

# Datenvorbehandlungsmethoden für präzisere Inline-Bestimmung von Stoffkonzentrationen im Bioreaktor aus NIR-Absorptions-Spektren

*Ulla Wenzel<sup>1</sup>, Sascha Princz, Linda Gröber, Rudolf Miller, Martin Heßling  
Hochschule Ulm, Institut für Medizintechnik und Mechatronik, Biotechnologie-Labor,  
Albert-Einstein-Allee 55, 89081 Ulm, Deutschland  
<sup>1</sup> wenzel@hs-ulm.de*

## Abstract

In Fermentationen mit Backhefe (*Saccharomyces cerevisiae*) sollen inline im Bioreaktor die Konzentrationen von Glucose, Ethanol und Biomasse aus NIR-Absorptionsspektren mithilfe chemometrischer PLS-Modelle bestimmt werden. Verwendet wird eine Transflexionssonde mit einer Spaltbreite von 0,25 mm. Die einzelnen aufgenommenen Absorptionsspektren gegen Wasser als Referenz schwanken teilweise stark sowohl im Offset, vermutlich verursacht durch Hefeklumpen, als auch in der Form, vermutlich wegen Gasblasen im Spalt der Sonde.

Die hier vorgestellte neuartige Datenvorbehandlungsmethode lässt über eine Zeitserie ungemittelter Spektren einen Filteralgorithmus laufen, der die korrumpierten Spektren verwirft und nur die restlichen „guten“ mittelt oder aber entscheidet, dass die gesamte Zeitserie nicht auswertbar ist. Diese Filterung bezieht sich erstens auf den Offset der Spektren und zweitens auf deren Form. Die Vorhersage der Glucose-, Ethanol- und Biomasse-Konzentrationen aus den PLS-Modellen kann dadurch noch deutlich verbessert werden.

**Keywords:** Datenvorbehandlung, Chemometrie, NIR-Absorptions-Spektroskopie, Inline-Messung im Bioreaktor, *Saccharomyces cerevisiae*

## Einleitung

In der Bioprozesstechnik ist eine kontinuierliche, schnelle, genaue und preiswerte Messtechnik zur Bestimmung von Stoffkonzentrationen in Echtzeit essentiell sowohl für die Qualitätssicherung als auch für eine Regelung des Bioprozesses. Alle diese Anforderungen erfüllt die optische Spektroskopie, kombiniert mit multivariater Datenanalyse (Chemometrie), wenn sie inline, also direkt im Prozess, erfolgt.

In der hier vorgestellten Arbeit wurden die Absorptionsspektren mit einer Transflexionssonde im NIR (Nahes Infrarot) im Wellenlängenbereich von 1100 bis 2100 nm aufgezeichnet. Die NIR-Strahlung wird auf ihrem Weg durch die Fermenterbrühe absorbiert, aber auch an Partikeln gestreut. Es wurden mehrere anaerobe Hefefermentationen mit *Saccharomyces cerevisiae* durchgeführt mit dem Ziel einer Vorhersage der Konzentrationen von Glucose, Ethanol und Biomasse aus chemometrischen PLS (Partial Least Squares Regression)-Kalibriermodellen.

Die Problematik bei Inline-Messung in einer Hefefermentation besteht in zweifacher Hinsicht: Erstens in Blasen, da bei der Umsetzung von Glucose in Ethanol durch die Hefezellen auch Kohlendioxid entsteht oder bei aeroben Fermentationen zusätzlich begast wird. Und zweitens in der Streuung durch die Hefezellen, speziell durch Hefeklumpen. Beim ersten Problem wurde bisher z.B. versucht, es mit einem Entgasungsmechanismus zu lösen [1], das zweite durch Filtrierung oder Zentrifugierung der Proben [2]. Beide Lösungsansätze stellen keine Inline-Messung mehr dar und benötigen zusätzliche Geräte. Daher wird hier ein neuartiger, rein datengestützter Ansatz vorgestellt.

