OL-Modelle werden hingegen dazu angewendet, die Interaktion von Prozessvariablen zu einzelnen Messzeitpunkten bzw. über den (zeitlichen) Verlauf eines Batches zu untersuchen. Diese Modelle eignen sich dann für eine Online-Anwendung.

Aus einem Pool von repräsentativen Batches lässt sich ein sogenanntes *Golden Batch*-Modell (Referenzmodell) erstellen. In einem *Golden Batch*-Modell wird der ideale Verlauf eines Prozesses, der ein Produkt mit hoher Qualität produziert, wiedergespiegelt. Gleichzeitig werden von diesem Modell statistische Grenzen bereitgestellt, gegen die neue Batches verglichen werden können [Nomikos and Macgregor, 1995].

Neue Batches, die sich innerhalb dieser Grenzen (*Golden Batch*-Limits) entwickeln, gelten als repräsentativ – die Prozessqualität ist hoch. Dagegen gelten neue Batches die vollständig oder teilweise außerhalb der Grenzen verlaufen als nicht repräsentativ und besitzen eine geringe Prozessqualität.

Wurde im Modell-Training ein Zusammenhang zwischen der Prozessqualität und der Produktqualität nachgewiesen, dann können diese Modelle dazu genutzt werden, die Produktqualität festzustellen und darüber hinaus weitere Analysen zur Bestimmung der Produktqualität drastisch reduzieren.

Zusätzlich zu dem Echtzeit-Monitoring mit MVDA lässt sich mit Hilfe von Imputations- und Regressionsverfahren eine Prozess-Prognose erstellen. Diese wird herangezogen, um frühzeitig Anpassungen an CPPs vorzunehmen, falls unter den gegebenen Bedingungen ein Herauslaufen des Prozesses aus seinem *Design Space* als wahrscheinlich anzusehen ist. Die Methode zur Echtzeit-Kontrolle eines Prozesses innerhalb seiner multivariaten Grenzen bezeichnet man als *Model Predictive Multivariate Control* (MPMC).

Mercier et al. stellten 2014 zusammenfassend fest, dass der höchste PAT-Level mit Prozesskontrolle innerhalb des *Design Space* in Echtzeit während eines laufenden Herstellungsprozesses noch nicht erreicht ist, MVDA aber für die Erreichung dieses letztendlichen Ziels eine zentrale Rolle spielen wird [Mercier et al., 2014].

In diesem Kapitel werden MVDA Methoden erprobt, um Wege dahin aufzuzeigen.

Der entwickelte sequentiell/parallele Prozess bietet dafür eine hervorragende Grundlage, da hier reproduzierbare Prozesse in kurzer Zeit durchführbar sind und auch bei gezielter Störung zur Erprobung von Online-Algorithmen jederzeit wieder in einen QbD-konformen Zustand zurückkehren.

## 8.2 Ermittlung des Golden Batch-Tunnels für den Upstream Prozess

Die PCA-Modellierung des Golden Batches in Abschnitt 7.6 wurde auf eine Reihe von Zyklen des gesamten Upstream-Prozesses, wie er bereits in Bild 6.5 gezeigt wurde, angewendet.

Dabei entstanden zwei getrennte PCA-Modelle für den jeweils 24-stündigen *cell breeding process* ( $p = 1$ ) und den *protein production process* ( $p = 2$ ). Deren Entwicklungsschritte sind in Bild 8.1 dargestellt.

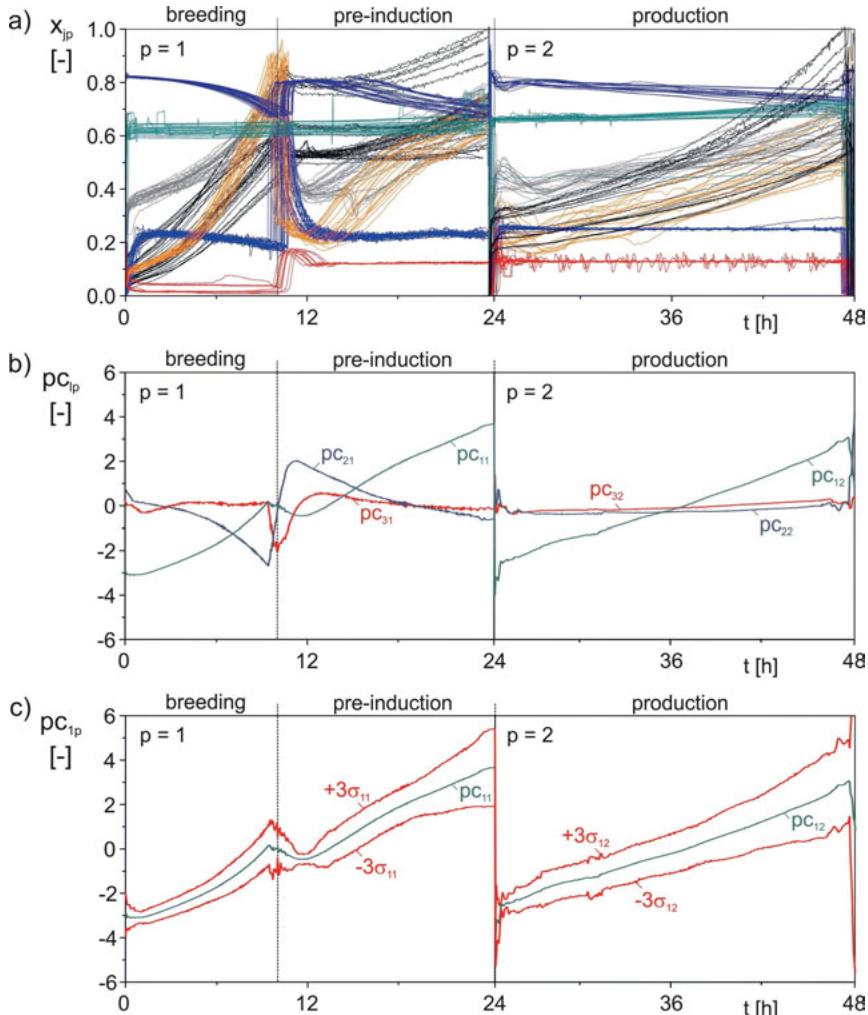


Bild 8.1: Offline-Entwicklung des Golden Batch-Tunnel für eine sequentiell/parallele Integrierte Proteinproduktion

Bild 8.1a zeigt, zwischen 0 und 1 normiert, ausgewählte Daten  $x_{jp}$ , die farblich wie folgt gekennzeichnet sind:  $N_{Stp}$  (grau),  $pO_{2p}$  (hellblau),  $c_{S2Mp}$  (rot),  $c_{XLp}$  (schwarz),  $x_{O2p}$  (dunkelblau),  $x_{CO2p}$  (orange),  $V_{Lp}$  (grün). Die normierte Form wurde hier nur für eine bessere Illustration gewählt. Für die Verarbeitung zum *Golden Batch* werden alle Online-Daten autoskaliert in eine Matrix  $X_M$  überführt. Obgleich die Daten von verschiedenen Kampagnen stammen, ist die Reproduzierbarkeit deutlich erkennbar.

In Bild 8.1b ist das Ergebnis der beiden PCA-Modelle ersichtlich. Die Mittelwerte der jeweils drei ersten Hauptkomponenten,  $pc_{1p}$  (grün),  $pc_{2p}$  (blau), und  $pc_{3p}$  (rot), beinhalten ca. 99 % der gesamten Prozessinformation.

Im *cell breeding process* ( $p = 1$ ) beschreiben  $pc_{11}$  und  $pc_{21}$  die Variation des prinzipiellen Kultivierungsprozesses. Dabei könnte  $pc_{11}$  als ein Maß für das Wachstum und  $pc_{21}$  als ein Maß des Verbrauches interpretiert werden. Die nur in der transienten Phase zwischen *breeding* und *pre-induction* von Null unterschiedliche Komponente  $pc_{31}$  beschreibt hingegen die Dynamik der Prozessänderung in Reaktor 1.

Im *protein production process* ( $p = 2$ ) sind  $pc_{22}$  und  $pc_{32}$  nahezu Null. Der gesamte Prozess ist daher nur über  $pc_{12}$  charakterisiert. Hierin sind vornehmlich die exponentiell leicht ansteigenden Zuwachsgrößen dominant.

In Bild 8.1c sind dann nur die beiden Mittelwerte der ersten Hauptkomponente,  $pc_{11}$  und  $pc_{12}$ , als Trajektorien für den *Golden Batch*-Tunnel des gesamten Upstream aufgezeichnet. Dieser ergibt sich über die Auftragung von drei Standardabweichungen  $\pm 3\sigma_{1p}$  (rot) um deren Mittelwerte  $pc_{1p}$  (grün).

Nachfolgende Prozesse, die innerhalb der  $\pm 3\sigma_{1p}$ -Bereiche verlaufen, werden als reproduzierter Zyklus eingestuft. Verlässt der Prozess den Tunnel, muss er kurzfristig zurückgeholt (s. Abschnitt 8.3) oder verworfen werden.

Die in Bild 8.1 gezeigten Ergebnisse wurden mit der MVDA-Software SIMCA® Version12 (MKS Umetrics, Schweden) erzielt [Eriksson et al., 2008].

Diese Software beinhaltet auch eine Reihe von Methoden zur Fehleranalyse beim Verlassen des Golden Batch-Tunnels. Diese werden hier nicht behandelt, sondern nur der Weg aufgezeichnet, wie mit den hergeleiteten Modellen auch ein Monitoring des Golden Batches und damit eine Online-Bewertung der Prozessreproduzierbarkeit ermöglicht wird.

### 8.3 Überwachung der Prozessreproduktion mit einem Golden Batch-Monitoring

Für den Online-Einsatz der Golden Batch-Methodik wird die in Kapitel 6 installierte Datenverarbeitung eingesetzt, die in Bild 8.2 noch einmal in ihrer Funktionalität verdeutlicht wird.

SIPAT sammelt alle Online-Daten und SIMCA stellt die Modelle sowie die Mittelwerte und Varianzbreiten der individuellen PCA-Variablen zur Verfügung.

Die MVDA-Modelle der Offline-Software SIMCA® sind in der Version 12 mit SIMCA® Q und SIPAT kompatibel. SIMCA® Q hat kein graphisches Interface und wird von SIPAT als numerische Maschine zur Online-Berechnung der Golden Batches genutzt.

Darüber hinaus ist MATLAB zur Berechnung von numerischen Online-Variablen wie  $c_{XLturb}$  und  $Q_{O2}$  auch in SIPAT eingebunden.

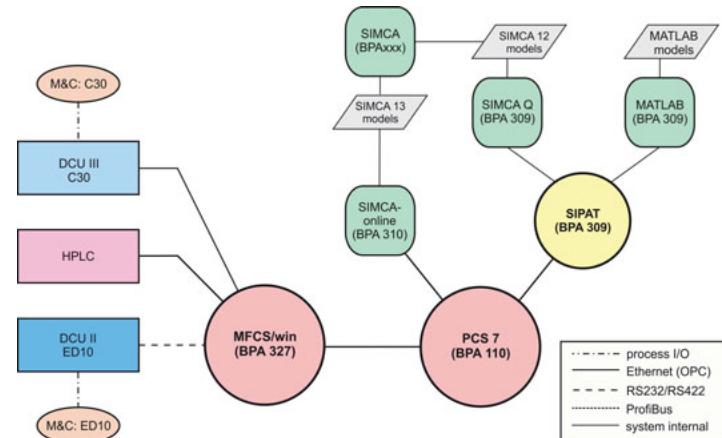


Bild 8.2: Software-Interaktionen während der Online-MVDA

In Bild 8.3 wird das Online-Monitoring eines parallel laufenden Upstream-Prozesses gezeigt. Die oberen Graphen in Bild 8.3a beinhalten Variablen des *k*-ten *cell breeding* in Reaktor 1 und die unteren in Bild 8.3b die der *k-1*-ten *protein production* in Reaktor 2.

Dabei sind drei unterschiedliche Hauptkomponenten dargestellt. Sie entstammen entweder den Mittelwerten von Prozesssignalen einer Datenquelle, wie die Zelldichte  $c_{XLP}$ , berechnet aus dem Trübungssignalen  $E_{LP}$  mit Gl. (3.1), oder aus mehreren Datenquellen, wie die bilanzierte Sauerstoffeintragsrate  $Q_{O2p}$ , bilanziert in Gl. (3.9) aus sechs Gasvariablen.

Die dritte Hauptkomponente ist die *Golden Batch*-Trajektorie  $pc_{1p}$ , die aus allen jeweiligen Online-Daten modelliert wurde. Im Online-Monitoring werden die Abläufe des aktuellen Prozesses mit diesen Zielvorgaben verglichen.

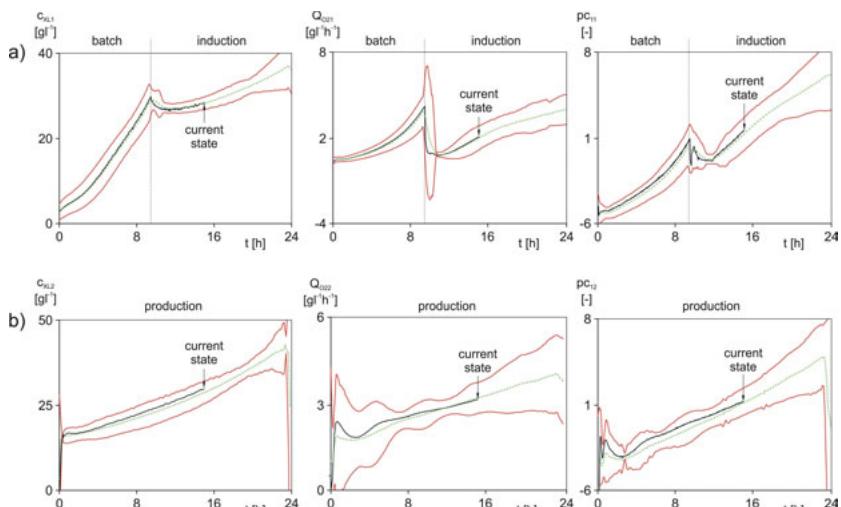


Bild 8.3: Online Monitoring eines parallelen Upstream-Prozesses  
a) *cell breeding* ( $p = 1$ ) b) *protein production* ( $p = 2$ )

Im vorliegenden Fall befinden sich beide Teilprozesse in Bild 8.3 bei  $t = 15$  h. Der *cell breeding* folgt allen drei Vorgaben nahezu deckungsgleich. Auch der Verlauf der *protein production* ist als ausgezeichnet zu bezeichnen.

Die gezeigte Methode verdeutlicht die hohe Reproduzierbarkeit der entwickelten Prozessstrategie und das hohe Potential, den *Golden Batch* als Online-Maß einer Bewertung der Prozessqualität einzusetzen. Sie hält daher immer mehr Einzug in die industrielle pharmazeutische Produktion [Menezes et al., 2009; Schaefer et al., 2014; Fonteyne et al., 2015].

Die nächsten Fragestellungen drängen sich allerdings auf.

1. Wie wird eine Verbindung zu den QbD-Vorgaben hergestellt?
2. Welche Maßnahmen sind zu ergreifen, wenn die Trajektorie  $pc_{1p}$  den Golden Batch-Tunnel verlässt?

Beide Probleme werden im nächsten Abschnitt behandelt.

## 8.4 Beobachtung, Prognose und Steuerung QbD-konformer Golden Batch-Verläufe

### 8.4.1 Erstellung QbD-konformer Golden Batch-Modelle

Im weiteren Verlauf dieser Arbeit wurden zwei von Umetrics entwickelte Methoden zur prädiktiven Vorausschau und zur Steuerung der Hauptkomponente  $t_1$  erprobt.

Dies erforderte zunächst die Erstellung von *QbD-konformen Golden Batch-Modellen*.

Hierzu wurden elf Produktionszyklen mit den von Fricke [Fricke, 2015] ermittelten, optimalen *Critical Process Parameter* (CPP)  $\vartheta_L = 22^\circ\text{C}$  und  $\text{pH} = 5.55$  gefahren. Daraus wurden die *Monitoringmodelle MM* erstellt, deren Ergebnisse Bild 8.4 zu entnehmen sind.

Die eingestellten Sollwerte  $\text{pH}_{2w}$  ( $X_{MV1}$ ) und  $\vartheta_{L2w}$  ( $X_{MV2}$ ) sind in Bild 8.4a als *osp-optimal set points* grün markiert.

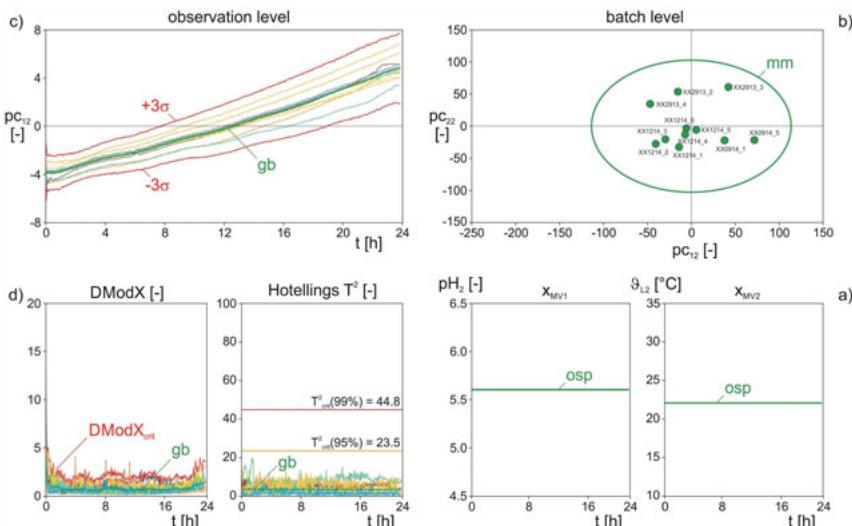


Bild 8.4: Erstellung der QbD-konformen Monitoringmodelle MM

Die zur Modellbildung herangezogenen optimalen Zyklen liegen methodenbedingt in der Hotelling  $T^2$ -Ellipse mm des *Batch Level Modells* (grün in Bild 8.4b).

Aus ihren Verläufen wurde die *Golden Batch*-Trajektorie gb und deren Tunnel  $\pm 3\sigma$ , grün und rot in Bild 8.4c, des *Observation Level Models* ermittelt. Darüber hinaus ergeben sich die Fehlerschranken für DModX (rot für  $DModX_{crit}$  in Bild 8.4d) und Hotelling  $T^2$  (rot für  $T^2_{crit,95\%}$ , gelb für  $T^2_{crit,99\%}$ , beide in Bild 8.4d).

Der nächste Schritt besteht darin, die Modelleinstellung mit neuen Produktionsprozessen zu validieren. Hierzu wurde eine weitere Kampagne XX2514 gefahren.

Die dabei vorgenommenen Manipulationen sind in Bild 8.5 dargestellt. Hierbei wird die Bewertungsstruktur von Bild 8.4 übernommen.

Das Experiment XX2514\_7 wurde mit den optimalen Sollwerten osp (braun) gefahren und lag bereits nach kurzer Zeit auf der *Golden Batch*-Trajektorie gb (grün). Der finale Punkt im Score Scatter Plot liegt innerhalb der  $T^2$ -Ellipse und die Fehlerschranken von  $DModX_{crit}$  und  $T^2_{crit,95\%}$  wurden ebenfalls erfüllt. Das zeugt von der Güte des Modells und von der Reproduzierbarkeit der Experimente.

Mit dem nächsten Experiment XX2514\_8 wurden die Sollwerte mit fsp – *false set points* absichtlich falsch eingestellt. Ein  $pH_{2w}$  von 5,25 und ein  $g_{L2w}$  von 26 °C führen dazu, dass der Score fsp (blau) den Goldenen Batch schnell verlässt, der Batch im Score Scatter Plot weit außerhalb der  $T^2$ -Ellipse liegt und die Fehlerschranken stark überschritten sind.

Im Weiteren werden nun Maßnahmen erläutert, wie diese Fehlverläufe im Voraus zu erkennen sind und wie ggf. dagegen anzusteuern ist.

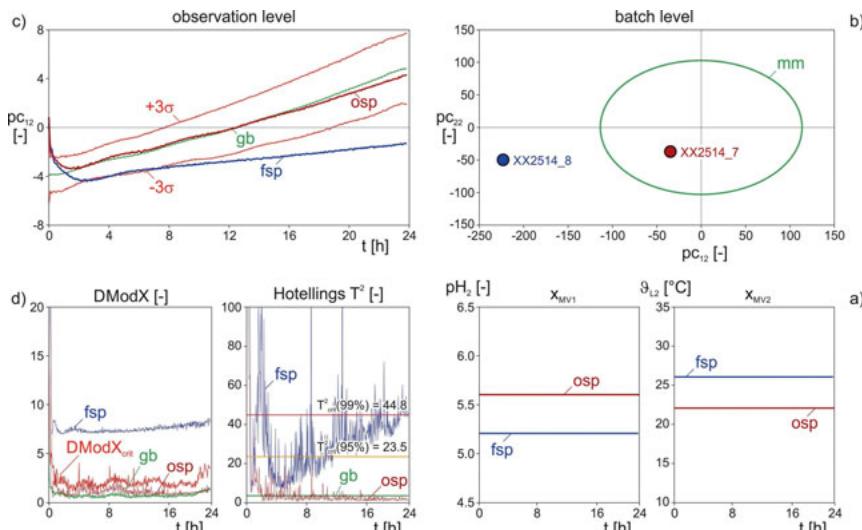


Bild 8.5: Validierung des Monitoring Modells MM

### 8.4.2 Prognose des Golden Batch-Verlaufes

Zur Prädiktion des *Golden Batch*-Verlaufes wurde von Umetrics eine Online-Methode auf der Basis des Algorithmus *Imputation By Regression* (IBR) entwickelt [Arteaga and Ferrer, 2002; McCready, 2013].

