

**Ultraschallgeschwindigkeitsmessungen
zum Bioprozessmonitoring**

**Von der Naturwissenschaftlichen Fakultät
der Universität Hannover**

zur Erlangung des Grades

DOKTOR DER NATURWISSENSCHAFTEN

- Dr. rer. nat. -

genehmigte Dissertation

von

Young-Lok Cha

Geboren am 3. 10. 1966 in Südkorea

2005

Referent : Prof. Dr. Bernd Hitzmann

Korreferent : Prof. Dr. Thomas Scheper

Tag der Promotion : Donnerstag, den 03. 11 2005

Hiermit versichere ich an Eides statt, die vorliegende Dissertation selbständig und nur unter Verwendung der angegebenen Hilfsmittel angefertigt zu haben. Ich versichere ferner, dass die Dissertation nicht schon als Diplomarbeit oder ähnliche Prüfungsarbeit verwendet worden ist.

Hannover, im November 2005

Young-Lok Cha

Danksagung

Herrn Prof. Dr. Bernd Hitzmann möchte ich zum Abschluß dieser Arbeit für das entgegengebrachte Vertrauen und für die geduldige fachliche Betreuung dieses interessanten Themas herzlichst danken.

Herrn Prof. Dr. Thomas Scheper danke ich für die Übernahme des Korreferats.

Der Arbeitsgemeinschaft industrieller Forschungsvereinigungen e.V. (AiF) danke ich für die finanzielle Förderung im Rahmen des Projekts mit dem Thema: *Online-Beobachter für biotechnische Prozesse*.

Mein Dank gilt der IFAK (Institut für Automation und Kommunikation) für die Zurverfügungstellung des Ultraschallmesssystems. Herrn Christop Daur danke ich für seine Unterstützung bei der Installation des Messsystems.

Herrn Dr. Egbert Stärk danke ich besonders für die Zusammenarbeit bei den Kultivierungen.

Den Mitarbeitern des Institutes für Technische Chemie in Hannover danke ich für die angenehme Atmosphäre, die anregenden Diskussionen und die gute Zusammenarbeit. Mein besonderer Dank geht an den Arbeitskreis Hitzmann für die Unterstützung und das hervorragende Arbeitsklima auch in schwierigen Zeiten.

Ein besonders herzlicher Dank geht an Frau Dipl.-Chem. Christine Klockow, Herrn Dipl.-Chem. David Geissler, Herrn Dipl.-Chem. Patrick Lindner, Frau M.Sc. Life-Science Cornelia Repenning und Dr. Michael Arndt für das gründliche Korrekturlesen dieser Arbeit.

Meine Familie danke ich für ihre Geduld und liebevolle Unterstützung.

Ultraschallgeschwindigkeitsmessungen zum Bioprozessmonitoring

In dieser Arbeit werden Messungen der Ultraschallgeschwindigkeit und deren Auswertungen zur Überwachung von *Saccharomyces cerevisiae* Kultivierungen vorgestellt. Die Ultraschallgeschwindigkeit ist abhängig von wesentlichen Prozessgrößen und kann daher zur Validierung des Prozesses genutzt werden. Das Ultraschallmesssystem wurde in einem Bypass-System im Stop-Flow-Verfahren online am Bioprozess betrieben, um den Einfluss von Blasen auf das Signal zu minimieren.

Die Kultivierungen sind durch unterschiedliche Prozessphasen, wie Glucose-, Ethanol- und Acetatabbau gekennzeichnet. Mittels der Ultraschallgeschwindigkeitsmessungen können die Prozessphasen identifiziert und Prozessgrößen online vorhergesagt werden.

Die Ultraschallgeschwindigkeit ist abhängig von der Biomasse-, Glucose- und Ethanolkonzentration. Als Modelle dienten für die Beschreibung und Simulation des Bioprozesses typische Monod-Modelle. Die Standardabweichung der simulierten Werte der Modellbildung betragen für Biomasse-, Glucose- und Ethanolkonzentration jeweils 0,5 g/L und 0,1 g/L sowie 0,5 g/L. Die simulierten Werte und die Messungen der Ultraschallgeschwindigkeit wurden für die Auswertung genutzt und daraus Prozessgrößen vorhergesagt. Die Vorhersage der verschiedenen Prozessgrößen erfolgte mit Hilfe sogenannter „Software-Sensoren“, die auf der multilinenen Regression (MLR), der Hauptkomponentenanalyse (PCA), des Kalman-Filters und neuronaler Netze beruhen. Dazu wurde die Software-Toolbox Unscrambler 7.6 für die PCA und SNNS 4.2 für das neuronale Netz sowie ein mit Delphi 5 geschriebenes Programm für das Kalman-Filter verwendet.

Ein Vergleich der Vorhersage mit den Messwerten zeigt eine gute Übereinstimmung. Die berechneten mittleren Fehler der Vorhersage ergaben für die Biomasse- und Glucosekonzentration jeweils 0,6 g/L und 1,7 g/L (ermittelt mit neuronalen Netzen) und für die Ethanolkonzentration 1,6 g/L (ermittelt mit MLR aus simulierten Werten) und waren niedriger als die Fehler, die mit den anderen „Software-Sensoren“ erzielt wurden.

Schlagworte: Ultraschallgeschwindigkeit, Bioprozessbeobachtung, Software-Sensor

Ultrasonic velocity measurements for bioprocess monitoring

Measurements and analysis of ultrasonic velocity for the monitoring and supervision of *Saccharomyces cerevisiae* cultivations are presented in this work. The dependence of the ultrasonic velocity on essential process variables is used to predict these for an optimal process validation. The ultrasonic measurements are performed in a by-pass with a stop-flow system to minimize the influence of bubbles on the signal.

Characteristically, yeast cultivations show different process phases like glucose, ethanol and acetate consumption which can be identified via ultrasonic velocity measurements. In addition, a prediction was carried out with measured and simulated data and thus the process variables were calculated.

The ultrasonic velocity depends on the concentration of biomass, glucose and ethanol. A typical Monod-Model was applied for the description and simulation to extend the databases for the evaluation. The standard deviation between measured and simulated process variables was 0.5 g/L for the biomass and ethanol concentration and 0.1 g/L for the glucose concentration. The simulated values and the measurements of ultrasonic velocity were used for the evaluation as well as the prediction of the process variables. The prediction of different process variables were performed by so-called software sensors, which are based on multilinear regression (MLR), principal component analysis (PCA), Kalman filter used as an observer and neural networks. For this purpose the software toolbox Unscrambler 7.6 for PCA and SNNS 4.2 for neural networks as well as a self-written program with Delphi 5 for the Kalman filter were used.

A comparison of the prediction with the measurements shows a good agreement. Mean error of prediction by the neural network is 0.6 g/L for biomass and 1.7 g/L for glucose. These errors calculated from simulated values by the MLR is 1.6 g/L for ethanol. These errors of the validation obtained was smaller compared to the values obtained by the other software sensors.

Key words: ultrasonic velocity, bioprocess monitoring, software sensor

Inhaltsverzeichnis

1.	Einleitung und Zielsetzung.....	1
2.	Grundlagen und Theorie.....	2
2.1.	Messverfahren des Ultraschallsensorsystems	2
2.2.	Hauptkomponentenanalyse	4
2.3.	Kontinuierlich-diskrete erweiterte Kalman-Filter	8
2.4.	Backpropagation-Netzwerk.....	10
3.	Material und Methoden	15
3.1.	Ultraschallmessgerät US 2100	15
3.2.	Messanordnung	16
3.3.	Verwendete Mikroorganismen.....	18
3.4.	Online-Analytik.....	19
3.4.1.	CO ₂ -, O ₂ -Abgas Messungen.....	19
3.4.2.	pO ₂ (Gelöstsauerstoff) Messungen.....	19
3.4.3.	Fluoreszenzspektroskopie	19
3.5.	Offline-Analytik	20
3.5.1.	Biotrockenmassemessung	20
3.5.2.	Glucosemessung.....	20
3.5.3.	Ethanolmessung	20
4.	Anwendung des Ultraschallsensorsystems.....	22
4.1.	Kalibrierung der Ultraschallsonde	22
4.2.	Ultraschallgeschwindigkeit im Stop-Flow-Verfahren	24
4.3.	Ultraschallgeschwindigkeit in Abhängigkeit einzelner Bioprozessgrößen.....	25
4.3.1.	Ultraschallgeschwindigkeit in Abhängigkeit der Glucosekonzentration	25
4.3.2.	Ultraschallgeschwindigkeit in Abhängigkeit der Ethanolkonzentration.....	26
4.3.3.	Ultraschallgeschwindigkeit in Abhängigkeit der Biomassekonzentration.....	27
4.4.	Typische Ultraschallmesssignale bei der Hefekultivierung	29
4.4.1.	Batch-Kultivierung.....	29
4.4.2.	Fed-Batch Kultivierung.....	30
5.	Multilineare Auswertung der Ultraschallgeschwindigkeit.....	32
6.	Prozessmodellierung der Hefekultivierung.....	37

6.1.	Mathematische Modellbildung.....	37
6.2.	Berechnung und Validierung von Modellparametern.....	38
7.	Vorhersage der Prozessgrößen mit diversen Software-Sensoren.....	42
7.1.	Vorhersage der Prozessgrößen mittels multilinearer Regression.....	43
7.2.	Hauptkomponentenregression zur Vorhersage der Prozessgrößen.....	47
7.2.1.	Vorhersage der Ultraschallgeschwindigkeit bei der Fed-Batch-Kultivierung.....	47
7.2.2.	Vorhersage der Bioprozessgrößen mit einer Hauptkomponentenregression.....	48
7.2.3.	Vorhersage der Bioprozessgrößen mit erweiterten Datensätzen.....	51
7.3.	Kalman-Filter zur Vorhersage der Bioprozessgrößen.....	52
7.4.	Neuronale Netze zur Vorhersage der Bioprozessgrößen.....	55
7.4.1.	Datenvorverarbeitung.....	55
7.4.2.	Vorhersage von Bioprozessgrößen mit dem „3-2-1“ Netz.....	57
7.4.3.	Vorhersage von Bioprozessgrößen mit dem „5-3-1“ Netz.....	59
7.4.4.	Vorhersage von Bioprozessgrößen mit dem „8-5-3“ Netz.....	60
7.5.	Diskussion.....	61
8.	Zusammenfassung.....	65
9.	Literaturverzeichnis.....	67
10.	Anhang.....	73
10.1.	Abkürzungs- und Symbolverzeichnis.....	73
10.2.	Blockdiagramm einer PCR mit Testset-Validierung.....	75
10.3.	Schatzmann-Medium.....	76
10.4.	Die trainierten Netze mit „3-2-1“ Struktur.....	76
10.5.	Die trainierten Netze mit „5-3-1“ Struktur.....	78
10.6.	Die trainierten Netze mit „8-5-3“ Struktur.....	81

1. Einleitung und Zielsetzung

Die Biotechnologie nimmt einen wichtigen Platz in der chemischen und pharmazeutischen Industrie sowie in der Lebensmittelindustrie ein. Für die optimale Prozessführung von biotechnischen Prozessen ist die zuverlässige Messung und Regelung möglichst vieler physikalischer und chemischer Größen erforderlich. Ein Problem der Bioprozesstechnik ist jedoch nach wie vor, dass wesentliche Prozessgrößen wie Substrat- und Metabolitkonzentrationen nicht direkt online gemessen werden können. Es besteht daher weiterhin ein Bedarf an Messsystemen zur Online-Charakterisierung von biotechnischen Prozessen, mit denen sich quantitative Informationen über spezielle Prozessgrößen oder relevante qualitative Informationen über den Prozesszustand erfassen lassen.

Neben den etablierten Online-Messungen wie z.B. pO_2 , pH, O_2 und CO_2 im Abgas stellen Ultraschallmessungen eine weitere Möglichkeit dar, Online-Daten über einen Bioprozess zu sammeln und damit die Online-Überwachung eines biotechnischen Prozesses zu verbessern. Auf der Basis der Ultraschallsensoren können die akustischen Eigenschaften eines komplexen biologischen Mediums analysiert werden. Durch die simultane und kontinuierliche Bestimmung der Schallgeschwindigkeit werden diverse Bioprozessgrößen vorhergesagt bzw. bestimmt.

Für eine effektive Bioprozessführung werden heutzutage nicht nur Hardware-Sensoren wie z.B. Biosensoren sondern auch Software-Sensoren zur Berechnung der nicht messbaren Prozessgrößen entwickelt. Mit Hilfe der Hauptkomponentenanalyse, dem erweiterten Kalman-Filter oder neuronaler Netze können Messdaten eines Bioprozesses zur Vorhersage der wichtigen Prozessgrößen genutzt werden, um einen Prozess online zu beobachten.

Ziel dieser Arbeit ist es, mit Hilfe von mathematischen Modellen und Ultraschallmessungen wesentliche Aussagen über den aktuellen Zustand eines biotechnischen Prozesses zu treffen. Die bisherigen Untersuchungen zeigten, dass die mit dem Ultraschallmesssystem erhaltenen Messsignale unspezifisch von einer Reihe von Bioprozessgrößen abhängen. Es soll ermittelt werden, ob die Ultraschallmessungen mit den gemessenen Größen korrelieren und ob theoretische und/oder chemometrische Modelle in der Lage sind, aus den Messwerten der Ultraschallgeschwindigkeit wesentliche Prozessgrößen zu berechnen. Da basierend auf Ultraschallmessungen später eine Prozessführung realisiert werden soll, werden die Untersuchungen an einem komplexen Bioprozess durchgeführt.

2. Grundlagen und Theorie

2.1. Messverfahren des Ultraschallsensorsystems

Die Geschwindigkeit von Ultraschallwellen in Flüssigkeiten hängt von der Dichte, der Temperatur, dem Druck und der Art der Flüssigkeit selber ab. Zur Messung der Abhängigkeiten des Ultraschallsystems kann ein modifiziertes Impulsverfahren verwendet werden. Dieses Verfahren nutzt die Abhängigkeit des Schallwechseldruckes von den stoffspezifischen Größen Schallgeschwindigkeit und Absorption aus. Die in dieser Arbeit für die Online-Überwachung des Bioprozesses gewählte akustische Stoffgröße ist die Schallgeschwindigkeit. Die Schallgeschwindigkeitsmessung wird bei konstantem Wandlerabstand durch eine Laufzeitmessung realisiert. Die Bestimmung der Absorption erfolgt durch einen Amplitudenvergleich zweier aufeinanderfolgender Empfangs- oder Echosignale in Abhängigkeit vom zwischenzeitlich zurückgelegten Weg des Ultraschallsignals. Als optimale Sendesignalgestaltung erwies sich die Ansteuerung der piezoelektrischen Schwinger mit Burstsignalen harmonischer Schwingungen. Die Variation der Anzahl der harmonischen Schwingungen je Burst kann dabei zu einer weiteren Verbesserung der Empfangssignalform beitragen und die Empfangssignalauswertung vereinfachen.

Mit Hilfe eines Ultraschallsenders wird in die zu untersuchende Flüssigkeit ein akustisches Signal eingestrahlt, das nach dem Passieren einer bestimmten Laufstrecke empfangen und ausgewertet wird. Dabei nutzt man die Tatsache, dass die Ausbreitung einer Ultraschallwelle in einer Flüssigkeit durch die Zustandsfunktion des Schallwechseldruckes beschrieben ist:

$$P(S,t) = P_0 e^{-j\omega(t-\frac{S}{C})} e^{-\alpha S} \quad (2-1)$$

Der Schallwechseldruck wird durch die Kreisfrequenz $\omega = 2\pi f$ und die Schallfrequenz f , die Zeit t , den Schallwechseldruck P_0 am Ort $S = 0$ und zur Zeit $t=0$, die Geschwindigkeit C und den Absorptionskoeffizient α , in Abhängigkeit von den stofflichen Eigenschaften bzw. der Zusammensetzung des durchstrahlten Mediums, beeinflusst [Bergman, 1954]. Der Absorptionskoeffizient $\alpha[\text{m}^{-1}]$ kann mit Hilfe der Empfangs- und Echoamplituden berechnet werden:

$$\alpha = \frac{1}{S} \ln \frac{A}{B} \quad (2-2)$$

Dabei sind A die Empfangsamplitude, B die Echoamplitude und S der Schallweg. Der Koeffizient α beschreibt die Dämpfung der Amplitude mit dem Abstand S vom Sender.

Basierend auf der Laufzeitmethode wird die Schallgeschwindigkeit C durch Quotientenbildung aus der direkt gemessenen Schalllaufzeit bei konstantem Wandlerabstand berechnet. Die erhaltene Schallgeschwindigkeit wird simultan bestimmt und gibt jederzeit Informationen zur Zusammensetzung der komplexen Flüssigkeit. Die Schallgeschwindigkeit C erlaubt Aussagen zur Änderung der spezifischen Dichte ρ oder der adiabatischen Kompressibilität β_{ad} und zeigt näherungsweise eine einfache Abhängigkeit vom Quotienten des Kompressionsmoduls K und der spezifischen Dichte in Gleichung (2-3).

$$C^2 = \frac{1}{\beta_{ad} \rho} = \frac{K}{\rho} \quad (2-3)$$

Eine weitere akustische Kenngröße ist die Schallkennimpedanz Z , die durch das Produkt aus der Schallgeschwindigkeit C und Dichte ρ berechnet wird. Es gilt nach Gleichung (2-4):

$$Z = C \cdot \rho \quad (2-4)$$

Die Schallkennimpedanz charakterisiert den Widerstand, den ein Stoff dem Schall entgegenbringt. Hieraus folgt, dass die Schallgeschwindigkeit nicht direkt aus der Dichte abgeleitet werden kann, wenn die Kompressibilität der Flüssigkeit nicht bekannt bzw. nicht konstant ist (was meist der Fall ist).

Abbildung 2-1 zeigt eine schematische Darstellung der Ultraschalmessung. Das eingesetzte Messprinzip basiert auf einem modifizierten Laufzeitverfahren. Die Ultraschallsignale werden von der Sendeeinheit erzeugt und einem Empfänger zugeführt.

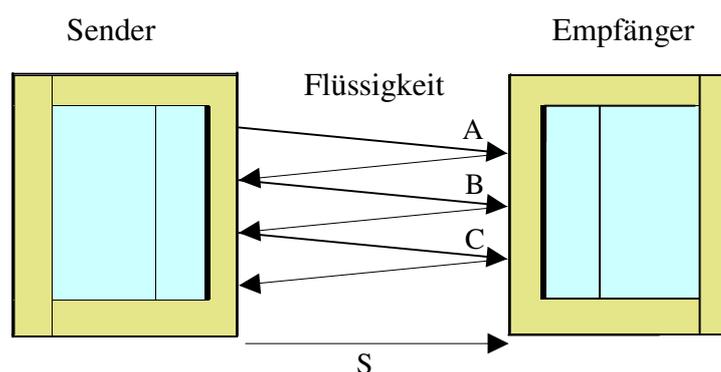


Abbildung 2-1: Schematische Darstellung der Ultraschalmessung (S ist der Abstand zwischen Sender und Empfänger, A direkte Empfangssignal, B und C sind an Empfänger und Sender reflektierte Signale)

Der Ultraschall wird in die zu untersuchende Kulturbrühe eingestrahlt und sowohl das direkte Signal (A in Abbildung 2-1) am Empfänger als auch das reflektierte Signal, das am Empfänger und Sender jeweils reflektiert wird (B in Abbildung 2-1), gemessen.

Das Sendesignal sowie die Empfangssignale der Ultraschallwellen ist in Abbildung 2-2 dargestellt. A ist die erste detektierte Amplitude und B das zweite Empfangssignal, das von dem Empfänger und Sender reflektiert wird, C und D das dritte bzw. vierte Echosignal von dem anfangs gesendeten Signal. Die Amplitude der reflektierten Empfangssignale sinkt exponentiell ab. t_1 und t_1+t_2 sind die Zeiten zu dem A bzw. B registriert werden.

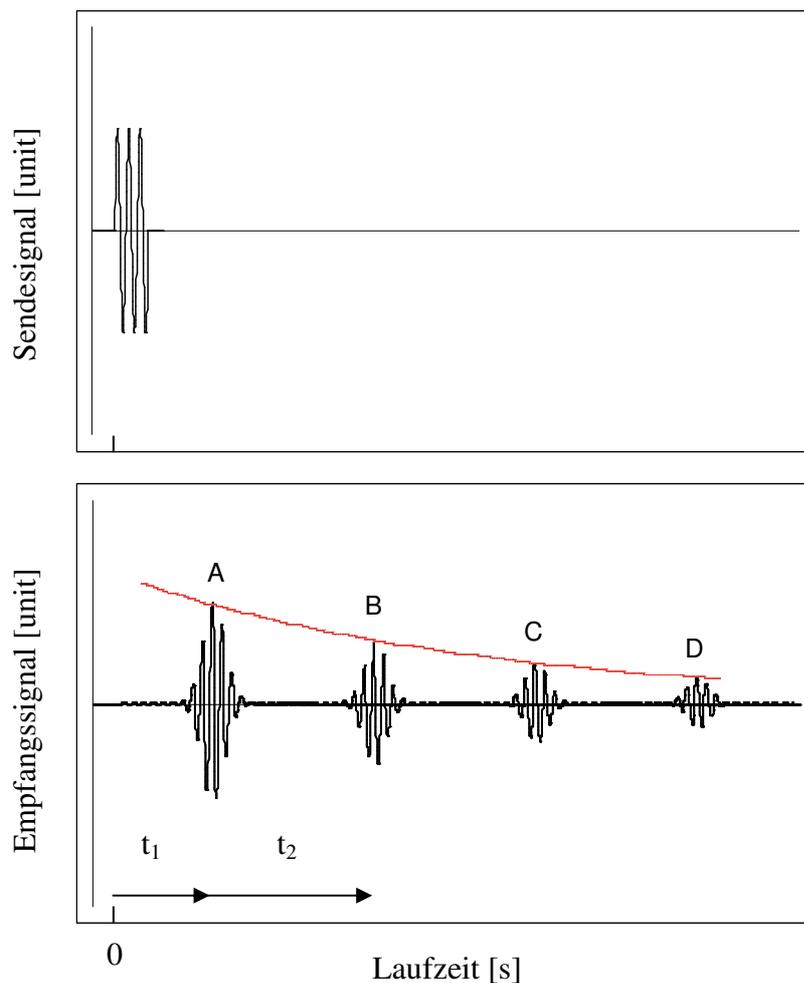


Abbildung 2-2: Sende- und Empfangssignal der Ultraschallwellen

2.2. Hauptkomponentenanalyse

Für die multivariate Auswertung von Messdaten wurde die Hauptkomponentenanalyse (PCA) entwickelt [Geladi, Kowalski, 1986]. Im Bereich der Bioprozessanalyse wurde die PCA als Auswertungsverfahren zum Beispiel für die Fluoreszenz-Messdaten angewandt [Hitzmann et

al., 1998]. Liegen Rohmesswerte z.B. aus einer Kultivierung vor, so können aus ihnen mit Hilfe der multilinearen Regression (MLR) Bioprozessgrößen berechnet werden. Die folgende Gleichung kann hierfür verwendet werden:

$$y_i = \sum_j b_j \cdot x_{ij} + b_0 \quad (2-5)$$

Dabei sind y_i die jeweiligen Bioprozessvariablen, x_{ij} die Messdaten. b_j ist der Koeffizientenvektor. Mit Hilfe der Gleichung (2-5) kann der Koeffizient b_0 berechnet werden:

$$b_0 = \bar{y} - \sum_j b_j \bar{x}_j \quad (2-6)$$

Dabei ist \bar{y} der Mittelwert von den Bioprozessvariablen y_i , \bar{x}_j sind die Mittelwerte von den x_{ij} Messdaten. Durch die Zentrierung der Messdaten kann man den Koeffizienten b_0 eliminieren.