## Material und Methoden

**Fermentation:** Es wurde die Backhefe „Omas Ur-Hefe“ (*Saccharomyces cerevisiae*-Stamm der Firma Fala GmbH, Bühl, Lesaffre-Gruppe, interne Bezeichnung L47) verwendet. Vier anaerobe Fermentationen im geregelten 7L-Bioreaktor der Hochschule Ulm wurden durchgeführt, bei denen Glucose zu Ethanol verstoffwechselt wurde. Die Temperatur der

Fermenterbrühe wurde auf 30°C geregelt und es wurde konstant mit 300 Umdrehungen pro Minute gerührt, um ein Absinken der Hefe zu vermeiden. Der Ansatz erfolgte mit einem Nährmedium mit 10 g/L Hefeextrakt und 20 g/L Pepton, beides von der Firma Merck KGaA, Darmstadt. Die Startbiomasse lag bei ca. 7,0...7,4 g/L Trockenmasse. Die eingewogene Startkonzentration von Glucose betrug zwischen 70 und 100 g/L. Am Ende der Fermentation, als die Glucose verbraucht war, wurde ein Spiking durchgeführt, d.h. eine nachträgliche Zugabe der drei interessierenden Substanzen: Bei den ersten beiden Fermentationen durch Zugaben von Ethanol und Glucose und bei den beiden letzten zusätzlich durch Zugaben von Hefe. Die hochkonzentrierte Hefelösung wurde ca. 30 min. vor Zugabe mit etwas Glucose angereichert.

NIR-Spektroskopie: Für die Aufnahme der Spektren im NIR (1100...2100 nm) wurde ein InGaAs-Dioden-Array-Spektrometer TIDAS S-1000 MS-T50/16 der Firma J&M Analytik GmbH in Essingen verwendet. Strahlungsquelle war eine luftgekühlte Halogenlampe AvaLight-Hal-S der Firma Avantes, Niederlande. Direkt im Bioreaktor, also inline im Prozess, befand sich eine faseroptische Transflexionssonde FDP-71R200-2-VAR der Firma Avantes. Bei der Messung nach dem Transflexionsprinzip tritt die einfallende Strahlung aus sechs im Kreis angeordneten Fasern aus, durchquert im optischen Spalt die Fermenterbrühe, wird an einer Fläche diffus reflektiert und durchquert ein zweites Mal die Brühe, bevor sie von einer mittigen Faser detektiert und zum Spektrometer geleitet wird (s. Abb. 1a)).

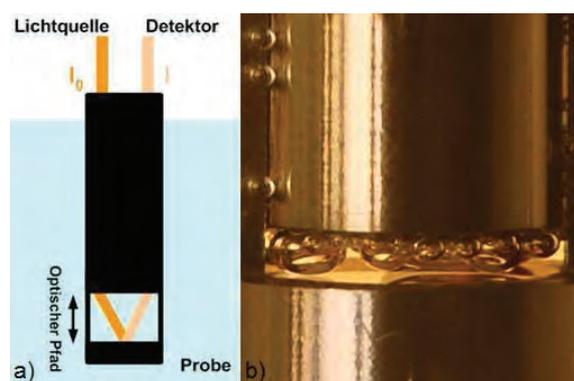


Abb. 1 a) Prinzip der Transflexionssonde, b) Blasen im Spalt der Sonde, aufgenommen in Bier

Der Spalt wurde auf 0,25 mm (Minimum) eingestellt, somit betrug der optische Weg durch die Brühe (ohne Streuung) 0,5 mm. Bei einem längeren Weg durch die wässrige Fermenterbrühe wäre aufgrund der starken Wasserabsorptionsbande um 1935 nm (s. Abb.

2) keine Strahlung dieser Wellenlängen mehr ins Spektrometer gelangt, außerdem wäre die Spaltbreite weniger reproduzierbar gewesen. Diese ist jedoch entscheidend für die Höhe der Absorption, d.h. ihre Reproduzierbarkeit wichtig für die Auswertbarkeit der Spektren. Auf diesem optischen Weg wird die Strahlung nicht nur durch Absorption der Substanzen in der Brühe abgeschwächt, sondern auch zu einem großen Teil an den Hefezellen gestreut. Da die Hefezellen mit einem Durchmesser von 3...10 µm bis zu einer Größenordnung über der Wellenlänge der betrachteten Strahlung liegen, handelt es sich bei der beobachteten Streuung um Mie-Streuung. Obwohl es sich nicht nur um Absorption handelt, wurden die Spektren als Absorptionsspektren gegen die Referenz deionisiertes Wasser derselben Temperatur wie die Proben (30°C) aufgenommen, weil die Brühe zum großen Teil aus Wasser besteht und dieses im verwendeten Wellenlängenbereich die dominante Absorption besitzt (Abb. 2). So kann für die Referenz dieselbe Integrationszeit wie für die Proben genommen werden. Auf die Übereinstimmung der Temperatur wurde geachtet, weil sich die Absorptionsbanden von Wasser im NIR sowohl von der Höhe als auch von der Wellenlänge her mit der Temperatur ändern.