Dieser Algorihmus wird am Prozess mit der Software SIMCA-online umgesetzt. Bild 8.2 ist zu entnehmen, dass diese Software ohne SIPAT direkt auf die Daten von PCS 7 zugreifen kann. Die Online-Exekution des IBR greift dabei auf *Golden Batch*-Modelle zurück, die offline mit SIMCA in der Version 13 erstellt werden.

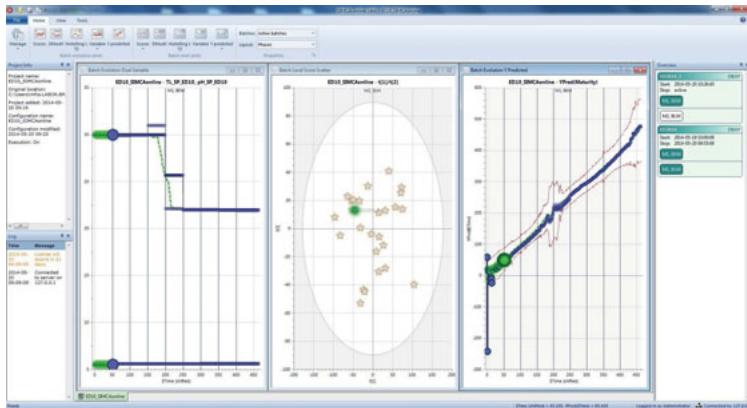


Bild 8.6: Beobachtung der Prozessprädiktion über den *Golden Batch*-Verlauf (OLM) und die Hotellings  $T^2$ -Ellipse (BLM)

Eine Online-Prädiktion wird in Bild 8.6 an Hand des Screenshots eines laufenden *cell breeding* Prozesses vorgestellt. Diese ist vor 50 Abtastintervallen (150 min) mit einem Temperatursollwert  $\vartheta_{L1w}$  von 30 °C und einem pH-Sollwert  $pH_{1w}$  von 5,25 gestartet. Die beobachteten Werte entsprechen den Sollwerttrajektorien, die im Weiteren mit *Manipulation Variable*  $X_{MV1}$  ( $pH_{1w}$ ) und  $X_{MV2}$  ( $\vartheta_{L1w}$ ) bezeichnet werden sollen.

Der IBR-Algorithmus schätzt auf Basis des Modells und der bereits vorhandenen Datenlage ab, wie der Prozess bis zum Ende der Zellanzucht verläuft. Die Schätzungen des Verlaufes des *Golden Batch*-Scores  $t_{11}$  und der Lage des *Score Scatter*-Vektors ( $t_{11}, t_{21}$ ) am Ende des Prozesses in der Hotellings  $T^2$ -Ellipse, werden dabei online aktualisiert.

Der augenblickliche Prozess erfüllt die Forderung eines reproduzierten Prozesses auszeichnet.

### 8.4.3 Modellgestützte multivariate Steuerung der Prozessqualität

In diesem Abschnitt wird die Erprobung einer neuen β-Version der Software SIMCA-online für den Einsatz einer *Model Predictive Multivariate Control* (MPMC) zur Rückholung des Scores  $pc_{12}$  in den *Golden Batch*-Tunnel der *protein production* vorgestellt.

Die Methode basiert auf der *Bound Optimization by Quadratic Approximation* (BOBYQA) [Powell, 2009].

Bei diesem Algorithmus wird ein *Process Performance Index* (PPI) in Form eines *Gütekriteriums J*,

$$\min_{X_{MV}} J \Rightarrow J = \theta_{X_{MV}} \cdot \sum_{i=1}^q (X_{MV_{\text{osp}} i} - X_{MV i})^2 + \theta_Y \cdot \sum_{h=1}^v (Y_{SP h} - Y_{preh})^2 + \theta_{DModX} \cdot (DModX)^2 + \theta_{T_2} \cdot (T^2)^2 \quad (8.1)$$

mit

- $\theta_{X_{MV}}$  := Wichtungsfaktor aller Manipulation Variables MV
- $X_{MV_{\text{osp}}}$  := ( $q \times 1$ ) Vektor der MV im Golden Batch des Monitoring Modells MM
- $X_{MV}$  := ( $q \times 1$ ) Vektor der über J mit dem Control Model CM geschätzten MV mit den Stellgrenzen  $X_{MV_{\text{min}}} \leq X_{MV} \leq X_{MV_{\text{max}}}$
- $q$  := Anzahl der manipulierten Prozesseingangsgrößen MV
- $\theta_Y$  := Wichtungsfaktor aller Prozessattribute Y
- $Y_{SP}$  := ( $v \times 1$ ) Vektor der zu berücksichtigenden Prozessqualitäts-Attribute, z.B. Golden Batch Scores T mit

$$Y = T = X \cdot P \quad (8.2)$$

- $Y_{pre}$  := ( $v \times 1$ ) Vektor der geschätzten Prozessqualitäts-Attribute, berechnet über ein PCA/PLS-Modell mit

$$Y_{prel} = f_i(X_{MV}) \quad (8.3)$$

- $v$  := Anzahl der verwendeten Attribute

- $\theta_{DModX}$  := Wichtungsfaktor des Modellfehlers DModX
- $DModX$  := Distance to Model (Modellfehler) mit

$$DModX = \sqrt{\frac{1}{m-r} \sum_{j=1}^m e_j^2}; DModX \leq a_{DModX} \cdot DModX_{crit} \quad (8.4)$$

- $DModX_{crit}$  := kritischer Grenzwert von DModX

- $a_{DModX}$  := Einstellschranke für  $DModX_{crit}$
- $e_j$  := Schätzfehler der Prozessvariablen  $x_j$  im Batch Level Model (BLM)
- $m$  := Anzahl der verwendeten Prozessvariablen  $x_j$  im BLM
- $r$  := Anzahl der Scores  $t_i$  im BLM

- $\theta_Y$  := Wichtungsfaktor für Hotellings  $T^2$  im BLM
- $T^2$  := statistische Abweichung im BLM mit

$$T^2 = \sum_{i=1}^r \frac{(t_i - \bar{t}_i)^2}{\sigma_{ti}^2}; T^2 \leq a_T \cdot T_{crit,95\%}^2 \quad (8.5)$$

- $t_i$  := Score Vektor von *principal component* l im BLM
- $\bar{t}_i$  := arithmetischer Mittelwert von  $t_i$
- $\sigma_{ti}^2$  := Varianz von  $t_i$
- $r$  := Anzahl der Scores im BLM
- $T_{crit,95\%}^2$  := 95 %-Vertrauensbereich für  $T^2$
- $a_T$  := Einstellschranke für die  $T^2$ -Grenze,

vorgegeben, das mit den quadratischen Abweichungen über  $X_{MV}$ ,  $Y_{SP}$ ,  $DModX$  und  $T^2$  aus vier Einflussgrößen aufgebaut ist.

Aus Bild 8.7 wird deutlich, dass hier die Berechnung der  $Y_{pre}$ , DModX und  $T^2$  aus einem BLM heraus nur erfolgreich sein kann, wenn eine vollständige Datenmatrix X vorliegt. Dies ist aber nur dann der Fall, wenn der Batch bereits bis zu seinem Ende hin abgelaufen ist, oder die Matrix X mit den Daten einer multivariaten Vorhersage des Batchverlaufes gefüllt ist [Muñoz et al., 2004].

Der BOBYQA-Algorithmus wird dazu in vorgebbaren Zeitintervallen gestartet. Er stellt dabei die augenblicklichen  $X_{MV}$  auf neue, bis zum Ende des Prozesses konstante, Werte ein und schätzt aus diesen und den bereits vorhandenen Prozessdaten  $X_K$  mit Hilfe von IBR-Methoden die fehlenden Prozessdaten  $X_D$  bis zum vorgegebenen Ende ab.

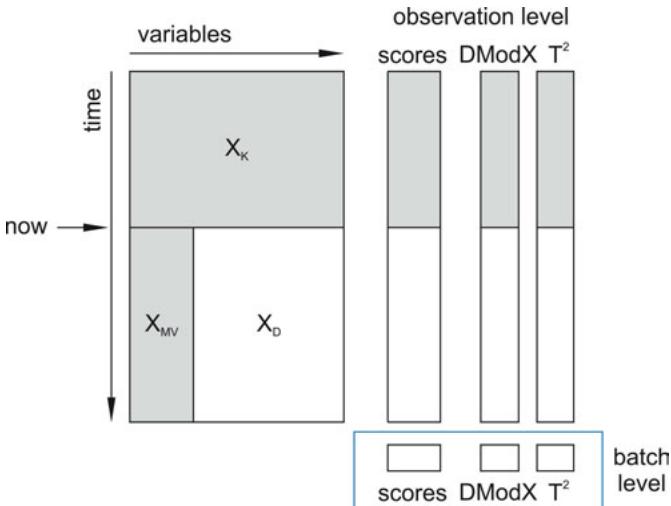


Bild 8.7: Auffüllen einer unvollständigen Prozessdatenmatrix X mit Hilfe des IBR-Algorithmus

Mit einer vollständigen Datenmatrix können dann die Bewertungsparameter scores, DModX und  $T^2$  für das BLM, unter den durch die  $X_{MV}$  vorgegebenen Bedingungen, leicht abgeschätzt und so wieder in den BOBYQA-Algorithmus mit einbezogen werden.

Durch Variation der gewählten q Manipulation-Variables  $X_{MV}$  wird für Gl. (8.1) ein Minimum von J gesucht. Der Algorithmus endet nach 1000 Iterationen. Die Software übergibt die neuen Sollwerte für  $pH_{2w}$  ( $X_{MV1}$ ) und  $\vartheta_{L2w}$  ( $X_{MV2}$ ) dann an das Prozessleitsystem MFCS/win, wo diese an die Prozessregler übergeben werden. Mit den veränderten Sollwerten der CPP soll der Prozess so beeinflusst werden, dass er wieder in den *Golden Batch*-Tunnel zurückkehrt.

Die berechneten Prozesseinflussgrößen  $X_{MV}$  werden dabei bis zum nächsten Schätzzeitpunkt nicht verändert. Dann wird der Algorithmus erneut gestartet.

Wurde das Modell explizit aus Prozessen erstellt, die eine hohe Produktqualität hervorgebracht haben, kann dann davon ausgegangen werden, dass ein Prozess, der innerhalb der multivariaten Grenzen des Modells verläuft, am Ende ebenfalls eine hohe Produktqualität produziert hat.

Im weiteren Verlauf dieser Arbeit wurde darauf aufbauend eine Studie zur Funktionalität der MPMC durchgeführt.

Zur ersten Erprobung wurde in Bild 8.8 ein drittes Experiment XX2514\_9 (mpmc) in der Validierungskampagne durchgeführt. Es begann, wie in Bild 8.8a in schwarz gezeigt, ebenfalls mit den falschen Sollwerten fsp und folgt in Bild 8.8c dem blauen (fsp) Zwillingsexperiment XX2514\_8 zunächst exakt.

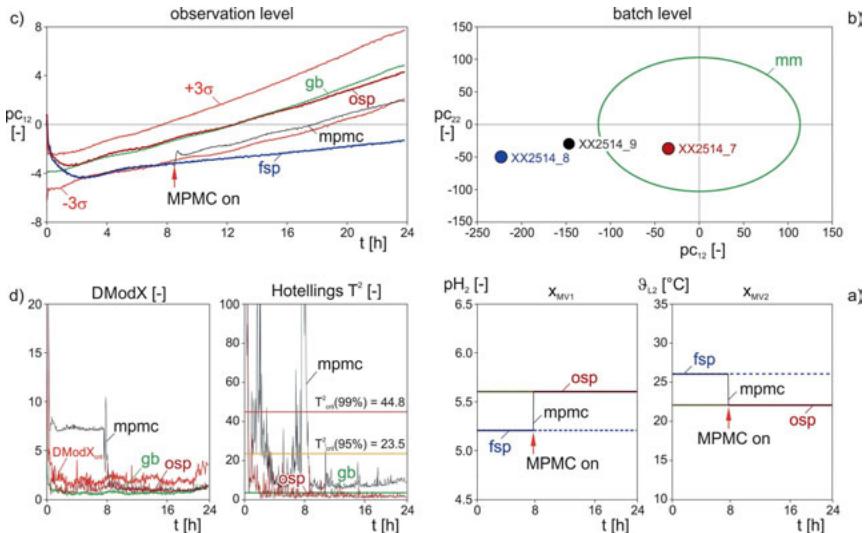


Bild 8.8: Erste Erprobung des *model predictive multivariate control* (MPMC)

Bei  $t = 8$  h wird der MPMC-Algorithmus gestartet, in dem das *monitoring model* (MM) auch als *control model* (CM) eingesetzt wurde. Dies bringt die *Manipulation Variables*  $X_{MV1}$  und  $X_{MV2}$  in Bild 8.8a in der ersten Iteration bereits auf die richtigen Werte osp zurück.

Der Score mpmc (schwarz) kehrt in Bild 8.8c sofort wieder in den *Golden Batch*-Tunnel zurück und die beiden Fehler, DModX und Hotellings  $T^2$ , fallen in Bild 8.8d wieder in den erlaubten Bereich.

Der Prozess XX2514\_9 liegt allerdings im Batch Level Modell, das in Bild 8.8b den gesamten Prozessverlauf einbezieht, dennoch außerhalb der  $T^2$ -Ellipse und würde somit als nicht reproduziert mit Hinblick auf den *Golden Batch* bewertet werden.

In einem weiteren Test erfolgte eine Erprobung des BOBYQA-Algorithmus mit einem Modell, dessen zugrundeliegenden Daten mit Experimenten erzielt wurden, die nicht nur auf Parametern des Design Space beruhten.

Die theoretische Vorgehensweise ist in Bild 8.9 gezeigt. Dabei wurde dem *monitoring model* MM ein *control model* CM, mit einer größeren Variationsbreite der Trainingsdaten, zur Seite gestellt.

Hierzu wurden Experimente mit optimalen und davon abweichenden DoE-Parametern für  $X_{MV1}$  und  $X_{MV2}$  verwendet.

Der Score Scatter Plot für das entwickelte CM ist mit der schwarzen Ellipse cm in Bild 8.9b gezeigt. Sie ist schmäler als die grüne Ellipse mm des Modells MM. Die *Control Batch*-Trajektorie (cb) in Bild 8.9c weicht ebenfalls leicht von der *Golden Batch*-Trajektorie (gb) ab, liegt aber in deren Variationsbereich  $\pm 3\sigma$ .

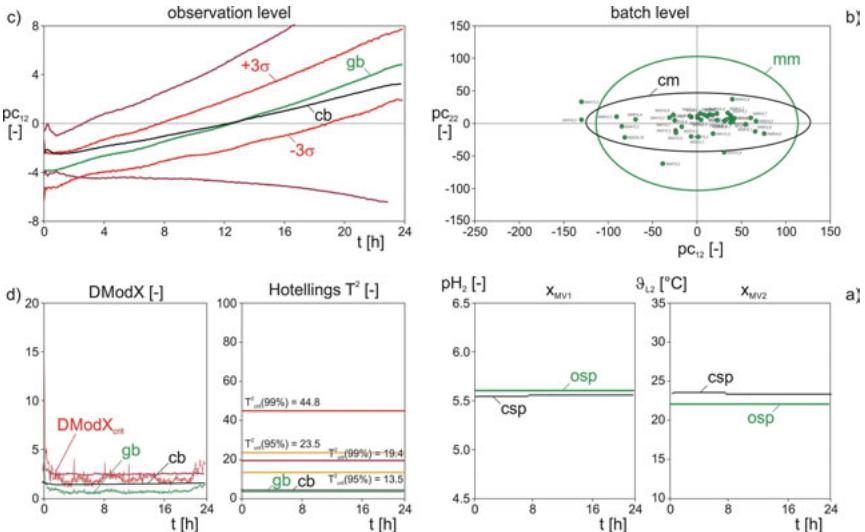


Bild 8.9: Erstellung des *Golden Batches* eines *Control Models* CM

Die Fehlergrenzen  $DModX_{crit}$  und  $T^2_{crit,95\%}$  des CM liegen in Bild 8.9d unterhalb der zulässigen Bereiche des MM, während die mittleren Verläufe der beiden Sollwerte (csp) in Bild 8.9a leicht von den optimalen Werten (osp) abweichen.

In Bild 8.10 ist das finale Ergebnis der experimentellen MPMC-Erprobung gezeigt.

Als Vergleich diente der bereits gezeigte Prozess XX2514\_8 mit den falschen Sollwerten (fsp), der in Bild 8.10b außerhalb (grün) der mm-Ellipse aber innerhalb (schwarz) der cm-Ellipse liegt.

Das neue Experiment XX3114\_4, beginnt mit den gleichen falschen Sollwerten (fsp), wird aber bei  $t = 1,5$  h vor dem Verlassen der  $\pm 3\sigma$ -Grenzen des *monitoring models* MM in Bild 8.10c mit der MPMC-Methode beaufschlagt. Diese berechnet in Zeitintervallen von 2 h jeweils die neuen Sollwerte  $pH_{2w}$  und  $9_{L2w}$  mit dem BOBYQA-Algorithmus und dem *control model* CM.

Die Prädiktionen sind Bild 8.10a zu entnehmen. Der Algorithmus führt die Variablen  $X_{MV1}$  und  $X_{MV2}$  stufenweise mit positiven und negativen Korrekturen (mpmc) auf die mittleren Sollwerte (csp) des CM.

Die zugehörige *principal component* mpmc (magenta) verläuft daher in Bild 8.10c innerhalb des *Golden Batch*-Tunnels  $\pm 3\sigma$ .

Daher liegen die Scores von XX3114\_4 im Scatter Plot in Bild 8.10b sowohl in der *control model*-Ellipse cm (schwarz) als auch in der *monitoring model*-Ellipse mm (grün). Letzteres weist darauf hin, dass der Prozess durch die MPMC-Aktionen als reproduziert anzusehen ist.

Da in Bild 8.10d die Fehler DModX und  $T^2$ , beide in Magenta, nach Einschalten des MPMC auch unter die zulässigen Fehlergrenzen fallen, ist die gewählte Prozessteuerung funktionsfähig.

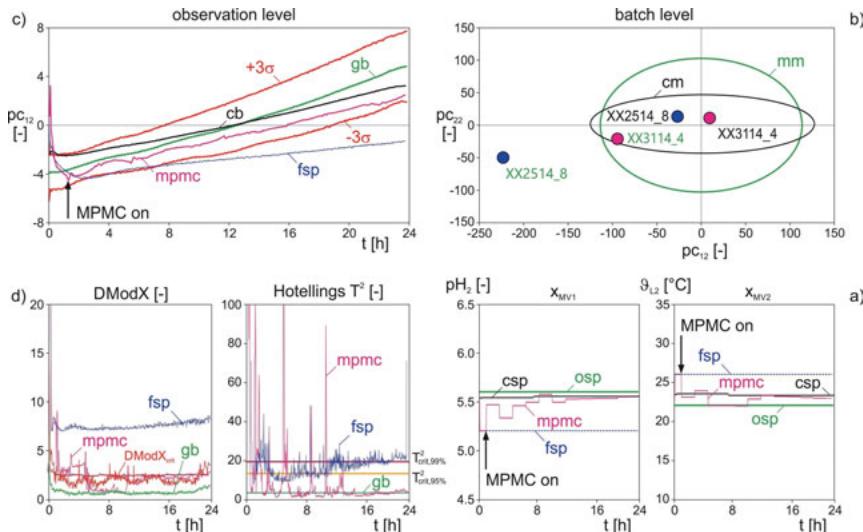


Bild 8.10: Erprobung des *control models* CM zur Steuerung des Scores  $t_1$  in den *Golden Batch*

Die gezeigten Ergebnisse sind allerdings trügerisch, denn in den Experimenten wurden lediglich falsch eingestellte Sollwerte durch den Algorithmus korrigiert.

Die Variationsbreite eines biotechnologisch pharmazeutischen Prozesses erfordert daher tiefere Erkenntnisse als nur die Abhängigkeit der beteiligten Reaktionen von pH und Temperatur, um eine verlässliche Rückführung von aus dem Ruder gelaufenen *Golden Batch*-Trajektorien bei exakt eingestellten Prozessparametern zu finden.

Hierzu bietet sich wiederum ein *Design of Experiments* an, um funktionelle Zusammenhänge in der Nähe von Reaktionsoptima zu finden.

Diese sollten dann in den BOBYQA-Algorithmus mit einfließen, um eine Varianz für die Prozessteuerung zu erlangen.

Im letzten Kapitel dieser Arbeit wird mit der Ermittlung des Verlaufs von Zellmasse, Glycerol und sekretiertem Zielprodukt, allein aus klassischen Online-Variablen, der Weg zur Online-Berechnung der zugehörigen zellspezifischen Reaktionsraten eröffnet. Diese Methoden bilden die Basis, um auch die Reproduktion des zellinternen Prozesses zu beobachten.

## 9. Bestimmung zellspezifischer Reaktionsraten über MVDA

### 9.1 Beteiligte Reaktionskomponenten

In Bild 9.1 ist das Reaktionsschema des zweistufigen Reaktionsprozesses aufgezeigt. Beide Kultivierungen mit der Zelldichte ( $c_{XL}$ ) erfolgen aerob, gesteuert über die Gelöstsauerstoffkonzentration ( $c_{OM}$ ) in der Medienphase.