Der Koeffizientenvektor der Steigung b der einzelnen Variablen x_{ij} wird für die zentrierten Daten wie folgt berechnet [Geladi, Kowalski, 1986]:

$$b = (X' \cdot X)^{-1} \cdot X' \cdot y \quad (2-7)$$

Dabei ist X' die transponierte Matrix von X . Bei der Durchführung der multilinearen Regression liegt ein Problem vor, wenn die Matrix von X kollinear ist. Dann kann die generalisierte Inverse der Datenmatrix nicht gebildet werden.

Die lineare Abhängigkeit der X Variablen wird durch eine Hauptkomponententransformation eliminiert. Für die Hauptkomponentenanalyse werden zuerst die Daten zentriert und dann gegebenenfalls skaliert. Für die Zentrierung werden folgende Gleichungen verwendet:

$$x_{ij}^{neu} = x_{ij} - \bar{x}_j \quad (2-8)$$

$$\text{mit } \bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij} \quad (2-9)$$

Dabei ist i ($i=1 \dots n$) der Index der Proben, n die Anzahl der Proben und x_{ij}^{neu} sind die zentrierten Werte. Durch die Skalierung wird die Varianz der Variablen x_{ij} transformiert. Für eine Variable j berechnet sich die autoskalierte Variable x_{ij}^s mit folgender Gleichung [Otto, 1997]:

$$x_{ij}^s = \frac{x_{ij} - \bar{x}_j}{s_j} \quad (2-10)$$

$$s_j = \sqrt{\frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}{n-1}} \quad (2-11)$$

Dabei ist s_j die Standardabweichung über alle n Wert der Variablen j .

Die Originaldaten werden durch die Transformation in eine neue Koordinate transformiert. Nach der Datenvorverarbeitung wird die Hauptkomponentenanalyse durchgeführt. Die Hauptkomponentenanalyse ist eine Methode, bei der jede Matrix X vom Rang r als Summe von r Matrizen M_i des Ranges eins dargestellt werden kann.

$$X = M_1 + M_2 + M_3 + \dots + M_r + E = \sum_{i=1}^r M_i + E \quad (2-12)$$

Dabei ist E die Residuen-Matrix. Jede dieser Matrizen lässt sich wiederum als ein Produkt zweier Vektoren ausdrücken (Scorevektor t_i und Loadingvektor p_i'):

$$X = t_1 p_1' + t_2 p_2' + t_3 p_3' + \dots + t_i p_i' + E = \sum_{i=1}^r t_i p_i' + E \quad (2-13)$$

Es lassen sich alle Vektoren t_i zur Matrix T und alle Vektoren p_i' zur Matrix P' zusammenfassen. Bei der Hauptkomponentenanalyse wird die ursprüngliche Datenmatrix X durch zwei Matrizen T und P' repräsentiert. Dabei wird T als Score- (bzw. Faktorenwerte-) und P als Loading- (bzw. Ladungs-) Matrix bezeichnet.

$$X = T \cdot P' + E \quad (2-14)$$

Die Spalten der Matrix T sind die Vektoren der Faktorenwerte und die Spalten der Matrix P' die Ladungsvektoren (Gewichtsfaktoren). Die Vektoren in T und in P' sind untereinander orthogonal. Diese Transformation der Daten führt zu neuen unkorrelierten Variablen.

Beim Einsatz des NIPALS-Algorithmus (Nonlinear Iterative Partial Least Squares-Algorithmus) werden die einzelnen Hauptkomponenten sukzessiv bestimmt [Kowalski et al., 1991]. Dieses wird iterativ durch wiederholte Regression von X und t erreicht, wobei es das Ziel ist, einen verbesserten Loadingvektor p zu erhalten. Mit X und p wird dann ein verbesserter Scorevektor t berechnet. Die Originaldatenmatrix wird als X_0 bezeichnet.

Die einzelnen Schritte sind im Folgenden beschrieben: Zuerst werden die Daten zentriert. Dann wird ein Scorevektor aus der Matrix X_{a-1} gewählt und dieser Vektor t_a genannt. Dabei bezeichnet a den Index der Hauptkomponente. Dieser repräsentiert im Prinzip die größte Varianz, d.h. $t_a' \cdot t_a$ ist maximal.

Im ersten Schritt des Iterationszyklus wird ein verbesserter Loadingvektor p_a' bestimmt. Dies wird durch Projektion der Matrix X_{a-1} auf t_a erreicht:

$$p_a' = (t_a' \cdot t_a)^{-1} t_a' \cdot X_{a-1} \quad (2-15)$$

Im zweiten Schritt wird nun p_a auf die Länge Eins normiert.

$$p_a = p_a' (p_a' \cdot p_a)^{-0.5} \quad (2-16)$$

Anschließend erhält man durch Projektion der Matrix X_{a-1} auf p_a einen verbesserten, genäherten Scorevektor t_a :

$$t_a = X_{a-1} \cdot p_a (p_a' \cdot p_a)^{-1} \quad (2-17)$$

Der verbesserte Eigenwert τ_a ergibt sich dann durch:

$$\tau_a = t_a' \cdot t_a \quad (2-18)$$

Jetzt wird die Konvergenz überprüft. Als Kriterium für die Konvergenz wird die Differenz zwischen den Eigenwerten des aktuellen und des vorherigen Iterationszyklus benutzt. Diese Iterationsschranke liegt z.B. bei $0,00001 \cdot \tau_a$. Wird die Iterationsschranke unterschritten, gilt die Hauptkomponente als bestimmt und wird aus der Datenmatrix eliminiert:

$$X_a = X_{a-1} - t_a \cdot p_a' \quad (2-19)$$

Aus dieser neuen Datenmatrix X_a wird nun derjenige Scorevektor gewählt, der die größte verbleibende Varianz repräsentiert und der iterative Algorithmus wird erneut durchlaufen. Dies wird so oft wiederholt, bis die vorgegebene Anzahl der Hauptkomponenten gefunden ist. Die Score-Werte werden zur Berechnung des multilinenen Regressionsmodells und die Loading-Werte für die Hauptkomponentenregression (Principal Component Regression, PCR) zur Vorhersage der Prozessgrößen benutzt.

$$T_{pred} = X \cdot P \quad (2-20)$$

Dabei ist T_{pred} die neu berechnete Score-Matrix, X die Datenmatrix und P die Loading-Matrix, die durch den NIPALS-Algorithmus berechnet wurde.

Mehrere unterschiedliche Größen y werden in der Matrix Y zusammengefasst. Man erhält so für jede Größe y einen Koeffizientenvektor b . Die Koeffizientenvektoren werden zur Matrix B zusammengefasst. Für die Vorhersage \hat{Y}_{pred} ergibt sich die folgende Gleichung:

$$\hat{Y}_{pred} = \bar{Y}_{kal} + T_{pred} \cdot B \quad (2-21)$$

Dabei ist \bar{Y}_{kal} der Vektor der Mittelwerte der jeweiligen y -Variablen (Prozessgrößen) der Kalibrationsmatrix.

2.3. Kontinuierlich-diskrete erweiterte Kalman-Filter

Zum Bioprozessmonitoring sind die Online- oder Offline-Bestimmungen der Zustandsgrößen des Prozesssystems notwendig. Bei der Bestimmung der Zustandsgrößen eines Systems ist davon auszugehen, dass es bei der Messung der Ausgangsgrößen zu Ungenauigkeiten aufgrund von stochastischen Störungen und Messfehlern kommt. Die Berücksichtigung und Minimierung des Einflusses der Ungenauigkeiten auf den geschätzten Zustand des Systems bezeichnet man als Filterung. Eine Methode, um die Filterungsaufgabe zu lösen, stellt der Kalman-Filter Algorithmus dar. Dieser bietet eine modellbasierte Schätzung des Systemzustands basierend auf diskreten Messwerten. Ein Kalman-Filter benötigt zur Filterung nur den jeweils aktuellen Messwert und nicht wie andere Filteralgorithmen eine Zeitreihe von Messwerten. Außerdem kann er je nach Auslegung auch eine Schätzung weiterer Prozessgrößen liefern. Das Kalman-Filter basiert auf einem Modell des eigentlichen Prozesses, in das stochastische Störgrößen einfließen. Sind die Modellgleichungen nicht linear, so spricht man von einem erweiterten Kalman-Filter (EKF). Das erweiterte Kalman-Filter verwendet eine kontinuierliche Prädiktion der Zustandsvariablen mit diskreter Filterung der eintreffenden Online-Messwerte auf Basis eines nichtlinearen Modells des Prozesses [Hitzmann et al., 2000].

Die kontinuierlichen Zustandsgleichungen für einen Prozess mit dem Zustandsvektor $x(t)$ und dem Prozessrauschen $z(t)$ sind für den allgemeinen Fall in Gleichung (2-22) dargestellt:

$$\frac{dx(t)}{dt} = f(x(t), t) + z(t) \quad (2-22)$$

Die Schätzung basiert auf Messungen und Zustands- und Messgleichungen (Ausgangsgleichungen). Die diskrete Ausgangsgleichung lautet:

$$y(t_i) = h(x(t_i)) + s(t_i) \quad (2-23)$$

Dabei steht $y(t_i)$ für den jeweiligen Messwert zu den diskreten Zeitpunkten t_i , $h(x(t_i))$ für den berechneten Messwert aus den Zustandsvariablen und $s(t_i)$ für das Messrauschen. Das Mess- und Prozessrauschen wird als weißes Rauschen angenommen, d.h. der Mittelwert ist gleich Null, während die Varianz r größer als Null ist. Das Messrauschen wird durch die Messrauschkovarianzmatrix R beschrieben, die sich wie folgt berechnet:

$$R = E\{s(t_i) \cdot s'(t_j)\} \quad (2-24)$$

Dabei repräsentiert E die Funktion zur Berechnung des Erwartungswerts der jeweiligen Größe. Die Stärke des Prozessrauschens beschreibt die Prozessrauschleistungsmatrix Q :

$$Q \cdot \delta(\tau) = E\{z(t) \cdot z'(t + \tau)\} \quad (2-25)$$

Da die Messungen als unabhängig angesehen werden, ist die Messrauschkovarianzmatrix R eine Diagonalmatrix mit den Varianzen des Messrauschens in der Hauptdiagonale. Durch die Vorgabe der Größe der Varianzen in R kann die Stärke des Messrauschens spezifiziert werden, wodurch das Wissen über die Messgüte in das Modell einfließt. Die Prozessrauschleistungsmatrix Q ist ein Maß für die Güte des verwendeten Prozessmodells. Die Schätzfehlerkovarianzmatrix P enthält die Erwartungswerte der quadratischen Abweichung der Schätzung des Zustandsvektors $x(t)$ von seinem wahren Wert:

$$P_{ii}(t_i) = E\{(x_i(t) - \hat{x}_i(t))^2\} \quad (2-26)$$

Dabei ist $x_i(t)$ der wahre Wert und $\hat{x}_i(t)$ der Schätzwert des Zustandsvektors. Die Schätzfehlerkovarianzmatrix $P(t)$ wird im Rahmen eines Kalman-Filters kontinuierlich nach Gleichung (2-27) berechnet solange kein Messwert eintrifft.

$$\frac{dP(t)}{dt} = F(t)P(t) + P(t)F'(t) + Q \quad (2-27)$$

Dabei ist $F(t)$ die Jacobi-Matrix von $f(x(t), t)$. Sobald ein neuer Messwert zum Zeitpunkt t_i eintrifft, findet die diskrete Filterung statt, d.h. die Werte der Zustandsvariablen und der Schätzfehlerkovarianz werden zum Zeitpunkt t_i neu berechnet. Dabei wird zunächst die Kalman-Verstärkungsmatrix $K(t_i)$ mit Gleichung (2-28) bestimmt.

$$K(t_i) = P(t_i) \cdot H'(t_i) \cdot \{H(t_i) \cdot P(t_i) \cdot H'(t_i) + R\}^{-1} \quad (2-28)$$

Dabei ist $P(t_i)$ die Schätzfehlerkovarianzmatrix der letzten Prädiktion und $H(t_i)$ die Jacobi-Matrix von $h(x(t_i))$. Die Kalman-Verstärkungsmatrix wird dabei so berechnet, dass die Schätzfehlerkovarianz minimal wird. Mit der neu bestimmten Kalman-Verstärkungsmatrix wird nun der neue Schätzwert durch die Gleichung (2-29) berechnet:

$$\hat{x}_{neu}(t_i) = \hat{x}(t_i) + K(t_i) \cdot \{y(t_i) - h(\hat{x}(t_i), t_i)\} \quad (2-29)$$

Dabei ist $\hat{x}_{neu}(t_i)$ der neue Schätzwert (gefilterter Wert) des Zustandsvektors und $y(t_i)$ der Messwert zu dem Zeitpunkt t_i .

Die neue Schätzfehlerkovarianz $P_{neu}(t_i)$ ergibt sich durch folgende Gleichung (2-30):

$$P_{neu}(t_i) = \{I - K(t_i) \cdot H(t_i)\} \cdot P(t_i) \cdot \{I - K(t_i) \cdot H(t_i)\}^T + K(t_i) \cdot R \cdot K^T(t_i) \quad (2-30)$$

Dabei ist I die Einheitsmatrix, $K(t_i)$ die Kalman-Filter-Verstärkungsmatrix und $H(t_i)$ die Jacobi-Matrix von $h(x(t_i))$. Jetzt wird mit den gefilterten Werten das Differenzialgleichungssystem für die Zustandsvariablen und die Schätzfehlerkovarianz gelöst, bis sich die Filterung bei Eintreffen eines neuen Messwertes wiederholt.

Zusammengefasst liefert das kontinuierlich-diskrete erweiterte Kalman-Filter (EKF) also eine kontinuierliche Schätzung des Zustandsvektors des Prozesssystems.

2.4. Backpropagation-Netzwerk

Die grundlegende Idee eines neuronalen Netzwerks entstammt den Erkenntnissen der Hirnforschung. Die menschliche Informationsverarbeitung ist im Wesentlichen auf die Übertragung von Erregungen zwischen Neuronen zurückzuführen. Künstliche neuronale Netzwerke versuchen nun, die Funktionsweise der Neuronen nachzuahmen. McCulloch und Pitts veröffentlichten ein erstes Modell in „Neuronale Netze“ [Brause, 1991], bei dem das biologische Neuron als logisches Schwellen-Element beschrieben wurde. Dieses Modell wurde später erweitert und verallgemeinert. Neuronale Netzwerke werden zur Zeit in verschiedenen Bereichen verwendet, wie z.B. in der Sprach- und Bildverarbeitung sowie bei der Modellierung, Optimierung, Steuerung und Regelung von Prozessen. Dabei kommen unterschiedliche Verfahren zur Anwendung.

Ein Backpropagation-Netzwerk ist meistens aus drei Schichten aufgebaut, die eine bestimmte Anzahl von Neuronen enthalten. Die Neuronenschichten sind untereinander verbunden.

Abbildung 2-3 zeigt den Aufbau eines dreischichtigen Netzes. Die Neuronen der Eingabeschicht, der inneren Schicht und der Ausgabeschicht werden mit dem Index k , j und i gekennzeichnet. Dabei ist w_{jk} das Gewicht zwischen der Eingabeschicht und der inneren Schicht und w_{ij} das Gewicht zwischen der inneren Schicht und der Ausgabeschicht.

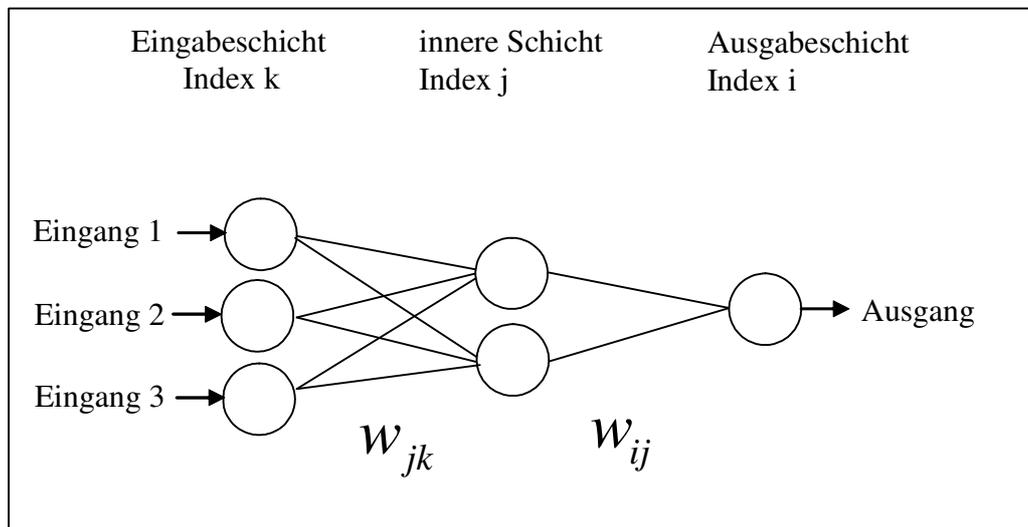


Abbildung 2-3: Ein Beispiel für ein neuronales Netz mit drei Schichten: Eingabeschicht mit drei Neuronen, innere Schicht mit zwei Neuronen und Ausgabeschicht mit einem Neuron

Die Neuronen der inneren Schichten bzw. der Ausgabeschicht verfügen über mehrere Eingänge und einen Ausgang. Die Eingänge sind mit den Neuronen der vorhergehenden Schicht verbunden. Jede Verbindung ist mit einem Gewicht w belegt, d.h. der Ausgangswert eines sendenden Neurons o wird mit dem Gewicht der Verbindung multipliziert, bevor er an das empfangende Neuron weitergegeben wird. Im empfangenden Neuron werden die gewichteten Ausgangswerte der Vorgängerneuronen zusammen mit einem Schwellenwert, dem Bias des Neurons j (Θ_j), zum Nettoinput net_j aufsummiert:

$$net_j = \sum_i w_{ij} \cdot o_i + \Theta_j \quad (2-31)$$

Der Bias Θ kann als Gewicht zu einem Neuron interpretiert werden, das ein Signal sendet, dessen Gewicht einen konstanten Wert eins besitzt. In Abbildung 2-4 ist das Blockschaltbild für ein einzelnes Neuron der Ausgabeschicht dargestellt, die Neuronen in den inneren Schichten haben den gleichen Aufbau. Das Blockschaltbild soll den Informationsfluss

innerhalb eines Neurons veranschaulichen. Die Aktivität a_j eines Neurons j ergibt sich aus dem Nettoinput net_j mit der sigmoiden Aktivierungsfunktion f_{sig} nach Gleichung (2-32) zu

$$a_j = f_{sig}(net_j) = \frac{1}{1 + e^{-net_j}} \quad (2-32)$$

Dabei steht der Index j für die innere Schicht. Die Aktivität ist gleichzeitig der Ausgabewert o_j des Neurons ($a_j = o_j$). Durch die sigmoide Aktivierungsfunktion werden die Ausgabewerte der Neuronen auf Werte zwischen Null und Eins beschränkt. Die Null und die Eins selbst werden dabei nur erreicht, wenn der Nettoinput gegen $-\infty$ bzw. $+\infty$ strebt [Ritter et al., 1990].

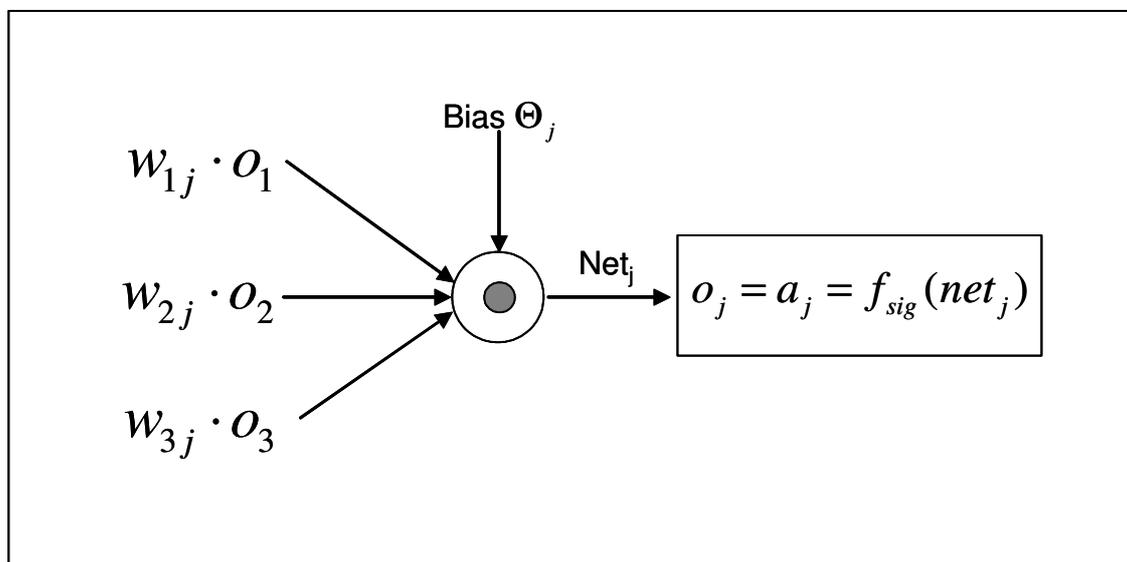


Abbildung 2-4: Blockschaltbild eines Neurons j mit drei Eingängen

Bevor ein Netzwerk trainiert werden kann, muss eine für die Aufgabe optimale Netzwerkstruktur gefunden werden. Dabei wird die Anzahl und Belegung der Ein- und Ausgänge weitgehend von der Problemstellung festgelegt.

Für das Training müssen erst Beispiele erstellt werden, die das Netzwerk lernen soll. Diese Beispiele werden Lernmuster genannt. Ein Lernmuster besteht aus einem Satz von Eingangswerten, dem Eingangsmuster, und einem Satz von Sollausgangswerten, dem Zielmuster. Jeder Ein- und Ausgang des neuronalen Netzes muss dabei im Lernmuster durch einen Wert belegt werden. Das Ziel des Lernprozesses ist es nun, die Verbindungsgewichte so einzustellen, dass das Netzwerk auf die Eingabe eines Eingangsmusters möglichst genau mit dem zugehörigen Zielmuster antwortet. Dabei wird die Summe der quadratischen

Abweichungen zwischen Soll- und Istausgangswerten des Netzes über alle Lernmuster, die auch die Fehlerquadratsumme SSE (Sum of Squared Errors) des Netzes genannt wird, als Gütekriterium gewählt. Es gilt für die Fehlerquadratsumme:

$$SSE = \sum_p E_p = \frac{1}{2} \sum_p \sum_j (s_{pj} - o_{pj})^2 \quad (2-33)$$

Die Indizes p und j laufen über alle Lernmuster bzw. alle Ausgabeneuronen. E_p ist der quadratische Fehler des p -ten Lernmusters, d.h. die Summe der quadratischen Abweichungen zwischen dem vorgegebenen Sollausgangswert s_{pj} des Lernmusters und dem tatsächlichen Istausgangswert o_{pj} des Netzes.