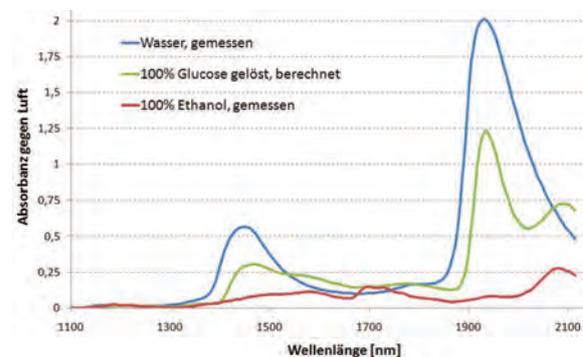


Abb. 2 NIR-Absorptionsspektren von Wasser, Glucose und Ethanol gegen Luft

Referenzanalytik: Für die Bestimmung der Glucose- und Ethanol-Konzentrationen der Proben wurden diese direkt nach der Entnahme aus dem Bioreaktor bei 10000 Umdrehungen pro Minute für 2 min. bei 4°C zentrifugiert. Anschließend wurde der Überstand zur späteren Auswertung eingefroren. Diese Auswertung wurde mit Enzym-UV-Tests der Firma R-Biopharm AG, Darmstadt, durchgeführt. Dazu mussten die Proben zuerst auf den Messbereich der Tests verdünnt werden. Die Biomasse wurde mit Hilfe der Trockenmassebestimmung mit Membranfiltration ermittelt, d.h. es wurde ein definiertes Volumen Fermenterbrühe entnommen und auf ein vorher gewogenes Filterpapier gegeben. Von einer Vakuumpumpe

wurde die Flüssigkeit nahezu vollständig abgesaugt. Diese Filter mit Hefe wurden zuerst bei Raumtemperatur und später im Ofen getrocknet, bis beim Wiegen keine weitere Abnahme messbar war.

Chemometrie / Datenvorbehandlung: Aus den Spektren wurden PLS-Modelle mit der Chemometrie-Software „The Unscrambler“ der Firma CAMO, Norwegen, erstellt. Vorher jedoch wurden verschiedene Datenvorbehandlungsmethoden angewandt, um die Vorhersagegenauigkeit der PLS-Modelle zu verbessern. Direkt vor Erstellung der Modelle wurden alle Spektren mit einer Savitzky-Golay-Glättung 2. Ordnung über neun Stützstellen geglättet. In diesem Zustand wurde der Bereich von 1100...1860 nm für das Biomasse-Modell herangezogen. Für die Ethanol- und Glucose-Modelle wurde zusätzlich ein Detrending 2. Ordnung des Bereichs von 1100...1900 nm durchgeführt, um die abfallende Kurve in den Spektren aufgrund der Hefestreuung herauszukorrigieren, die die Absorption der anderen Substanzen überlagert. Diese beiden Vorbehandlungen sind im „Unscrambler“ fertig enthalten. Noch vor diesen Methoden wurde eine neuartige Datenvorbehandlung angewandt, um den abträglichen Effekt von den zwei im Folgenden beschriebenen Problemen bei der Inline-Messung in Hefefermentationen zu verringern.