Die Zellen werden auf dem Substrat Glycerol ( $c_{S1M}$ ) im *cell breeding* Reaktor angezüchtet und mit Methanol ( $c_{S2M}$ ) vorinduziert. Die eigentliche Produktbildung ( $c_{P1M}$ ) erfolgt im *protein production* Reaktor, weiterhin auf Methanol. Von großem Interesse sind dabei die beteiligten zellspezifischen Reaktionsraten  $q_{I/x}$ ,

$$q_{I/x}(t) = \pm \frac{\dot{m}_{ILrea}(t)}{m_{XL}(t)}, \quad (9.1)$$

die sich aus dem Quotienten der Massenreaktionsrate  $\dot{m}_{ILrea}$  einer betrachteten Komponente I und der Biotrockenmasse  $m_{XL}$  im Reaktionsraum ergibt.

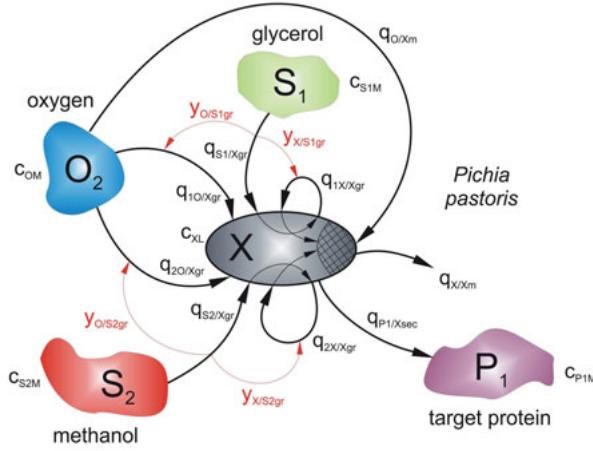


Bild 9.1: Reaktionsschema eines aeroben Herstellungsprozesses mit *Pichia pastoris* [Martens, 2014]

Der in Bild 9.1 gezeigte Reaktionsansatz ist bei Martens [Martens, 2014] näher beschrieben. Er beinhaltet die zellspezifische Zellreaktionsrate  $q_{xx}$ ,

$$q_{x/x}(t) = q_{1x/Xgr}(t) + q_{2x/Xgr}(t) - q_{x/Xm} \quad (9.2)$$

mit

$q_{1x/Xgr}$  := zellspezifische Zuwachsrate auf Glycerol

[ $h^{-1}$ ]

$q_{2x/Xgr}$  := zellspezifische Zuwachsrate auf Methanol

[ $h^{-1}$ ]

$q_{x/Xm}$  := zellspezifische Zellverlustrate durch Maintenance

[ $h^{-1}$ ],

die zellspezifische Glycerolaufnahmerate  $q_{S1/X}$ ,

$$q_{S1/X}(t) = q_{S1/Xgr}(t) = \frac{q_{1X/Xgr}(t)}{y_{X/S1gr}} \quad (9.3)$$

mit

$q_{S1/Xgr} :=$  zellspezifische Substrataufnahmerate bei Zuwachs auf Glycerol  $[h^{-1}]$   
 $y_{X/S1gr} :=$  Zellzuwachsausbeutekoeffizient bei Aufnahme von Glycerol  $[ - ]$ ,

die zellspezifische Methanolaufnahmerate  $q_{S2/X}$ ,

$$q_{S2/X}(t) = q_{S2/Xgr}(t) = \frac{q_{2X/Xgr}(t)}{y_{X/S2gr}} \quad (9.4)$$

mit

$q_{S2/Xgr} :=$  zellspezifische Substrataufnahmerate bei Zuwachs auf Methanol  $[h^{-1}]$   
 $y_{X/S2gr} :=$  Zellzuwachsausbeutekoeffizient bei Aufnahme von Methanol  $[ - ]$ ,

die zellspezifische  $O_2$ -Aufnahmerate  $q_{O/X}$ ,

$$q_{O/X}(t) = q_{1O/Xgr}(t) + q_{2O/Xgr}(t) + q_{O/Xm} \\ = y_{O/S1gr} \cdot q_{S1/Xgr}(t) + y_{O/S2gr} \cdot q_{S2/Xgr}(t) + q_{O/Xm} \quad (9.5)$$

mit

$q_{1O/Xgr} :=$  zellspezifische Sauerstoffaufnahme bei Zuwachs auf Glycerol  $[h^{-1}]$   
 $q_{2O/Xgr} :=$  zellspezifische Sauerstoffaufnahme bei Zuwachs auf Methanol  $[h^{-1}]$   
 $q_{O/Xm} :=$  zellspezifische Sauerstoffaufnahme durch Maintenance  $[h^{-1}]$   
 $y_{O/S1gr} :=$  stöchiometrischer  $O_2/S_1$ -Koeffizient bei Zuwachs auf Glycerol  $[ - ]$   
 $y_{O/S2gr} :=$  stöchiometrischer  $O_2/S_2$ -Koeffizient bei Zuwachs auf Methanol  $[ - ]$ ,

und die beobachtbare zellspezifische Produktbildungsrat  $q_{P1/X}$ ,

$$q_{P1/X}(t) = q_{P1/Xsec}(t) \quad (9.6)$$

mit

$q_{P1/Xsec} :=$  zellspezifische Produktsekretionsrate in das Medium  $[h^{-1}]$ .

Im Weiteren sollen die beobachtbaren Reaktionsraten  $q_{X/X}$ ,  $q_{S1/X}$ ,  $q_{S2/X}$ ,  $q_{O/X}$  und  $q_{P1/X}$  ohne direkte Messung der beteiligten Medienkomponenten  $c_{XL}$ ,  $C_{S1M}$ ,  $C_{S2M}$ ,  $C_{OM}$  und  $C_{P1M}$  allein aus den zur Verfügung stehenden klassischen Online-Daten über eine multivariate Datenverarbeitung online bestimmt werden.

Hierzu sind zunächst entsprechende Auswertegleichungen herzuleiten.

## 9.2 Herleitung von Massen- und Konzentrationsbilanzen

Die beiden Teilprozesse *cell breeding* und *protein production* stellen gekoppelte Batch/Fed Batch Operationen dar, in denen eine Reihe von Pumpen und Waagen zur Steuerung und Bilanzierung der Reaktionsabläufe eingesetzt werden.

Die in Bild 9.2 gezeigten Reaktoren beinhalten die Zufütter- und Ernteoperationen, die anteilig an beiden Prozessen *cell breeding*,  $p = 1$ , und *protein production*,  $p = 2$ , erfolgen.

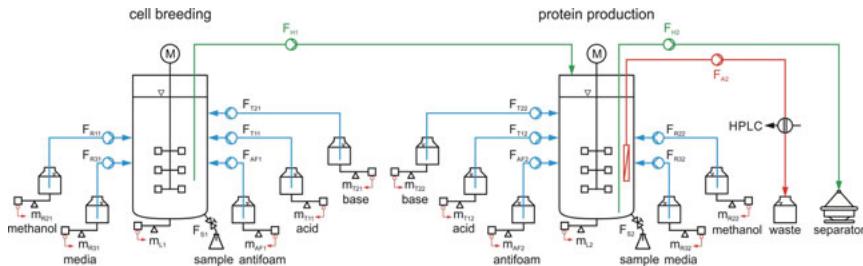


Bild 9.2: Zufütter- und Ernteoperationen in einem zweistufigen, zyklischen Bioprozess  
 $F_{Rip}$  := Zufütterrate i in Reaktor p,  $F_{Tjp}$  := pH-Titrationsrate j,  $F_{Vp}$  := Verdampfungsrate  
 $F_{Sp}$  := Probenahme,  $F_{AFp}$  := Antischaumzugabe,  $F_{Hp}$  := Ernterate,  $F_{Ap}$  := Analyserate

Die Berechnung der unbekannten zellspezifischen Reaktionsraten  $q_{i/x}$  soll mit Hilfe von Massenbilanzen erfolgen. In dieser Arbeit wird der Dreiphasenansatz [Cornelissen, 2004; Peuker, 2006; Martens, 2014] verwendet.

Dieser ist in Bild 9.3 angedeutet. Dabei durchströmt die Gasphase  $G_p$  die Reaktionsphase  $L_p$  und tauscht über die  $O_2$ -Transferrate  $CTR_p$  und die  $CO_2$ -Transferrate  $CTR_p$  die Komponenten  $O_2$  und  $CO_2$  mit der Flüssigphase  $L_p$  aus.

Diese ist in die zellfreie Medienphase  $M_p$  und die Biofeuchtpause  $Z_p$  mit

$$V_{Lp}(t) = V_{Mp}(t) + V_{Zp}(t), \quad (9.7)$$

unterteilt. Die Biofeuchtpause nimmt bei Hochzelldichtekultivierungen einen nicht zu vernachlässigenden Anteil der Reaktionsphase ein.

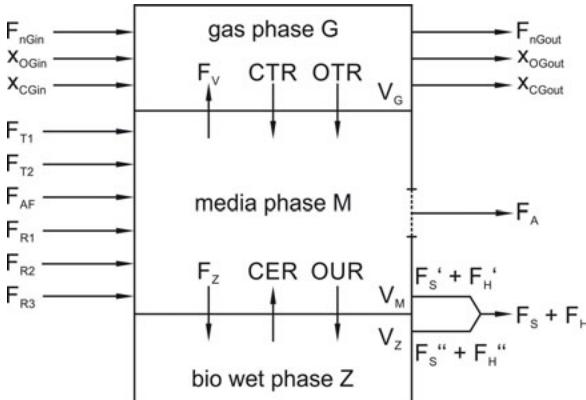


Bild 9.3: Makroskopisches Volumenelement eines Dreiphasenmodells

$T_1$  := Titration 1 (acid),  $T_2$  := Titration 2 (base), AF := Anti Foam,  
 $R_1$  := Feed 1 (glycerol),  $R_2$  := Feed 2 (methanol),  $R_3$  := Feed 3 (fresh media),  
S := Sampling, H := Harvest, A := cell-free Analysis

Das Kernstück der weiteren Untersuchung bildet die Biotrockenmasse  $m_{x_{Lp}}$ ,

$$m_{x_{Lp}}(t) = V_{Lp}(t) \cdot c_{x_{Lp}}(t) \quad (9.8)$$

mit

$V_{Lp}$  := Reaktionsraum (Flüssigphase) in Reaktor p

$c_{x_{Lp}}$  := (Trocken-) Zelldichte, bezogen auf  $V_{Lp}$

[l]

[g l<sup>-1</sup>]

die in der Biofeuchtmasse  $m_{z_{Lp}}$ ,

$$m_{z_{Lp}}(t) = \alpha_{z/x} \cdot m_{x_{Lp}}(t) \quad (9.9)$$

mit

$\alpha_{z/x}$  := Biofeucht/Biotrocken-Verhältnis ( $\alpha_{z/x} = 3,54$ ),

enthalten ist.

Die Zellmassenbilanz in der Flüssigphase  $L_p$ ,

$$\dot{m}_{x_{Lp}}(t) = \left[ q_{x/xp}(t) - \frac{F_{Sp}(t) + F_{Hp}(t)}{V_{Lp}(t)} \right] \cdot m_{x_{Lp}}(t) \quad (9.10)$$

mit

$q_{x/xp}$  := zellspezifische Wachstumsrate in  $V_{Lp}$

[h<sup>-1</sup>]

$F_{Sp}$  := Probenahmerate an Reaktor p

[l h<sup>-1</sup>]

$F_{Hp}$  := Ernterate an Reaktor p

[l h<sup>-1</sup>]

beinhaltet die Zellreaktionsrate  $q_{x/xp}$ , die bei Zuwachs positiv, bei Substratmangel aber negativ wird, und den Massenverlust über die Probenahme  $F_{Sp}$  und die Ernte  $F_{Hp}$ .

Will man eine Zelldichtebilanz nutzen, wird über die Auflösung der Zellmassenänderung  $\dot{m}_{x_{Lp}}$ ,

$$\dot{m}_{x_{Lp}}(t) = \dot{V}_{Lp}(t) \cdot c_{x_{Lp}}(t) + V_{Lp}(t) \cdot \dot{c}_{x_{Lp}}(t), \quad (9.11)$$

die Beschreibung weitaus komplexer, da nunmehr eine Volumenbilanz für  $V_{Lp}$ ,

$$\dot{V}_{Lp}(t) = F_{inp}(t) - F_{outp}(t), \quad (9.12)$$

benötigt wird.

Der Eingangsvolumenstrom  $F_{inp}$ ,

$$F_{inp}(t) = \sum_{i=1}^3 F_{Rip}(t) + \sum_{j=1}^2 F_{Tjp}(t) + F_{AFp}(t) - F_{vp}(t) \quad (9.13)$$

mit

$F_{Rip}$  := Zufütterrate aus Reservoir R<sub>i</sub> an Reaktor p

[l h<sup>-1</sup>]

$F_{Tjp}$  := pH-Titrationsrate aus Tank T<sub>j</sub> an Reaktor p

[l h<sup>-1</sup>]

$F_{AFp}$  := Antischaum-Zugabe aus Vorlage AF an Reaktor p

[l h<sup>-1</sup>]

$F_{vp}$  := Verdampfungsrate durch Begasung von Reaktor p

[l h<sup>-1</sup>],

beinhaltet auch den Wasserdampfaustrag  $F_{vp}$ , da diese Rate nur das Volumen  $V_{Lp}$  ändert, aber keine Komponenten austrägt.

Der Ausgangsvolumenstrom  $F_{\text{outp}}(t)$ ,

$$F_{\text{outp}}(t) = F_{\text{Sp}}(t) + F_{\text{Hp}}(t) + F_{\text{Ap}}(t) \quad (9.14)$$

mit

$F_{\text{Ap}}$  := zellfreier Analysestrom aus Reaktor p  $[\text{lh}^{-1}]$ ,

trägt die Zellmasse und Medienkomponenten aus dem Reaktor p aus.

Aus diesen Ansätzen ergibt sich die Bilanz für die Zelldichte  $c_{XLp}(t)$ ,

$$\dot{c}_{XLp}(t) = \left[ q_{X/Xp}(t) - \frac{F_{\text{inp}}(t) - F_{\text{Ap}}(t)}{V_{Lp}(t)} \right] \cdot c_{XLp}(t), \quad (9.15)$$

in der, im Gegensatz zu Gl. (9.10), nunmehr die Zufütterungen, der Verdampfungsverlust  $F_{Vp}$  und der Analysenstrom  $F_{Ap}$  zu berücksichtigen sind. Letztere konzentrieren beide die Zell-dichte  $c_{XLp}$  auf, was besonders in kleinen Reaktoren zum Tragen kommt.

Die Medienkomponenten  $c_{IMp}$  werden in der Medienphase  $M_p$  gemessen. Das zugehörige Volumen  $V_{Mp}$ ,

$$V_{Mp}(t) = V_{Lp}(t) - V_{Zp}(t), \quad (9.16)$$

ist gegenüber  $V_{Lp}$  um die Biofeuchtphase verringert.

Das Biofeuchtvolumen  $V_{Zp}$ ,

$$V_{Zp}(t) = \frac{\alpha_{Z/X} \cdot c_{XLp}(t)}{\rho_Z} \cdot V_{Lp}(t) \quad (9.17)$$

mit

$\rho_Z$  := mittlere Dichte der feuchten Zellen ( $\rho_Z = 1.149 \cdot 10^3 \text{ gl}^{-1}$ ),

wird über die Zelldichte  $c_{XLp}$  beeinflusst und führt auf den Bilanzraum  $V_{Mp}$ ,

$$V_{Mp}(t) = \left( 1 - \frac{\alpha_{Z/X} \cdot c_{XLp}(t)}{\rho_Z} \right) \cdot V_{Lp}(t), \quad (9.18)$$

der die gemessenen Konzentrationen  $c_{IMp}$  enthält.

Die Bilanzierung dieser Komponenten  $I \neq X$  kann im Reaktionsraum  $L_p$  oder im Medienraum  $M_p$  erfolgen. Hierbei können gegebenenfalls gemessene Werte  $c_{IMp}$  einfach in  $c_{ILp}$ ,

$$c_{ILp}(t) = \frac{V_{Mp}(t) \cdot c_{IMp}(t)}{V_{Lp}(t)} = \left( 1 - \frac{\alpha_{Z/X} \cdot c_{XLp}(t)}{\rho_Z} \right) \cdot c_{IMp}(t), \quad (9.19)$$

überführt werden.

Zur Auswertung kann nun die Konzentrationsbilanz in der Flüssigphase  $L_p$ ,

$$\begin{aligned} c_{ILp}(t) &= \frac{\sum F_{Jp}(t) \cdot c_{Ijp} - F_{\text{inp}}(t) \cdot c_{ILp}(t)}{V_{Lp}(t)} \pm q_{I/Xp}(t) \cdot c_{XLp}(t) + ITR_p(t) \\ &\quad - \frac{\alpha_{Z/X} \cdot c_{XLp}(t)}{\rho_Z - \alpha_{Z/X} \cdot c_{XLp}(t)} \cdot \frac{F_{Ap}(t) \cdot c_{ILp}(t)}{V_{Lp}(t)}, \quad I \neq X, \end{aligned} \quad (9.20)$$

mit

$c_{ILp}$  := Konzentration der Komponente I in Flüssigphase  $L_p$

[ $gl^{-1}$ ]

$F_{Jp}$  := Zufütterrate aus Reservoir  $J_p$  an Reaktor  $p$

[ $lh^{-1}$ ]

$c_{Up}$  := Konzentration der Komponente I in Reservoir  $J_p$

[ $gl^{-1}$ ]

$ITR_p$  := volumetrische Transferrate der Komponente I aus der Gasphase  $G_p$  in die Flüssigphase  $L_p$  in Reaktor  $p$

[ $gl^{-1}h^{-1}$ ],

oder die Konzentrationsbilanz in der Medienphase  $M_p$ ,

$$\dot{c}_{IMp}(t) = \left[ \frac{\sum F_{Jp}(t) \cdot c_{IJp} - F_{inp}(t) \cdot c_{IMp}(t)}{V_{Lp}(t)} \pm q_{I/Xp}(t) \cdot c_{Xlp}(t) + ITR_p(t) \right. \\ \left. + \frac{\alpha_{Z/X} \cdot c_{XLp}(t)}{\rho_z} \cdot q_{X/Xp}(t) \cdot c_{IMp}(t) \right] \cdot \frac{\rho_z}{\rho_z - \alpha_{Z/X} \cdot c_{XLp}(t)}, \quad I \neq X,$$
(9.21)

neben Gl. (9.15) zur Bestimmung der zellspezifischen Reaktionsraten  $q_{I/Xp}$  in beiden Reaktoren herangezogen werden.

### 9.3 Herleitung der $q_{I/X}$ -Auswertungen des zyklischen Prozesses

Im Weiteren wurden Auswertegleichungen der zellspezifischen Reaktionsraten  $q_{I/Xp}$  für einen vollständigen Produktionszyklus hergeleitet. Dieser ist in Bild 9.4 gezeigt.

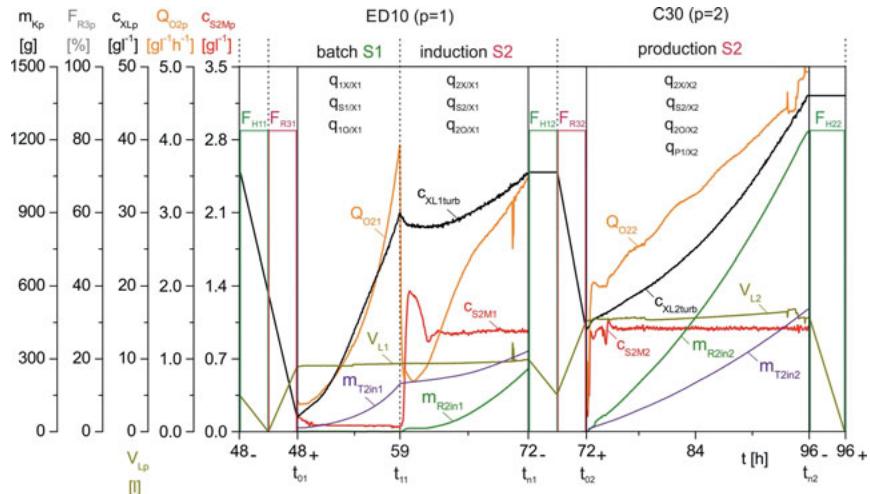


Bild 9.4: Vollständiger *cell breeding* und *protein production* Zyklus

Der *cell breeding* Prozess startet mit der Justierung der Anfangszelldichte  $c_{XL10}$  und des Volumens  $V_{L10}$  durch eine Restentleerung über  $F_{H11}$  und einem Refresh über  $F_{R31}$  mit der Vorgabe von  $c_{S1L10}$ . Danach erfolgt zwischen  $t_{01} = 48+$  h und  $t_{11} = 59$  h eine Batchphase auf Glycerol zur schnellen Zellanzucht.

Hier werden online die Zelldichte  $c_{XL1turb}$  über das Trübungssignal (Gl. (3.1)) und die O<sub>2</sub>-Eintragsrate  $Q_{O21}$  über die Abgasbilanzierung (Gl. (3.9)) beobachtet sowie die Zugabe an Lauge  $m_{T2in1}$  und das Volumen  $V_{L1}$  gemessen. In dieser Phase liegt weder eine Probenahme noch eine Ernte oder ein Analysenstrom vor.