Durch die Variation jedes einzelnen Gewichts w_{jk} und w_{ij} ändert sich der Ausgabewert der Neuronen in der Ausgabeschicht und damit auch der Gesamtfehler E des Netzwerks. D.h. E ist eine Funktion aller Gewichte w_{jk} bzw. w_{ij} . Der am häufigsten benutzte Ansatz, um den Fehler zu minimieren, ist das Gradientenabstiegsverfahren. Das Gradientenabstiegsverfahren ermöglicht es, für jedes einzelne Verbindungsgewicht w_{ij} Werte zu finden, die den Lernfehler für ein vorgegebenes Muster minimieren. Dabei wird ein Eingangsmuster durch das Netzwerk propagiert (Forwardpropagation-Lauf). Dann werden in einem Backpropagation-Lauf die Änderungen der Verbindungsgewichte Δw_{ij} entsprechend den Ausgangsfehlern nach folgender Gleichung (2-34) berechnet:

$$\Delta w_{ij} = -\alpha \cdot \frac{\partial E_p}{\partial w_{ij}} \quad (2-34)$$

Dabei ist α die Lernrate und E_p der quadratische Fehler des Lernmusters. Das Lernverfahren wird Backpropagation genannt, weil die Verbindungsgewichte entgegen der normalen Richtung der Aktivierungsausbreitung mit Δw_{ij} modifiziert werden.

Die Gewichtsänderungen erfolgen in Richtung des stärksten Abfalls der Fehlerfunktion E_p mit der Lernrate α . Durch die Auswertung der Gleichung (2-34) kann man die Gewichtsänderungen explizit bestimmen, dabei müssen allerdings Fallunterscheidungen vorgenommen werden [Kruse et al., 1991]. Für die Änderung der Verbindungsgewichte Δw_{jk} bzw. Δw_{ij} gelten:

$$\Delta w_{jk} = \alpha \cdot \sum_i (s_{pi} - o_{pi}) \cdot o_k \cdot o_i \cdot (1 - o_i) \cdot w_{ij} \cdot o_j \cdot (1 - o_j) \quad (2-35)$$

$$\Delta w_{ij} = \alpha \cdot (s_{pi} - o_{pi}) \cdot o_j \cdot o_i \cdot (1 - o_i) \quad (2-36)$$

Die Herleitung findet man in der Literatur [Ritter et al., 1990]. Die Gewichtsänderungen können nach jedem Lernmuster durchgeführt werden (learn by pattern) oder aufsummiert und erst nach dem Durchlauf von allen Lernmustern abgearbeitet werden (learn by epoch). In dieser Arbeit wird die musterweise Gewichtsanzpassung gewählt. Der komplette Durchlauf aller Muster wird Lernzyklus genannt.

Eine Erweiterung des Backpropagation-Lernverfahrens ist das Lernen mit Momentum. Hier wird der alte Wert der Gewichtsveränderung bei der Berechnung eines neuen Wertes mit dem Momentum η nach Gleichung (2-37) berücksichtigt. Die Änderung des Gewichts Δw_{ij} zur Iteration $r+1$ erfolgt dann in Abhängigkeit vom Wert der Gewichtsänderung zur Iteration r und kann durch die folgende Gleichung berechnet werden:

$$\Delta w_{ij}(r+1) = -\alpha \cdot \frac{\partial E_p}{\partial w_{ij}} + \eta \cdot \Delta w_{ij}(r) \quad (2-37)$$

Dabei ist r die Iteration für die Gewichtsänderung, η das Momentum und α die Lernrate. Das Momentum soll verhindern, dass die Fehlerfunktion im Verlauf des Lernprozesses um ein Minimum pendelt. Dies geschieht dann, wenn aufeinanderfolgende Gewichtsänderungen mit entgegengesetztem Vorzeichen durchgeführt werden. Eine Einbeziehung des alten Wertes der Gewichtsänderung verbessert die Konvergenz der Fehlerfunktion im Minimum.

Der Abbruch eines Lernprozesses kann nach unterschiedlichen Kriterien erfolgen. Es kann z.B. nach dem Erreichen einer bestimmten maximalen Lernzyklenzahl oder eines maximal tolerierbaren quadratischen Fehlers abgebrochen werden. Daneben kann auch die Veränderung des quadratischen Fehlers oder die Erkennung einer minimalen Musterzahl als Abbruchkriterium verwendet werden. Dabei muss allerdings eine Toleranzschwelle für die Fehler der Ausgänge festgelegt werden, ab der ein Lernmuster als „erkannt“ gilt. Es sind auch Kombinationen der einzelnen Abbruchkriterien möglich. Eine passende Abbruchbedingung muss für jede Problemstellung neu gefunden werden, allgemeingültige Regeln für ein optimales Beenden des Lernprozesses existieren nicht.

3. Material und Methoden

3.1. Ultraschallmessgerät US 2100

Das verwendete akustische Online-Sensorsystem US 2100 wurde für die Untersuchung biotechnischer Prozesse am Institut für Automation und Kommunikation (IFAK, Barleben) entwickelt. Abbildung 3-1 zeigt die Ultraschallsonde im Bypass des Bioreaktors. Das System besteht aus dem akustischen Sensor und einer Kontrolleinheit. Die akustische Sensorsonde kann vom System getrennt werden, um die Sonde leicht sterilisieren zu können. Die Sonde ist für die Online-Messung spezifisch als Bypass-Sonde entwickelt worden. Die Kontrolleinheit des Ultraschallsensorsystems ist in Abbildung 3-2 dargestellt.



Abbildung 3-1: Ultraschallsonde mit Anschlussvorrichtung für den Bypass



Abbildung 3-2: Kontrolleinheit des Ultraschallsensorsystems

Die Kontrolleinheit steuert den gesamten Messablauf und realisiert die Messdatenverarbeitung sowie die Ausgabe der Messinformationen. An die Kontrolleinheit ist der Messrechner angeschlossen, der die Automation des Systems erlaubt. Abbildung 3-3 zeigt den Versuchsbioreaktor mit dem angeschlossenen Ultraschallmesssystem.

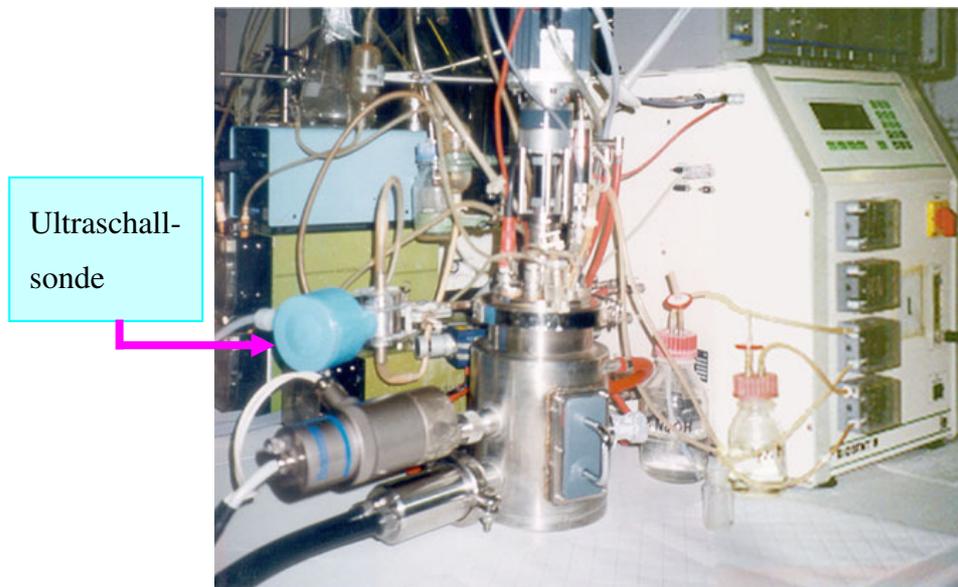


Abbildung 3-3: Bioreaktor mit angeschlossenen Messsystem

3.2. Messanordnung

Auf der Basis des Ultraschallsensorsystems werden die Ultraschallmessungen bei Kultivierungen von *Saccharomyces cerevisiae* durchgeführt, die in einem kombinierten Batch- / Fed-batch-Verfahren betrieben werden. Abbildung 3-4 zeigt schematisch den Messaufbau für die Ultraschallmessungen bei Kultivierungen. Um Blasen in der Messzone zu eliminieren, wird die Ultraschallmessung im Bypass-System durchgeführt. Mit Hilfe einer Pumpe wird die Kulturbrühe durch den Bypass gepumpt. Die Strömungsgeschwindigkeit beträgt 65 mL/s. Dabei dient ein Ventil dazu, die Kulturbrühe zeitweise in der Messzelle zu stoppen (Stop-Flow). Der Messrechner der Ultraschallsonde dient auch zur Steuerung des Stop-Flow-Betriebs.

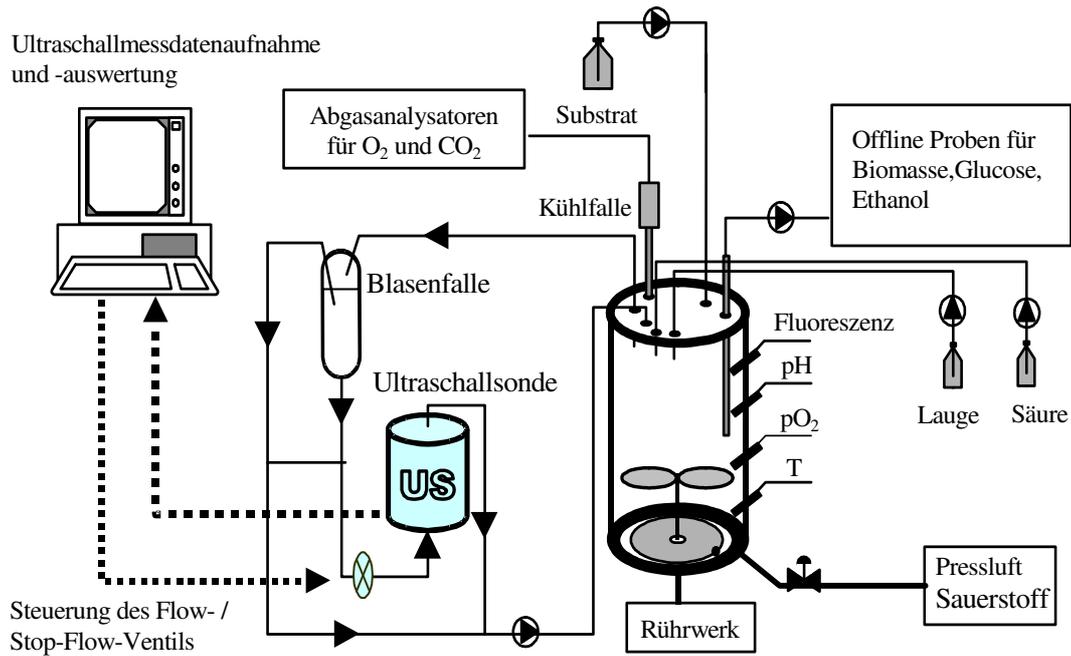


Abbildung 3-4: Messaufbau für die Ultraschallgeschwindigkeitsmessung

Die Software zur Automation des Ultraschallsystems wurde im IFAK entwickelt. Abbildung 3-5 zeigt die Bedienoberfläche für die Ultraschallmessungen, mit der die Funktionstüchtigkeit des Sensorsystems überprüft werden kann. Der Datentransfer von dem Ultraschallsensorsystem zum Computer wird über eine serielle Schnittstelle realisiert.

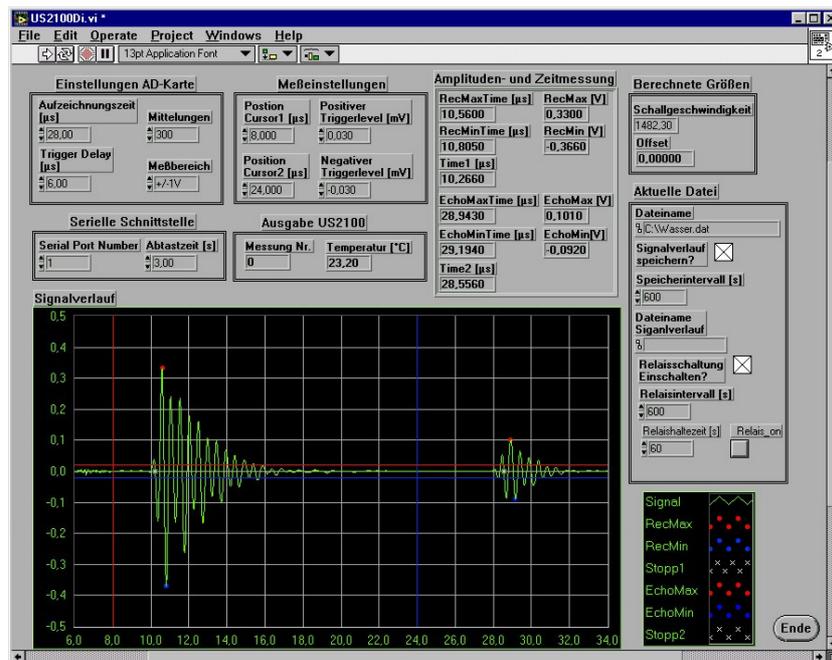


Abbildung 3-5: Bedienoberfläche des Ultraschallsensorsystems

3.3. Verwendete Mikroorganismen

Die in dieser Arbeit untersuchte Hefe *Saccharomyces cerevisiae* H620 gehört als Protoacnomycet zu den Sprosspilzen. Sie ist eine eukaryontische Zelle mit hoch entwickelten Zellorganellen (Zellkern, Mitochondrien) und besitzt eine widerstandsfähige Zellwand [Bellgardt, 1984].

Die Kultivierungsbedingungen, die bei der Untersuchung der Ultraschallsignale verwendet werden, sind in Tabelle 3-1 aufgelistet. Die Zusammensetzung des Schatzmann-Mediums ist im Anhang 10.3 aufgeführt [Schatzman, 1975]. Es werden eine Mineralsalzlösung und eine Vitaminlösung mit der jeweils 1000-fachen Konzentration in destilliertem Wasser angesetzt. 2,2 L Leitungswasser werden mit konzentrierter H_2SO_4 auf $\text{pH} = 2$ eingestellt. $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ und $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ werden gelöst, bevor die Mineralsalzlösung und die übrigen Komponenten außer der Glucose-Lösung, die extra autoklaviert wurde, zugegeben werden. Anschließend wird der pH-Wert mit konzentriertem H_2SO_4 auf $\text{pH} = 5,5$ eingestellt. Die Glucose-Lösung und die Vitaminlösung sind vor dem Animpfen steril injiziert worden. Natrium-Citrat wird nur in der Vorkultur (Volumen 50 mL) verwendet, die bei 30°C in einem 300 mL Schüttelkolben mit 120 UpM für 18,5 h inkubiert wird.

Die Online-Messdaten sowie daraus abgeleitete Größen werden mit dem RISP-System (Real Time Integration Software Platform, TCI) aufgenommen.

Tabelle 3-1: Kultivierungsbedingungen

Betriebsweise	Batch		Fed-Batch	
Bezeichnung	Hefe 22/23	Hefe 27/30	Fed-Batch 2	Fed-Batch 11
Organismus	<i>Saccharomyces cerevisiae</i>			
Bioreaktor	2L-Reaktor (Typ Biostat M, B. Braun Biotech International, Melsungen)			
Arbeitsvolumen	1,5 L		1,0 L	
Medium	Schatzmann-Medium			
Temperatur	30 °C			
pH	5,0	5,5	5,5	5,5
Regelung des pH-Werts	4% HCl, 10% NaOH			
Rührerdrehzahl [UpM]	1000	800	800	800
Begasungsrate	2,25 L/min (mit Druckluft)			

3.4. Online-Analytik

Der Reaktor ist mit verschiedenen Sensoren und Messgeräten bestückt: pO_2 , CO_2 -Abgas, O_2 -Abgas, pH, Säurezugabe, Laugezugabe, Temperatur, Waage (zur Bestimmung der in den Reaktor gepumpten Menge der Glucoselösung) und 2-D-Fluoreszenzspektrometer [Marose et al., 1998].

3.4.1. CO_2 -, O_2 -Abgas Messungen

Der Sauerstoffgehalt des Abgases wird elektrochemisch über eine Elektrolysezelle bestimmt. Ein auf Kohlendioxid sensibilisiertes, einstrahliges IR-Photometer wird für die quantitative Bestimmung durch Messung der IR-Absorption bei charakteristischen Wellenlängen genutzt. Die Kalibrierung der Analysegeräte erfolgt automatisch oder manuell mit Prüfgas der Zusammensetzung 1,98 % CO_2 , 19,0 % O_2 , Rest N_2 .

3.4.2. pO_2 (Gelöstsauerstoff) Messungen

Der Gelöstsauerstoff wird mit Hilfe einer autoklavierbaren Elektrode (Mettler-Toledo, Schweiz) als Sauerstoffpartialdruck nach dem polarographischen Prinzip (Clark-Prinzip) gemessen. Dabei diffundiert der Sauerstoff aus der Lösung durch eine gasdurchlässige Teflonmembran, die die Messlösung von der Ag-Anode und Pt-Kathode trennt, und wird an der Kathode reduziert. An der Anode wird Silber oxidiert. Die Elektrode wird nach dem Autoklavieren im Fermenter kalibriert (Zwei-Punkt-Kalibration). Dabei wird zunächst durch Begasung mit Stickstoff der 0-%-Wert für pO_2 festgelegt. Der 100-%-Wert wird durch Begasung mit Druckluft unter den jeweiligen Kultivierungsbedingungen (Rührerdrehzahl und Begasungsrate) eingestellt.

Die während der Kultivierung gemessenen Werte geben die relative Sauerstoffsättigung (%-rel. Sättigung) an und werden in dieser Arbeit als Gelöstsauerstoffgehalt [%] bezeichnet.

3.4.3. Fluoreszenzspektroskopie

Das verwendete Fluorimeter ist ein Fluoreszenz-Spektralphotometer F-4500 (Hitachi, Japan). Im F-4500 werden die Wellenlängen von Exzitation und Emission unabhängig voneinander mit Hilfe eines Gittermonochromators (Konkavgitter mit 900 Linien/mm) mit wählbarer Spaltbreite (2,5 nm; 5 nm; 10 nm oder 20 nm) eingestellt. Ein Strahlteiler lenkt einen Teil des Anregungslichts auf einen Referenzphotomultiplier, wodurch Intensitätsschwankungen an der Lichtquelle kompensiert werden können. Ein Verschluß (*shutter*) sorgt dafür, dass die Probe nicht ständig angeregt wird. Die Gerätesteuerung und die Datenaufnahme erfolgen über einen

Personal-Computer (PC). Für die kontinuierliche Datenaufnahme von Fluoreszenzspektren wurde ein *Quick-Basic*-Programm verwendet.

3.5. Offline-Analytik

Über Offline-Proben werden folgende Größen bestimmt: Biotrockenmasse-, Glucose- und Ethanolkonzentration.

3.5.1. Biotrockenmassemessung

Die Bestimmung der Biotrockenmasse (BTM) erfolgt gravimetrisch. Dazu werden je Probe 10 mL Kulturlösung in sogenannte Falcon-Gefäße (50 ml, Greiner) mit bekanntem Leergewicht gefüllt, mit 10 ml 0,9-%-iger NaCl-Lösung (aq.) verdünnt und in einer Zentrifuge (Modell J2-21, Rotor JA-14, Beckmann) 20 min bei 6000 g und 4°C abzentrifugiert, um möglichst alle löslichen Bestandteile in den Überstand zu überführen. Der Überstand wird verworfen. Nach Trocknung der Zell-Pellets im Trockenschrank bei 80°C bis zur Gewichtskonstanz (24 h) wird das Falcon-Gefäß erneut ausgewogen. Die Biotrockenmasse ergibt sich aus der Massendifferenz zwischen dem gefüllten und dem leeren Falcon-Gefäß.

3.5.2. Glucosemessung

Die Bestimmung der Glucosekonzentration erfolgt mit dem YSI Analysator Modell 2700 Select (Yellow Springs Instruments, USA). Der Analyt wird durch das auf einer Membran immobilisierte Enzym Glucoseoxidase umgesetzt, wobei in stöchiometrischer Menge Wasserstoffperoxid (H₂O₂) entsteht. Dieses wird an einer Platinelektrode elektrochemisch zu Sauerstoff oxidiert. Die Stromstärke des resultierenden Stromflusses ist proportional zur Analytkonzentration. Interferenzen durch andere oxidierbare Substanzen werden durch eine Zelluloseacetat-Membran vor der Elektrode verhindert. Die Kalibrierung des Analysators erfolgt mit Hilfe einer Standardlösung der Glucosekonzentration 2 g/L.

3.5.3. Ethanolmessung

Die Ethanolkonzentration wird in einem Bereich von 0-15 g/L mit einem Shimadzu Gaschromatographen (GC-14B) bestimmt. Die Alkohole werden über eine Chromosorb-101-Säule getrennt und mit einem Flammenionisationsdetektor detektiert. Die Daten werden mit der Class-VP-Software von Shimadzu (Version 4.2) auf einem Rechner aufgenommen und ausgewertet. Es wird mit vier Standards kalibriert. Die Auswertung der Proben erfolgt über eine Kalibration mit einem internen Standard. Die Parameter der Messung wurden wie folgt gewählt (Tabelle 3-2):

Tabelle 3-2: Parameter der Ethanolmessungen

Parameter	Bedingungen
Trärgas	Stickstoff (Carrier Flow 38 mL/min, 220 kPa)
Brenngas	Wasserstoff (65kPa), synth. Luft (50 kPa)
Säule	Chromosorb 101, 80/100, Edelstahlsäule 6' x 1/8"
T _{säule}	160°C
T _{injektor}	180 °C
T _{detektor}	220 °C
Probenvolumen	1 µL
Interner Standard	n-Propanol (8 g/L)
Probenvorbereitung	500 µL Probe (abzentrifugiert) + 50 µL Standard

4. Anwendung des Ultraschallsensorsystems

4.1. Kalibrierung der Ultraschallsonde

Die Temperaturabhängigkeiten von Dichte und Kompressibilitätskoeffizienten bewirken eine Abhängigkeit der Schallgeschwindigkeit von der Temperatur. Bei Flüssigkeiten verringert sich die Schallgeschwindigkeit häufig linear mit zunehmender Temperatur bei konstantem Druck. Die Schallgeschwindigkeit in Wasser zeigt aber mit steigender Temperatur einen parabolischen Verlauf, wobei die maximale Schallgeschwindigkeit von 1557 m/s bei 74 °C erreicht wird [Bergman, 1954]. Der Zusammenhang zwischen Schallgeschwindigkeit und Temperatur für eine Messung mit Wasser wird nach der Willard-Beziehung mit Gleichung (4-1) [Börnstein, 1967] bestimmt:

$$US(T) = 1557 - 0,0245 (74 - T)^2 \quad (4-1)$$

Dabei ist US die temperaturabhängige Schallgeschwindigkeit [m/s] und T die Temperatur [°C].