### Problematik bei Inline-Messung

Während der Messung der Spektren im sehr kleinen Spalt der Transflexionssonde werden zwei Probleme deutlich: Erstens wird bei der Fermentation von Glucose zu Ethanol Kohlendioxid gebildet, auch im Spalt, wo es sich zu Blasen formt, bis diese den Spalt verlassen. Dies wird exemplarisch in Abb. 1b) für Bier gezeigt, da die Hefebürhe zu trübe ist. Diese Blasen stören die Spektren grundlegend, da zum Teil statt der Fermenterbrühe die geringere Absorption der Blasen gemessen wird. Dadurch ändert sich die Form der Spektren. Zweitens löst sich die Backhefe nie vollständig in der Brühe, so dass auch kleine Hefebrocken den Spalt passieren und die Spektren durch höhere Streuung verfälschen. Dies resultiert in einem höheren Offset der abfallenden, durch die Streuung bedingten Kurve.

### Neuartige Datenvorbehandlungsmethode (5%-Quantil-Methode)

Normalerweise wird direkt bei der Aufnahme eine Anzahl von Spektren gemittelt, um den Einfluss des Rauschens zu verringern. Der neue Ansatz hingegen nimmt alle 5 min. eine

Zeitreihe von Spektren (z.B. über 3 min. ein Spektrum jede Sekunde) auf, die in einem nachgeschalteten Schritt gefiltert werden, d.h. auf Plausibilität geprüft und nur bei „Bestehen“ in die Mittelung aufgenommen werden. Diese Prüfung setzt sich aus zwei Kriterien zusammen, eins für die Blasen und eins für die Streuung durch Hefeklumpen.

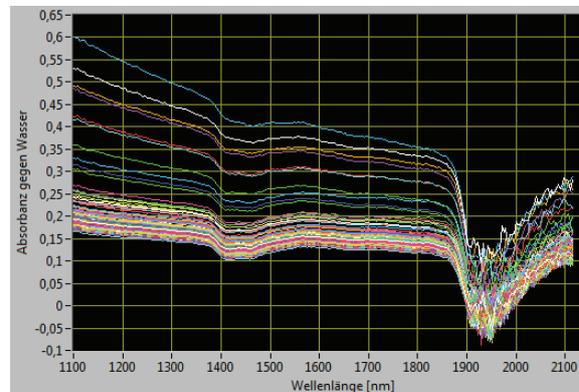


Abb. 3 Ungefilterte Zeitserie mit großer statistischer Streuung aus den ersten 12...15 min. einer Fermentation. Zu sehen ist die durch die Hefestreuung überlagerte abfallende Kurve. Ohne Hefezellen wäre der Bereich zwischen 1100...1300 nm ungefähr auf der Nulllinie, da dort kaum Absorption stattfindet (s. Abb. 1). Diese überlagerte Kurve schwankt stark in Offset und Krümmung, aber hat vor allem nach oben große Ausreißer, vermutlich aufgrund von Hefeklumpen, die den Spalt passieren. Im Normalfall sind die meisten Spektren am unteren Rand der vorkommenden Spektren angesiedelt.

1. Streu-Kriterium: Als entscheidender Wert wird hier der Offset der Spektren bei 1300 nm gewählt, weil dort wenig Absorption stattfindet. Da beobachtet werden kann, dass sich die meisten Spektren im Normalfall am unteren Rand der vorkommenden Spektren befinden, wird das 5%-Quantil gewählt, von dem die „guten“ Spektren nur um einen Schwellwert (0,03 AU) abweichen dürfen. Das Minimum (des Offsets bei 1300 nm) wird gemieden, weil immer Ausreißer nach unten denkbar sind, wenn sie auch hier nicht beobachtet werden können. Die Spektren am unteren Rand, also um das 5%-Quantil, fügen sich am besten in den Fermentationsverlauf ein, wenn man eine Zeitserie mit großer statistischer Streuung wie in Abb. 3 mit zeitlich benachbarten Zeitserien mit weniger Störungen vergleicht (s. Ausreißer im Verlauf in Abb. 6a)).

Dieser und die folgenden Schwellwerte und die Wahl des 5%-Quantils sind empirisch gefundene Werte, die sich in den durchgeführten Hefefermentationen bewährt haben.