Bei idealer Abgasrückkühlung ( $F_{Vp} = 0 \text{ l h}^{-1}$ ), keiner Atline-Analyse an Reaktor 1 ( $F_{A1} = 0 \text{ l h}^{-1}$ ) und unter Berücksichtigung der Laugezugabe durch den pH-Regler ( $F_{T21}$ ) ist im Zeitraum  $t \in [t_{01}, t_{11}]$  über Gl. (9.15) somit die erste Bestimmungsgleichung für die zellspezifische Wachstumsrate  $q_{1X/X1}$  auf Glycerol,

$$q_{1X/X1}(t) = \frac{\dot{c}_{XL1}(t)}{c_{XL1}(t)} + \frac{F_{in1}(t)}{V_{L1}(t)} = \frac{\dot{c}_{XL1}(t)}{c_{XL1}(t)} + \frac{F_{T21}(t)}{V_{L1}(t)}, \quad (9.22)$$

festgelegt. Zur Auswertung sind allerdings der Verlauf der Zelldichte und deren 1. Ableitung erforderlich.

Die Ermittlung der zugehörigen zellspezifischen O<sub>2</sub>-Aufnahmerate  $q_{1O/X1}$  auf Glycerol,

$$q_{1O/X1}(t) = \frac{Q_{O21}(t)}{c_{XL1}(t)}, \quad (9.23)$$

gestaltet sich noch einfacher, da  $Q_{O21}$  online vorliegt und alle konvektiven O<sub>2</sub>-Terme vernachlässigbar sind.

Für die zellspezifische Glycerolaufnahmerate  $q_{S1/X1}$ ,

$$q_{S1/X1}(t) = -\frac{\dot{c}_{S1L1}(t)}{c_{XL1}(t)} - \frac{F_{T21}(t) \cdot c_{S1L1}(t)}{V_{L1}(t) \cdot c_{XL1}(t)}, \quad (9.24)$$

ist für den pH-geregelten Batch-Betrieb die Bestimmungsgleichung leicht über Gleichung (9.20) herleitbar. Hierzu werden später die Glycerolkonzentration  $c_{S1L1}$  und deren 1. Ableitung ebenfalls abgeschätzt.

Die durch den pH-Regler pulsförmig gesteuerte Laugetitrationsrate  $F_{T21}$  ist nicht messbar. Sie ist allerdings über die gewogene Basezufüttermenge  $m_{T2in1}$  mittels

$$F_{T21}(t) = \frac{\dot{m}_{T2in1}(t)}{\rho_{T21}}, \quad (9.25)$$

durch eine numerische Differentiation berechenbar.

In der zweiten Phase des *cell breeding*, einem Methanol Fed Batch-Prozess, werden ab  $t = t_{11}$  die Zellen für die Produktion auf Methanol vorinduziert. In dieser Phase stehen die Messwerte für  $c_{XL1turb}$ ,  $Q_{O21}$ ,  $V_{L1}$ ,  $m_{R2in1}$ ,  $m_{T2in1}$  und  $c_{S2M1}$  zur Verfügung.

Im Gegensatz zum konstitutiven Metabolismus auf Glycerol werden die zellspezifischen Reaktionsraten hier erst durch Bildung von Alkoholoxidase aufgebaut.

Die zellspezifische Wachstumsrate  $q_{2X/X1}$  und die zellspezifische O<sub>2</sub>-Aufnahmerate  $q_{2O/X1}$  werden analog zu Gl. (9.22) bzw. Gl. (9.23) berechnet. Die Berechnung von  $q_{2X/X1}$  beinhaltet allerdings mit

$$F_{in1}(t) = F_{T21}(t) + F_{R21}(t) = \frac{\dot{m}_{T2in1}(t)}{\rho_{T21}} + \frac{\dot{m}_{R2in1}(t)}{\rho_{R21}}, \quad (9.26)$$

nunmehr eine gesteuerte Zugabe von Lauge ( $F_{T21}$ ) und Methanol ( $F_{R21}$ ).

Mit deren Hilfe und der gemessenen Methanolkonzentration  $c_{S1M1}$  kann die zellspezifische Methanolaufnahmerate  $q_{S2/X1}$ ,

$$q_{S2/X1}(t) = \frac{F_{R21}(t) \cdot c_{S2R21} - [F_{R21}(t) + F_{T21}(t)] \cdot c_{S2M1}(t)}{V_{L1}(t) \cdot c_{XL1}(t)} \\ + \frac{\rho_Z \cdot S_2 \cdot TR_1(t) + \alpha_{Z/X} \cdot c_{XL1}(t) \cdot [q_{2X/X1}(t) \cdot c_{S2M1}(t) + \dot{c}_{S2M1}(t)] - \rho_Z \cdot \dot{c}_{S2M1}(t)}{\rho_Z \cdot c_{XL1}(t)}, \quad (9.27)$$

über Gl. (9.21) leicht hergeleitet werden.

Der Methanolverlust durch den Begasungsauftrag wird über die volumetrische Austragsrate  $S_2TR_1$ ,

$$S_2TR_1(t) = -\frac{K_{S2vol} \cdot F_{nG1}(t) \cdot c_{S2M1}(t)}{V_{L1}(t)} \quad (9.28)$$

mit

$K_{S2vol}$  := Methanol-Verdampfungskonstante

[ $\text{-}$ ]

$F_{nG1}$  := Begasungsrate von Reaktor 1

[ $\text{lh}^{-1}$ ],

berücksichtigt und durch das Raoult'sche Gesetz [Adam and Stark, 2009] beschrieben.

Kurz nach  $t_{n1} = 72$  h endet der *cell breeding* Prozess ( $p = 1$ ) mit einer definierten Teilernte  $F_{H1}$ , bei der 225 g Zellmasse in den *protein production* Reaktor ( $p = 2$ ) übergeben werden.

Hier beginnt nach einem Medienrefresh mit  $F_{R32}$  die eigentliche Herstellung des Zielproteins zwischen  $t_{02} = 72$  h und  $t_{n2} = 96$  h. In diesem Reaktor werden die äquivalenten Meßgrößen wie in der Vorinduktionsphase in Reaktor 1 bereitgehalten. Hinzu kommt allerdings die At-line-Messung des Zielproteins  $c_{P1M2}$  über eine IMAC Aufreinigungsmethode mit einer HPLC über den zellfreien Analysenstrom  $F_{A2}$ .

Somit stehen für den zweiten Reaktor im Zeitraum  $t \in [t_{02}, t_{n2}]$  folgende Auswertegleichungen für die zellspezifische Wachstumsrate auf Methanol  $q_{2X/X2}$ ,

$$q_{2X/X2}(t) = \frac{\dot{c}_{XL2}(t)}{c_{XL2}(t)} + \frac{F_{R22}(t) + F_{T22}(t) - F_{A2}(t)}{V_{L2}(t)}, \quad (9.29)$$

die zellspezifischen  $O_2$ -Aufnahmerate  $q_{2O/X2}$  auf Methanol,

$$q_{2O/X2}(t) = \frac{Q_{O22}(t)}{c_{XL2}(t)}, \quad (9.30)$$

die zellspezifische Methanolaufnahmerate  $q_{S2/X2}$ ,

$$q_{S2/X2}(t) = \frac{F_{R22}(t) \cdot c_{S2R22} - [F_{R22}(t) + F_{T22}(t)] \cdot c_{S2M2}(t)}{V_{L2}(t) \cdot c_{XL2}(t)} \\ + \frac{\rho_Z \cdot S_2 \cdot TR_2(t) + \alpha_{Z/X} \cdot c_{XL2}(t) \cdot [q_{2X/X2}(t) \cdot c_{S2M2}(t) + \dot{c}_{S2M2}(t)] - \rho_Z \cdot \dot{c}_{S2M2}(t)}{\rho_Z \cdot c_{XL2}(t)}, \quad (9.31)$$

und die zellspezifische Produktbildungsrates  $q_{P1/X2}$ ,

$$\begin{aligned} q_{P1/X2}(t) &= \frac{[F_{R22}(t) + F_{T22}(t)] \cdot c_{P1M2}(t)}{V_{L2}(t) \cdot c_{XL2}(t)} \\ &\quad - \frac{\alpha_{Z/X} \cdot c_{XL2}(t) \cdot [q_{2X/X2}(t) \cdot c_{P1M2}(t) + \dot{c}_{P1M2}(t)] + p_Z \cdot \dot{c}_{P1M2}(t)}{p_Z \cdot c_{XL2}(t)}, \end{aligned} \quad (9.32)$$

zur Verfügung.

Im weiteren Teil dieser Arbeit wird gezeigt, wie die zur Lösung dieser Gleichungen benötigten Zelldichten sowie Substrat- und Produktkonzentrationen mit Hilfe eines MVDA Ansatzes über klassische, leicht messbare Prozessvariablen abgeschätzt werden können.

## 9.4 Die Prädiktion der Zelldichte mit Hilfe eines PLS-Modells

### 9.4.1 Die Bereitstellung von Trainingsdaten

Die erforderlichen Reaktionskomponenten  $c_{XLp}$ ,  $c_{S1M1}$ ,  $c_{S2Mp}$  und  $c_{P1Mp}$  liegen in den beiden Reaktoren nur teilweise als Online-Daten vor.

In Bild 9.5 ist die Datenlage für die Entwicklung einer Prädiktion mit der MVDA gezeigt. Die Zelldichte  $c_{XLp}$  wird in jedem Reaktor online über das Trübungssignal  $E_{Lp}$  über Gl. (3.1) berechnet. Darüber hinaus wird sie in beiden Reaktoren jeweils zu Beginn ( $t_{0p}$ ) und weiteren Zeitpunkten ( $t_{ip}$ ) offline bestimmt.

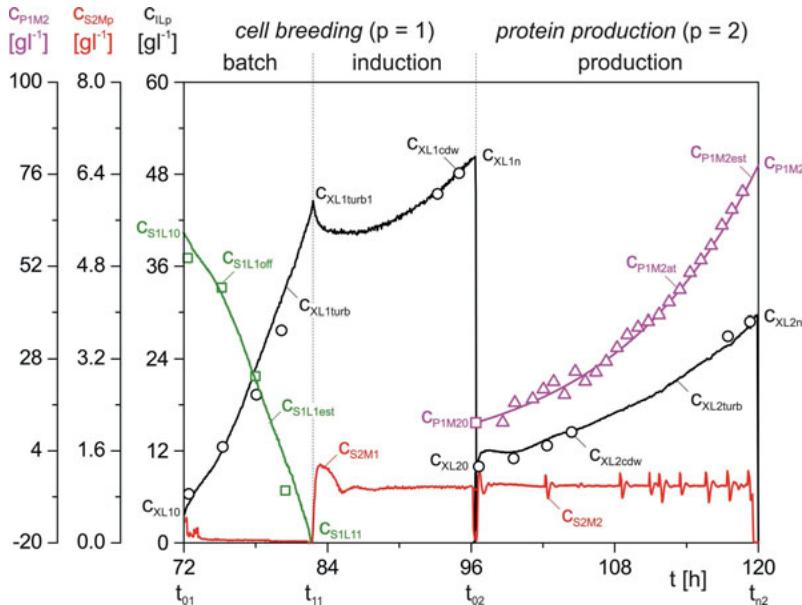


Bild 9.5: Verlauf der Reaktionskomponenten im zweistufigen Produktionsprozess

Die Methanolkonzentrationen  $c_{S2Mp}$  werden jeweils über eine Tubingsonde inline gemessen und dann geregelt. Die Glycerolkonzentration  $c_{S1L10}$  wird zu Beginn ( $t_{01}$ ) eingewogen und am Ende der Batchphase ( $t_{11}$ ), das online identifiziert wird, zu  $0 \text{ gl}^{-1}$  angenommen. Die Zielproduktkonzentration  $c_{P1M2at}$  wird in Reaktor 2 atline über das SEC/IMAC-Aufreinigungsverfahren mit einer HPLC detektiert.

Die Verläufe von  $c_{S2Mp}$  werden direkt zur späteren Berechnung von  $q_{S2/xp}$  verwendet. Die Verläufe von  $c_{XLp}$ ,  $c_{S1L1}$  und  $c_{P1Mp}$  sollen hingegen über Softsensoren mittels MVDA aus den restlichen Online-Daten ermittelt werden und als Grundlage der Berechnung von zugehörigen zellspezifischen Reaktionsraten dienen.

Die Zelldichte  $c_{XLpturb}$  ist direkt als Trainingsset  $c_{XLpest}$ ,

$$c_{XLpest}(t) = c_{XLpturb}(t), \quad (9.33)$$

einsetzbar.

In der Batchphase auf Glycerol kann ein mittlerer Glycerolausbeutekoeffizient  $\bar{y}_{X/S1}$ ,

$$\bar{y}_{X/S1} = \frac{V_{L11} \cdot c_{XL1est1} - V_{L10} \cdot c_{XL10}}{V_{L10} \cdot c_{S1L10}}, \quad (9.34)$$

über den Batchbereich  $[t_{01}, t_{11}]$  berechnet werden.

Hiermit ist der Verlauf der geschätzten Substratkonzentration  $c_{S1L1est}$  in der Flüssigphase L<sub>1</sub>,

$$c_{S1L1est}(t) = \frac{V_{L10} \cdot (\bar{y}_{X/S1} \cdot c_{S1L10} + c_{XL10}) - V_{L1}(t) \cdot c_{XL1est}(t)}{\bar{y}_{X/S1} \cdot V_{L1}(t)}, \quad (9.35)$$

leicht zu berechnen.

Die Vorgabe von Trainingsdatensätzen  $c_{P1Mpest}$  für das Zielprodukt gestaltet sich schwieriger. Wie Bild 9.5 zu entnehmen ist, liegen im Produktionsbereich  $[t_{02}, t_{n2}]$  zu bestimmten Zeitpunkten m Atline-Messwerte  $c_{P1M2atj}$  vor.

Da der Produktionsreaktor eine Anfangsmenge des Zielproduktes enthält und die Zellen bereits vorinduziert sind, ist ein exponentieller Verlauf zu beobachten. Durch diesen wird eine Ausgleichskurve zur Beschreibung von  $c_{P1M2est}$ ,

$$c_{P1M2est}(t) = c_{P1M20} \cdot e^{k_{P1exp} \cdot (t - t_{02})}, \quad (9.36)$$

gelegt, in der die beiden Parameter  $c_{P1M20}$  ( $p_1$ ) und  $k_{P1exp}$  ( $p_2$ ) unbekannt sind.

Diese werden durch eine nichtlineare Least Squares-Kurvenanpassung,

$$J_p = \min_{p_1, p_2} \sum_{j=1}^m (c_{P1M2est}(p_1, p_2, t_{j2}) - c_{P1M2at}(t_{j2}))^2, \quad (9.37)$$

mit

$c_{P1M2at}$  := Atline-Messwerte des Zielproduktes P<sub>1</sub> in Reaktor 2 [gl<sup>-1</sup>]

$c_{P1M2est}$  := geschätzter zeitlicher Verlauf von  $c_{P1M2}$  [gl<sup>-1</sup>]

$c_{P1M20}$  := identifizierter Anfangswert von  $c_{P1M2est}$  [gl<sup>-1</sup>]

$k_{P1exp}$  := identifizierte Produktbildungsgeschwindigkeit [h<sup>-1</sup>]

$t_{j2}$  := Produkt-Messzeitpunkt in Reaktor 2 [h],

mit MATLAB für jeden Zyklus identifiziert.

#### 9.4.2 Prädiktion der Zelldichte

Die Zelldichte ist der Schlüsselparameter bei der Berechnung zellspezifischer Reaktionsraten. Sie soll nunmehr ohne Trübungssignal oder andere spektroskopischen Inline-Messungen allein aus den klassischen Variablen eines Bioreaktionsprozesses bestimmt werden.

Bild 9.6 zeigt dafür den Verlauf von Prozessdaten in einem Produktionszyklus im BIORSTAT C30. Die gezeigten Daten stellen eine Untermenge der Prozessdatenmatrix  $X_2$  dar und beinhalten vornehmlich die Dynamik des Wachstumsprozesses. Die  $O_2$ - und  $CO_2$ -Abgasmolenbrüche  $x_{O2}$  und  $x_{CO2}$  korrelieren bei konstanter Begasung mit Luft direkt mit der  $O_2$ -Eintragsrate  $Q_{O22}$  und der  $CO_2$ -Austragsrate  $Q_{CO22}$ .

Diese steuern zusammen mit der Methanolzufütterrate  $F_{R22}$  die volumetrische Zellwachstumsrate. Die gebildete Zellmasse ist hingegen aus den Waagensignalen  $m_{R2in2}$  und  $m_{T2in2}$  rekonstruierbar. Der Anstieg der Drehzahl  $N_{S12}$  gewährleistet eine Vermeidung einer  $O_2$ -Transportlimitierung.

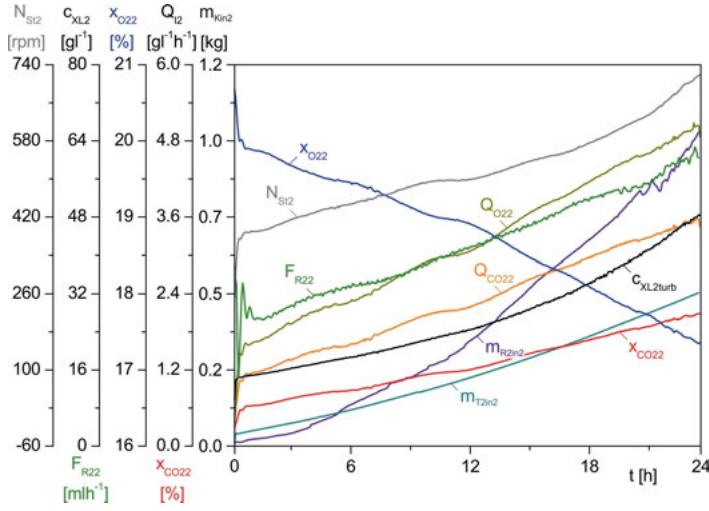


Bild 9.6: Charakteristischer Verlauf von Online-Prozessdaten einer Produktionsphase

Um den Einfluss aller Variablen, z.B. zur Bestimmung der Zelldichte  $c_{XLp}$ , zu erkennen, wird zunächst die Matrix  $X_p$  um den Spaltenvektor  $Y_i$ , der die abzuschätzende Variable  $c_{XLpest}$  enthält, ergänzt.

Diese Matrix wird einer Kreuzkorrelation,

$$C_{pi} = \frac{[X_p \ Y_i]^T \cdot [X_p \ Y_i]}{n - 1}, \quad (9.38)$$

mit

$X_p$  := (n×m) vollständige Prozessdatenmatrix von Reaktor p

$Y_i$  := (n×1) Spaltenvektor der abzuschätzenden Variable

$C_{pi}$  := ((m+1)×(m+1)) Kreuzkorrelationsmatrix,

unterzogen.

Die daraus entstehenden Korrelationskoeffizienten geben nun in der letzten Spalte von  $C_{pi}$  Aufschluss darüber, ob die zugehörige Zeilenvariable einen signifikanten Einfluss auf die Prädiktion von  $Y_i$  hat oder nicht. Prozessgrößen mit geringem Einfluss werden aus Matrix  $X_p$  entfernt, aus der dann die Untermatrix  $X_{pi}$  entsteht.

Derartige Operationen sind für die Prozessphasen *batch*, *induction* (cell breeding Matrix  $X_1$ ) und *production* (protein production Matrix  $X_2$ ) anzuwenden. In Tabelle 9.1 sind die an der Prädiktion von  $c_{XLp}$ ,  $c_{S1L1}$  und  $c_{P1M1}$  beteiligten Matrizen  $X_{pi}$  und  $Y_i$  zusammengefasst.

Tabelle 9.1: An den PLS-Operationen beteiligte Matrizen

$Y_i$	$C_{XL1}$	$C_{S1L1}$	$C_{P1M1}$	$C_{XL2}$	$C_{P1M2}$
batch	$X_{11}, Y_1$	$X_{12}, Y_2$			
induction	$X_{13}, Y_3$		$X_{14}, Y_4$		
production				$X_{25}, Y_5$	$X_{26}, Y_6$

Die jeweils an den Matrizen  $X_{pi}$  und  $Y_i$  beteiligten Variablen sind Tabelle 9.3 in Abschnitt 9.4.5 zu entnehmen.

Beispielhaft sind im Folgenden die wichtigsten Einflussgrößen auf  $c_{XL2}$  und ihre Korrelationskoeffizienten in Tabelle 9.2 aufgeführt.

Tabelle 9.2: Korrelationstabelle der verwendeten Variablen für das PLS-Modell von  $c_{XL2pre}$

	$F_{R22}$	$m_{R2in2}$	$m_{T2in2}$	$N_{St2}$	$Q_{CO22}$	$Q_{O22}$	$x_{CO22}$	$x_{O22}$	$c_{XL2turb}$
$F_{R22}$	1.00	0.80	0.82	0.93	0.90	0.90	0.91	-0.90	0.73
$m_{R2in2}$		1.00	0.97	0.88	0.88	0.88	0.89	-0.89	0.96
$m_{T2in2}$			1.00	0.91	0.93	0.93	0.94	-0.94	0.97
$N_{St2}$				1.00	0.95	0.96	0.96	-0.96	0.86
$Q_{CO22}$					1.00	1.00	1.00	-1.00	0.90
$Q_{O22}$						1.00	1.00	-1.00	0.90
$x_{CO22}$							1.00	-1.00	0.91
$x_{O22}$								1.00	-0.91
$c_{XL2turb}$									1.00

Die beiden Abgasdaten  $x_{O2}$  und  $x_{CO2}$  korrelieren perfekt mit ihren zugeordneten Bilanzwerten  $Q_{O22}$  und  $Q_{CO22}$ . Alle vier korrelieren mit 0.9 mit der Zelldichte  $c_{XL2turb}$ . Einen ähnlichen Wert zeigt die Drehzahl  $N_{St2}$  als pO<sub>2</sub>-Stellgröße auf, da sowohl die O<sub>2</sub>-Transferrate OTR<sub>2</sub> als auch der O<sub>2</sub>-Eintrag  $Q_{O22}$  der O<sub>2</sub>-Aufnahmerate OUR<sub>2</sub> bei konstantem pO<sub>2</sub> folgen und die Drehzahl den k<sub>La</sub>-Wert dafür entsprechend anheben muss.