Bei den Hefekultivierungen wird die Temperatur bei 30 °C konstant gehalten. Daher wird die Ultraschallgeschwindigkeitsmessung im Bereich von 25-35 °C für die Kalibration der Schallsonde durchgeführt. Mit den Messwerten wird die temperaturabhängige Ultraschallgeschwindigkeit durch ein Polynom zweiter Ordnung bezüglich der Temperatur angenähert. Folgende Gleichung wird für das Medium ermittelt:

$$US(T) = -0,0217 T^2 + 3,465 T + 1427,7 \quad (4-2)$$

In Abbildung 4-1 ist die Temperaturabhängigkeit der Schallgeschwindigkeit in Wasser und im Medium dargestellt. Im Bereich 25-35 °C verläuft die Ultraschallgeschwindigkeit in Wasser und im Medium parallel mit einem Abstand von ca. 2,5 m/s.

Die Bestimmung der Ultraschallgeschwindigkeit erfolgt bei konstantem Abstand von Sender und Empfänger durch die Differenzbildung der Laufzeit des am Empfänger und Sender reflektierten Signals (t_2) mit der Laufzeit des ursprünglich am Empfänger registrierten Signals (t_1) nach folgender Beziehung:

$$C = \frac{2S}{t_2 - t_1} \quad (4-3)$$

Dabei ist C die Schallgeschwindigkeit [m/s], S der Schallweg und $t_2 - t_1$ die Laufzeitdifferenz der beiden Signale.

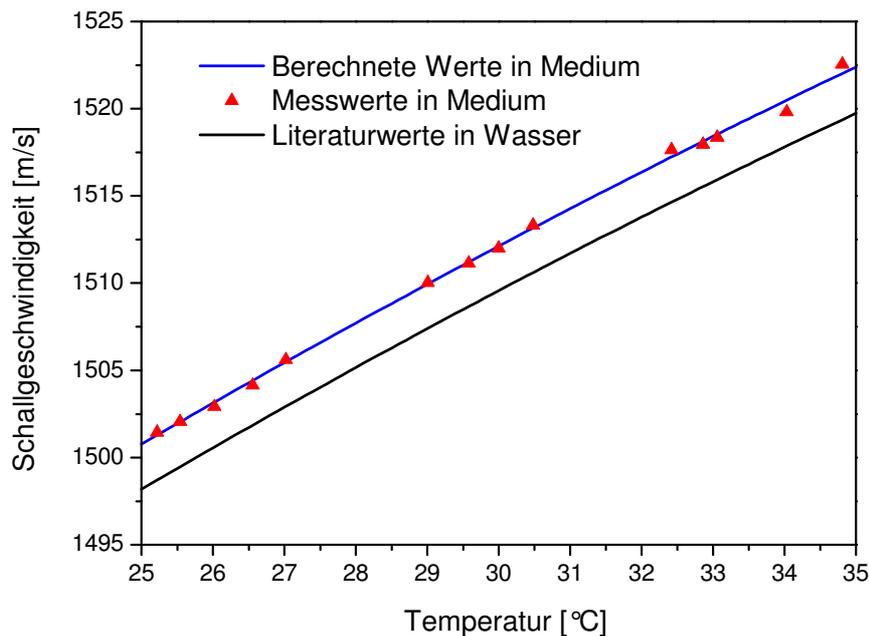


Abbildung 4-1: Vergleich der Temperaturabhängigkeit der Schallgeschwindigkeit in Wasser und im Medium

Um den Schallweg zu berechnen, wird die Laufzeitdifferenz bei einer konstanten Temperatur von 30 °C in Wasser gemessen. Nach der Willard-Beziehung für Wasser wird die Schallgeschwindigkeit zu $C(30\text{ °C}) = 1509,6\text{ m/s}$ berechnet (siehe Gleichung (4-1)). Durch Umformen von Gleichung (4-3) kann der Schallweg mit folgender Gleichung (4-4) bestimmt werden:

$$S(T) = \frac{1}{2} C(T)(t_2 - t_1) \quad (4-4)$$

Mit Hilfe der Kalibrierungswerte wird der Schallweg zu $S = 13,956\text{ mm}$ bei der konstanten Temperatur von 30 °C bestimmt.

Abbildung 4-2 zeigt den Verlauf der Amplituden von den gemessenen Signalen. Die Laufzeiten werden über die in Abbildung 4-2 gekennzeichneten Maxima A und B ermittelt.

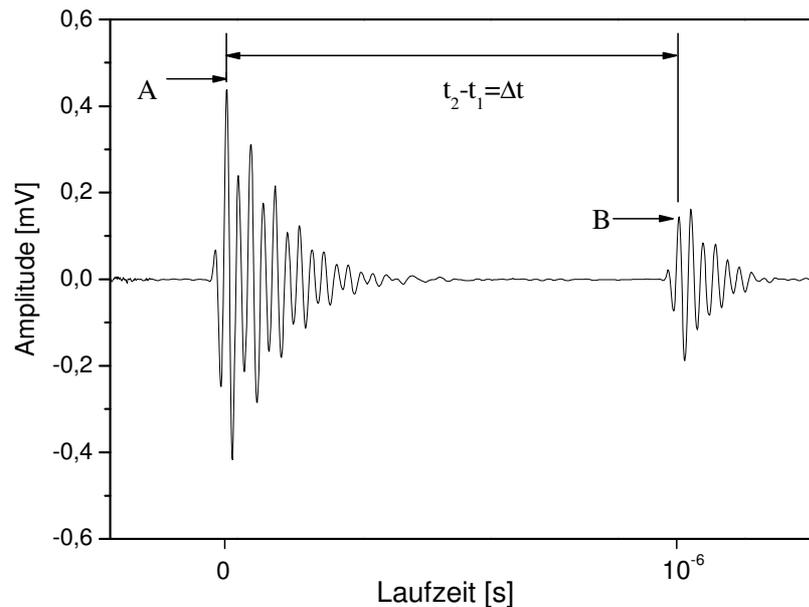


Abbildung 4-2: Messsignal von der gemessenen Amplitude

4.2. Ultraschallgeschwindigkeit im Stop-Flow-Verfahren

Für die Online-Messung des Ultraschallsignals im Bioreaktor entsteht ein Problem durch Blasen in der Kulturbrühe. Im Rahmen dieser Arbeit werden die Kultivierungen mit einer Begasungsrate von 2,25 L/min und einer Rührerdrehzahl von 800 - 1000 UpM betrieben (siehe Tabelle 3-1). Dadurch bilden sich im Reaktor Blasen. In blasenhaltiger Kulturbrühe werden die Ultraschallsignale instabil und der Empfang der Echosignale wird abgeschwächt. Durch den Einsatz des Stop-Flow-Verfahrens wird die Ultraschallgeschwindigkeit ohne Störung durch Blasen stabil gemessen. Mit dem in Abschnitt 3.2 beschriebenen Messaufbau werden bei unterschiedlichen Kultivierungen Ultraschallmesssignale aufgenommen. In Abbildung 4-3 ist der typische zyklische Verlauf dieser Messungen dargestellt. An der Stelle A wird das Ventil geschlossen. An der Stelle B wird das Ventil geöffnet. Die Zeitdauer zwischen A und B beträgt 40 s und zwischen B und A 140 s. Nach Schließen des Ventils sinkt die Ultraschallgeschwindigkeit. In dieser Zeit sank die Temperatur in der Messzelle um ca. 1 °C. Diese Temperaturreduktion entspricht eigentlich einer Erniedrigung der Ultraschallgeschwindigkeit um 2 m/s. Dies wird jedoch von der Systemsoftware der Ultraschallsonde kompensiert, so dass vom Messsystem nur temperaturunabhängige Werte ausgegeben werden. Zum Zeitpunkt B ist das Messsystem optimal blasenfrei. Für die Auswertung der Ultraschallgeschwindigkeit werden diese Messpunkte verwendet.

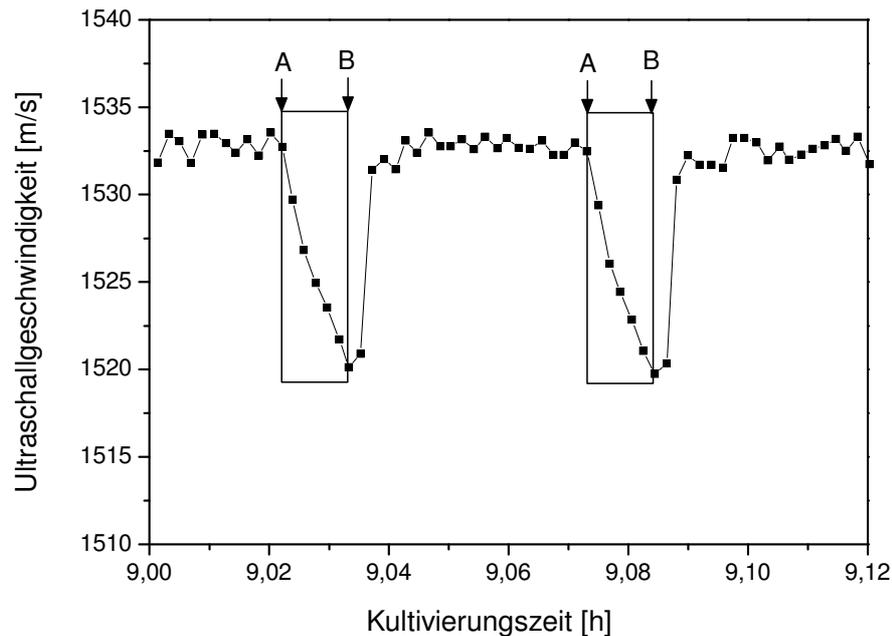


Abbildung 4-3: Ultraschallmesssignal im Stop-Flow-Verfahren

4.3. Ultraschallgeschwindigkeit in Abhängigkeit einzelner Bioprozessgrößen

Das Ziel dieser Arbeit besteht darin, bei Kultivierungen zu überprüfen, ob signifikante Zusammenhänge zwischen der Ultraschallgeschwindigkeit und den interessierenden Prozessgrößen (z.B. Biomasse, Glucose und Ethanol) existieren. Dafür wird zuerst die Abhängigkeit der Ultraschallgeschwindigkeit von den einzelnen Bioprozessgrößen untersucht.

Für die Ultraschallmessungen werden identische Versuchsbedingungen wie bei den Hefekultivierungen gewählt: Temperatur 30 °C, Rührerdrehzahl 800 UpM, pH 5,5 sowie Begasung 2,25 L/min in Schatzmann-Medium.

4.3.1. Ultraschallgeschwindigkeit in Abhängigkeit der Glucosekonzentration

Die Glucosekonzentration beträgt zu Beginn der Hefekultivierung ca. 30 g/L. Daher wird die Ultraschallmessung mit Glucosekonzentrationen im Bereich von 1-40 g/L durchgeführt. Abbildung 4-4 zeigt die Abhängigkeit der Ultraschallgeschwindigkeit von der Glucosekonzentration. Mittels einer linearen Regression werden die Messwerte ausgewertet. Demnach nimmt die Ultraschallgeschwindigkeit mit steigender Glucosekonzentration um

0,37 m·L/s·g zu. Bei der Kultivierung kann die Ultraschallgeschwindigkeit um maximal 11,1 m/s abnehmen, während die 30 g/L Glucose vollständig verbraucht wird.

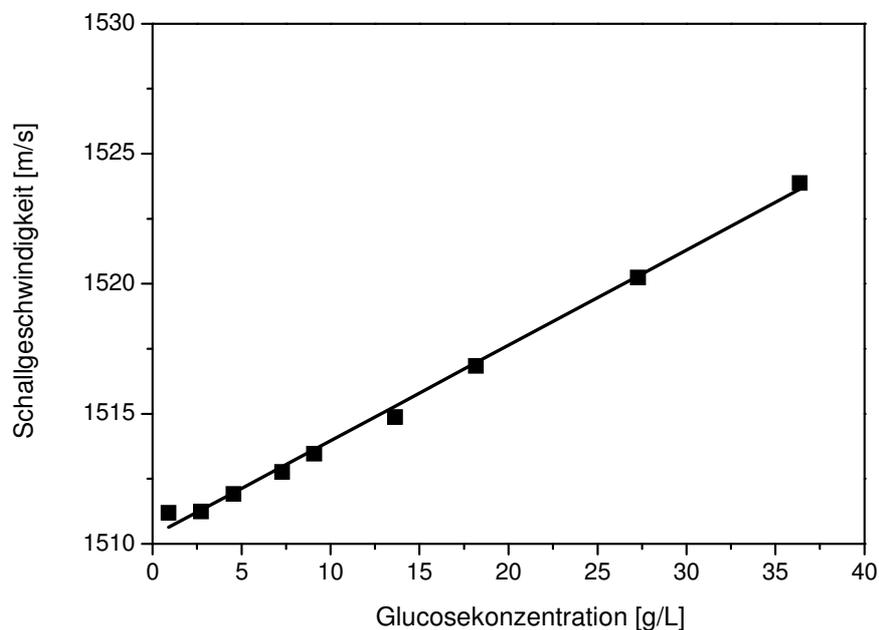


Abbildung 4-4: Ultraschallgeschwindigkeit in Abhängigkeit der Glucosekonzentration

4.3.2. Ultraschallgeschwindigkeit in Abhängigkeit der Ethanolkonzentration

Die Ultraschallmessungen zur Bestimmung der Ethanolkonzentration wurden mit einem Ethanol-Wasser-Gemisch bereits am IFAK durchgeführt [Henning et al., 1998b]. Die Ultraschallgeschwindigkeit steigt im Konzentrationsbereich von 0-30 % Ethanol. Bei über 30 % Ethanolkonzentrationen sinkt die Ultraschallgeschwindigkeit. Bei Hefekultivierungen mit einer Startglucosekonzentration von 30 g/L werden maximal 15 g/L Ethanol produziert. Daher wird hier die Kalibrationsgleichung für die Ethanolkonzentrationen im Bereich von 1-30 g/L verwendet. In Abbildung 4-5 sind die gemessenen Ultraschallgeschwindigkeiten für verschiedene Ethanolkonzentrationen dargestellt. Die Ultraschallgeschwindigkeit steht in einer linearen Abhängigkeit zur Ethanolkonzentration mit einer Steigung von 0,64 m·L/s·g.

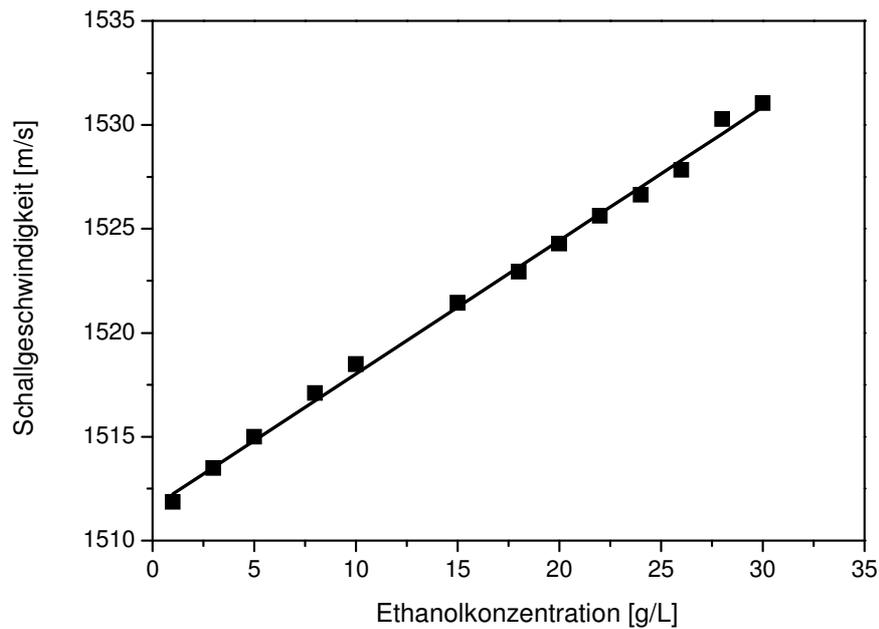


Abbildung 4-5: Ultraschallgeschwindigkeit in Abhängigkeit der Ethanolkonzentration

4.3.3. Ultraschallgeschwindigkeit in Abhängigkeit der Biomassekonzentration

Für die Bestimmung der Biomassekonzentration wird normale Bäckerhefe im Kulturmedium verwendet. Untersuchungen zum Einfluss der Biomassekonzentration auf die Ultraschallgeschwindigkeit zeigen, dass eine Biomassezunahme eine Erhöhung der Ultraschallgeschwindigkeit hervorruft. Abbildung 4-6 zeigt eine lineare Regression der Messwerte mit einer Steigung von 1,09 m·L/s·g.

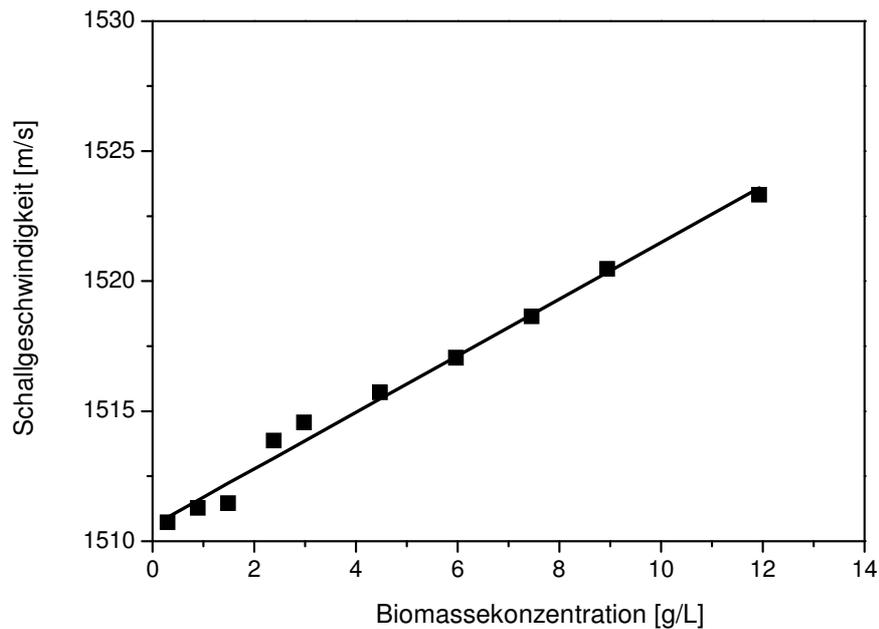


Abbildung 4-6: Ultraschallgeschwindigkeit in Abhängigkeit der Biotrockenmassekonzentration

Zusammenfassend dient die Einzelkomponentenmessung dazu, die Zusammenhänge zwischen den Bioprozessgrößen Biomasse-, Glucose- und Ethanolkonzentration und der Schallgeschwindigkeit zu ermitteln. Dies ist die Grundlage für das Verständnis und die Einschätzung der Ergebnisse der Online-Ultraschallmessungen. In Tabelle 4-1 sind die Abhängigkeiten der Ultraschallgeschwindigkeit von den einzelnen Bioprozessgrößen dargestellt. Die Ultraschallgeschwindigkeit steigt mit zunehmender Biomasse-, Glucose- und Ethanolkonzentration. Diese Untersuchung zeigen, dass die Ultraschallgeschwindigkeit bei Änderungen der Biomassekonzentration am stärksten steigt.

Tabelle 4-1: Regressionsparameter für die lineare Abhängigkeit der einzelnen Bioprozessgrößen

Bioprozessgrößen	Steigung [m·L/s·g]	Achsenabschnitt [m/s]
Biotrockenmasse	1,09	1510,6
Glucose	0,37	1510,3
Ethanol	0,64	1511,6

4.4. Typische Ultraschallmesssignale bei der Hefekultivierung

4.4.1. Batch-Kultivierung

Der typische Verlauf der Ultraschallgeschwindigkeit und der CO₂-Abgaswerte ist in Abbildung 4-7 dargestellt. Im normalen Betrieb des Bypasssystems wurden in dem Zustand „Flow“ und „Stop-Flow“ Messwerte aufgenommen. Abbildung 4-7 zeigt deutlich das in Kapitel 4.2 beschriebene Absinken der Ultraschallgeschwindigkeit um ca. 10 m/s im „Stop-Flow-Zustand“. Bei dem hier gezeigten Prozess handelt es sich um eine Hefekultivierung im Batch-Betrieb. Die Hefekultivierung kann in drei Phasen eingeteilt werden, die für den Prozess typisch sind [Locher et al., 1993]. Am Anfang wird ca. neun Stunden lang die gesamte Glucose verstoffwechselt, es wird Ethanol und Biomasse produziert (siehe Abbildung 4-8). In dieser Zeit bleibt die Ultraschallgeschwindigkeit nahezu konstant. Danach wird das gebildete Ethanol abgebaut. In dieser zweiten Phase erniedrigt sich die Ultraschallgeschwindigkeit. In der dritten Phase stimmen die Ultraschallgeschwindigkeitssignale mit dem Acetat-Abbau überein. Dies wird ebenfalls in Abbildung 4-8 deutlich, in der die Messwerte der anderen Bioprozessgrößen dargestellt sind.