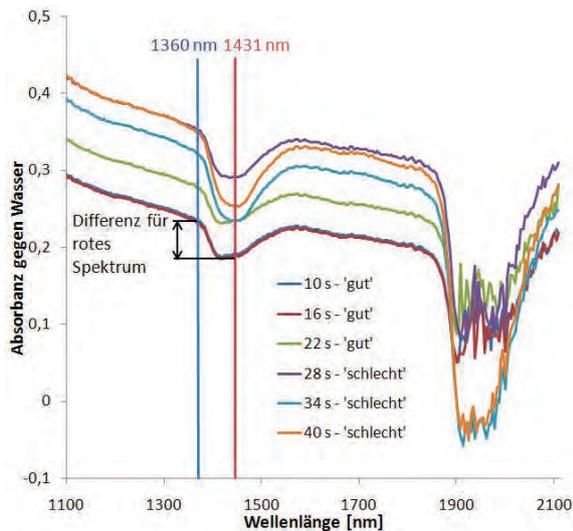


Abb. 4 Verschiedene Spektren einer Zeitserie ohne Blasen ("gut") und mit Blasen ("schlecht"). Als Kriterium für Blasen im Spektrum dient die Differenz der Absorbanz gegen Wasser bei 1360 und 1431 nm.

2. Blasen-Kriterium: Hier ist der entscheidende Wert die Differenz der Absorbanz gegen Referenz Wasser bei 1360 und 1431 nm (s. Abb. 4). Diese Differenz entspricht im Wasserspektrum gegen Luft (Abb. 2) der ersten Wasserabsorptionsbande, somit in den Spektren der Fermenterbrühe gegen Wasser der ersten negativen Bande, die durch das in der Brühe in geringerem Anteil vorhandene Wasser entsteht. Wasser gegen Wasser gemessen wäre die Nulllinie, die hier durch die Hefestreuung einer abfallenden Kurve entspricht. Daher zeigt ein höherer Wert dieser Differenz eine größere Abweichung vom Wasserspektrum an, d.h. dass statt der Absorption der wasserhaltigen Fermenterbrühe zum Teil die des Gases (Blasen) gemessen wurde. Da auch hier die geringsten Werte die „guten“ Spektren anzeigen, wird ein ähnliches Verfahren angewandt wie beim Offset: Es werden die Spektren als „gut“ beurteilt, deren oben definierte Differenz höchstens um einen bestimmten Schwellwert (0,003 AU) vom 5%-Quantil abweicht. Das Ergebnis der Filterung wird in Abb. 5 für eine Zeitserie mit vielen Blasen gezeigt.

Erst wenn beide Bedingungen erfüllt sind, wird das Spektrum zur Mittelwertbildung herangezogen. Diese Prozedur wird für jede Zeitserie durchgeführt.

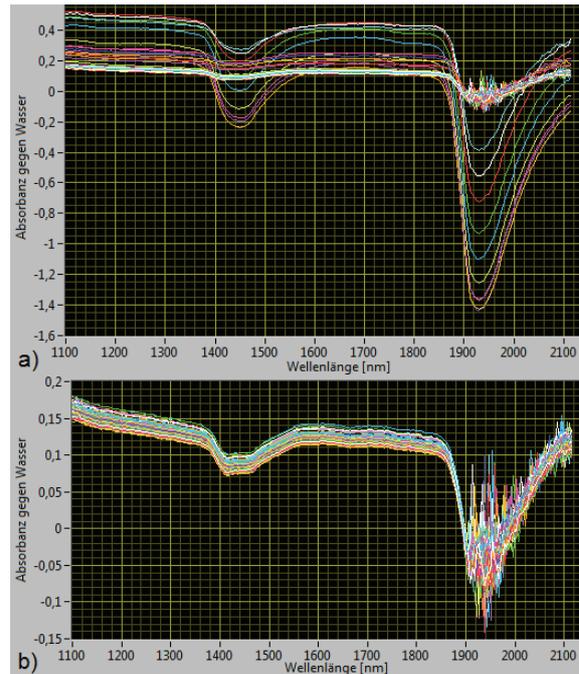


Abb. 5 Zeitserie von Spektren mit vielen Blasen a) ungefiltert, b) gefiltert; es wurden ca. 9% der Spektren als „schlecht“ beurteilt und aussortiert. Man beachte die unterschiedliche Skalierung der Absorbanz-Achse!