Die höchste Korrelation mit  $c_{XL2turb}$  haben mit 0.96 bzw. 0.97 die beiden zugefütterten Mengen an Methanol ( $m_{R2in2}$ ) und Lauge ( $m_{T2in2}$ ).

Im nächsten Schritt wird das Prädiktions-Modell der Zelldichte  $c_{XL2}$  mit der in Abschnitt 7.7 erläuterten PLS-Methode,

$$Y_5 = X_{25} \cdot W_{25} \cdot (P_{25}^T \cdot W_{25})^{-1} \cdot Q_5^T + G_5 = X_{25} \cdot B_5 + G_5 \quad (9.39)$$

mit

$Y_5 := (n \times 1)$  Vektor des Attributes Zelldichte  $c_{XL2est}$

$X_{25} := (n \times m)$  Messdatenmatrix zur Prädiktion von  $c_{XL2est}$

$P_{25} := (r \times m)$  Loadingmatrix für  $X_{25}$

$Q_5 := (r \times 1)$  Loadingmatrix für  $Y_5$  ( $c_{XL2est}$ )

$W_{25} := (m \times r)$  Wichtungsmatrix für  $P_{25}$  und  $Q_5$

$B_5 := (m \times 1)$  Matrix der Regressionskoeffizienten bei der Prädiktion von  $c_{XL2est}$

$G_5 := (n \times 1)$  Residuenmatrix bei der Prädiktion von  $c_{XL2est}$ ,

erstellt.

Die Datenmatrix  $X_{25}$  enthält dabei, ohne Zelldichte, die in Tabelle 9.2 aufgezählten Prozessvariablen von 16 Zyklen aus 4 Produktionskampagnen, während die Matrix  $Y_5$  die zugehörige Zelldichte  $c_{XL2est}$  beinhaltet.

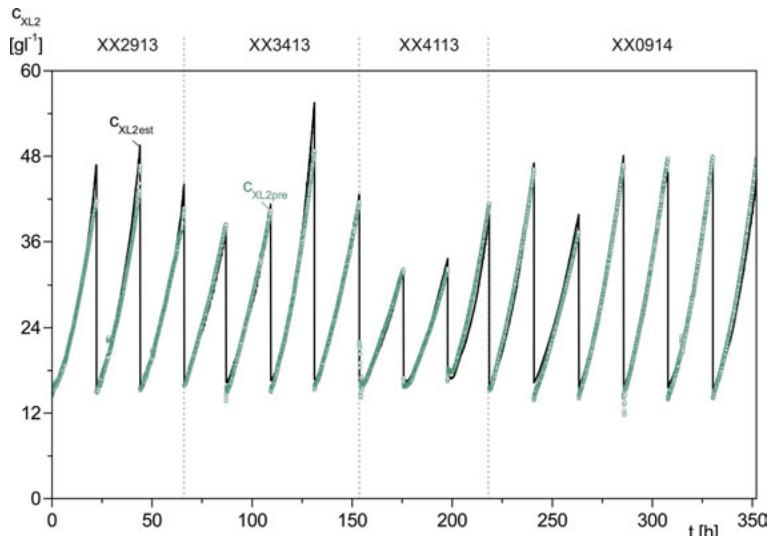


Bild 9.7: Verlauf der Zelldichte  $c_{XL2pre}$  nach Anpassung an das Online-Trainingsset  $c_{XL2est}$

Die Trainingsdaten in Bild 9.7 zeigen bei Anwendung des Modells eine gute Übereinstimmung der Prädiktion  $c_{XL2pre}$ ,

$$c_{XL2pre} = X_{25} \cdot B_5, \quad (9.40)$$

mit der Messung  $c_{XL2est}$ .

Die MVDA erfordert dann allerdings eine Validierung des Modells, die in Bild 9.8 dargestellt ist. Der obere Teil enthält die Validierungsdaten von 6 Zyklen, während sich im unteren Teil

die berechnete MVDA-Variable  $c_{XL2pre}$  (grün), die Messwerte  $c_{XL2est}$  (schwarz) und die offline ausgewogene Zelldichte  $c_{XLcdw}$  ( $\circ$ ) befinden.

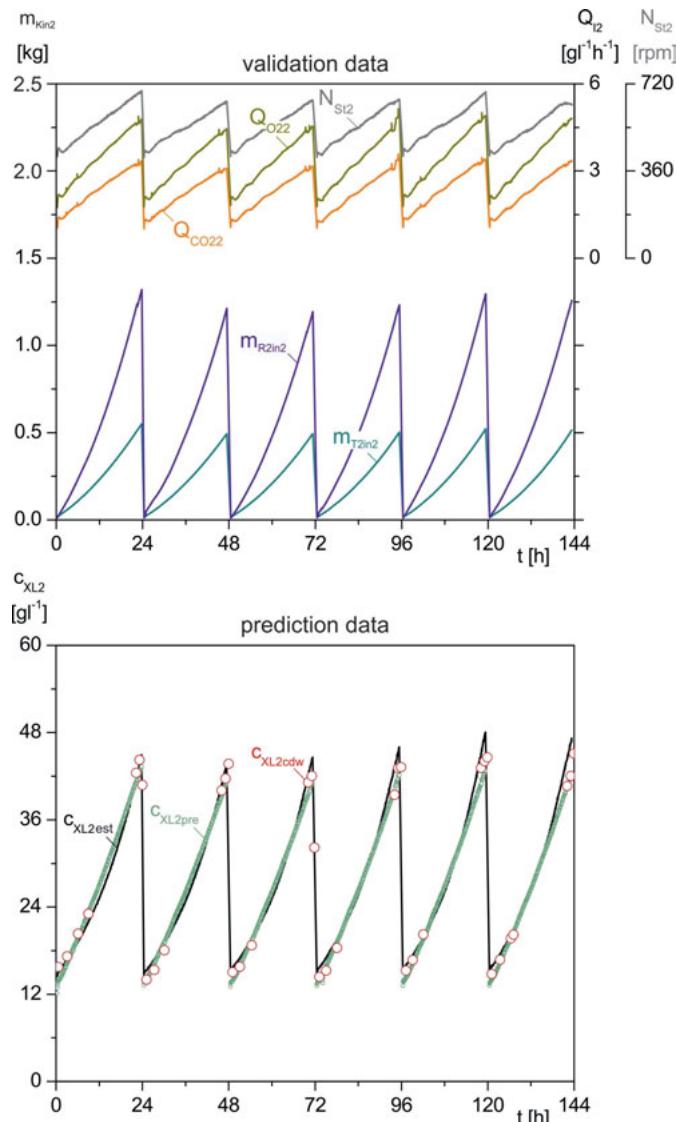


Bild 9.8: Validierung des Zelldichtemodells im *protein production* Prozess

Die Offline-Daten, die Online-Messungen mit Trübungssonde und die Prädiktion aus Online-Daten ohne Sonde zeigen eine sehr gute Übereinstimmung. Dies ermutigt, die MVDA-Methode auch für die komplexere Zielgröße Zielprodukt  $C_{P1M2}$  anzuwenden.

#### 9.4.3 Prädiktion des sekretierten Zielproduktes

In Bild 9.9 sind die Trainingsdaten  $C_{P1M2est}$  (schwarz) gezeigt, die aus Atline-Messungen  $C_{P1M2at}$  ( $\Delta$ ) mittels Gl. (9.36) ermittelt wurden. Für diese Daten wurden 19 Zyklen mit hoher Variationsbreite aus Produktionskampagnen, DoE-Versuchen aber auch Fehlversuchen ausgewählt.

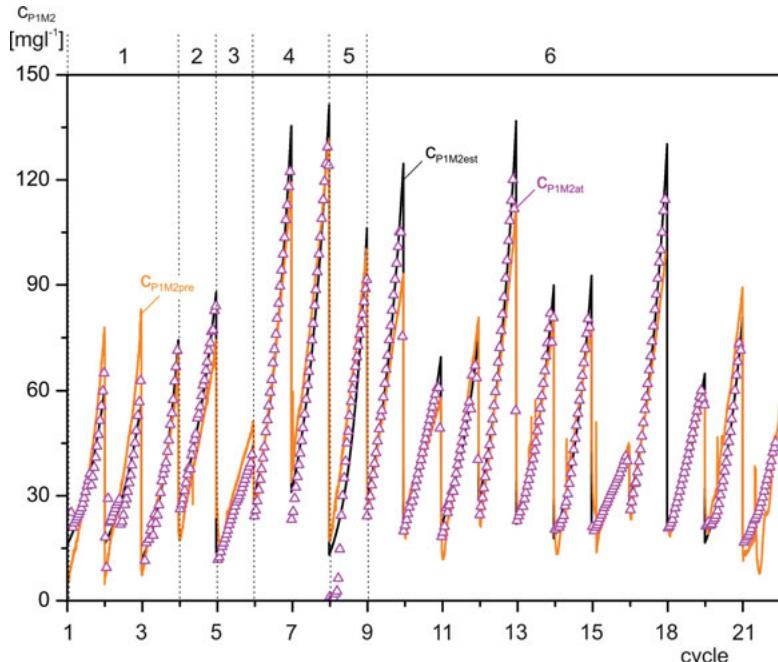


Bild 9.9: Verlauf der Atline-Messwerte  $C_{P1M2at}$  ( $\Delta$ ), Trainingsset  $C_{P1M2est}$  (schwarz) und Prädiktion  $C_{P1M2pre}$  (braun) der PLS-Produktanpassung im *protein production* Prozess

Die  $C_{P1M2est}$ -Daten bilden den Spaltenvektor  $Y_6$  für die Erstellung eines PLS-Modells für  $C_{P1M2pre}$  mit einer reduzierten Prozessmatrix  $X_{26}$ . Deren Komponenten sind ebenfalls Tabelle 9.3 zu entnehmen.

Auf  $X_{26}$  und  $Y_6$  wird analog zu Gl. (9.39) eine PLS-Modellierung angewendet, deren Ergebnis ebenfalls in Bild 9.9 als  $C_{P1M2pre}$  (braun) gezeigt ist. Wenn man bedenkt, welche Daten zur Berechnung des Produktverlaufes herangezogen wurden, meistert die Anpassung erstaunlicherweise auch die Experimente mit geringer Produktausbeute hervorragend.

Allerdings erfordert auch das PLS-Produktmodell eine nachfolgende Validierung. Diese ist mit 8 Produktionszyklen in unterschiedlichen Kampagnen in Bild 9.10 gezeigt.

Die Prädiktion  $C_{P1M2pre}$  (braun) trifft die Validierungsdaten  $C_{P1M2est}$  (schwarz) recht gut, allerdings weisen Abweichungen in Zyklus 4 und 6 auch auf Mängel der MVDA-Methode hin.

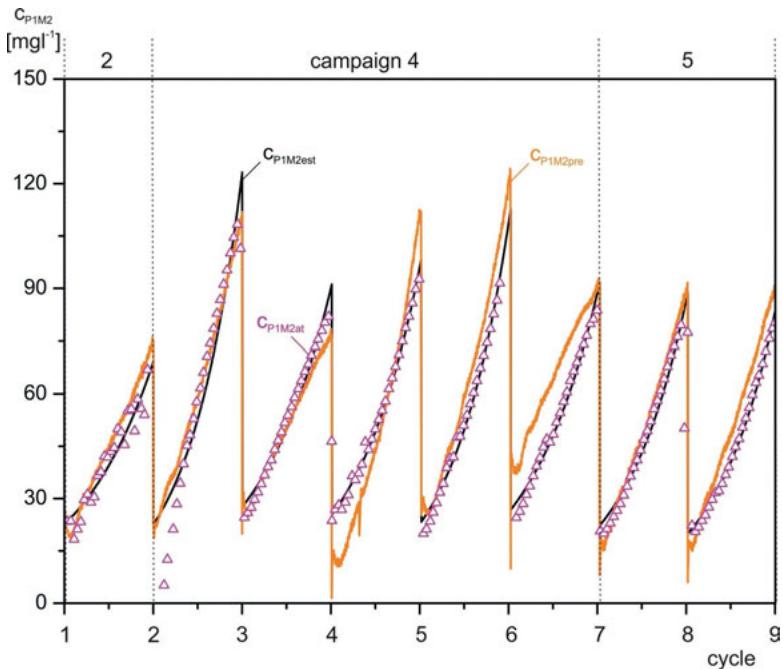


Bild 9.10: Verlauf einer Validierung des PLS-Produktmodells aus Bild 9.9

#### 9.4.4 Prädiktion der Komponenten im cell breeding Prozess

Die Prädiktionsmodelle im *cell breeding* Prozess entsprechen prinzipiell der Modellbildung im *protein production* Prozess.

Für die Zelldichte  $c_{XL1}$  im Batch ( $X_{11}, Y_1$ ) und in der Induktion ( $X_{13}, Y_3$ ) sowie die Glycerolkonzentration  $c_{S1L1}$  im Batch ( $X_{12}, Y_2$ ) stehen entsprechende Trainingsdatensätze zur Verfügung.

Der Produktverlauf  $C_{P1M1}$  in der Induktion wurde weder offline noch atline gemessen. Daher wird die Matrix  $X_{26}$  um die äquivalenten im *cell breeding* Prozess gemessenen Daten ( $X_{14}$ ) ergänzt und die Matrix  $Y_6$  als  $Y_4$  weiterhin für die PLS-Anpassung verwendet. Die fehlenden Werte in  $Y_4$  für die *cell breeding* Zyklen werden entsprechend den in der Software SIMCA implementierten Methoden zur Behandlung fehlender Werte berücksichtigt. Die Fa. Umetrics verweist für die verwendete IBR-Methode auf Arteaga und Ferrer [Arteaga and Ferrer, 2002].

Unter der Annahme, dass in seiner späteren Phase der Induktionsprozess sehr stark dem Produktionsprozess ähnelt, kann auf diese Weise ein Modell für den Produktverlauf im *cell breeding* Prozess erstellt werden. Da sich beide Prozesse aber in ihrem Reaktionsvolumen unterscheiden, ist es unbedingt notwendig diesen Faktor mit in die Modellbildung einzubeziehen. So kann mit dem erstellten Modell dann auch  $C_{P1M1pre}$  aus den ergänzten Daten von Reaktor 1 in  $X_{14}$  berechnet werden.

In Bild 9.11 ist die Prädiktion der *cell breeding* Variablen für drei Tageszyklen dargestellt. Die Zelldichte  $C_{XL1pre}$  und die Glycerolkonzentration  $c_{S1L1pre}$  beschreiben den Prozess sehr genau, während das Zielprodukt  $C_{P1M1pre}$  jeweils bei  $t_{n1}$  zu gering ausfällt. Trotzdem hat sich auch das Modell einer Produkt-Prädiktion ohne eigentliches Trainingsset als hinreichend erwiesen.

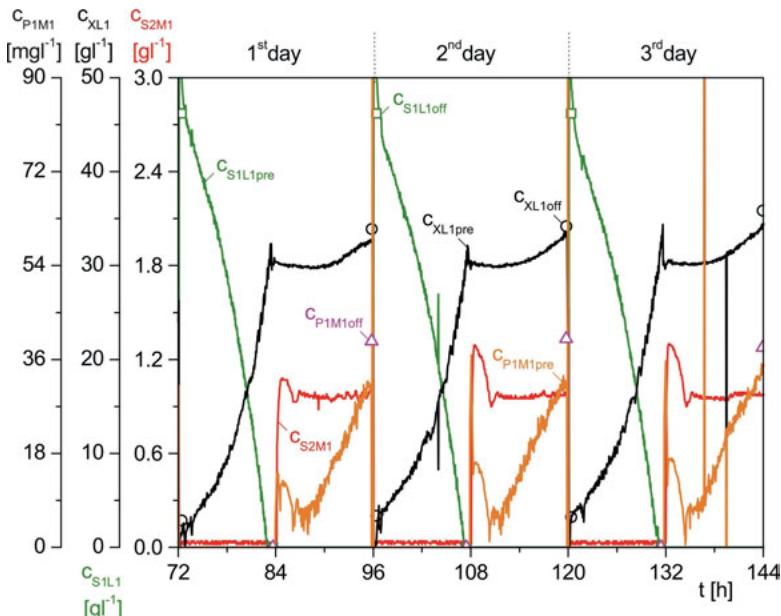


Bild 9.11: Online-Prädiktionen der Variablen  $c_{XL1}$ ,  $c_{S1L1}$  und  $c_{P1M1}$  in *cell breeding* Zyklen

#### 9.4.5 PLS-Koeffizienten der entwickelten Prädiktionsmodelle

Abschließend zur Prädiktion der Reaktionskomponenten sind in nachfolgender Tabelle 9.3 die ermittelten Komponenten der Prädiktorektoren  $B_i$  aufgeführt. Die Spalte  $X_p$ -values beinhaltet dabei die Online-Prozessgrößen der Reaktoren 1 und 2, die für eine MVDA zur Verfügung standen und aus denen die entsprechenden Untermatrizen  $X_{pi}$  gebildet wurden.

Bei der Modellbildung wurden folgende Kenngrößen eingesetzt bzw. ermittelt:

- $n_{TS}$  := Anzahl der Observationen im Trainingsset
- $o_{TS}$  := Anzahl der Batches im Trainingsset
- $n_{PS}$  := Anzahl der Observationen im Validierungsset
- $o_{PS}$  := Anzahl der Batches im Validierungsset
- $R^2_{YTS}$  := *goodness of fit* für das Modell aus dem Trainingsset
- $Q^2_{YTS}$  := *goodness of prediction* für das Modell aus dem Trainingsset.

Für eine quantitative Bewertung der erstellten Modelle sind in Tabelle 9.4 die Zahlenwerte dieser gängigen statistischen Parameter aus den Gln. (7.23) bis (7.39) (siehe Abschnitt 7.8) aufgelistet.

Die *root-mean-square error* (RMSE) sind dabei jeweils für das Trainingsset (TS) und das Validierungsset (PS) aufgeführt und normiert, um einen Vergleich der Modelle untereinander zu ermöglichen. Die Normierung erfolgte dabei durch Division durch den Mittelwert des TS bzw. PS (CV) oder durch Division durch die Differenz des maximalen und minimalen Wertes des TS bzw. PS (N).