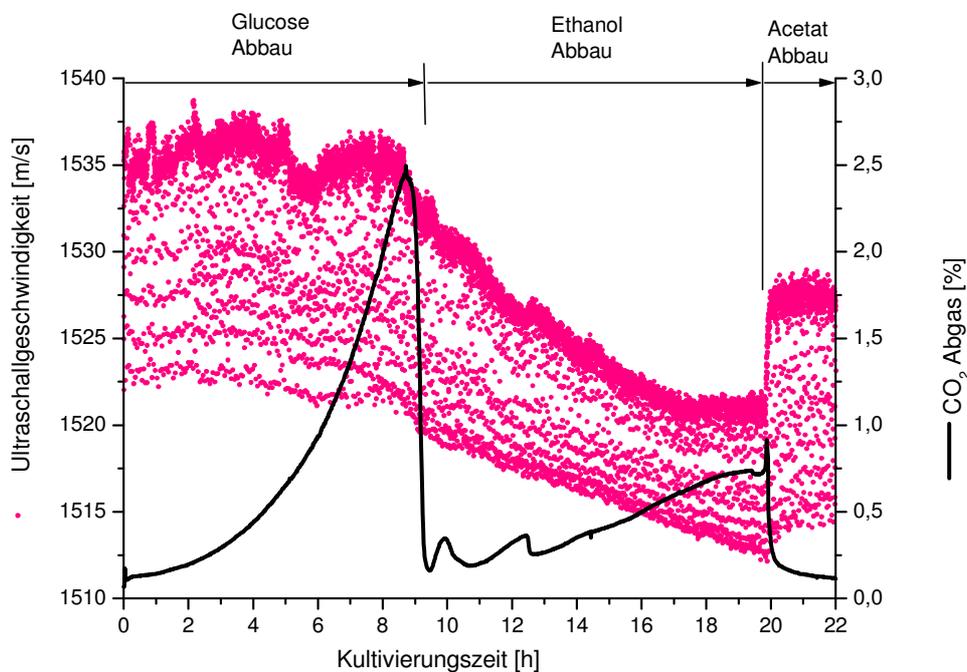


Abbildung 4-7: Ultraschallgeschwindigkeit im Vergleich zu CO₂-Abgaswerten

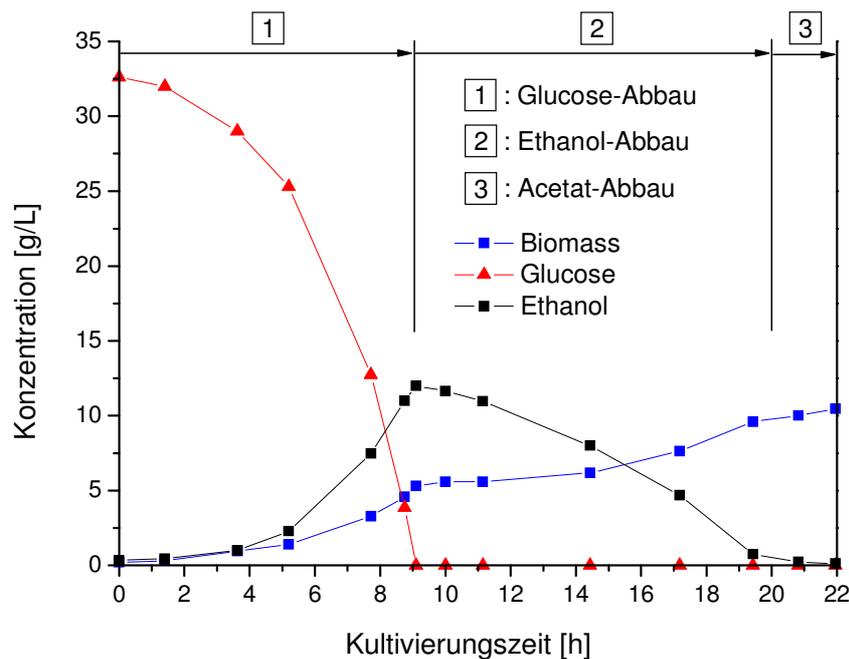


Abbildung 4-8: Offline-Messwerte von Biomasse-, Glucose- und Ethanolkonzentration

4.4.2. Fed-Batch Kultivierung

Abbildung 4-9 zeigt die systematische Korrelation der Ultraschallgeschwindigkeit bei einer Hefekultivierung im Fed-Batch-Betrieb. Diese Kultivierung (Fed-Batch 11) wird mit einer Glucosekonzentration von 2,5 g/L gestartet. Nach der Batch-Phase (ca. 3,5 h) wird eine Glucosekonzentration von 0,5 g/L erreicht. Zu diesem Zeitpunkt wird der Fed-Batch-Betrieb begonnen und die Glucosekonzentration geregelt. Die Glucosekonzentration in der Feedlösung (Schatzmannmedium) beträgt 50 g/L. Die Ultraschallgeschwindigkeit steigt, solange die Glucose zugegeben wird. Danach sinkt die Ultraschallgeschwindigkeit in der Ethanolabbauphase. Mittels der Ultraschallgeschwindigkeit kann der Zeitpunkt, der das Ende der der Glucosezugabe markiert, deutlich erkannt werden.

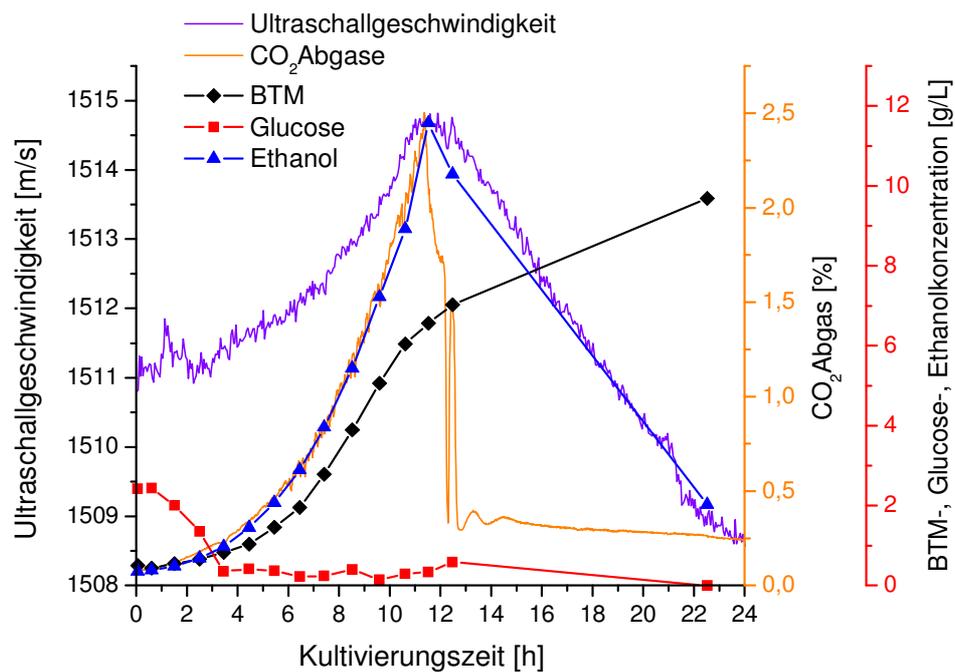


Abbildung 4-9: Ultraschallgeschwindigkeit bei einer Hefekultivierung im Fed-Batch-Betrieb (Fed-Batch 11) im Vergleich zu den CO₂-Abgaswerten und den offline gemessenen Bioprozessgrößen

5. Multilineare Auswertung der Ultraschallgeschwindigkeit

In diesem Kapitel werden die Ergebnisse einer multilineareren Regression (MLR), eines sogenannten „Software-Sensors“ zur Auswertung der Ultraschallgeschwindigkeitsmesswerte und Bioprozessgrößen vorgestellt. Zunächst wird untersucht, ob eine multilineare Abhängigkeit der Ultraschallgeschwindigkeit von den Bioprozessgrößen gefunden werden kann. Dazu wird zur Berechnung der Ultraschallgeschwindigkeit ein multilineares Regressionsmodell mit den Bioprozessgrößen Biomasse-, Glucose- und Ethanolkonzentration aufgestellt.

$$V_{US} = K_0 + K_{BTM} C_{BTM} + K_{Glucose} C_{Glucose} + K_{Ethanol} C_{Ethanol} \quad (5-1)$$

Hierbei ist V_{US} die Ultraschallgeschwindigkeit, K_0 der Achsenabschnitt der Ultraschallgeschwindigkeit und sind K_{BTM} , $K_{Glucose}$ und $K_{Ethanol}$ die Koeffizienten des multilineareren Regressionsmodells. Der Index BTM steht für die Biotrockenmasse. C_{BTM} , $C_{Glucose}$ und $C_{Ethanol}$ sind die Konzentration der jeweiligen Bioprozessgröße.

Tabelle 5.1 zeigt die berechneten Koeffizienten. Abbildung 5-1 stellt sowohl die berechneten als auch die gemessenen Ultraschallgeschwindigkeiten dar. Der mittlere Fehler (Standardabweichung) der Vorhersage des multilineareren Regressionsmodells beträgt 1,09 m/s.

Tabelle 5-1: Regressionskoeffizienten des multilineareren Regressionsmodells

Koeffizienten	Werte
K_0 [m/s]	1510
K_{BTM} [m·L/s·g]	0,51
$K_{Glucose}$ [m·L/s·g]	0,47
$K_{Ethanol}$ [m·L/s·g]	0,52

Für die Fehlerberechnung wird die Standardabweichung (MEP_x) verwendet.

$$MEP_x = \sqrt{\frac{\sum_{i=1}^N (\hat{Y}_i - Y_i)^2}{N-1}} \quad (5-2)$$

Dabei sind MEP_x die Standardabweichungen der Vorhersage (mean error of prediction) mit dem Index der Variablen x , der Vorhersagewerter \hat{Y} sowie den Messwerten bzw. simulierten Werten Y_i . N ist die Anzahl der Messungen von x .

Ein Vergleich der Werte von Tabelle 5-1 mit den Ultraschallabhängigkeiten von einzelnen Bioprozessgrößen (in Tabelle 4-1 dargestellt) zeigt, dass der Regressionskoeffizient für die Glucose im multilinearen Modell etwas größer ist (Glucose linear: 0,37 m·L/s·g, Glucose multilinear: 0,47 m·L/s·g). Hingegen hat sich der Wert für Biomasse und Ethanol reduziert (Biomasse linear: 1,09 m·L/s·g, Biomasse multilinear: 0,51 m·L/s·g; Ethanol linear: 0,64 m·L/s·g, Ethanol multilinear: 0,52 m·L/s·g). Da die Ultraschallgeschwindigkeit von verschiedenen Kultivierungen auch zu Beginn der Messung sehr unterschiedlich ist (Hefe 23: 1527 m/s, Hefe 22: 1525 m/s, Hefe 27: 1523 m/s, Hefe 30: 1520 m/s), wird der Wert K_0 für die jeweiligen Kultivierungen korrigiert. Dieser Wert wird für jede Kultivierung mit dem Anfangswert der Ultraschallgeschwindigkeit abgeglichen.

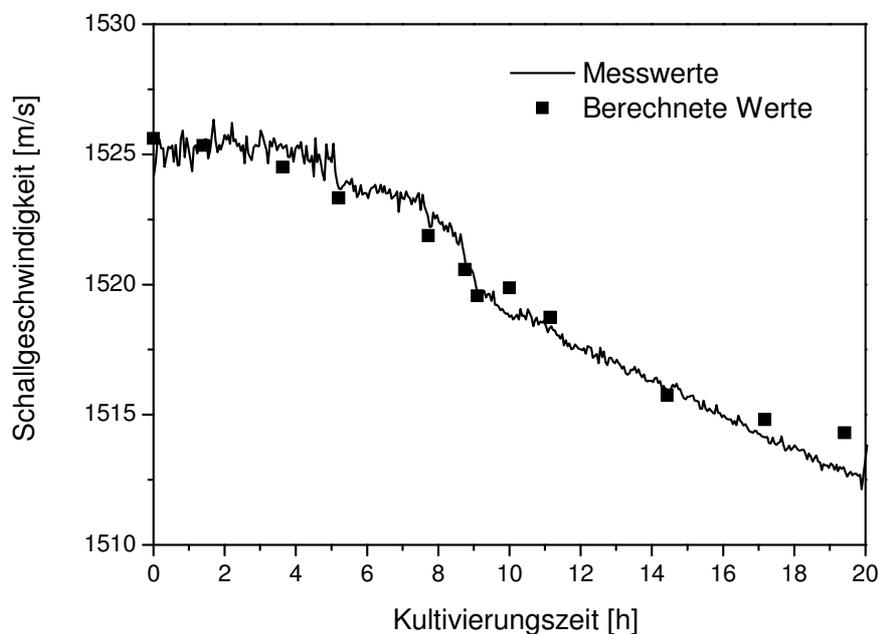


Abbildung 5-1: Vorhersage der Ultraschallgeschwindigkeit aus Biotrockenmasse, Glucose und Ethanol mit einem multilinearen Regressionsmodell, dessen Werte zur Berechnung des multilinearen Regressionsmodells verwendet wurden

Das erstellte Regressionsmodell kann auch dazu verwendet werden, eine Bioprozessgröße vorherzusagen, wenn die anderen Größen bekannt sind. So kann beispielsweise die Biomasse vorhergesagt werden, wenn Glucose, Ethanol und die Ultraschallgeschwindigkeit bekannt sind. Für diese Vorhersagen werden die in Tabelle 5-2 aufgeführten Gleichungen verwendet.

Tabelle 5-2: Gleichungen von Biotrockenmasse-, Glucose- und Ethanolkonzentration

Bioprozessgrößen	Gleichungen
Biotrockenmasse	$C_{BTM} = \frac{1}{K_{BTM}} [V_{US} - K_0 - K_{Glucose} C_{Glucose} - K_{Ethanol} C_{Ethanol}]$
Glucose	$C_{Glucose} = \frac{1}{K_{Glucose}} [V_{US} - K_0 - K_{BTM} C_{BTM} - K_{Ethanol} C_{Ethanol}]$
Ethanol	$C_{Ethanol} = \frac{1}{K_{Ethanol}} [V_{US} - K_0 - K_{BTM} C_{BTM} - K_{Glucose} C_{Glucose}]$

Abbildung 5-2 zeigt die Vorhersage der einzelnen Bioprozessgrößen Biotrockenmasse-, Glucose- und Ethanolkonzentration für den Bioprozess, dessen Werte zur Berechnung des Regressionsmodells dienen. Die jeweiligen mittleren Fehler der Vorhersage (Hefe 23) sind in Tabelle 5-3 dargestellt.

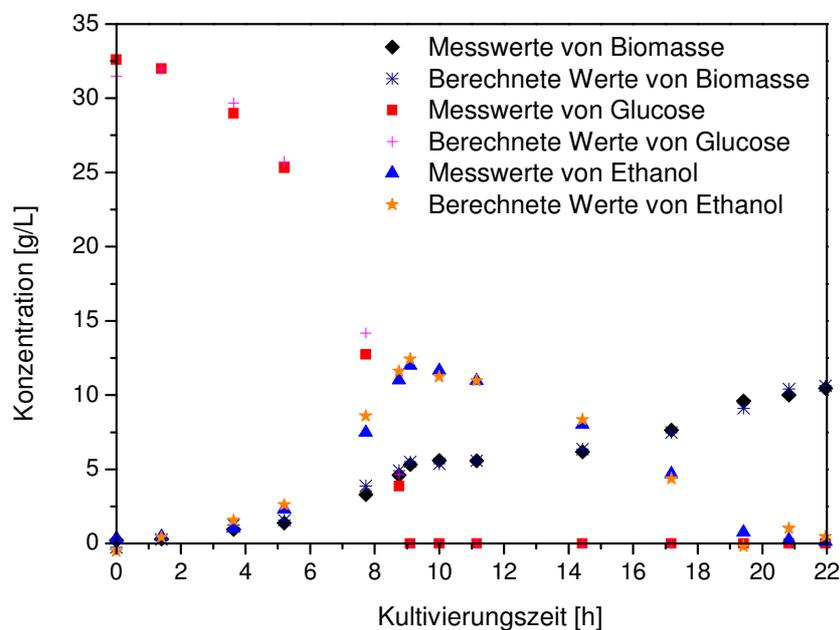


Abbildung 5-2: Messwerte und berechnete Werte des Prozesses, der zur Berechnung des multilinearen Regressionsmodells herangezogen wird

Die Validierung dieses Modells mit Daten einer anderen Kultivierung (Hefe 22) zeigt Abbildung 5-3. Die jeweiligen mittleren Fehler sind in Tabelle 5-3 mit den Ergebnissen weiterer Prozesse (Hefe 27, Hefe 30) dargestellt.

Tabelle 5-3: Mittlere Fehler der Vorhersagewerte von Bioprozessgrößen

Mittlere Fehler	Biomasse [g/L]	Glucose [g/L]	Ethanol [g/L]
Hefe 23 Modellbildung	0,3	0,9	0,6
Hefe 22 Validierung	1,0	2,0	2,0
Hefe 27 Validierung	1,0	1,3	2,0
Hefe 30 Validierung	1,1	2,7	2,1
Mittelwerte Validierung	1,0	2,0	2,0

Die Werte von verschiedenen Validierungen weichen um etwa das Doppelte bis Dreifache von den Fehlern der Modellbildung ab. Die Ursache ist die geringe Anzahl an Messwerten der Bioprozessgrößen. Um die Fehler zu minimieren, werden in dem folgenden Kapitel simulierte Werte der Bioprozessgrößen mit einem Modell berechnet und verschiedene Analyseverfahren wie die multilineare Regression, die Hauptkomponentenanalyse, das Kalman-Filter und das neuronale Netz verwendet.

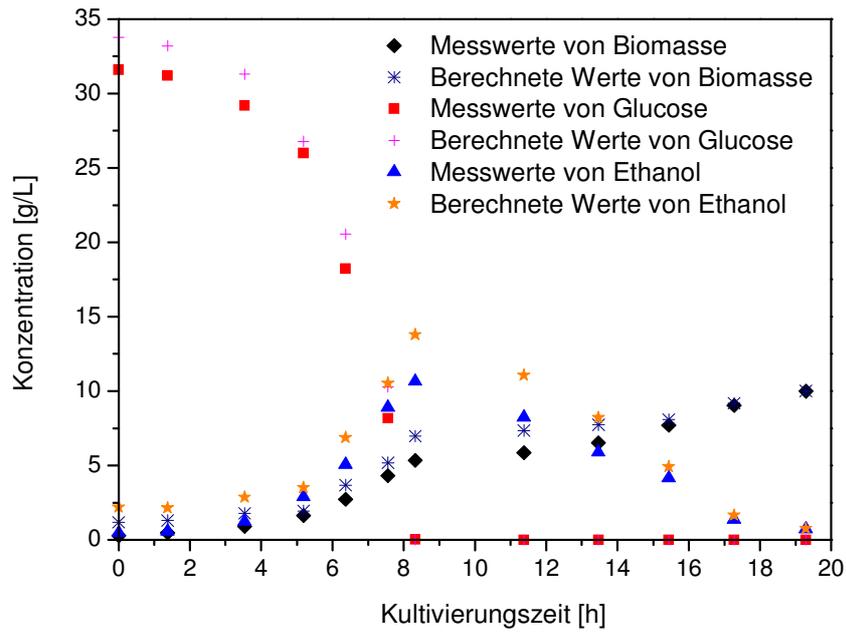


Abbildung 5-3: Messwerte und die vorhergesagten Bioprozessgrößen einer Kultivierung (Hefe 22), deren Daten nicht zur Berechnung des Regressionsmodells verwendet wurden

6. Prozessmodellierung der Hefekultivierung

6.1. Mathematische Modellbildung

Aus den Abbildungen 5-2 bis 5-4 wird ersichtlich, dass das multilineare Regressionsmodell nicht in allen Bereichen eine korrekte Vorhersage liefert. Insbesondere gibt die Vorhersage der Bioprozessgrößen den Glucoseabbau nicht korrekt wieder (siehe Abbildung 5-3). Dies liegt unter anderem an dem begrenzten Datenmaterial. Um die Vorhersage zu verbessern, werden mathematische Modelle für die Bioprozessgrößen Biomasse-, Glucose- und Ethanolkonzentration aufgestellt, mit denen Prozessdaten simuliert werden können. Dadurch werden sowohl die nicht systematischen Messfehler eliminiert, als auch das Problem der mangelnden Messpunkte gelöst.

Das Prozessmodell beschreibt das Wachstum von Hefe unter aeroben Bedingungen im Batch-Reaktor. Es wird sowohl das oxidative-fermentative Wachstum auf Glucose, als auch das oxidative Wachstum auf Ethanol erfasst. Im Folgenden wird die spezifische Wachstumsrate für das Wachstum auf Glucose beschrieben:

$$\mu_S = \mu_S^{\max} \frac{S}{S + K_S} \quad (6-1)$$

In dieser Monod Gleichung ist μ_S^{\max} die maximale Wachstumsrate und K_S die Limitierungskonstante.

Das Wachstum auf Ethanol ergibt sich durch:

$$\mu_E = \mu_E^{\max} \frac{E}{E + K_E} \quad (6-2)$$

Für die spezifische Wachstumsrate beim Wachstum auf zwei Metaboliten werden die Wachstumsraten auf Glucose und auf Ethanol addiert, wobei μ_E^{\max} solange Null ist, wie die Glucosekonzentration größer als Null ist.

$$\mu_E^{\max} = \begin{cases} 0 & \text{für } S > 0 \\ \mu_{E_0}^{\max} & \text{für } S = 0 \end{cases} \quad (6-3)$$

$$\mu = \mu_S + \mu_E \quad (6-4)$$

Der zeitliche Konzentrationsverlauf der Biotrockenmasse wird durch folgende Differentialgleichung beschrieben:

$$\text{Biomasse (X)} : \frac{dX}{dt} = \mu X \quad (6-5)$$

Die Differentialgleichung für die Glucosekonzentration ist:

$$\text{Glucose (S)} : \frac{dS}{dt} = -\frac{\mu_S}{Y_{XS}} X \quad (6-6)$$

Dabei ist Y_{XS} der Substratausbeutekoeffizient für die Bildung der Biomasse aus Glucose. Beim fermentativen Wachstum auf Glucose wird Ethanol produziert und danach beim oxidativen Wachstum auf Ethanol verbraucht. Den Prozessverlauf der Ethanolkonzentration beschreibt folgende Gleichung:

$$\text{Ethanol (E)} : \frac{dE}{dt} = (\mu_S Y_{EX} - \frac{\mu_E}{Y_{XE}}) X \quad (6-7)$$

Dabei ist Y_{XE} der Ausbeutekoeffizient für die Bildung von Biomasse aus Ethanol und Y_{EX} der Ausbeutekoeffizient für die Bildung von Ethanol aus Glucose.

Die optimierten Parameterwerte sind in Tabelle 6-1 dargestellt. Sie werden zur Simulation des Prozesses verwendet.

Tabelle 6-1: Modellparameter der Hefekultivierung

Größe	Wert	Beschreibung
K_S [g/L]	0,4	Limitierungskonstante für Glucose
K_E [g/L]	0,05	Limitierungskonstante für Ethanol
μ_S^{\max} [1/h]	0,3	Maximale Wachstumsrate auf Glucose
μ_E^{\max} [1/h]	0,06	Maximale Wachstumsrate auf Ethanol
Y_{XS} [gBTM/gGlucose]	0,12	Ausbeutekoeffizient für Biomasse aus Glucose
Y_{XE} [gBTM/gEthanol]	0,4	Ausbeutekoeffizient für Biomasse aus Ethanol
Y_{EX} [gEthanol/gBTM]	3,5	Ausbeutekoeffizient für Ethanol aus Glucose

6.2. Berechnung und Validierung von Modellparametern

Für die Simulation und Berechnung der Parameter sind die Anfangswerte der Bioprozessgrößen notwendig. Da es bei jeder Kultivierung jedoch andere Anfangsbedingungen gibt,

werden unterschiedliche Anfangswerte aufgestellt. Die zugehörigen Anfangswerte und die Anzahl der Messungen sind in Tabelle 6-2 aufgeführt. Diese Werte werden wie die Parameter aus den Messdaten ermittelt. Abbildung 6-1 zeigt die Ergebnisse der Kultivierung (Hefe 23), deren Daten für die Parameterberechnung verwendet werden.