Um zu entscheiden, ob eine Zeitserie insgesamt auswertbar ist, wird ein selbst definiertes Maß verwendet, das an das robuste Streuungsmaß MAD (Median der absoluten Abweichungen vom Median, median absolute deviation) [3] angelehnt ist. Der MAD ist folgendermaßen definiert:

$$MAD = \text{med} |x_i - \text{med } x_i| \quad (1)$$

Das neue, an den MAD angelehnte Streuungsmaß nutzt statt des „inneren“ Median das 5%-Quantil, da wie beschrieben die hinsichtlich beider Kriterien untersten Spektren die „guten“ zu sein scheinen.

$$MADQ_{0,05} = \text{med} |x_i - Q_{0,05}(x_i)| \quad (2)$$

Dieser Median der absoluten Abweichungen um das 5%-Quantil wird für beide Kriterien (Offset bei 1300 nm und Differenz zwischen 1360 und 1431 nm) gebildet. Nur wenn der  $MADQ_{0,05}$  für den Offset kleiner als ein Schwellwert (0,04 AU) und der für die Differenz kleiner als ein anderer Schwellwert (0,004 AU) ist, wird die Zeitserie als auswertbar bezeichnet. Wenn eine nicht auswertbare Zeitserie zu einer Probenahme gehört, kann das Spektrum für die PLS-Modell-Erstellung aus den zeitlich benachbarten Zeitserien gemittelt werden.

Als einfachere Vergleichsmethode wird der Median der Absorbanz aller Spektren bei jeder

Wellenlänge zu einem „Medianspektrum“ zusammengesetzt.

Beide Methoden sind automatisiert in der grafischen Programmierumgebung LabVIEW der Firma National Instruments, USA.

### Ergebnisse und Diskussion

Durch die Filterung der Spektren jeder Zeitserie kann der Verlauf der Spektren über eine Fermentation eine plausible Form annehmen, was ohne die Filterung bzw. auch mit der Vergleichsmethode „Median“ nicht gelingt (s. Abb. 6a) und b)).

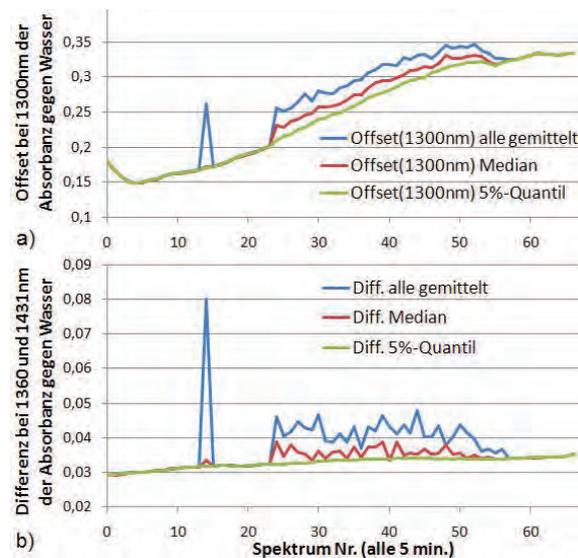


Abb. 6 Verlauf der Spektrenkriterien (a) Offset bei 1300 nm, b) Differenz bei 1360 und 1431 nm über eine Fermentation (ohne Spiking), bei der besonders starke statistische Streuung auftrat. Vergleich der neuen „5%-Quantil-Methode“ zu „Mittelung aller Spektren“ und Vergleichsmethode „Median aller Spektren“. Gut zu sehen ist der glatte Verlauf bei der 5%-Quantil-Methode.

Auch bei den Ergebnissen der PLS-Modelle nach den beschriebenen Datenvorbehandlungen kann eine deutliche Verbesserung beobachtet werden (s. Tab.1). Zur Validierung der Modelle wurde eine volle Kreuzvalidierung durchgeführt, d.h. eine Probe wird herausgenommen, das Modell mit allen

anderen erstellt und die fehlende Probe damit vorhergesagt. Der SEP (Standard Error of Prediction) der vollen Kreuzvalidierung als Maß für die Güte des Modells kann durch die 5%-Quantil-Methode gegenüber Mittelung aller Spektren bei Biomasse um 31% und bei Ethanol um 24% gesenkt werden. Nur bei Glucose ist lediglich eine kleine Verbesserung um 7% zu sehen. Bei Biomasse und Ethanol liegt die Vergleichsmethode „Median“ zwischen den beiden anderen Methoden, nur bei Glucose ist der SEP sogar schlechter als ohne Filterung. Auch beim Bestimmtheitsmaß  $R^2$  lässt sich ein ähnliches Ergebnis ablesen.