Tabelle 9.3: Verwendete Prozessdaten  $X_{ki}$  und daraus ermittelte PLS-Regressionskoeffizienten  $B_i$  der Prädiktionsmodelle

		PLS Regressionskoeffizienten ( $B_i$ )								
$X_p$ -values	units	CxLppre [ $gl^{-1}$ ]			Cs1L1pre [ $gl^{-1}$ ]	Cp1Mppre [ $gl^{-1}$ ]			units	
		$Y_1$	$Y_3$	$Y_5$		$Y_2$	$Y_4$	$Y_6$		
constant	[ $-$ ]	232,66	-26,68	39,68	-6,30	-81,90	-344,69	[ $gl^{-1}$ ]		
Cs2Mp	[ $gl^{-1}$ ]	-	-	-	-	6,09	-	[ $-$ ]		
CxLppre	[ $gl^{-1}$ ]	-	-	-	-3,49·10 <sup>-1</sup>	2,55·10 <sup>-1</sup>	1,89·10 <sup>-1</sup>	[ $-$ ]		
F <sub>R2p</sub>	[ $lh^{-1}$ ]	-	12,09	-81,47	-	241,51	244,95	[ $ghl^{-2}$ ]		
F <sub>R2rel1</sub>	[%]	-	-	-	-	2,47	-	[ $gl^{-1\%^{-1}}$ ]		
m <sub>R2p</sub>	[g]	-	1,33·10 <sup>-2</sup>	1,19·10 <sup>-2</sup>	-	9,81·10 <sup>-3</sup>	2,09·10 <sup>-2</sup>	[ $l^{-1}$ ]		
m <sub>T2p</sub>	[g]	3,06·10 <sup>-2</sup>	2,26·10 <sup>-2</sup>	2,39·10 <sup>-2</sup>	-4,67·10 <sup>-2</sup>	3,95·10 <sup>-2</sup>	2,25·10 <sup>-2</sup>	[ $l^{-1}$ ]		
N <sub>Stp</sub>	[rpm]	1,73·10 <sup>-3</sup>	1,13·10 <sup>-3</sup>	-1,26·10 <sup>-3</sup>	-1,70·10 <sup>-2</sup>	5,60·10 <sup>-3</sup>	9,12·10 <sup>-2</sup>	[ $gl^{-1} rpm^{-1}$ ]		
pH <sub>2</sub>	[ $-$ ]	-3,50·10 <sup>1</sup>	8,63	-	-	-	33,53	[ $gl^{-1}$ ]		
pO <sub>2p</sub>	[%]	1,80·10 <sup>-2</sup>	2,07·10 <sup>-2</sup>	-	-2,79·10 <sup>-1</sup>	-	-9,91·10 <sup>-1</sup>	[ $gl^{-1\%^{-1}}$ ]		
Q <sub>CO2p</sub>	[ $gl^{-1} h^{-1}$ ]	1,49	5,07·10 <sup>-4</sup>	9,51·10 <sup>-1</sup>	-1,63	-3,83	-1,27	[h]		
Q <sub>O2p</sub>	[ $gl^{-1} h^{-1}$ ]	1,42	1,54·10 <sup>-1</sup>	7,30·10 <sup>-1</sup>	-1,57	-4,19	69,37	[h]		
RQ <sub>p</sub>	[ $-$ ]	-1,01	1,78	-	21,48	-	2,20	[ $gl^{-1}$ ]		
9 <sub>DJp</sub>	[ $^{\circ}C$ ]	-2,86·10 <sup>-1</sup>	1,15·10 <sup>-1</sup>	-	4,22·10 <sup>-1</sup>	8,92·10 <sup>-1</sup>	1,87·10 <sup>-1</sup>	[ $gl^{-1} C^{-1}$ ]		
9 <sub>Lp</sub>	[ $^{\circ}C$ ]	-	-	-	-	1,11	2,09·10 <sup>-2</sup>	[ $gl^{-1} C^{-1}$ ]		
V <sub>L1</sub>	[l]	-	-	-	-	-2,21	-	[ $gl^{-2}$ ]		
X <sub>CO2p</sub>	[%]	1,80	1,38·10 <sup>-1</sup>	2,09	-2,18	4,74·10 <sup>-1</sup>	-	[ $gl^{-1\%^{-1}}$ ]		
X <sub>O2p</sub>	[%]	-1,33	-2,16·10 <sup>-1</sup>	-1,26	1,63	1,78	1,45	[ $gl^{-1\%^{-1}}$ ]		

Tabelle 9.4: Bewertungsparameter für die erstellten PLS-Modelle

	CxLppre			Cs1L1pre batch	Cp1Mppre	
	batch	induction	production		induction	production
n <sub>TS</sub>	7255	9011	7053	5424	16408	8736
o <sub>TS</sub>	38	38	16	29	44	19
n <sub>PS</sub>	2433	2812	2766	777	-	3711
o <sub>PS</sub>	11	11	6	4	-	8
R <sub>YTS</sub> <sup>2</sup>	0.927	0.878	0.978	0.971	0.774	0.904
Q <sub>YTS</sub> <sup>2</sup>	0.926	0.878	0.978	0.971	0.774	0.904
RMSECV <sub>TS</sub>	0.156	0.032	0.047	0.087	0.658	0.167
RMSEN <sub>TS</sub>	0.056	0.082	0.034	0.043	0.087	0.060
RMSECV <sub>PS</sub>	0.226	0.082	0.060	0.165	-	0.161
RMSEN <sub>PS</sub>	0.102	0.067	0.048	0.082	-	0.080

## 9.5 Ermittlung der zellspezifischen Reaktionsraten

### 9.5.1 Offline-Auswertung der drei Prozessphasen

Im Anschluss an die Prädiktion von Zelldichte ( $c_{XLP}$ ), Glycerol- ( $c_{S1L1}$ ) und Zielproduktkonzentration ( $c_{P1MP}$ ) wurden aus 53 *cell breeding* und 55 *protein production* Kultivierungen die zell-spezifischen Reaktionsraten  $q_{IXP}$  bestimmt.

Hierbei wurden nur Prozesszyklen mit den optimalen kritischen Produktionsparametern ausgewählt, deren Sollwerte in Tabelle 9.5 aufgelistet sind.

Tabelle 9.5: Optimale Produktionsparameter in den drei Prozessphasen

	batch (p=1)	induction (p=1)	production (p=2)
pH <sub>pw</sub> [-]	5.55	5.55	5.55
θ <sub>Lpw</sub> [°C]	30.0	22.0	22.0
C <sub>S2Mpw</sub> [g l <sup>-1</sup> ]	-	1.0	1.0

In Bild 9.12 ist der Verlauf der  $q_{IXP}$  in den drei Prozessphasen gezeigt. Der jeweilige Mittelwert ist in schwarz eingezeichnet.

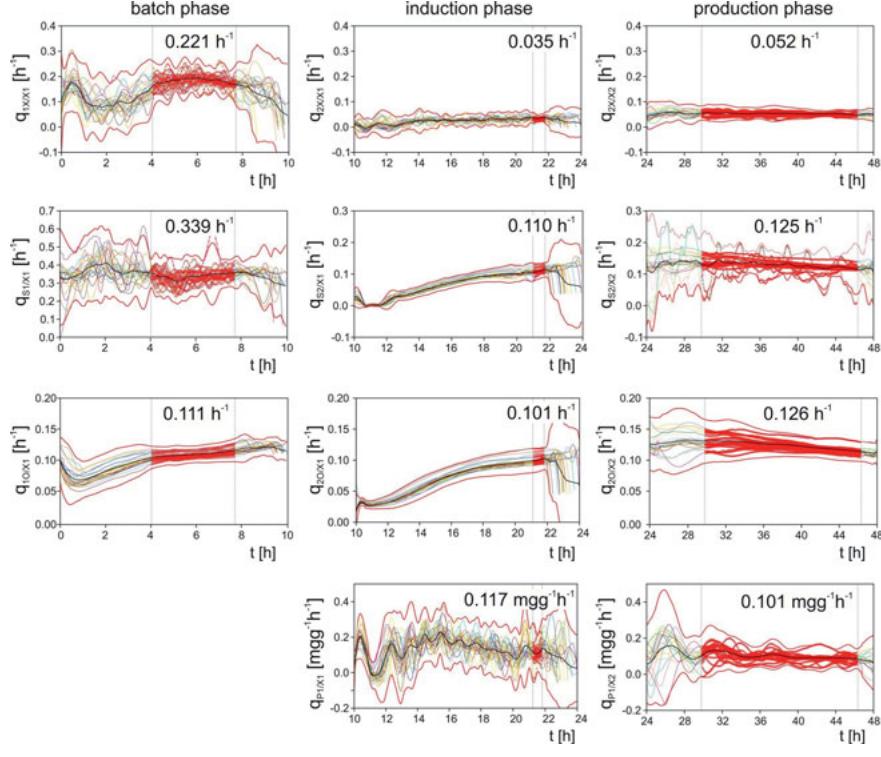


Bild 9.12: Verlauf der zellspezifischen Reaktionsraten in den drei Prozessphasen

Das Ziel dieser Untersuchung war es, den zeitlichen Verlauf der  $q_{I/X_p}$  zu bestimmen und aus ausgewählten Bereichen die maximalen Raten  $q_{I/X_{max}}$  für die Phasen *cell breeding* auf Glycerol und *protein production* auf Methanol zu identifizieren. Zur Berechnung wurden für Reaktor 1 die Gln. (9.22) bis (9.27) und für Reaktor 2 die Gln. (9.29) bis (9.32) herangezogen.

Das Problem einer unbekannten 1. Ableitung einer Konzentration der Komponente I in oder einer unbekannte Zufütterrate aus Teilsystem J wird über die zentralen Differenzenquotienten,

$$\dot{c}_{I,J_p}(t_j) = \frac{c_{I,J_p}(t_{j+1}) - c_{I,J_p}(t_{j-1})}{t_{j+1} - t_{j-1}}, J = L, M \quad (9.41)$$

und

$$F_{J_p}(t_j) = -\frac{m_{J_p}(t_{j+1}) - m_{J_p}(t_{j-1})}{\rho_J \cdot (t_{j+1} - t_{j-1})}, J = T2, R2, \quad (9.42)$$

gelöst.

### 9.5.2 Identifikation der maximalen zellspezifischen Reaktionsraten

In den rot gekennzeichneten, relativ stationären Bereichen der Glycerol-Batchphase in Reaktor 1 und der Methanol Produktionsphase in Reaktor 2 wird der jeweilige Mittelwert als Maximalwert  $q_{I/X_{max}}$  angesehen.

Diese sind in Bild 9.13 gezeigt und mit Werten einer Einreaktorstrategie verglichen, die von Martens [Martens, 2014] bei hoher Probenahme aus Offline-Daten über eine DoE-Strategie gefunden wurden.

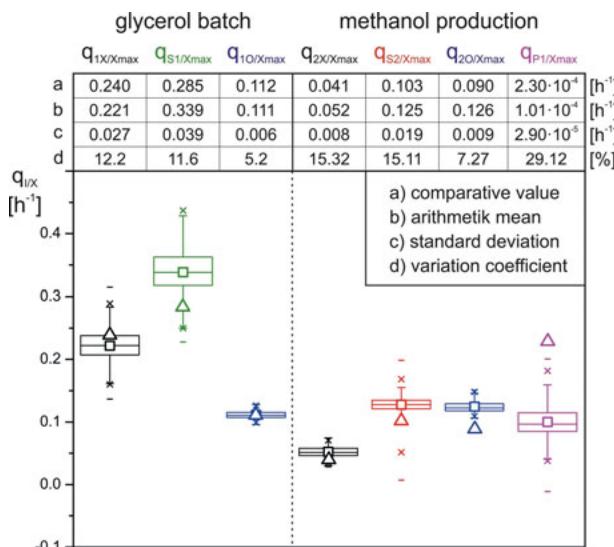


Bild 9.13: Vergleich der maximalen zellspezifischen Reaktionsraten  $q_{I/X_{max}}$   
a) Vergleichswerte Martens ( $\Delta$ ), b) arithmetischer MVDA-Mittelwert ( $\square$ ),  
c) Standardabweichungen, d) Variationskoeffizienten

Der dargestellte Box-Plot beinhaltet Informationen über die Verteilung der für die Berechnung des Mittelwertes herangezogenen Stichprobe. Der Median der Stichprobe ist durch den Strich in der Box gekennzeichnet. Der untere Teil der Box beschreibt das untere, der obere Teil der Box das obere Quantil der Stichprobe, d.h. 25 % der Stichprobenwerte liegen unterhalb des unteren Quantils und 25 % liegen oberhalb des oberen Quantils.

Die Länge der Antennen wurde zu beiden Seiten mit dem Wert der Stichprobe festgelegt, der noch in dem Bereich von 1,5-mal dem Interquartilsabstand liegt. Das x-Symbol im unteren bzw. oberen Teil des Whiskers markiert den 1 % bzw. 99 %-Bereich der Stichprobe, das Minus-Zeichen den Minimal- bzw. den Maximalwert.

Der arithmetische Mittelwert der jeweiligen zellspezifischen Reaktionsrate liegt in allen Fällen auf dem des Median der Stichprobe, so dass davon ausgegangen werden kann, dass kein Ausreißer das Ergebnis verfälscht hat. Im Folgenden wurden daher diese Werte für einen Vergleich herangezogen.

Mit Hilfe der DoE-Modelle lassen sich die bei Martens [Martens, 2014] gefundenen Parameter leicht mit den vorliegenden Bedingungen für pH<sub>pw</sub> und θ<sub>Lpw</sub> vergleichen.

Dargestellt sind dafür die korrigierten Maximalwerte von Martens, die mit Hilfe der MVDA Prädiktion c<sub>i,p</sub> berechneten arithmetischen Mittelwerte  $\bar{q}_{i/X_{max}}$  sowie deren zugeordnete Standardabweichung  $\sigma_{i/X}$  und der zugehörige Variationskoeffizient  $V_{i/X}$ . Zu bedenken ist, dass die Parameter der Batchphase nur für pH<sub>1w</sub> = 5.5 und θ<sub>L1w</sub> = 30 °C und die der Produktionsphase nur für pH<sub>2w</sub> = 5.5 und θ<sub>L2w</sub> = 22 °C gelten.

Für das Wachstum ( $q_{i/X_{max}}$ ) auf Glycerol (i = 1) als auch Methanol (i = 2) besteht eine hohe Übereinstimmung mit den Werten von Martens. Die Methanolaufnahme bei voller Induktion ( $q_{S2/X_{max}}$ ) ist identisch. Die Glycerolaufnahme ( $q_{S1/X_{max}}$ ) und die Produktbildung ( $q_{P1/X_{max}}$ ) weichen leicht (S<sub>1</sub>) bis stark (P<sub>1</sub>) voneinander ab.

### 9.5.3 Online-Bestimmung der zellspezifischen Raten

Im Weiteren wurden die Prozeduren zur Berechnung von  $q_{i/X_1}$  auch im Online-Betrieb am *cell breeding* Reaktor erprobt. Hierzu wurde die bereits in Bild 8.2 gezeigte Infrastruktur der erweiterten Datenverarbeitung genutzt.

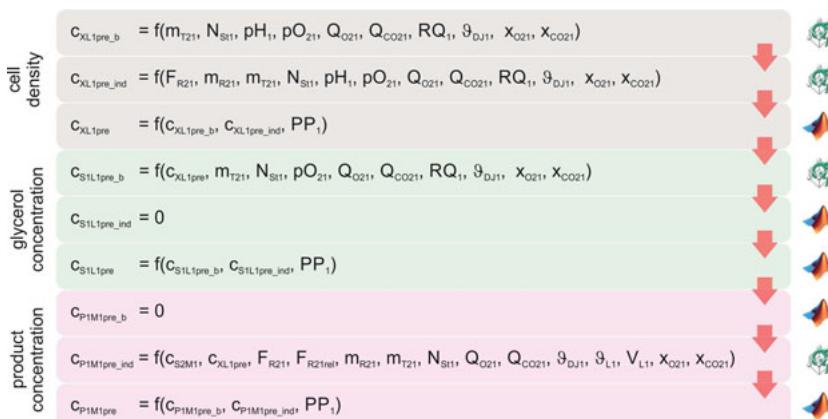


Bild 9.14: An der Prädiktion beteiligte Variablen und Softwarepakete MATLAB und SIMCA Q

Die Prozessdaten der beiden Bioreaktoren wurden mit MFCS/win aufgenommen und über PCS 7 in SIPAT in eine MVDA-gerechte Form gebracht. Die PLS-Modelle zur Prädiktion der drei Medienkomponenten und deren Reaktionsraten wurden mit SIMCA erstellt und online mit SIMCA Q exekutiert. Dabei laufen die Modelle für die beiden beteiligten Phasen *glycerol batch* und *methanol pre-induction* parallel mit.

In einem ebenfalls online mitlaufenden MATLAB-Skript wird anhand einer Online-Variablen PP<sub>1</sub> (*process phase*) entschieden, welche Prozessphase gerade läuft und dafür die entsprechenden Raten q<sub>1XX1</sub> und q<sub>S1/X1</sub> bzw. q<sub>2XX1</sub>, q<sub>S2/X1</sub> und q<sub>P1/X1</sub> online berechnet.

Die an der Prädiktion von c<sub>XL1</sub>, c<sub>S1L1</sub> und c<sub>P1M1</sub> beteiligten Prozessdaten sind Bild 9.14 zu entnehmen. Die zugehörigen Online-Abschätzungen sind in Bild 9.15 gezeigt.

Aus diesen und der Methanolkonzentration c<sub>S2M1</sub> wurden die zellspezifischen Reaktionsraten q<sub>i/X1</sub> mit den Gln. (9.22 bis 9.27) und Gl. (9.32) ebenfalls online berechnet. Wie man erkennt ist nur der Verlauf von q<sub>10/X1</sub> zufriedenstellend gelungen. Die Bestimmungen von q<sub>1XX1</sub> und q<sub>S1/X1</sub> sind sehr verrauscht. Hier bedarf es noch des Einsatzes von leistungsfähigen Online-Filtermethoden für die geschätzten Konzentrationen, um glatte Verläufe der Reaktionsraten zu erhalten.

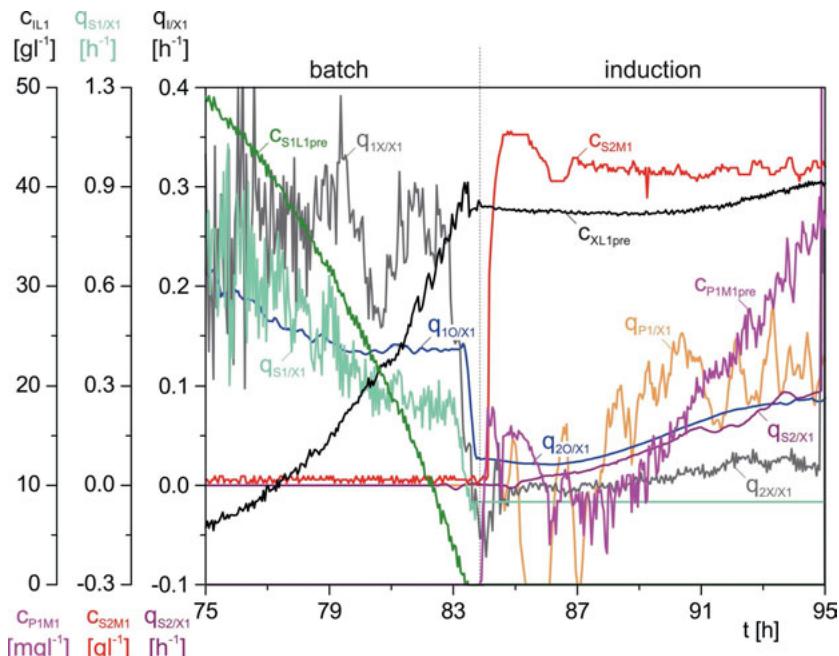


Bild 9.15: Ein erster Versuch einer Online-Bestimmung von q<sub>iXX1</sub>, q<sub>S1/X1</sub>, q<sub>10/X1</sub> und q<sub>P1/X1</sub>

In der Anlaufphase der Vorinduktion ab t = 84,1 h liegt noch kein AOX in der Zelle vor. Die Zelldichte c<sub>XL1</sub> nimmt ab, da die Morphologie stark umgebaut wird. Die Wachstumsrate q<sub>2XX1</sub> wird dabei zunächst negativ.

Mit Umschalten der Prozessphase schwingen einige der an der Prädiktion beteiligten Prozessvariablen auf die neuen Prozessbedingungen ein. Dieser Vorgang erzeugt auch starke Schwingungen in der Produktbestimmung  $c_{P1M1pre}$ . Insbesondere die noch nicht funktionsfähige Methanolregelung bei fehlendem AOX und die damit verbundene Aufladung der Mediaphase M<sub>1</sub> mit Methanol erzeugt eine starke Fehlinformation.

Daher ist die Prädiktion von  $q_{P1/X1}$  erst ab  $t = 88$  h glaubhaft. Deren Anstieg verläuft parallel zu  $q_{2wx_1}$ ,  $q_{2ox_1}$  und  $q_{s2x_1}$ , die von dem zellinternen AOX-Gehalt  $g_{P2/X1}$  gesteuert werden. Dieser wird allerdings erst mit einem Verzögerungsverhalten höherer Ordnung aufgebaut [Luttmann et al., 2010].

Leider konnten durch Auslauf des Projektes die Arbeiten an dieser Thematik nicht fortgesetzt werden. Dennoch zeigen die Ergebnisse das große Potential eines MVDA-begleiteten Prozessmonitorings auf. Dies sollte in weiteren Forschungsvorhaben genutzt werden.

## 10. Zusammenfassung und Ausblick

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

Im Zuge der experimentellen und theoretischen Untersuchungen wurden quasi-kontinuierliche Prozesse aus herkömmlichen *Batch-Unit Operations* zur Prozessintensivierung aufgebaut.

Dabei standen die Anwendungen von erweiterten mess- und Regelungstechnischen Maßnahmen sowie systemtheoretischer Methoden im Vordergrund, die allgemein unter PAT-*Process Analytical Technology* subsumiert werden.

Die damit durchgeführten Prozessentwicklungen dienen insbesondere dazu, die von der FDA 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. An dieser Anlage wurde dann eine sequentiell/integrierte 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 identifiziert.

Für den EBA-Chromatographieschritt wurde dabei ein mehrstufiges Optimierungsverfahren entwickelt.

In einem industriell orientierten Bereich wurde dann eine vorhandene quasi-kontinuierlich arbeitende integrierte Scale-down Produktionsanlage für die Produktion eines pharmazeutischen Produktes ergänzt und intensiv erprobt.

Bei der eingesetzten Prozessführung wurden das *cell breeding*, die *protein production* und der Downstream mit *cell clarification*, *micro*-, *ultra*- und *dia-crossflow filtration* sowie *purification* sequentiell, aber auch auf drei Tage verteilt, parallel gefahren.

Hieraus entwickelte sich eine quasi-kontinuierliche Prozessführung von zyklischer Wiederholung eines 56-stündigen Prozesses, der über eine Reihe von Tagen alltäglich geerntet wurde.

Der Einsatz der integrierten Konzepte führte zu einer Verdreifachung der Produktivität im Vergleich zur klassischen *Batch-to-Batch*-Produktion.

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.

Aufbauend auf einer Erläuterung verschiedener MVDA-Methoden wie PCA, MPCA und PLS wurden diese zur Analyse, dem Monitoring und der Qualitätssteuerung von zyklischen Prozessen eingesetzt.

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 aus klassischen Online-Messungen mit PLS – *Partial Least Squares* Regression eingesetzt.

Hieraus konnte damit online der Verlauf zugehöriger zellspezifischer Reaktionsraten für Zell-dichte, Glycerol, Methanol und Zielprodukt beobachtet werden.

Die Anwendung der MVDA eröffnete somit ein weites Feld eines erweiterten PAT Einsatzes und brachte das Ziel eines *Online-Release* von pharmazeutischen Produkten aus MVDA-beobachteten Prozessen aus der Wunschecke in eine realitätsnahe Zukunft.

Daher besitzt diese Arbeit das Potential zu dem Ziel einer routinemäßigen Anwendung der MVDA im industriellen Alltag beizusteuern.