Tabelle 6-2: Anfangswerte der Bioprozessgrößen und Anzahl der Messungen

Kultivierungen	Anzahl der Messungen	Biomasse [g/L]	Glucose [g/L]	Ethanol [g/L]
Hefe 23 Modellbildung	433	0,20	32,60	0,35
Hefe 22 Modellvalidierung	393	0,27	31,60	0,47
Hefe 27 Modellvalidierung	393	0,26	30,00	0,10
Hefe 30 Modellvalidierung	433	0,25	29,40	0,30

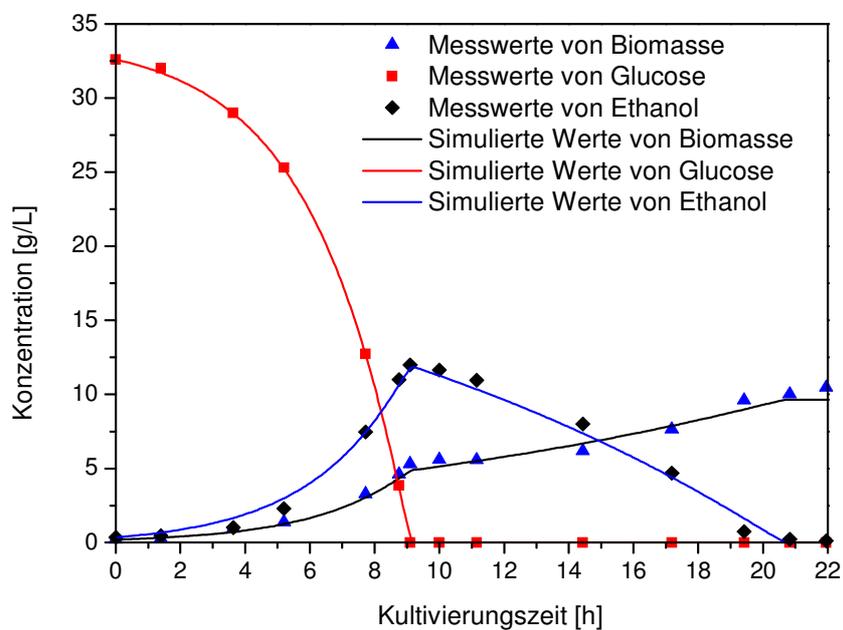


Abbildung 6-1: Vergleich der simulierten Bioprozessgrößen mit den Messwerten, mit denen die Modellparameter berechnet wurden

Die jeweiligen mittleren Fehler sind 0,5 g/L für Biomassekonzentration, 0,1 g/L für Glucosekonzentration und 0,5 g/L für Ethanolkonzentration (siehe Tabelle 6-3). Mit diesen Parameterwerten werden weitere Prozesse simuliert, die mit einer Startglucosekonzentration von 31,6 g/L (Hefe 22), 30,0 g/L (Hefe 27) und 29,4 g/L (Hefe 30) durchgeführt werden (siehe Tabelle 6-2). Die erhaltenen Fehler sind ebenfalls in der Tabelle 6-3 aufgeführt.

Tabelle 6-3: Mittlere Fehler der simulierten Werte bezüglich der Messwerte

Kultivierung	Biomasse [g/L]	Glucose [g/L]	Ethanol [g/L]
Hefe 23 Modellbildung	0,5	0,1	0,5
Hefe 22 Validierung	0,4	1,7	1,1
Hefe 27 Validierung	0,3	0,2	0,7
Hefe 30 Validierung	0,3	0,4	0,9
Mittelwerte Validierung	0,3	0,7	0,9

Die simulierten Werte sind beispielhaft für eine Kultivierung (Hefe 22) in Abbildung 6-2 dargestellt. Die jeweiligen Mittelwerte der mittleren Fehler sind 0,3 g/L für die Biomassekonzentration, 0,7 g/L für Glucosekonzentration und 0,9 g/L für die Ethanolkonzentration (siehe Tabelle 6-3). Der Fehler der Biomassekonzentration ist kleiner als der Fehler, der bei der Parameterberechnung ermittelt wurde. Hingegen ist der Fehler für Ethanolkonzentration hier doppelt so groß. Die Anpassung für die Glucosekonzentration weicht stark von den Messwerten ab, woraus ein sehr großer Fehler folgt. Im nächsten Kapitel werden die Auswertungen mit Hilfe der verschiedenen Auswerteverfahren aus den Datensätzen von simulierten Werten und online gemessenen Ultraschallgeschwindigkeiten durchgeführt.

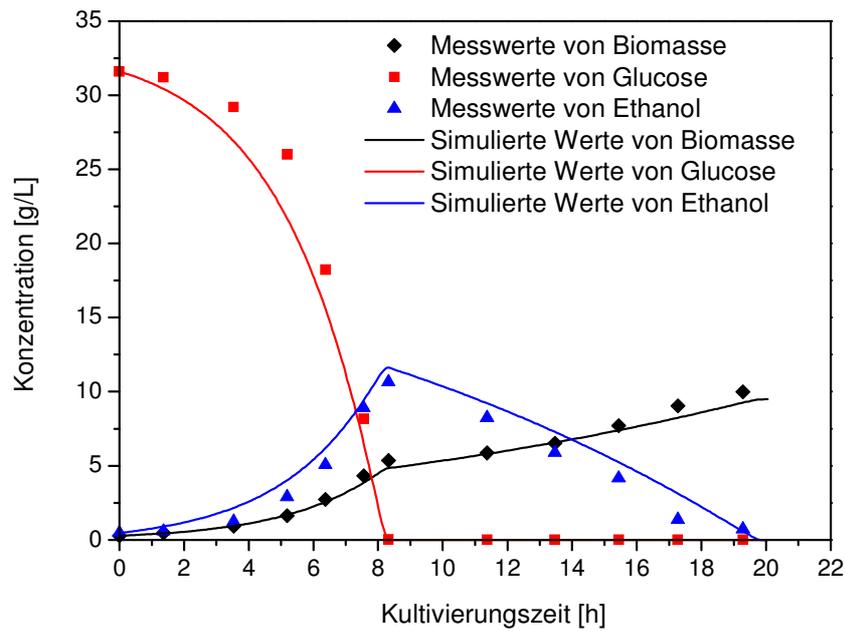


Abbildung 6-2: Die simulierten Bioprozessgrößen im Vergleich zu den Messwerten, die nicht zur Berechnung der Modellparameter dienen (Hefe 22)

7. Vorhersage der Prozessgrößen mit diversen Software-Sensoren

In diesem Kapitel wird die Vorhersage der Bioprozessgrößen mit verschiedenen Analyseverfahren (Software-Sensoren) behandelt. Mit dem kinetischen Modell (Monod-Kinetik) werden Bioprozessgrößen wie Biomasse und Glucose sowie Ethanol simuliert, wie in Kapitel 6 beschrieben. Um die Vorhersage dieser Bioprozessgrößen zu verbessern, werden in diesem Kapitel die simulierten Werte und die online gemessenen Bioprozessgrößen sowie die Ultraschallgeschwindigkeiten zur Berechnung verwendet. Für die Modellbildung und Validierungen werden vier Kultivierungsdatensätze (Hefe 22, Hefe 23, Hefe 27, Hefe 30) benutzt. Da der Fehler der Modellanpassung bei den Daten von Hefe 23 am geringsten ist, dient diese Kultivierung als Referenz. Die anderen Kultivierungen (Hefe 22, Hefe 27 und Hefe 30) werden zur Validierung verwendet. Allerdings wird zur Vorhersage der Ultraschallgeschwindigkeit der Kultivierungsdatensatz (Fed-Batch 2) genutzt. Bei der Berechnung dieser Vorhersage werden zusätzlich 2D-Fluoreszenzspektren verwendet.

Zur Auswertung der Kultivierungsdaten werden unterschiedlichen Verfahren genutzt. Abbildung 7-1 stellt eine Übersicht der angewandeten Auswerteverfahren dar. Für die Auswertung werden zweierlei Datensätze, nämlich Messwerte und simulierte Werte, verwendet. Zur Bewertung wird jeweils der Fehler der Vorhersage bezogen auf die Offline-Messwerte (V-M) und, soweit verfügbar, bezogen auf die Simulation (V-S) berechnet. Da die Ergebnisse der Auswertung von Ultraschallgeschwindigkeitsmesswerten, die unter entgasten Bedingungen (Minimum der Messwerte bei geschlossenem Ventil) aufgezeichnet werden, besser sind als die, die mit offenem Ventil erzielt werden, werden nur die Ergebnisse unter entgasten Bedingungen im Folgenden weiter behandelt. Die hier verwendeten Bioprozessgrößen sind die Biotrockenmasse, Glucose und Ethanol, die bei der Kultivierung offline gemessen werden, sowie pO_2 , O_2 und CO_2 im Abgas und zweidimensionale Fluoreszenzmesssignale, die bei der Kultivierung online gemessen werden. Simulierte Werte werden mit einem mathematischen Modell der Kultivierung berechnet (siehe Kapitel 6.1).

Um die Ultraschallgeschwindigkeit sowie die Biomasse-, Glucose- und Ethanolkonzentration vorhersagen zu können, werden verschiedene Verfahren wie die multilineare Regression (MLR), die Hauptkomponentenregression (PCR), das Kalman-Filter und das neuronale Netz verwendet.

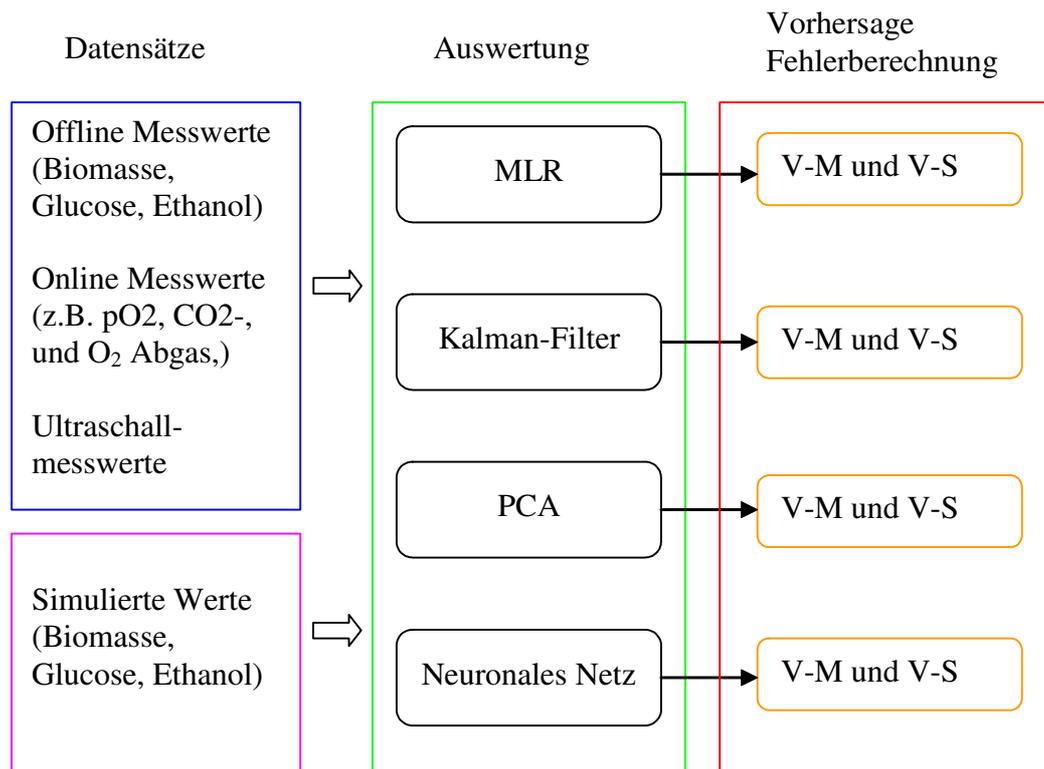


Abbildung 7-1: Übersicht über die Auswerteverfahren

Durch die Modellanpassung der Daten der Hefekultivierungen wird zusätzliches Datenmaterial erzeugt, das zur weiteren Auswertung genutzt werden kann. Da mehr simulierte Werte als online gemessene Ultraschallwerte vorliegen, ist eine Vorverarbeitung der Datensätze notwendig. Mit einem Delphi-Programm werden den Ultraschallwerten die entsprechenden simulierten Werte zugeordnet. Dies gilt auch für die im Folgenden noch beschriebenen Auswertungen mittels PCR, Kalman-Filter und neuronaler Netze.

7.1. Vorhersage der Prozessgrößen mittels multilinearer Regression

Mit Hilfe der simulierten Bioprozesswerte (von Hefe 23) wird für die Ultraschallgeschwindigkeit ein multilineares Regressionsmodell entsprechend dem in Kapitel 5 vorgestellten berechnet. Die multilinearen Regressionskoeffizienten sind in Tabelle 7-1 dargestellt. Diese Werte weichen von den in Tabelle 5-1 dargestellten Werten ab. Bei der multilineareren Regression von simulierten Werten ist der Wert $K_0=1457,55$ m/s optimal. Er ist damit deutlich kleiner als der entsprechende Wert, der nur mit Messwerten ermittelt wird. Dagegen sind die multilinearen Regressionskoeffizienten deutlich größer.

Tabelle 7-1: Multilineare Regressionskoeffizienten ermittelt mit simulierten Prozesswerten

Koeffizienten	Werte
K_0 [m/s]	1457,55
K_{BTM} [m·L/s·g]	5,78
K_{Glucose} [m·L/s·g]	2,03
K_{Ethanol} [m·L/s·g]	2,77

Dieses Modell wird zur Berechnung der Ultraschallgeschwindigkeit genutzt. Das Ergebnis ist in Abbildung 7-2 dargestellt. Der mittlere Fehler der Vorhersage der Ultraschallgeschwindigkeit beträgt 1,0 m/s und ist somit identisch mit dem Wert, der nur mit Messwerten ermittelt wird (siehe Kapitel 5).

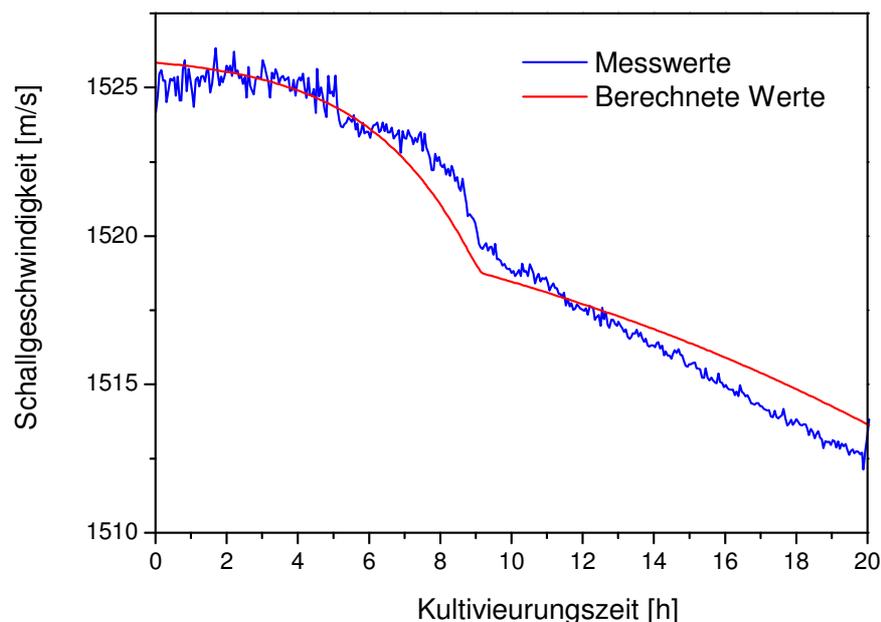


Abbildung 7-2: Gemessene und berechnete Ultraschallgeschwindigkeiten einer Kultivierung (Hefe 23)

Das Modell kann auch wie in Kapitel 5 zur Berechnung einer Prozessgröße (z.B. Ethanol) verwendet werden, wenn die beiden anderen Prozessgrößen (z.B. Biomasse- und Glucosekonzentration) sowie die Ultraschallgeschwindigkeit bekannt sind. Abbildung 7-3 zeigt die vorhergesagten Bioprozessgrößen im Vergleich zu den gemessenen und simulierten Werten. Die simulierten und vorhergesagten Werte stimmen sehr gut mit den Messwerten überein.

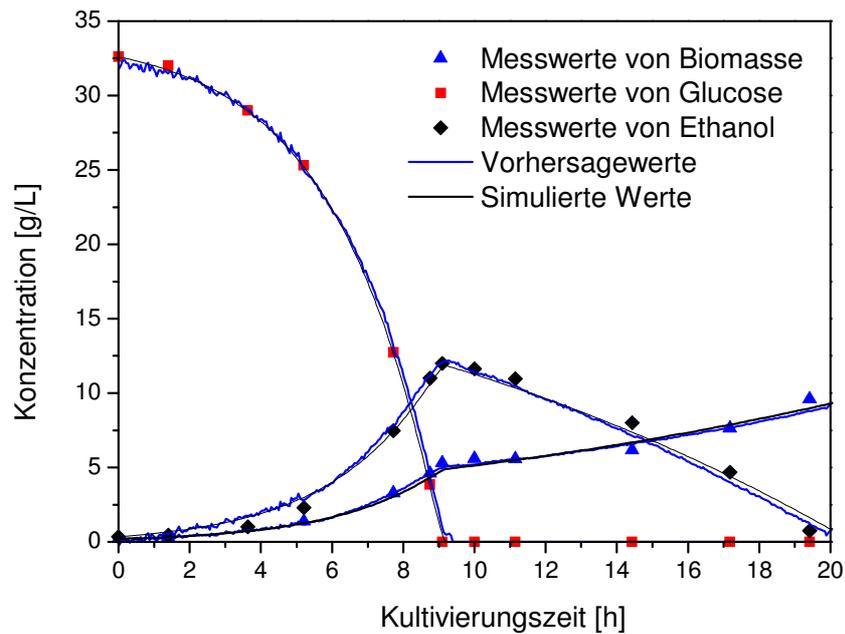


Abbildung 7-3: Gemessene und simulierte Werte im Vergleich zu den vorhergesagten Bioprozessgrößen einer Kultivierung (Hefe 23)

Um die Güte des Modells zu bewerten, wird eine externe Validierung mit drei weiteren Kultivierungen (Hefe 22, Hefe 27 und Hefe 30) durchgeführt. Beispielhaft sind in Abbildung 7-4 die Messwerte und die vorhergesagten Bioprozessgrößen aus simulierten Werten für Hefe 22 dargestellt. Die Vorhersage zeigt beim Biomasseverlauf eine gute Übereinstimmung mit den Messwerten, dagegen liegt der Glucoseverlauf leicht unter und der Ethanolverlauf etwas über den Messwerten.

Die jeweiligen mittleren Fehler der Vorhersage der Bioprozessgrößen bezüglich der gemessenen und simulierten Werte sind in Tabelle 7-2 dargestellt. Die mittleren Fehler V-S der Kultivierung von Hefe 23, die zur Modellanpassung verwendet wurden, sind 0,2 g/L für die Biomasse, 0,3 g/L für Glucose und 0,4 g/L für Ethanol. Im Vergleich dazu ist der Fehler V-M bei allen Prozessvariablen größer. Dieser Fehler ist ebenfalls kleiner als der Fehler der Modellanpassung mit den Messwerten (Tabelle 5-3). Der mittlere Fehler der Validierung wurde ebenfalls durch die Verwendung von simulierten Werten reduziert. Insgesamt wurde die Vorhersage der Prozessgrößen durch die Verwendung simulierter Werte verbessert, weshalb für die weiteren Auswertungen die simulierten Werte verwendet werden.

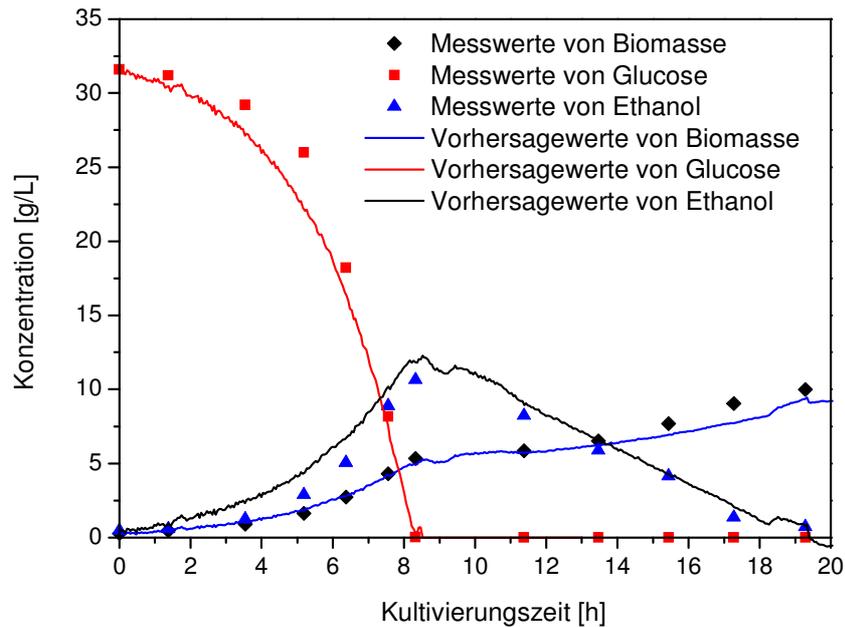


Abbildung 7-4: Messwerte und vorhergesagte Bioprozessgrößen aus simulierten Werten einer Kultivierung (Hefe 22)

Tabelle 7-2: Mittlere Fehler der Vorhersage der Bioprozessgrößen bezüglich der gemessenen und simulierten Werte

Mittlere Fehler	Biomasse [g/L]		Glucose [g/L]		Ethanol [g/L]	
	V-M	V-S	V-M	V-S	V-M	V-S
Hefe 23 Modellbildung	0,3	0,2	0,5	0,3	0,6	0,4
Hefe 22 Validierung	0,5	0,3	1,6	0,4	1,1	0,6
Hefe 27 Validierung	1,9	0,6	2,6	0,7	1,9	1,3
Hefe 30 Validierung	0,7	0,5	1,5	1,4	1,9	1,0
Mittelwerte Validierung	1,0	0,5	1,9	0,8	1,6	1,0

7.2. Hauptkomponentenregression zur Vorhersage der Prozessgrößen

7.2.1. Vorhersage der Ultraschallgeschwindigkeit bei der Fed-Batch-Kultivierung

Bei der durchgeführten Fed-Batch-Kultivierung (Fed-Batch 2) wurden die Zellen anfangs in der Batch-Phase in einem Medium (modifiziertes Schatzmann-Medium mit 10 g/L Glucose) gezüchtet, das ein Volumen von 1 L besaß. Die Dauer der Batch-Phase betrug ca. 13 h. In dieser Zeit verstoffwechselten die Zellen die Glucose sowie das daraus produzierte Ethanol und die organischen Säuren. Nachdem die CO₂-Produktion sowie die O₂-Aufnahme der Zellen ungefähr Null waren, wurde die Fed-Batch-Phase gestartet. Dabei wurde die Glucosepumprate einer 250 g/L Glucose Feedlösung exponentiell gesteigert, um einen Übergang der Hefezellen von der oxidativen zu der oxido-reduktiven Glucoseverstoffwechslung zu bewirken. Dies wurde mit der vorhandenen Analytik aufgezeichnet.