Es muss allerdings eine Einschränkung bei der Biomasse-Vorhersage gemacht werden: Die Proben, bei denen auch mit der gelösten Hefe ein Spiking durchgeführt wurde und nicht nur mit Ethanol und Glucose, passen nicht in das PLS-Modell hinein, genauso wie die jeweils erste Probe der Fermentation. D.h. es kann kein optimales Modell gefunden werden, z.B. schlägt „Unscrambler“ eine Anzahl von 9 PLS-Komponenten vor, was vermuten lässt, dass nur noch Rauschen in den Daten modelliert wird. Es wurde mehrfach beobachtet, dass die Spektren vom Offset her (z.B. bei 1300 nm) in den ersten 20...30 min. nach Inokulation zuerst absinken (s. Abb. 6a)), wohingegen der Offset im Rest der Fermentation mit der Biomasse-Konzentration ansteigt. D.h. die Hefezellen ändern ihre optischen Eigenschaften, vor allem ihr spektrales Streuverhalten in der sogenannten Lag-Phase, in der sie sich an die neuen Umgebungsbedingungen (Temperatur, Nährstoffangebot, pH-Wert usw.) anpassen und ihr Stoffwechsel anläuft. Vor dem Spiking mit Hefe wurde versucht, diese Lag-Phase durch Ansetzen der Hefe-Wasser-Lösung mit Glucose vorwegzunehmen, doch dieses ist offensichtlich nicht gelungen. Deshalb wurden einige Proben aus dem PLS-Modell, aber nur aus dem für Biomasse, herausgenommen: Die jeweils erste Probe und die, nachdem mit Hefe verfälscht wurde, sowie einige, bei denen von vornherein keine Biomasse-Referenz-Analyse durchgeführt wurde.

Tab. 1: Ergebnisse der vollen Kreuzvalidierung der PLS-Modelle, Vergleich der Datenvorbehandlungsmethoden

Analyt	Konzentrationen	Anzahl Proben	Anzahl PLS-Komponenten	5%-Quantil-Methode		Median-Methode		Alle Spektren gemittelt	
	[g/L]			SEP	$R^2$	SEP	$R^2$	SEP	$R^2$
Biomasse	7...14	57	4	0,3089	0,9750	0,3447	0,9689	0,4471	0,9477
Ethanol	0...80	86	5	2,168	0,9890	2,289	0,9877	2,871	0,9806
Glucose	0...167	86	5	3,412	0,9941	3,951	0,9921	3,663	0,9932