## 11. Anhang

### 11.1 Kultivierungsmedien

Als Vorkultur-, Kultivierungs- und Refreshmedium wurde ein Grundsatzmedium (Tabelle 11.1) gemäß Tabelle 11.2 mit dem Primärsubstrat Glycerol versetzt und bei 121 °C für 20 min autoklaviert. Nach Abkühlen werden 8 ml<sup>-1</sup> der Biotinvorlage und 4 ml<sup>-1</sup> der PTM4-Spuren elementlösung zugefügt. Diese Stocklösungen wurden zuvor nach Tabelle 11.4 angesetzt, mit 0.2 µm Porendurchmesser sterilfiltriert und bei 4 °C gelagert.

Tabelle 11.1: Grundsatzmedien

Komponente	Konzentration [gl <sup>-1</sup> ]	
	BIOSTAT® ED10/C30 Medium 1	BIOSTAT® ED5 Medium 2
KH <sub>2</sub> PO <sub>4</sub>	25.7	7.0
(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	5.0	5.0
K <sub>2</sub> SO <sub>4</sub>	8.6	2.0
CaSO <sub>4</sub> · 2H <sub>2</sub> O	1.4	0.8
MgSO <sub>4</sub>	16.4	8.0
Na <sub>2</sub> -citrate · 2H <sub>2</sub> O	5.9	5.0

Tabelle 11.2: Glycerolvorlagen

Komponente	Konzentration [gl <sup>-1</sup> ]		
	Vorkultur	Batch	Refresh
Glycerol	30	30 bzw. 55	0

Bei dem Grundsatzmedium 1 handelt es sich um ein modifiziertes FM22 Medium nach Stratton [Stratton et al., 1998]. Um das Ausfallen der Salze bei leicht sauren pH-Werten zu reduzieren, wurde als *Low Salt* Variante Medium 2 konzipiert. Dieses wurde ausschließlich in den Kultivierungen mit dem BIOSTAT® ED5 verwendet, alle anderen erfolgten mit Medium 1.

Die Konzentrationen in den verwendeten Zufüttervorlagen sind in Tabelle 11.3 und die in den Stocklösungen in Tabelle 11.4 gegeben.

Tabelle 11.3: Feedlösungen

Feedlösung	Konzentration [gl <sup>-1</sup> ]
Glycerol	500
Methanol	792

Tabelle 11.4: Stocklösungen

PTM4		Biotinvorlage	
Komponente	Konzentration	Komponente	Konzentration
CuSO <sub>4</sub> · 5 H <sub>2</sub> O	2.00 g l <sup>-1</sup>	C <sub>10</sub> H <sub>16</sub> N <sub>2</sub> O <sub>3</sub> S	0.20 g l <sup>-1</sup>
NaI	0.08 g l <sup>-1</sup>		
MnSO <sub>4</sub> · H <sub>2</sub> O	3.00 g l <sup>-1</sup>		
Na <sub>2</sub> MoO <sub>4</sub> · 2 H <sub>2</sub> O	0.20 g l <sup>-1</sup>		
H <sub>3</sub> BO <sub>3</sub>	0.02 g l <sup>-1</sup>		
CaSO <sub>4</sub>	0.50 g l <sup>-1</sup>		
CoCl <sub>2</sub> · 6 H <sub>2</sub> O	0.50 g l <sup>-1</sup>		
ZnSO <sub>4</sub> · 7 H <sub>2</sub> O	7.00 g l <sup>-1</sup>		
FeSO <sub>4</sub> · 7 H <sub>2</sub> O	22.00 g l <sup>-1</sup>		
H <sub>2</sub> SO <sub>4</sub>	1.00 mill <sup>-1</sup>		

## 11.2 Pufferlösungen für Filtration und Chromatographie

In Tabelle 11.5 sind die verwendeten Pufferlösungen der EBA-Chromatographie und der Atline-HPLC zur Detektion von D1M1 aufgelistet. Hier wurden in beiden Systemen die identischen Ansätze für die Puffer in unterschiedlichen Mengen vorgenommen.

Tabelle 11.5: Verwendete Pufferlösungen am integrierten EBA-Prozess

Komponente	Konzentration [mM]				
	Lauf (pH = 7.4)	Elution (pH = 7.4)	Stripping (pH = 7.4)	Acetat (pH = 5)	Recharge (pH = 5)
NaCl	500	500	500	-	-
NaH <sub>2</sub> PO <sub>4</sub> · 2H <sub>2</sub> O	20	20	20	-	-
Imidazol	8	100	-	-	-
NaAc-Trihydrat	-	-	-	50	-
CuSO <sub>4</sub> · 5H <sub>2</sub> O	-	-	-	-	100
EDTA	-	-	30	-	-

Tabelle 11.6 enthält die Konzentrationen der Komponenten des PBS-Puffers, die bei der SEC zur Feinreinigung von D1M1 erforderlich waren.

Tabelle 11.6: PBS-Puffer zur Feinreinigung und Umpufferung von D1M1 mittels einer SEC

Komponente	Konzentration [mM]
	Lauf (pH = 7.4)
Na <sub>2</sub> HPO <sub>4</sub>	10
NaCl	140
KH <sub>2</sub> PO <sub>4</sub>	1.8
KCl	2.7

Für die Atline- und die Offline-HPLC zur Detektion von D1M1H wurden die in Tabelle 11.7 aufgeführten Puffer angesetzt.

Tabelle 11.7: Verwendete Pufferlösungen zur Atline- und Offline-Detektion von D1M1H

Komponente	Konzentration [mM]					
	Lauf (pH = 7.4)	Elution 1 (pH = 7.4)	Elution 2 (pH = 7.4)	Stripping (pH = 7.4)	Acetat (pH = 5)	Recharge (pH = 5)
NaCl	500	500	500	500	-	-
NaH <sub>2</sub> PO <sub>4</sub> · 2H <sub>2</sub> O	20	20	20	-	-	-
Imidazol	0	30	300	-	-	-
NaAc-Trihydrat	-	-	-	-	50	-
NiSO <sub>4</sub> · 6H <sub>2</sub> O	-	-	-	-	-	15
EDTA	-	-	-	30	-	-

Bei der kombinierten Ultra- und Diafiltration wurde der in Tabelle 11.8 zusammengefasste Puffer zur Umpufferung von D1M1H eingesetzt.

Tabelle 11.8: Verwendete Pufferlösungen bei der Diafiltration

Komponente	Konzentration [mM]	
	Lauf (pH = 7.4)	
NaCl	500	
NaH <sub>2</sub> PO <sub>4</sub> · 2H <sub>2</sub> O	20	
Imidazol	5	

Eine Feinreinigung von D1M1H mit dem ÄKTA™purifier 100 erfolgte mit den in Tabelle 11.9 aufgelisteten Puffern.

Tabelle 11.9: Verwendete Pufferlösungen zur Feinreinigung von D1M1H

Komponente	Konzentration [mM]				
	Lauf (pH = 7.4)	Elution (pH = 7.4)	Stripping (pH = 7.4)	Acetat (pH = 5)	Recharge (pH = 5)
NaCl	500	500	500	-	-
NaH <sub>2</sub> PO <sub>4</sub> · 2H <sub>2</sub> O	20	20	-	-	-
Imidazol	5	300	-	-	-
NaAc-Trihydrat	-	-	-	50	-
NiSO <sub>4</sub> · 6H <sub>2</sub> O	-	-	-	-	15
EDTA	-	-	30	-	-

## 11.3 Offline-Messungen

### 11.3.1 Bestimmung der optischen Dichte, der Biofeucht- und Biotrockenmasse

Die photometrische Bestimmung der optischen Dichte erfolgte in Einmalküvetten mit 1 cm Schichtdicke in einem Absorptionsmessbereich von 0.2 bis 0.6 bei einer Wellenlänge von 600 nm im Spektrophotometer.

Zur Ermittlung der Biotrockenmasse wurden 1 ml der zu bestimmenden Probe in vorgewogenen Mikroreaktionsgefäße zentrifugiert (5000 rpm, 3 min), der Überstand entnommen und

das Pellet bei 80 °C für mindestens 3 Tage getrocknet. Die erneute Aufnahme des Gewichts ergab abzüglich des Leergewichts die Biotrockenmasse.

### 11.3.2 Bestimmung des Gesamtproteingehalts (Bradford-Test)

Zur Bestimmung der Gesamtproteinkonzentration wurde ein Bradford-Test in Mikrotiterplatten durchgeführt.

Lösungen:

- A Standards: Quick Start™ Bovine  $\gamma$ -Globulin Standard Set (Bio-Rad Laboratories)
- B Reaktionslösung: Quick Start™ Bradford Dye Reagent (Bio-Rad Laboratories)

Unter Verwendung von A wurde im Bereich von 12.5 bis 200  $\mu\text{g ml}^{-1}$  BGG in PBS eine Kalibriergerade erstellt. Es wurden 20  $\mu\text{l}$  Leerprobe (PBS), Standard oder Probe (ggf. verdünnt in PBS) mit 300  $\mu\text{l}$  B gemischt, wobei jeweils eine Dreifachbestimmung erfolgte.

Bei der Bindung des in der Reaktionslösung enthaltenen Farbstoffs *Coomassie brilliant blue G-250* an Proteine verschiebt sich dessen Absorptionsmaximum [Bradford, 1976]. In einem Mikroplattenphotometer (Sunrise, Tecan) wurde die Extinktion bei  $\lambda = 595 \text{ nm}$  gegen PBS gemessen, die somit ein Maß für die Gesamtproteinkonzentration  $c_{\text{ProteK}}$  in der Lösung war. Diese konnte dann mit Hilfe einer Kalibriergeraden bestimmt werden.

### 11.3.3 SDS-PAGE zur Reinheitsbestimmung

Zur Elektrophorese wurden 4-12 % Polyacrylamid-Gele (Criterion XT Bis-Tris, Bio-Rad Laboratories) verwendet.

Lösungen:

- A Probenpuffer: XT Sample Buffer, 4x (Bio-Rad Laboratories)
- B Marker für SDS-PAGE: PageRuler Unstained Protein Ladder (Thermo Scientific)
- Marker für Western Blot: PageRuler Prestained Protein Ladder (Thermo Scientific)
- C Laufpuffer Stock: 10x Tris/Glycin/SDS Buffer (Bio-Rad Laboratories)
- D Laufpuffer: 10 % C
- E Coomassielösung: Coomassie R350 (PhastGel™ Blue R, GE Healthcare) in 170 ml H<sub>2</sub>O
- F Färbelösung: 200 ml Methanol, 30 ml Essigsäure (Eisessig) in 0.2  $\mu\text{m}$  filtrierter Lösung E
- G Entfärbelösung: 10 % Isopropanol, 10 % Essigsäure

Zur Bestimmung der Reinheit P<sub>P1K</sub> des Zielproteins im jeweiligen Teilsystem K wurde zunächst die Gesamtproteinkonzentration nach Bradford ermittelt und die Proben mit PBS auf eine Konzentration von 250  $\text{mg l}^{-1}$  verdünnt.

Die Probe wurde mit A versetzt und bei 70 °C für 10 min inkubiert. Nach Auftragen von 30 µl Probe bzw. 5 µl B erfolgte die Trennung bei 150 V für ca. 75 Minuten unter Verwendung von D. Das Gel wurde ca. 60 min in E gefärbt und mehrfach in G inkubiert bis der Hintergrund vollständig entfärbt war.

Die Digitalisierung erfolgte mit einem Fluor-S Multi-Imager (Bio-Rad Laboratories) über die Software Quantity One (Bio-Rad Laboratories). Die Auswertung der Proteinbanden wurde anhand der Software TotalLab Quant (TotalLab) durchgeführt.

## 12. Literaturverzeichnis

### 12.1 Literatur

- Adam, G., Stark, G., 2009. *Physikalische Chemie und Biophysik*, Springer-Verlag, Berlin.
- Aksu, B., De Beer, T., Folestad, S., Ketolainen, J., Lindén, H., Lopes, J.A., De Matas, M., Oostra, W., Rantanen, J., Weimer, M., 2012. Strategic funding priorities in the pharmaceutical sciences allied to Quality by Design (QbD) and Process Analytical Technology (PAT). *Eur. J. Pharm. Sci.*, 47, 402–405.
- Angarita, M., Müller-Späth, T., Baur, D., Lievrouw, R., Lissens, G., Morbidelli, M., 2015. Twin-column CaptureSMB: a novel cyclic process for protein A affinity chromatography. *J. Chromatogr. A*, 1389, 85–95.
- Anspach, F.B., Curbelo, D., Hartmann, R., Garke, G., Deckwer, W.D., 1999. Expanded-bed chromatography in primary protein purification. *J. Chromatogr. A*, 865, 129–144.
- Arteaga, F., Ferrer, A., 2002. Dealing with missing data in MSPC: Several methods, different interpretations, some examples. *J. Chemom.*, 16, 408–418.
- Au, R., 2014. The paradigm shift to an “open” model in drug development. *Appl. Transl. Genomics*, 3, 86–89.
- Billeter, E.P., 1972. Grundlagen der erforschenden Statistik; statistische Testtheorie. Springer-Verlag, Berlin.
- Bondi, R.W., Drennen, J.K., 2011. Quality by Design and the Importance of PAT in QbD. *Sep. Sci. Technol.*, 10, 195–224.
- Borchert, S.-O., Paul, J., Schuetzmeier, F., Voß, T., Luttmann, R., Cornelissen, G., 2013. Development of an integrated bioprocess for production of potential Malaria vaccines with *Pichia pastoris*. *Comput. Appl. Biotechnol.*, 12, 6–11.
- Brereton, R.G., 2003. Chemometrics: Data Analysis for the Laboratory and Chemical Plant. Wiley-VCH Verlag, New Jersey.
- Cereghino, G.P.L., Cereghino, J.L., Ilgen, C., Cregg, J.M., 2002. Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*. *Curr. Opin. Biotech.*, 13, 329–332.
- Cereghino, J.L., Cregg, J.M., 2000. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol. Rev.*, 24, 45–66.
- Chang, Y.K., Chase, H.A., 1996. Development of Operating Conditions for Protein Purification Using Expanded Bed Techniques : The Effect of the Degree of Bed Expansion on Adsorption Performance. *Biotechnol. Bioeng.*, 49, 512–526.
- Cornelissen, G., 2004. Integrierte Bioprozessentwicklung zur Herstellung pharmakologischer wirksamer Proteine mit *Pichia pastoris*. Dissertation, Universität Hannover.
- Cornelissen, G., Bertelsen, H.-P., Lenz, K., Hahn, B., Schultz, M., Scheffler, U., Werner, E., Leptien, H., Krüß, S., Jansen, A.-K., Elsholz, O., Gliem, T., Wilhelm, B.-U., Sowa, E., Radeke, H.H., Luttmann, R., 2003. Production of recombinant proteins with *Pichia pastoris* in integrated processing. *Eng. Life Sci.*, 3, 361–370.

- Cregg, J.M., 1999. Expression in the Methylotrophic Yeast *Pichia pastoris*, in *Gene Expression Systems*. Elsevier B.V., Amsterdam, pp. 157–191.
- Derringer, G., Suich, R., 1980. Simultaneous optimization of several response variables. *J. Qual. Technol.*, 12, 214–219.
- Ellert, A., 2015. Automatisierung und Optimierung von Herstellungsprozessen rekombinanter pharmazeutischer Proteine. Dissertation, Universität Hannover.
- Ellis, S.B., Brust, P.F., Koutz, P.J., Waters, A.F., Harpold, M.M., Gingeras, T.R., 1985. Isolation of alcohol oxidase and two other methanol regulatable genes from the yeast *Pichia pastoris*. *Mol. Cell. Biol.*, 5, 1111–1121.
- Eriksson, L., Johansson, E., Kettaneh-Wold, N., Trygg, J., Wikström, C., Wold, S., 2006. Multi- and Megavariate Data Analysis, Part I, second ed. MKS Umetrics AB, Umeå.
- Eriksson, L., Johansson, L., Kettaneh-Wold, N., Wikström, C., Wold, S., 2008. Design of Experiments: Principles and Applications, third ed. MKS Umetrics AB, Umeå.
- Faber, B.W., Remarque, E.J., Morgan, W.D., Kocken, C.H.M., Holder, A.A., Thomas, A.W., 2007. Malaria vaccine-related benefits of a single protein comprising *Plasmodium falciparum* apical membrane antigen 1 domains I and II fused to a modified form of the 19-kilodalton C-terminal fragment of merozoite surface protein. *Infect. Immun.*, 75, 5947–5955.
- Faber, B.W., Younis, S., Remarque, E.J., Rodriguez Garcia, R., Riasat, V., Walraven, V., van der Werff, N., van der Eijk, M., Cavanagh, D.R., Holder, A.A., Thomas, A.W., Kocken, C.H.M., 2013. Diversity covering AMA1-MSP1-19 fusion proteins as malaria vaccines. *Infect. Immun.*, 81, 1479–90.
- Faber, K., Kowalski, B.R., 1997. Propagation of measurement errors for the validation of predictions obtained by principal component regression and partial least squares. *J. Chemom.*, 11, 181–238.
- FDA, 2004a. Pharmaceutical cGMPs for the 21st Century — a risk-based approach. URL: <http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/Manufacturing/QuestionsandAnswersonCurrentGoodManufacturingPractices-cGMPforDrugs/UCM176374.pdf> (accessed 7.9.15).
- FDA, 2004b. Guidance for industry: PAT — a framework for innovative pharmaceutical development, manufacturing, and quality assurance. URL: <http://www.fda.gov/downloads/Drugs/Guidances/ucm070305.pdf> (accessed 7.9.15).
- Fonteyne, M., Vercruyse, J., De Leersnyder, F., Van Snick, B., Vervaet, C., Remon, J.P., De Beer, T., 2015. Process Analytical Technology for continuous manufacturing of solid-dosage forms. *Trends. Anal. Chem.*, 67, 159–166.
- Fricke, J., 2015. Optimierung eines Herstellungsprozesses artifizieller Malaria-Vakzine mittels Design of Experiments. Dissertation, Universität Hannover.
- Fricke, J., Pohlmann, K., Jonescheit, N.A., Ellert, A., Joksch, B., Luttmann, R., 2013. Designing a fully automated multi-bioreactor plant for fast DoE-optimization of pharmaceutical protein production. *Biotechnol. J.*, 8, 738–747.

- Fricke, J., Pohlmann, K., Tatge, F., Lang, R., Faber, B., Luttmann, R., 2011. A multi-bioreactor system for optimal production of malaria vaccines with *Pichia pastoris*. *Bio-technol. J.*, 6, 437–451.
- Grubbs, F.E., 1950. Sample Criteria for Testing Outlying Observations. *Ann. Math. Stat.*, 21, 27–58.
- Hotelling, H., 1931. The Generalization of Student's Ratio. *Ann. Math. Stat.*, 2, 360–378.
- Hyndman, R.J., Koehler, A.B., 2006. Another look at measures of forecast accuracy. *Int. J. Forecast.*, 22, 679–688.
- ICH, 2009. Pharmaceutical Development Q8, ICH Harmonised Tripartite Guideline. URL: [http://www.ich.org/fileadmin/Public\\_Web\\_Site/ICH\\_Products/Guidelines/Quality/Q8\\_R1/Step4/Q8\\_R2\\_Guideline.pdf](http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q8_R1/Step4/Q8_R2_Guideline.pdf) (accessed 7.9.15).
- ICH, 2005. Quality Risk Management Q9, ICH Harmonised Tripartite Guideline. URL: [http://www.ich.org/fileadmin/Public\\_Web\\_Site/ICH\\_Products/Guidelines/Quality/Q9/Step4/Q9\\_Guideline.pdf](http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q9/Step4/Q9_Guideline.pdf) (accessed 7.9.15).
- ICH, 2008. Pharmaceutical Quality Systems Q10, ICH Harmonised Tripartite Guideline. URL: [http://www.ich.org/fileadmin/Public\\_Web\\_Site/ICH\\_Products/Guidelines/Quality/Q10/Step4/Q10\\_Guideline.pdf](http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q10/Step4/Q10_Guideline.pdf) (accessed 7.9.15).
- Ishikawa, K., 1981. *What is total quality control : The Japanese way*. Prentice Hall, New York.
- Kaiser, C., 2010. Methoden zur Prozessüberwachung und Modellierung von Herstellungsprozessen rekombinanter Proteine mit *Escherichia coli*. Dissertation, Universität Hannover.
- Kessler, W., 2007. *Multivariate Datenanalyse für die Pharma, Bio- und Prozessanalytik*, first ed. WILEY-VCH Verlag, Weinheim.
- Konstantinov, K.B., Cooney, C.L., 2014. White Paper on Continuous Bioprocessing. URL: <https://iscmp.mit.edu/> (accessed 7.9.15).
- Kourti, T., Nomikos, P., MacGregor, J.F., 1995. Analysis, monitoring and fault diagnosis of batch processes using multiblock and multiway PLS. *J. Process Control*, 5, 277–284.
- Langer, E.S., Rader, R.A., 2014a. Introduction to Continuous Manufacturing : Technology Landscape and Trends, in *Continuous Bioprocessing – Current Practice & Future Potential*. URL: <http://www.continuous-bioprocessing.com/> (accessed 7.9.15).
- Langer, E.S., Rader, R.A., 2014b. Continuous Bioprocessing and Perfusion: Wider Adoption Coming as Bioprocessing Matures. *Bioprocess. J.*, 13, 43–49.
- Lögering, K., 2015. Automatisierung, Optimierung und Analyse Integrierter Produktionsprozesse zur Herstellung rekombinanter Proteine mit *Pichia pastoris*. Dissertation, Universität Hannover.
- Lögering, K., Müller, C., Fricke, J., Bertelsen, H.-P., Scheffler, U., Luttmann, R., 2010. Optimization of Recombinant Enzyme Production with *Pichia Pastoris* in an Integrated Industrial Scale-Down Production Plant. *Comput. Appl. Biotechnol.*, 11, 293–298.
- Lögering, K., Müller, C., Voss, J.-P., Wagenfuehrer, C., Zahn, D., Bertelsen, H.-P., Scheffler, U., Luttmann, R., 2011. An integrated scale-down plant for optimal recombinant enzyme production by *Pichia pastoris*. *Biotechnol. J.*, 6, 428–436.