Um zu testen, ob insgesamt die online gemessenen Größen mit der Ultraschallgeschwindigkeit korrelieren, wird eine Hauptkomponentenregression durchgeführt. Dazu werden alle Online-Messgrößen zusammengefasst, die zu einem Zeitpunkt registriert werden und als ein multidimensionaler Messwertvektor für die Hauptkomponentenregression genutzt. Ein 2D-Fluoreszenzspektrum besteht aus 1085 einzelnen Messpunkten. Darüber hinaus werden die Online-Messwerte von pO₂, O₂-, und CO₂-Abgas, pH, Säure sowie Base als Messvariablen benutzt. Daraus resultiert der Messwertvektor mit insgesamt 1091 Komponenten. Für die Hauptkomponentenanalyse werden insgesamt 691 solcher Messwertvektoren verwendet. Zur Berechnung des multilinearen Regressionsmodells werden 5 Hauptkomponenten verwendet, die sich als signifikant herausgestellt haben. Die Messwerte und die Vorhersage der Schallgeschwindigkeit, die mit einer internen Validierung erzeugt werden, sind in Abbildung 7-5 dargestellt. Deutlich ist eine Übereinstimmung der Werte zu erkennen. Dies zeigt, dass die Information der Ultraschallgeschwindigkeit in den Online-Messungen der Prozessgrößen enthalten sein muss. Dabei weichen einzelne Vorhersagewerte von der Messung stark ab. Dies resultiert aus dem Rauschen der Fluoreszenzdaten.

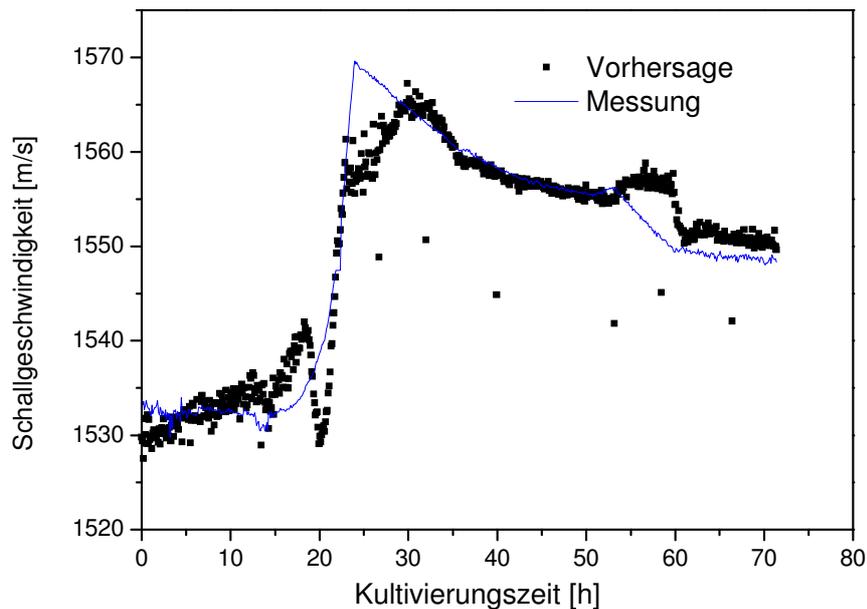


Abbildung 7-5: Messwerte und Vorhersage der Ultraschallgeschwindigkeit aus Fluoreszenzmesdaten und den online gemessenen Prozessgrößen

7.2.2. Vorhersage der Bioprozessgrößen mit einer Hauptkomponentenregression

Die Vorhersage der Bioprozessgrößen Biomasse-, Glucose- sowie Ethanolkonzentration geschieht mittels Hauptkomponentenregression aus drei bekannten Prozessgrößen. Zum Beispiel werden für die Vorhersage der Biomassekonzentration simulierte Werte der Glucose- und Ethanolkonzentration sowie der online gemessene Ultraschallgeschwindigkeit benutzt. Für die Hauptkomponentenregression wird der Unscrambler 7.6 (Camo, Norwegen) verwendet, da mit diesem Softwaretool eine schnelle Berechnung der Modelle möglich ist. Die Bestimmung der Anzahl der Hauptkomponenten erfolgt durch die Analyse des minimalen quadratischen Fehler der Validierung. Die interne Validierung der Modelle erfolgt durch eine Kreuzvalidierung [Otto, 1997]. Zur Modellbildung wird die Kultivierung Hefe 23 herangezogen. Für die Vorhersage werden zwei Hauptkomponenten zur Berechnung der multiliniaren Regression verwendet.

Beispielhaft für die Vorhersagewerte der Biomasse zeigt Abbildung 7-6 die resultierende Vorhersage sowie die simulierten und die Offline-Messwerte der Biomasse-, Glucose-, Ethanolkonzentration und gemessene Ultraschallgeschwindigkeit. Entsprechend wurden auch Modelle für die Vorhersage von Ethanol bzw. Glucose erstellt.

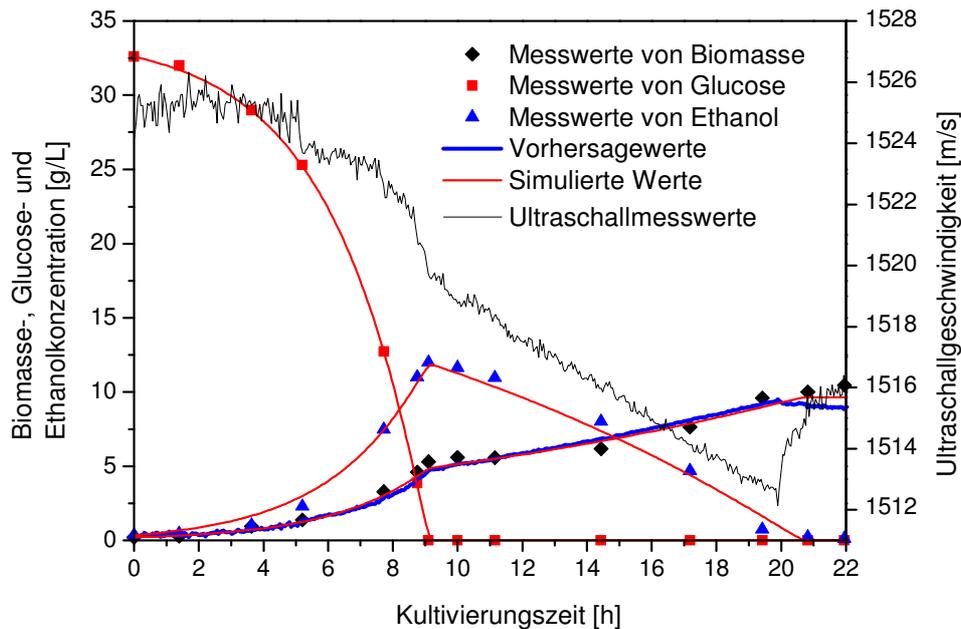


Abbildung 7-6: Vorhersage der Biomasse der Modellkultivierung Hefe 23 aus simulierten Glucose- und Ethanolwerten sowie der Ultraschallgeschwindigkeit

Eine externe Validierung mit den Kultivierungsdaten Hefe 22 ist in Abbildung 7-7 dargestellt. Die Vorhersage der Bioprozessgrößen Biomasse, Glucose und Ethanol werden zusammen mit den Offline- und simulierten Werten gezeigt.

Die mittleren Fehler aller Vorhersagen, sowohl der Modellbildung Hefe 23 als auch der Validierungen (Hefe 22, Hefe 27 und Hefe 30), sind in Tabelle 7-3 dargestellt. Angegeben sind die mittleren Fehler zwischen Offline-Messwerten und vorhergesagten Werten (V-M) und zwischen simulierten Werten und vorhergesagten Werten (V-S). Die mittleren Fehler der Vorhersage (V-S) der Kultivierung Hefe 23 sind 0,2 g/L für Biomasse und jeweils 0,1 g/L für Glucose und Ethanol. Diese sind kleiner als die Fehler der Vorhersage (V-S) von Hefe 23 mittels MLR basierend auf den simulierten Werten (Tabelle 7-2).

Bei der Vorhersage der Bioprozessgrößen sind die Validierungsmittelwerte der mittleren Fehler (V-S) 0,6 g/L für Biomasse, 1,4 g/L für Glucose und 1,0 g/L für Ethanol. Verglichen mit den Fehlern aus Tabelle 7-2 ist der Fehler für Ethanol mit 1,0 g/L gleich geblieben. Die Fehler bei der Vorhersage der Biomasse und der Glucose sind jedoch leicht gestiegen. Diese Fehler der Validierungen mittels PCR können auch mit den Fehlern der Validierung, die mit den Messwerten mittels MRL berechnet wurden (vergl. Tabelle 5-3), verglichen werden. Im Vergleich der Mittelwerte von drei Validierungen ist der mittlere Fehler von 1,0 g/L auf 0,6 g/L für Biomasse, von 2,0 g/L auf 1,4 g/L für Glucose und von 2,0 g/L auf 1,0 g/L für Ethanol

gesunken. Insgesamt wird der Prozess sowohl von der Simulation als auch durch die Vorhersage gleichermaßen gut beschrieben.

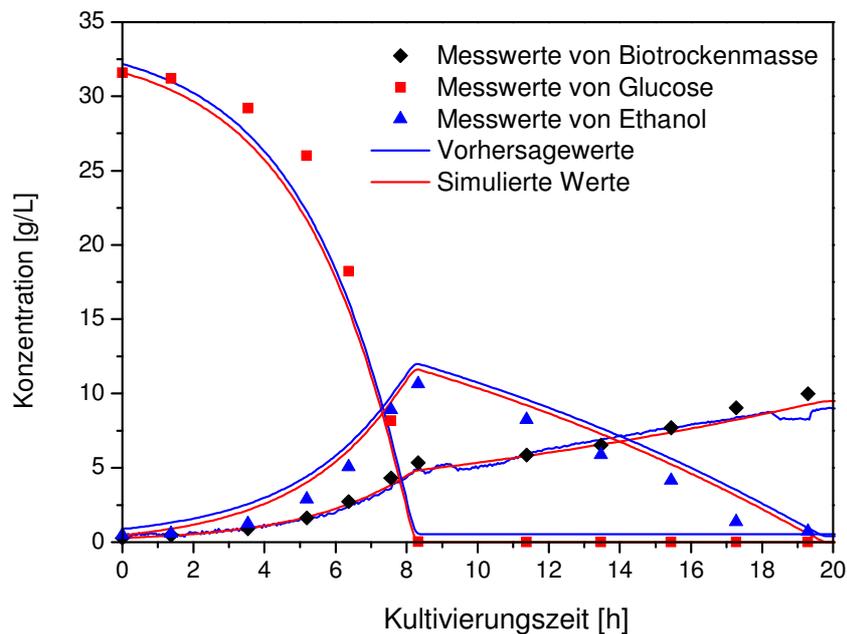


Abbildung 7-7: Vorhersage der Kultivierung Hefe 22

Tabelle 7-3: Mittlere Fehler der Vorhersagewerte der Bioprozessgrößen, ermittelt mit simulierten Prozesswerten und der Ultraschallgeschwindigkeit

Mittlere Fehler	Biomasse [g/L]		Glucose [g/L]		Ethanol [g/L]	
	V-M	V-S	V-M	V-S	V-M	V-S
Hefe 23 Modellbildung	0,5	0,2	0,1	0,1	0,5	0,1
Hefe 22 Validierung	0,6	0,3	1,5	0,6	1,4	0,4
Hefe 27 Validierung	1,3	1,1	2,7	2,7	2,6	1,9
Hefe 30 Validierung	1,7	0,3	2,6	0,9	2,9	0,6
Mittelwerte Validierung	1,2	0,6	2,3	1,4	2,3	1,0

7.2.3. Vorhersage der Bioprozessgrößen mit erweiterten Datensätzen

Anschließend wird eine weitere Vorhersage aus fünf Prozessgrößen berechnet. Die Vorhersagen der einzelnen Bioprozessgrößen Biomasse-, Glucose- und Ethanolkonzentration werden hier aus simulierten Werten von Biomasse-, Glucose- und Ethanolkonzentration und online gemessenen Abgaswerten von CO₂ und O₂ sowie der Ultraschallgeschwindigkeit durchgeführt. So wird zum Beispiel die Biomassekonzentration basierend auf simulierten Werten der Glucose- und Ethanolkonzentration, online gemessenen Abgaswerten von CO₂ und O₂ und der Ultraschallgeschwindigkeit vorhergesagt. Zur Berechnung der multilineareren Regression werden drei Hauptkomponenten genutzt.

Die berechneten mittleren Fehler sind in Tabelle 7-4 dargestellt. Im Vergleich zu den Vorhersagen (V-M) basierend auf drei Prozessgrößen (siehe Tabelle 7-3) ist der Mittelwert der mittleren Fehler über die drei Validierungen für Glucose um 0,2 g/L gesunken. Dagegen sind der mittlere Fehler für Biomasse und Ethanol jeweils um 0,2 g/L und 0,1 g/L größer geworden. Im Vergleich zu der Vorhersage (V-S) ist der mittlere Fehler von 0,6 g/L auf 1,0 g/L für Biomasse, von 1,4 g/L auf 1,9 g/L für Glucose und von 1,0 g/L auf 1,3 g/L für Ethanol gestiegen. Insgesamt wird durch diesen Versuch aber keine signifikante Verbesserung erzielt.

Tabelle 7-4: Mittlere Fehler der Vorhersagewerte von Bioprozessgrößen ermittelt mit fünf Prozessgrößen (simulierte Werte, Abgasmesswerte von CO₂, O₂ und Ultraschallgeschwindigkeit)

Mittlere Fehler	Biomasse [g/L]		Glucose [g/L]		Ethanol [g/L]	
	V-M	V-S	V-M	V-S	V-M	V-S
Hefe 23 Modellbildung	0,6	0,2	0,6	0,3	0,7	0,3
Hefe 22 Validierung	0,7	0,4	1,5	1,1	1,6	0,6
Hefe 27 Validierung	1,2	1,1	2,1	2,3	2,2	1,7
Hefe 30 Validierung	2,3	1,4	2,8	2,2	3,3	1,6
Mittelwerte Validierung	1,4	1,0	2,1	1,9	2,4	1,3

7.3. Kalman-Filter zur Vorhersage der Bioprozessgrößen

Die kontinuierlichen Zustandsgleichungen für ein Prozessmodell sind in Gleichung (7-1) dargestellt. Der Zustandsvektor beinhaltet die Differentialgleichungen von Biomasse-, Glucose- und Ethanolkonzentration. Das Prozessrauschen $z(t)$ ist ebenfalls durch einen Vektor beschrieben:

$$f(x(t), t) + z(t) = \begin{pmatrix} \left(\frac{\mu_S^{\max} S}{S + K_S} + \frac{\mu_E^{\max} E}{E + K_E} \right) X \\ - \left(\frac{\mu_S^{\max} S}{S + K_S} \right) \frac{X}{Y_{XS}} \\ - \left(\frac{\mu_E^{\max} E}{E + K_E} \right) \frac{X}{Y_{XE}} + \left(\frac{\mu_S^{\max} S}{S + K_S} \right) XY_{EX} \end{pmatrix} + \begin{pmatrix} z_X \\ z_G \\ z_E \end{pmatrix} \quad (7-1)$$

Hierbei ist X die Biomassekonzentration, S die Glucosekonzentration und E die Ethanolkonzentration. μ_S^{\max} ist die maximale Wachstumsrate auf Glucose und μ_E^{\max} die maximale Wachstumsrate auf Ethanol. Dabei gilt μ_E^{\max} nach Gleichung (6-3). K_S ist eine Limitierungskonstante für Glucose und K_E eine Limitierungskonstante für Ethanol. Y_{XS} ist der Ausbeutekoeffizient für die gebildete Biomasse aus Glucose, Y_{XE} der Ausbeutekoeffizient für die gebildete Biomasse aus Ethanol sowie Y_{EX} der Ausbeutekoeffizient für das gebildete Ethanol aus Glucose.

Mit Hilfe des Kalman-Filters wird versucht, die Vorhersage der Bioprozessgrößen zu verbessern. Mit dem Datensatz einer Kultivierung (Hefe 23) werden die optimalen Parameter des Kalman-Filters berechnet. Die erhaltenen Parameterwerte sind im Folgenden beschrieben:

Die Messrauschkovarianz R beträgt $0,169 \text{ m}^2/\text{s}^2$ und optimale Werte der Prozessrauschleistungsmatrix Q sind:

$$Q = \begin{pmatrix} 0,00001 \frac{\text{g}^2}{\text{L}^2 \text{h}} & 0 & 0 \\ 0 & 0,000085 \frac{\text{g}^2}{\text{L}^2 \text{h}} & 0 \\ 0 & 0 & 0,000105 \frac{\text{g}^2}{\text{L}^2 \text{h}} \end{pmatrix}$$

Anfangswerte für die Schätzfehlerkovarianzmatrix $P(t=0)$ sind:

$$P(t) = \begin{pmatrix} 0,0005 \frac{g^2}{L^2} & 0 & 0 \\ 0 & 0,001 \frac{g^2}{L^2} & 0 \\ 0 & 0 & 0,001 \frac{g^2}{L^2} \end{pmatrix}$$

Für die Prädiktion des Kalman-Filters wird das Runge-Kutta-Verfahren vierter Ordnung verwendet. Als Messgleichung dient ein multilineares Regressionsmodell, das die Ultraschallgeschwindigkeit in Abhängigkeit von der Biomasse-, Glucose- und Ethanolkonzentration beschreibt. Das kontinuierlich-diskrete erweiterte Kalman-Filter ist in Borland Delphi 5 implementiert.

Mit dem Kalman-Filter werden die Bioprozessgrößen Biomasse, Glucose und Ethanol vorhergesagt. Die Abbildung 7-8 zeigt Mess- und Modellwerte des Prozesses, die für die Optimierung der Parameter des Kalman-Filters verwendet werden. In Abbildung 7-9 ist die Validierung mit einem Prozess dargestellt, dessen Werte nicht zur Berechnung der Parameter des Kalman-Filters genutzt werden.

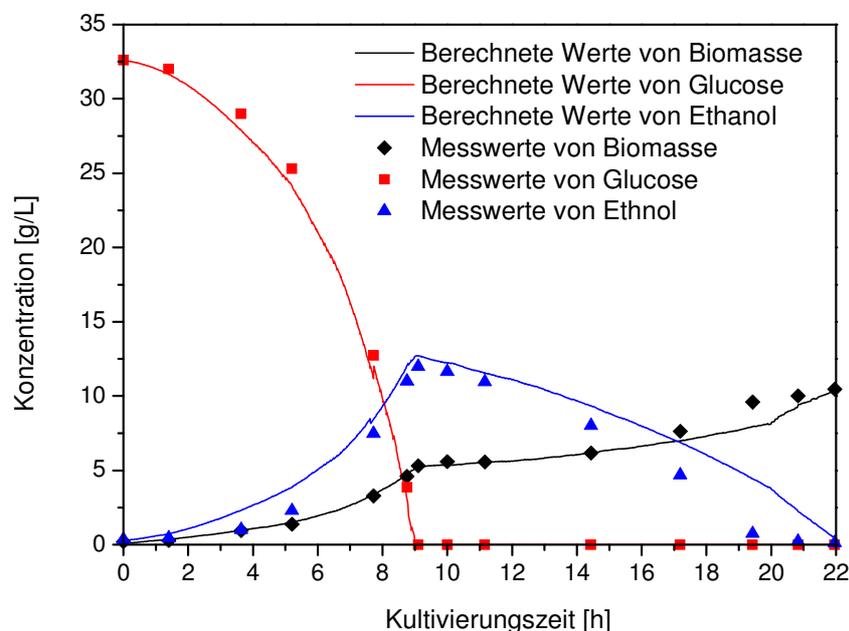


Abbildung 7-8: Messwerte und mit Kalman-Filter berechnete Werte des Prozesses (Hefe 23), der zur Berechnung der Parameter des Kalman-Filters verwendet wurde

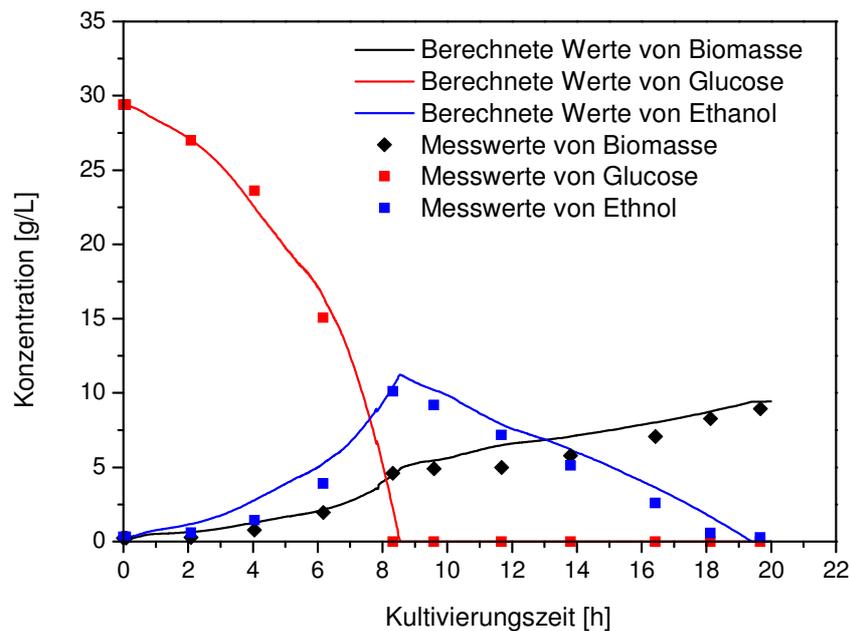


Abbildung 7-9: Messwerte und mit Kalman-Filter berechnete Werte eines Prozesses (Hefe 30), der nicht zur Berechnung der Parameter des Kalman-Filters verwendet wurde

Eine Übersicht über die mittleren Fehler der jeweiligen Vorhersagen ist in Tabelle 7-5 dargestellt. Die Ergebnisse können mit den Resultaten der Auswertung der PCR verglichen werden, die mit drei Prozessgrößen berechnet wird. Die mittleren Fehler (V-M) der Modellbildung der PCR sind 0,5 g/L für Biomasse, 0,1 g/L für Glucose und 0,5 g/L für Ethanol (siehe Tabelle 7-3) und die Fehler aus der Kalman-Filter-Auswertung sind 1,3 g/L für Biomasse, 2,1 g/L für Glucose und 3,3 g/L für Ethanol (siehe Tabelle 7-5). Dies zeigt, dass bei der Auswertung mit dem Kalman-Filter ein wesentlich größerer Fehler als bei der Auswertung mittels PCR auftritt. Im Vergleich zum Mittelwert (V-M) der Validierungen zeigt sich jedoch eine kleine Verbesserung der Vorhersagegüte durch die Berechnung mit dem Kalman-Filter. Der Mittelwert der Fehler bei der Validierung wird für Biomasse und Glucose jeweils um 0,1 g/L reduziert.