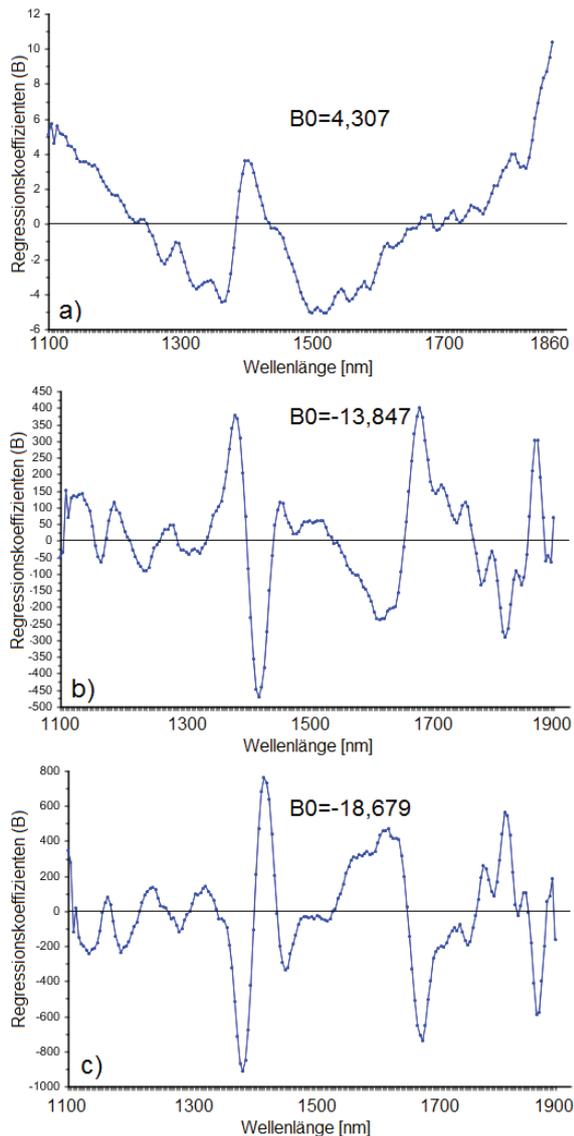


Abb. 7 Regressionskoeffizienten der PLS-Modelle für a) Biomasse, b) Ethanol, c) Glucose

An den in Abb. 7 gezeigten Regressionskoeffizienten der PLS-Modelle kann man sehen, dass der für Ethanol wichtige Absorptionspeak um 1690 nm und die Glucose-Bande von 1580...1650 nm jeweils hoch bewertet werden, d.h. die Modelle entsprechen insofern den chemisch-physikalischen Eigenschaften der Stoffe. Trotzdem fällt auf, dass die Regressionskoeffizienten dieser beiden Stoffe sich zueinander fast wie an der x-Achse gespiegelt verhalten, also relativ stark negativ korreliert zu sein scheinen. Diese negative Korrelation ist in der Fermentation zwar zu erwarten, da Glucose in Ethanol umgesetzt wird, aber es bestand die begründete Hoffnung, durch das durchgeführte umfangreiche Spiking die Korrelation zu „brechen“.

Es existieren keine reinen Biomasse-Absorptionsmessungen, mit denen die Regressionskoeffizienten verglichen werden könnten, da es schwierig ist, die Hefe im Zustand des laufenden Stoffwechsels, zu dem Glucose benötigt wird, aber ohne die Absorption eben dieser Glucose bzw. des produzierten Ethanols zu messen.

### Schlussfolgerungen

Unter den gegebenen schwierigen Bedingungen a) Inline-Messung direkt in der stark hefehaltigen Fermenterbrühe ohne vorherige Filterung oder Entgasung und b) mit Spiking durch Ethanol, Glucose und Biomasse werden in den PLS-Modellen mit nur 4 oder 5 PLS-Komponenten sehr gute Ergebnisse erzielt. Der SEP der vollen Kreuzvalidierung liegt zwischen 2,0 und 2,7% der maximalen Konzentration. Dazu trägt die neuartige Datenvorbereitungsmethode bei, die das 5%-Quantil der beiden Kriterien „Offset bei 1300 nm“ (Störungen durch Hefeklumpen) und „Differenz zwischen Absorbanz bei 1360 und 1431 nm“ (Störungen durch Blasen) nutzt, um verfälschte Spektren vor der Mittelwertbildung auszusortieren. Durch die Beurteilung der Auswertbarkeit der gesamten Zeitserie von Spektren anhand des neuen Streuungsmaßes „MADQ<sub>0,05</sub>“ können weiterhin Ausreißer in den gesamten Zeitserien eliminiert werden, bevor die chemometrischen Modelle erstellt werden.

### Danksagung

Diese Arbeit wurde im Rahmen des ZIM-Projektes „BIOsens“ finanziell gefördert vom Bundesministerium für Wirtschaft und Technologie (Förderkennzeichen KF 2186204CL1).

### Literatur

- [1] G. Mazarevica, J. Diewok, J. R. Baena, E. Rosenberg, B. Lendl, On-Line Fermentation Monitoring by Mid-infrared Spectroscopy, *Applied Spectroscopy*, Volume 58, Issue 7 (2004), 804-810; doi: 10.1366/0003702041389229
- [2] B. Liebmann, A. Friedl, K. Varmuza, Applicability of near-infrared spectroscopy for process monitoring in bioethanol production, *Biochemical Engineering Journal*, 52 (2010), 187-193; doi: 10.1016/j.bej.2010.08.006
- [3] F. R. Hampel, The Influence Curve and Its Role in Robust Estimation, *Journal of the American Statistical Association*, Volume 69, Issue 346 (1974), 383-393; doi: 10.1080/01621459.1974.10482962