- Luttmann, R., Borchert, S.-O., Müller, C., Lögering, K., Aupert, F., Weyand, S., Kober, C., Faber, B., Cornelissen, G., 2015. Sequential/parallel production of potential Malaria vaccines – A direct way from single batch to quasi-continuous integrated production. *J. Biotechnology*, 213, 83–96.
- Luttmann, R., Hukelmann, E., Ellert, A., Kazemi, A., Cornelissen, G., 2010. A Cybernetic Approach for Modelling of Complex Dynamic Expression in Recombinant Protein Production with *Pichia pastoris*. *Comput. Appl. Biotechnol.*, 383–388.
- MacGregor, J., Cinar, A., 2012. Monitoring, fault diagnosis, fault-tolerant control and optimization: Data driven methods. *Comput. Chem. Eng.*, 47, 111–120.
- Mahajan, E., George, A., Wolk, B., 2012. Improving affinity chromatography resin efficiency using semi-continuous chromatography. *J. Chromatogr. A*, 1227, 154–162.
- Martens, S., 2014. Analyse und Synthese von zyklischen Herstellungsprozessen artifizieller Diversity Covering (DiCo) Malariavakzinkandidaten. Dissertation, Universität Hannover.
- Martens, S., Borchert, S.-O., Faber, B.W., Cornelissen, G., Luttmann, R., 2011. Fully automated production of potential Malaria vaccines with *Pichia pastoris* in integrated processing. *Eng. Life Sci.*, 11, 429–435.
- McCready, C., 2013. Imputation by Regression – Imputation methods for predictive monitoring and model predictive control. MKS Umetrics AB, Umeå.
- Menezes, J.C., 2011. Process Analytical Technology in Bioprocess Development and Manufacturing, in *Comprehensive Biotechnology*, second ed. Elsevier B.V., Amsterdam.
- Menezes, J.C., Ferreira, A.P., Rodrigues, L.O., Brás, L.P., Alves, T.P., 2009. Chemometrics Role within the PAT Context: Examples from Primary Pharmaceutical Manufacturing, in *Comprehensive Chemometrics*. Elsevier B.V., Amsterdam, pp. 313–355.
- Mercier, S.M., Diepenbroek, B., Wijffels, R.H., Streefland, M., 2014. Multivariate PAT solutions for biopharmaceutical cultivation: Current progress and limitations. *Trends Biotechnol.*, 32, 329–336.
- Muñoz, S.G., Kourti, T., MacGregor, J.F., 2004. Multivariate forecasting of batch evolution for monitoring and fault detection, in *Dynamics and Control of Process Systems 2004*. Elsevier B.V., Amsterdam, pp. 71–76.
- Nomikos, P., Macgregor, J.F., 1995. Multivariate SPC Charts for Batch Monitoring Processes. *Technometrics*, 37, 41–59.
- Nomikos, P., Macgregor, J.F., 1994. Monitoring Batch Processes Using Multiway Principal Component Analysis. *AIChE J.*, 40, 1361–1375.
- Paul, J., Jensen, S., Dukart, A., Cornelissen, G., 2014. Optimization of a preparative multimodal ion exchange step for purification of a potential malaria vaccine. *J. Chromatogr. A*, 1366, 38–44.
- Peifer, T., Profos, P., 2008. *Handbuch der industriellen Messtechnik*. Oldenbourg Industrieverlag, München.
- Pestieau, A., Krier, F., Thoorens, G., Dupont, A., Chavez, P.F., Ziemons, E., Hubert, P., Evrard, B., 2014. Towards a real time release approach for manufacturing tablets using NIR spectroscopy. *J. Pharm. Biomed. Anal.*, 98, 60–67.

- Peuker, T., 2006. Modellgestützte Verfahren zur Beobachtung und Steuerung bioverfahrenstechnischer Prozesse. Dissertation, Universität Hannover.
- Potvin, G., Ahmad, A., Zhang, Z., 2012. Bioprocess engineering aspects of heterologous protein production in *Pichia pastoris*: A review. *Biochem. Eng. J.*, 64, 91–105.
- Powell, M.J.D., 2009. The BOBYQA algorithm for bound constrained optimization without derivatives. *Dep. Appl. Math. Theor. Physics*, Cambridge University.
- Rathore, A.S., 2009. Roadmap for implementation of quality by design (QbD) for biotechnology products. *Trends Biotechnol.*, 27, 546–553.
- Rathore, A.S., Winkle, H., 2009. Quality by design for biopharmaceuticals. *Nat. Biotechnol.*, 27, 26–34.
- Reifsnyder, D.H., McKnight, N.L., Kelley, B., 2011. Process Validation, in: *Comprehensive Biotechnology*. Elsevier B.V., Amsterdam, pp. 511–519.
- Reinfeldt, M., Tränkle, U., 1976. *Signifikanztabellen statistischer Testverteilungen. Binomial-, Poisson-, Normal-, x<sub>2</sub>-, t-, r-Tabellen, F-Tabellen für acht Signifikanzniveaus*. Oldenbourg Industrieverlag, München.
- Remarque, E.J., Faber, B.W., Kocken, C.H.M., Thomas, A.W., 2008. A diversity-covering approach to immunization with *Plasmodium falciparum* apical membrane antigen 1 induces broader allelic recognition and growth inhibition responses in rabbits. *Infect. Immun.*, 76, 2660–2670.
- S Clinical Trials Partnership, 2015. Efficacy and safety of RTS,S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: final results of a phase 3, individually randomised, controlled trial. *Lancet*, 6736.
- Schaefer, C., Clicq, D., Lecomte, C., Merschaert, A., Norrant, E., Fotiadu, F., 2014. A Process Analytical Technology (PAT) approach to control a new API manufacturing process: Development, validation and implementation. *Talanta*, 120, 114–125.
- Schneider, A., 2014. Knowledge Management in QbD Environment. *PAT & QbD Forum: Enabling Efficient Bioprocess Development and Robust Manufacturing 2014*.
- Schneider, G., 2007. Grundlagen der Bioinformatik. Universität Frankfurt am Main.
- Stratton, J., Chiruvolu, V., Meagher, M., 1998. High cell-density fermentation. *Methods Mol. Biol.*, 103, 107–20.
- Tolle, R., Bujard, H., Cooper, J.A., 1995. *Plasmodium falciparum*: variations within the C-terminal region of merozoite surface antigen-1. *Exp. Parasitol.*, 81, 47–54.
- Trautmann, H., 2011. Quality by Design – Paradigmenwechsel bei der Bioprozessentwicklung. URL: <https://www.gesundheitsindustrie-bw.de/de/fachbeitrag/aktuell/quality-by-design-paradigmenwechsel-bei-der-bioprozessentwicklung/> (accessed 7.9.15).
- Valero, F., 2013. Bioprocess Engineering of *Pichia pastoris*, an Exciting Host Eukaryotic Cell Expression System. *Protein Eng.-Technol. Appl.*, 3–32.
- Van Hecke, W., Kaur, G., De Wever, H., 2014. Advances in in-situ product recovery (ISPR) in whole cell biotechnology during the last decade. *Biotechnol. Adv.*, 32, 1245–1255.

- Voss, J.-P., Mittelheuser, N.E., Maiwald, M., Luttmann, R., 2014. Off-line Analysis and Online Monitoring of Biopharmaceutical Production Processes with Spectroscopic Methods and Multivariate Data Analysis. *Third European Conference on Process Analytics and Control Technology*.
- Wold, H., 1966. Estimation of Principal Components and Related Models by Iterative Least squares, in *Multivariate Analysis*, Academic Press, New York, 391–420.
- Wold, S., Kettaneh, N., Fridén, H., Holmberg, A., 1998. Modelling and diagnostics of batch processes and analogous kinetic experiments. *Chemom. Intell. Lab. Syst.*, 44, 331–340.
- Wold, S., Kettaneh-Wold, N., MacGregor, J.F., Dunn, K.G., 2009. Batch Process Modeling and MSPC, in *Comprehensive Chemometrics*. Elsevier B.V., Amsterdam, pp. 163–197.
- Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.*, 58, 109–130.
- Yoon, S., MacGregor, J.F., 2004. Principal-component analysis of multiscale data for process monitoring and fault diagnosis. *AICHE J.*, 50, 2891–2903.
- Zijlstra, G., 2013. New Approaches in Continuous BioManufacturing: Continuous XD® cell cultures (around 100 mln cells/mL) coupled to the Rhobust® EBA integrated clarification and purification technology. DSM Biologics, The Netherlands.  
URL: <http://www.engconf.org/staging/wp-content/uploads/2013/12/3.-Gerben-Zijlstra-02.pdf> (accessed 7.9.15).

## 12.2 Veröffentlichungen des Autors

Die nachfolgende Liste enthält die Veröffentlichungen des Autors, die während seines Studiums und der nachfolgenden experimentellen Forschung zur Doktorarbeit entstanden sind.

### Publikationen:

- [1] Ellert, A., Thiesing, D., Tatge, F., Borchert, S., Ganzlin, M., Luttmann, R.: Optimization of Fluorescent Fusion Protein Expression in *Escherichia coli* via Design of Experiments. *Computer Applications in Biotechnology*, Volume: 11, Part: 1, Pages: 287–292, DOI: 10.3182/20100707-3-BE-2012.0068, 2010.
- [2] Martens, S., Borchert, S.-O., Faber, B. W., Cornelissen, G., Luttmann, R.: Fully automated production of potential Malaria vaccines with *Pichia pastoris* in integrated processing. *Engineering in Life Sciences*, Volume 11, Pages: 429–435, 2011.
- [3] Borchert, S.-O.: Einsatz einer Expanded Bed Adsorptionsanlage am Bioreaktor zur direkten Abtrennung von Malariavakzinen aus Fermentationsbrühen. Final report for the Max-Buchner-Forschungsstiftung Kennziffer 2890, 2012, [http://www.dechema.de/dechema\\_media/AB\\_2890-p-4554.pdf](http://www.dechema.de/dechema_media/AB_2890-p-4554.pdf).
- [4] Borchert, S.-O., Paul, J., Schützmeier, F., Voß, T., Luttmann, R., Cornelissen G.: Development of an Integrated Bioprocess with Expanded Bed Adsorption Chromatography for production of potential Malaria vaccines with *Pichia pastoris*. *Computer Applications in Biotechnology*, Volume: 12, Part: 1, Pages: 6–11, DOI: 10.3182/20131216-3-IN-2044.00037, 2013.

- [5] Luttmann, R., Borchert, S.-O., Müller, C., Lögering, K., Aupert, F., Weyand, S., Kober, C., Faber, B., Cornelissen, G.: Sequential/Parallel Production of Potential Malaria Vaccines - a Direct Way from Single Batch to Quasi-Continuous Integrated Production. *J. Biotechnol.*, Volume 213, Pages: 83–96, 2015, <http://dx.doi.org/10.1016/j.biotec.2015.02.022>.
- [6] Borchert, S.-O., Voß, T., Schützmeier, F., Paul, J., Cornelissen G., Luttmann, R.: Development and monitoring of an integrated bioprocess for production of a potential malaria vaccine with *Pichia pastoris*. *Journal of Process Control*, Volume 35, Pages: 113–126, 2015, DOI: 10.1016/j.jprocont.2015.08.006

### Tagungsbeiträge:

- [7] Martens, S., Borchert, S.-O., Hartwich, P., Cornelissen, G., Anspach, B., Faber, B. W., Luttmann, R.: Production and Purification of Complex Malaria Vaccines from High Cell Density Cultivations with *Pichia pastoris*. Poster on the Pichia 2009 – Protein Expression Conference, Tucson, USA, 18. - 21. October 2009.
- [8] Borchert, S.-O., Luttmann, R., Cornelissen, G.: Herstellung rekombinanter Malariavakzin-Kandidaten – Überwachung und Reinigung mit IMAC. Presentation on the Chrom Forum, Berlin, Germany, 17. March 2010.
- [9] Borchert, S.-O., Martens, S., Paul, J., Voß, T., Faber, B., Luttmann, R., Cornelissen, G.: Optimierung einer Zellabtrennung und Reinigung von Malariavakzinen in einer Expanded Bed Adsorptions Chromatographie mit Design of Experiments. Poster on the 8. Hamburger Studententagung zur Medizin- und Biotechnologie, Hamburg, Germany, 04. May 2011.
- [10] Borchert, S.-O., Martens, S., Faber, B., Luttmann, R., Cornelissen, C.: Optimierung einer Zellabtrennung und Reinigung von Malariavakzinen in einer Integrierten Expanded Bed Adsorptions Chromatographie mit Design of Experiments. Presentation on the Dechema Vortrags- und Diskussionstagung: Bioverfahrenstechnik an Grenzflächen, Potsdam, Germany, 30. May - 01. June 2011.
- [11] Borchert, S.-O., Martens, S., Luttmann, R.: Aufbau und Optimierung eines Integrierten Gesamtprozesses zur Malariavakzin-Herstellung. Presentation on the Downstream Days Summer School 2011, Wetter an der Ruhr, Germany, 13. - 16. September 2011.
- [12] Borchert, S.-O., Paul, J., Martens, S., Faber, B.W., Luttmann, R., Cornelissen, G.: Optimization of an Integrated Bioprocess for production of potential Malaria vaccines with *Pichia pastoris* via Design of Experiments. Presentation on the Biotecnología Habana, Havanna, Cuba, 05. - 08. March 2012.
- [13] Borchert, S.-O.: Entwicklung eines Integrierten Bioprozesses mit einer Expanded Bed Adsorptionsanlage. Poster auf dem HAW-Forschungstag, Hamburg, Germany, 15. June 2012. Ausgezeichnet mit dem 3. Platz für die Posterpräsentation.
- [14] Borchert, S.-O., Martens, S., Paul, J., Faber, B., Luttmann, R., Cornelissen, G.: Optimization of combined downstream operations in an Integrated Bioprocess for potential Malaria vaccine production. Presentation on the ACHEMA 2012, Frankfurt am Main, Germany, 18. - 22. June 2012.

- [15] Borchert, S.-O.: Development of an Integrated Bioprocess with Expanded Bed Adsorption-Technology for the production of a potential Malaria vaccine with *Pichia pastoris*. Presentation on the 9th PhD Seminar on Chromatographic Separation Science, Weggis-Vierwaldstättersee, Switzerland, 24. - 26. February 2013.
- [16] Borchert, S.-O.: Development of quasi-continuous production of recombinant proteins. Presentation on the 12th Annual Biological Production Forum, Neuss, Germany, 23. - 25. September 2013.
- [17] Borchert, S.-O., Paul, J., Faber, B. W., Luttmann, R., Cornelissen, G.: Pseudo-continuous Production of Potential Malaria Vaccines by Integration of Bioreaction, Expanded Bed Adsorption and Fixed Bed Chromatography. Poster on the ECI – Integrated Continuous Bio-manufacturing Conference, Castelldefels, Spain, 20. - 24. October 2013. Awarded with the 2. price for poster presentation.
- [18] Borchert, S.-O., Paul, J., Schützmeier, F., Voß, T., Luttmann, R., Cornelissen G.: Development of an Integrated Bioprocess with Expanded Bed Adsorption Chromatography for production of potential Malaria vaccines with *Pichia pastoris*. Keynote-Lecture on the CAB 12 – Computer Applications in Biotechnology, Bombay, India, 16.-18. December 2013.
- [19] Borchert, S.-O., Kober, C., Weyand, S., Aupert, F., Luttmann, R.: Application of Multivariate Data Analysis in Production of Recombinant Pharmaceutical Proteins. Poster on the Sartorius PAT & QbD Forum 2014, Göttingen, Germany, 17. - 19. Februar 2014. Awarded with the 1. price for poster presentation.
- [20] Borchert, S-O., Paul, J., Faber, B., Luttmann, R. and Cornelissen, G.: Pseudo-continuous Production of Potential Malaria Vaccines by Integration of Bioreaction, Expanded Bed Adsorption and Fixed Bed Chromatography. Poster on the Pichia 2014 – Protein Expression Conference, San Diego, USA, 2. - 5. March 2014.



## **Wissenschaftlicher Werdegang**

### **Persönliche Angaben**

Name: Sven-Oliver Borchert  
Geburtstag: 26.07.1981  
Geburtsort: Göttingen

### **Ausbildung**

- 09/2011 – 12/2015 Promotion im Bereich Bioprozessautomatisierung  
HAW - Hochschule für Angewandte Wissenschaften Hamburg  
Forschungsschwerpunkt Bioprozess- und Analysentechnik  
in Kooperation mit dem Institut für Technische Chemie der  
Leibniz Universität Hannover  
mündliche Prüfung zum Dr. rer. nat. am 21. Dezember 2015
- 09/2009 - 08/2011 Aufbaustudium im Masterprogramm Biotechnology  
Hochschule für Angewandte Wissenschaften Hamburg  
Master Thesis im Forschungs- und Transferzentrum für Bioprozess-  
und Analysentechnik an der HAW Hamburg  
Abschluss: M.Sc., Note: sehr gut (1,4)
- 09/2003 - 08/2009 Abschluss des Diplomstudienganges Biotechnologie  
Studienschwerpunkt: Bioprozessautomatisierung  
Hochschule für Angewandte Wissenschaften Hamburg  
Diplomarbeit im Forschungs- und Transferzentrum für Bioprozess-  
und Analysentechnik an der HAW Hamburg  
Abschluss: Dipl.-Ing. (FH), Note: gut (2,0)
- 10/2008 - 03/2009 Praxissemester am Instituto de Ingeniería UNAM, Mexiko
- 10/2002 - 08/2003 Studiengang: Bioverfahrenstechnik  
Technische Universität Hamburg Harburg

### **Berufstätigkeit**

- Seit 10/2015 Entwicklungsingenieur bei der INVITE GmbH
- 09/2010 – 05/2015 Wissenschaftlicher Mitarbeiter im Forschungs- und Transferzentrum  
Bioprozess- und Analysentechnik an der HAW - Hochschule für  
Angewandte Wissenschaften Hamburg

### **Forschungsstipendium**

- 07/2011 – 07/2012 Max-Buchner-Forschungsstipendium der DECHEMA (Gesellschaft für  
Chemische Technik und Biotechnologie e.V.) mit der Förderkennziffer:  
2890



# Online-Buchshop für Ingenieure

■■■ VDI nachrichten

Online-Shops



Fachliteratur und mehr -  
jetzt bequem online recher-  
chieren & bestellen unter:  
[www.vdi-nachrichten.com/](http://www.vdi-nachrichten.com/)  
Der-Shop-im-Ueberblick



Täglich aktualisiert:  
Neuerscheinungen  
VDI-Schriftenreihen



## BUCHSHOP

Im Buchshop von vdi-nachrichten.com finden Ingenieure und Techniker ein speziell auf sie zugeschnittenes, umfassendes Literaturangebot.

Mit der komfortablen Schnellsuche werden Sie in den VDI-Schriftenreihen und im Verzeichnis lieferbarer Bücher unter 1.000.000 Titeln garantiert fündig.

Im Buchshop stehen für Sie bereit:

### VDI-Berichte und die Reihe **Kunststofftechnik**:

Berichte nationaler und internationaler technischer Fachtagungen der VDI-Fachgliederungen

### Fortschritt-Berichte VDI:

Dissertationen, Habilitationen und Forschungsberichte aus sämtlichen ingenieurwissenschaftlichen Fachrichtungen

### Newsletter „Neuerscheinungen“:

Kostenfreie Infos zu aktuellen Titeln der VDI-Schriftenreihen bequem per E-Mail

### Autoren-Service:

Umfassende Betreuung bei der Veröffentlichung Ihrer Arbeit in der Reihe Fortschritt-Berichte VDI

### Buch- und Medien-Service:

Beschaffung aller am Markt verfügbaren Zeitschriften, Zeitungen, Fortsetzungsreihen, Handbücher, Technische Regelwerke, elektronische Medien und vieles mehr – einzeln oder im Abo und mit weltweitem Lieferservice

VDI nachrichten

BUCHSHOP [www.vdi-nachrichten.com/Der-Shop-im-Ueberblick](http://www.vdi-nachrichten.com/Der-Shop-im-Ueberblick)

## Die Reihen der Fortschritt-Berichte VDI:

- 1 Konstruktionstechnik/Maschinenelemente
- 2 Fertigungstechnik
- 3 Verfahrenstechnik
- 4 Bauingenieurwesen
- 5 Grund- und Werkstoffe/Kunststoffe
- 6 Energietechnik
- 7 Strömungstechnik
- 8 Mess-, Steuerungs- und Regelungstechnik
- 9 Elektronik/Mikro- und Nanotechnik
- 10 Informatik/Kommunikation
- 11 Schwingungstechnik
- 12 Verkehrstechnik/Fahrzeugtechnik
- 13 Fördertechnik/Logistik
- 14 Landtechnik/Lebensmitteltechnik
- 15 Umwelttechnik
- 16 Technik und Wirtschaft
- 17 Biotechnik/Medizintechnik
- 18 Mechanik/Bruchmechanik
- 19 Wärmetechnik/Kältetechnik
- 20 Rechnerunterstützte Verfahren (CAD, CAM, CAE CAQ, CIM ...)
- 21 Elektrotechnik
- 22 Mensch-Maschine-Systeme
- 23 Technische Gebäudeausrüstung

ISBN 978-3-18-329217-2