Tabelle 7-5: Mittlere Fehler der Vorhersage von Bioprozessgrößen mit dem Kalman-Filter

Mittlere Fehler	Biomasse [g/L]		Glucose [g/L]		Ethanol [g/L]	
	V-M	V-S	V-M	V-S	V-M	V-S
Hefe 23 Modellbildung	1,3	0,5	2,1	0,7	3,3	2,0
Hefe 22 Validierung	2,0	1,1	1,9	2,5	4,4	2,3
Hefe 27 Validierung	0,6	0,5	3,5	2,8	1,5	1,3
Hefe 30 Validierung	0,8	0,6	1,3	0,8	0,9	0,4
Mittelwerte Validierung	1,1	0,7	2,2	2,0	2,3	1,3

7.4. Neuronale Netze zur Vorhersage der Bioprozessgrößen

Für die Berechnung der Bioprozessgrößen mit einem neuronalen Netz wird die Simulations-toolbox „SNNS v4.2“ verwendet [Zell, 1995]. Der Stuttgart Neural Network Simulator (SNNS) ist ein Programm zur Simulation neuronaler Netze. Für die Vorhersage der Bioprozessgrößen werden Backpropagation-Netzwerke mit verschiedenen Netzstrukturen genutzt. Die Netzwerkstrukturen bestehen aus drei Eingabeneuronen, zwei Neuronen in der versteckten Schicht und einem Ausgabeneuron. Dies wird im Folgenden als Netz „3-2-1“ beschrieben. Weitere genutzte Strukturen sind „5-3-1“ und „8-5-3“. Als Eingabewerte stellen die Ultraschallgeschwindigkeitsmesswerte und die simulierten Werte der Bioprozessgrößen Biomasse-, Glucose- und Ethanolkonzentration sowie online gemessene Werte von O₂-, CO₂-Abgas, Säure, Base, RQ, OTR und CPR zur Verfügung. Zum Training der Netze werden die Daten von der Kultivierung Hefe 23 verwendet. Zur Validierung werden die drei Kultivierungen Hefe 22, Hefe 27 und Hefe 30 benutzt.

7.4.1. Datenvorverarbeitung

Die Bestimmung der Ein- und Ausgabewerte der zu erlernenden Muster hat einen großen Einfluss auf das Lernverhalten des Netzwerkes. Durch eine ungünstige Auswahl der Musterwerte kann es zu einer Konvergenz des Fehlerwertes gegen ein lokales Minimum der Fehlerfunktion oder sogar zu einem Abbruch des Backpropagation-Lernverfahrens kommen.

Wenn an einem Backpropagation-Netzwerk die Eingabewerte von 0 und oder 1 anliegen, so werden die Gewichte zu den Eingabeneuronen nicht trainiert. Im Falle einer Null liegt darin begründet, dass die Veränderung der Gewichte zu den Eingabeneuronen proportional zum Ausgabewert der Eingabeneuronen erfolgt. Wenn ein Neuron hingegen den Eingabewert Eins hat, wird wegen des Sättigungseffekts der Aktivierungsfunktion das entsprechende Gewicht nicht trainiert. Aus diesen Gründen wird in den folgenden Anwendungsfällen das Intervall in einen anderen Wertebereich [$0,2 \leq X_{i,neu} \leq 0,8$] verschoben:

$$X_{i,neu} = \left[\frac{x_i - x_{min}}{\Delta} \cdot 0,6 \right] + 0,2 \quad (7-2)$$

$$\text{mit } \Delta = x_{max} - x_{min} \quad (7-3)$$

Dabei sind x_{max} und x_{min} jeweils der maximale und minimale Wert von X.

Diese Skalierung wird durchgeführt, bevor die Daten in das Netzwerk eingespeist werden. Die maximalen (x_{max}) und minimalen (x_{min}) Werte der Bioprozessgrößen und der Ultraschallgeschwindigkeit sind in Tabelle 7-6 dargestellt. Es handelt sich um die minimalen und maximalen Werte aller Kultivierungen.

Tabelle 7-6: Maximale und minimale Werte der Bioprozessgrößen und der Ultraschallgeschwindigkeit

Prozessgrößen	Minimum	Maximum
US	1510 [m/s]	1535 [m/s]
BTM	0,1 [g/L]	15,0 [g/L]
Glucose	0,0 [g/L]	35,0 [g/L]
Ethanol	0,0 [g/L]	15,0 [g/L]

Zum Training des Netzes wird der epochale Lernzyklus angepasst und die Lernrate sowie das Momentum definiert. Typische Parameter, die zum Training des Netzes verwendet werden, sind in Tabelle 7-7 zusammengefasst.

Tabelle 7-7: Typische Parameter für das Training der Netze

Parameter	Werte
Epochale Lernzyklen	400
Lernrate	0,2
Momentum	0,8

7.4.2. Vorhersage von Bioprozessgrößen mit dem „3-2-1“ Netz

Die einzelnen Bioprozessgrößen Biomasse-, Glucose- und Ethanolkonzentration werden mit dem „3-2-1“ Netz vorhergesagt. Das trainierte Netz enthält für die Vorhersage der Biomassekonzentration als Eingabeneuronen die Ultraschallgeschwindigkeit, die Glucose- und Ethanolkonzentration, für die Vorhersage der Glucosekonzentration als Eingabeneuronen die Ultraschallgeschwindigkeit, die Biomasse- und Ethanolkonzentration sowie für die Vorhersage der Ethanolkonzentration als Eingabeneuronen die Ultraschallgeschwindigkeit, die Biomasse- und Glucosekonzentration (siehe Anhang 10.4)

Die Vorhersage der Bioprozessgrößen mit dem „3-2-1“ Netz ist in Abbildung 7-10 dargestellt. Das Netz wird mit den Werten der Kultivierung Hefe 23 trainiert. Zur Validierungen werden die weiteren Kultivierungen Hefe 22, Hefe 27 und Hefe 30 mit dem trainierten Netz vorhergesagt.

Abbildung 7-11 zeigt beispielhaft die Vorhersage für die Kultivierung Hefe 22. Die mittleren Fehler der Vorhersage sind in Tabelle 7-8 dargestellt. Die Ergebnisse der Vorhersage zeigen eine kleine Verbesserungen gegenüber den Auswertungen mittels der PCR (siehe Tabelle 7-4) oder dem Kalman-Filter-Verfahren (siehe Tabelle 7-5).

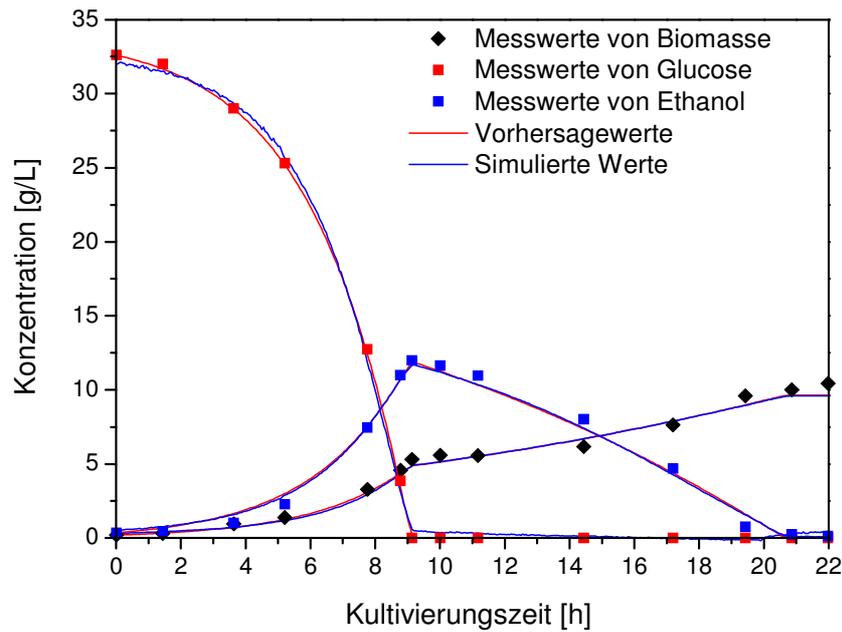


Abbildung 7-10: Vorhersage der Prozessgrößen Biotrockenmasse, Glucose und Ethanol mit einem „3-2-1“ Netz (Training des Netzes mit Hefe 23)

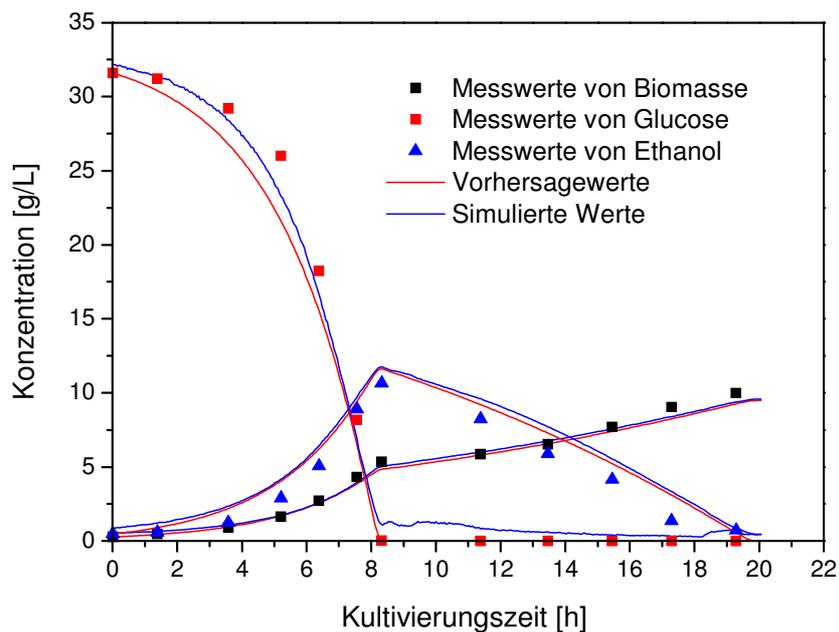


Abbildung 7-11: Vorhersage der Prozessgrößen mit einem „3-2-1“ Netz (Zur Validierung des Netzes mit Hefe 22)

Tabelle 7-8: Mittlere Fehler der Vorhersage der Bioprozessgrößen mit einem „3-2-1“ Netz

Mittlere Fehler	Biomasse [g/L]		Glucose [g/L]		Ethanol [g/L]	
	V-M	V-S	V-M	V-S	V-M	V-S
Hefe 23 Modellbildung	0,4	0,1	0,3	0,3	0,5	0,1
Hefe 22 Validierung	0,3	0,2	1,2	1,0	1,3	0,3
Hefe 27 Validierung	0,7	0,7	1,9	1,9	2,6	1,9
Hefe 30 Validierung	0,9	0,8	1,9	1,9	3,0	2,2
Mittelwerte Validierung	0,6	0,6	1,7	1,6	2,3	1,5

7.4.3. Vorhersage von Bioprozessgrößen mit dem „5-3-1“ Netz

In diesem Abschnitt werden die Ergebnisse der Vorhersage mit dem „5-3-1“ Netz für die einzelnen Bioprozessgrößen Biomasse-, Glucose- und Ethanolkonzentration vorgestellt. Die Eingabeneuronen umfassen die simulierten Werte von den Bioprozessgrößen Biomasse-, Glucose- und Ethanolkonzentration (jedoch nicht die jeweilige Vorhersagegröße), die Ultraschallgeschwindigkeitsmesswerte sowie die online gemessenen Abgaswerte von CO₂ und O₂. Die trainierten Netze sind in Anhang 10.5 dargestellt. Die mittleren Fehler der Vorhersage werden in Tabelle 7-9 gezeigt.

Tabelle 7-9: Mittlere Fehler der Vorhersage von Bioprozessgrößen mit einem „5-3-1“ Netz

Mittlere Fehler	Biomasse [g/L]		Glucose [g/L]		Ethanol [g/L]	
	V-M	V-S	V-M	V-S	V-M	V-S
Hefe 23 Modellbildung	0,4	0,1	0,5	0,3	0,5	0,1
Hefe 22 Validierung	0,4	0,1	1,0	1,3	1,5	0,5
Hefe 27 Validierung	0,9	0,8	1,2	0,4	2,7	1,9
Hefe 30 Validierung	1,2	1,0	0,9	0,9	3,0	2,0
Mittelwerte Validierung	0,8	0,6	1,0	0,9	2,4	1,5

Diese Auswertungen sind vergleichbar mit den Vorhersagen beim „3-2-1“ Netz. Jedoch sind die mittleren Fehler der Modellbildung vom „5-3-1“ Netz größer als bei der Auswertung mit dem „3-2-1“ Netz. Die Mittelwerte der Fehler der Validierung zeigen ebenfalls keine Verminderung bei Biomasse und Ethanol, obwohl der Fehler der Glucose bei V-M von 1,7 g/L auf 1,0 g/L und bei V-S von 1,6 g/L auf 0,9 g/L verringert wird.

7.4.4. Vorhersage von Bioprozessgrößen mit dem „8-5-3“ Netz

Um eine Verbesserung der Vorhersage der Bioprozessgrößen zu erzielen, wird ein „8-5-3“ Netz verwendet. Die acht Eingabeneuronen sind die Ultraschallgeschwindigkeit und die online gemessenen Bioprozessgrößen (CO₂ und O₂ im Abgas, Säure, Base, RQ, OTR, CPR). Die drei Ausgabeneuronen sind die Biomasse-, Glucose- und Ethanolkonzentration. Tabelle 7-10 zeigt die mittleren Fehler der Vorhersage mit dem „8-5-3“ Netz. Die Bioprozessgrößen von Biomasse, Glucose, Ethanol sind hier als Eingabeneuronen nicht beteiligt. Es resultiert jedoch keine verbesserte Vorhersage trotz der Verwendung von acht Eingabeneuronen für die Vorhersage der Bioprozessgrößen.

Tabelle 7-10: Mittlere Fehler der Vorhersage der Bioprozessgrößen mit einem „8-5-3“ Netz

Mittlere Fehler	Biomasse [g/L]		Glucose [g/L]		Ethanol [g/L]	
	V-M	V-S	V-M	V-S	V-M	V-S
Hefe 23 Modellbildung	0,4	0,1	0,7	0,4	0,6	0,3
Hefe 22 Validierung	0,8	0,4	2,5	2,9	3,6	2,8
Hefe 27 Validierung	0,7	0,8	2,4	2,3	2,4	2,0
Hefe 30 Validierung	1,0	0,9	2,3	2,5	2,9	2,1
Mittelwerte Validierung	0,8	0,7	2,4	2,6	2,9	2,3

Zusammenfassend werden in Tabelle 7-11 die mittleren Fehler der Validation der Bioprozessgrößen bei verschiedenen Netzstrukturen gezeigt. Insgesamt sind die Werte der Fehler bei Biomasse und Ethanol vom „3-2-1“ Netz und bei der Glucose vom „5-3-1“ Netz

am kleinsten. Der Vergleich der mittleren Fehler mit den Fehlern anderer Analyseverfahren wird im folgenden Abschnitt diskutiert.

Tabelle 7-11: Mittelwerte der Fehler der Validation der Bioprozessgrößen

Netzstruktur	Biomasse [g/L]		Glucose [g/L]		Ethanol [g/L]	
	V-M	V-S	V-M	V-S	V-M	V-S
„3-2-1“	0,6	0,6	1,7	1,6	2,3	1,5
„5-3-1“	0,8	0,6	1,0	0,9	2,4	1,5
„8-5-3“	0,8	0,7	2,4	2,6	2,9	2,3

7.5. Diskussion

Um die Bioprozessgrößen Biomasse-, Glucose- und Ethanolkonzentration vorherzusagen, wurden insgesamt fünf verschiedene Analyseverfahren getestet und die Ergebnisse verglichen.

Zunächst wurde eine multilineare Regression mit Offline-Messwerten von Biomasse, Glucose und Ethanol sowie mit Online-Ultraschallmesswerten durchgeführt. Dabei wurden jeweils drei Prozessgrößen für die Vorhersage eingesetzt, um die vierte Prozessgröße zu berechnen. Tabelle 7-12 zeigt die mittleren Fehler der Validierungen beim Einsatz dieses und der vier weiteren Verfahren. Die Ergebnisse der MLR (in Tabelle 7-12 mit * markiert) basieren ausschließlich auf den am Prozess gemessenen Werten. Im Gegensatz dazu beruhen die folgenden vier Auswerteverfahren auf simulierten Werten (in Tabelle 7-12 mit ** markiert). Bei jeder Kultivierung wurden nur ca. 12-14 Offline-Proben gezogen, da eine häufigere Probenahme zu einem zu großen Volumenverlust führen würde. Die simulierten Werte wurden mit Hilfe eines theoretischen Modells durch Anpassung an Messwerte erzeugt, um die Auswertung zu verbessern.

Die Vorhersagewerte aus simulierten Werten wurden sowohl mit den Messwerten als auch mit den simulierten Werten verglichen und daraus die mittleren Fehler der Vorhersage bezüglich der Messwerte (V-M) und der Vorhersage bezüglich der simulierten Werte (V-S) berechnet. Ein Vergleich der ersten beiden Verfahren, beides multilineare Regressionen, zeigt, dass durch den Einsatz simulierter Werte eine Verbesserung erzielt werden konnte. Die

Hauptkomponentenregression und die Verwendung eines Kalman-Filters haben zu keiner weiteren Verbesserung geführt. Allerdings konnten die Fehler bei der Vorhersage von Biomasse und Ethanol durch ein neuronales Netz weiter gesenkt werden.

Daraus folgt, dass für die Vorhersage von Glucose und Ethanol das relativ einfache Verfahren MLR aus simulierten Werten ausreicht, zur Auswertung der Biomasse hingegen sollten neuronale Netze herangezogen werden.

Tabelle 7-12: Übersicht über alle mittleren Fehler der Validierung bezüglich der Offline-Messwerte (V-M) und der simulierten Werte (V-S)

Auswerteverfahren	Mittlere Fehler der Validierung					
	Biomasse [g/L]		Glucose [g/L]		Ethanol [g/L]	
	V-M	V-S	V-M	V-S	V-M	V-S
MLR*	1,0	-	2,0	-	2,0	-
MLR**	1,0	0,5	1,9	0,8	1,6	1,0
PCR**	1,2	0,6	2,3	1,4	2,3	1,0
Kalman-Filter**	1,1	0,7	2,2	2,0	2,3	1,3
„3-2-1“ Neuronales Netz**	0,6	0,6	1,7	1,6	2,3	1,5

Abgesehen von der Güte der Vorhersagen ist für eine vollständige Bewertung der verschiedenen Auswerteverfahren, der sogenannten „Software-Sensoren“, ebenfalls der Aufwand und die erforderliche Zeit zu berücksichtigen. Zur Berechnung der multiliniaren Regression wurden vier Koeffizienten angepasst, unter deren Verwendung die gewünschten Prozessgrößen berechnet werden können. Für die Hauptkomponentenregression (PCR) wurde zuerst eine Hauptkomponentenanalyse durchgeführt und die erforderliche Anzahl an Hauptkomponenten ermittelt. Mit diesen Hauptkomponenten wurde die multiliniare Regression durchgeführt. Insgesamt gehen 5 Parameter in dieser Auswertung ein. Für die Berechnung der PCR wurden sowohl ein Programm in Delphi 5 als auch das Softwaretool „Unscrambler“ verwendet. Die Umsetzung des Kalman-Filter-Algorithmus in ein Delphi-Programm erfordert diverse Vorarbeiten. Zunächst wurden aus den Messdaten die Prozessrauschleistungsmatrix und die Schätzfehlerkovarianzmatrix in Simulationen ermittelt. Für diese Auswertung mussten 14 Parameter berechnet werden. Insgesamt war der Aufwand für die Implementierung des Kalman-Filters von allen Auswerteverfahren am höchsten. Zur

Verwendung des neuronalen Netzes wurde ein Softwaretool, der Stuttgart Neural Network Simulator (SNNS), verwendet. In dieser Arbeit wurden drei verschiedene Netzstrukturen „3-2-1“ (8 Parameter), „5-3-1“ (18 Parameter) und „8-5-3“ (55 Parameter) zur Vorhersage der Prozessgrößen genutzt. Die Parameter werden bei der Anwendung des neuronalen Netzes durch das Training des Netzes bestimmt und haben als Wichtungsfaktoren keine physikalische Bedeutung, da die zur Auswertung datengetriebenen Modelle genutzt wurden. Insgesamt wurden sieben Netze trainiert und für Vorhersagen herangezogen. PCR und neuronale Netze sind vergleichbar aufwendig, im Gegensatz zum Kalman-Filter weisen sie einen geringeren Aufwand auf, da bei der Implementierung des Kalman-Filters mehr Informationen erforderlich sind. Von Vorteil ist jedoch, dass beim Kalman-Filter keine weiteren Messungen notwendig sind. Bei der Anwendung des Kalman-Filters müssen also die Prozessgleichungen bekannt sein. Bei der PCR und den neuronalen Netze hingegen ist kein Prozesswissen erforderlich.

Zur Berechnung der Modelle wurden vier Kultivierungen durchgeführt. Im Hinblick auf die Nutzung verschiedener Software-Sensoren wäre es vorteilhaft, weitere Kultivierungen durchzuführen, um die Validierung zu präzisieren. Jedoch stand das Ultraschallmesssystem leider nicht mehr zur Verfügung.

Um eine blasenfreie Messung der Ultraschallgeschwindigkeit zu ermöglichen, wurde ein Stop-Flow-Betrieb im Bypasssystem durchgeführt. Diese Messung erwies sich als stabil und vorteilhaft für die Auswertung. Der störende Einfluss von Blasen wurde so vermieden. Nachteilig wirkt sich dabei jedoch die unzureichende Versorgung der Mikroorganismen mit Sauerstoff sowie ein Temperaturabfall um ca. 1°C aus, der durch den zeitweisen Verschluss des Ventils zustande kommt. Um diese Nachteile zu vermeiden, wäre es zukünftig empfehlenswert ein Online-Ultraschallmesssystem zu entwickeln, bei dem das Probenahmemodul direkt im Reaktor eingesetzt werden kann (In-situ-Verfahren).

Neben der Behebung des Problems der Sauerstoffversorgung und des Temperaturabfalls bei der Bypassmessung liefert das In-situ-Verfahren außerdem den Vorteil der Reduktion der Verweilzeiten. Bei Verwendung der Einstecksonde müssen allerdings die Durchströmungsbedingungen berücksichtigt werden, da bei der Kultivierung durch die hohe Rührerdrehzahl und die Begasung Blasen gebildet werden. Die Konstruktion der Einstecksonde muss im Fermenter so realisiert werden, dass die Probe in die Sonde eingesaugt wird, um dort solange zu verbleiben, bis gewährleistet ist, dass blasenfrei gemessen werden kann. Im Gegensatz zu

einem Bypasssystem, welches in der industriellen Praxis nicht gerne gesehen wird, würde ein solches In-situ-System die Akzeptanz in der Industrie erhöhen.

8. Zusammenfassung

In dieser Arbeit wird untersucht welche Information aus Ultraschallgeschwindigkeitsmessungen für die Beurteilung von Kultivierungen gewonnen werden kann. Hierzu werden verschiedene Auswerteverfahren entwickelt, die auf den Messungen der Ultraschallgeschwindigkeit sowie weiterer Bioprozessgrößen basieren. Dabei soll die Abhängigkeit der Ultraschallgeschwindigkeit von wesentlichen Prozessgrößen genutzt werden, um diese Größen für eine optimale Validierung vorherzusagen. Als Modellbioprozess wird die Kultivierung von Hefezellen (*Saccharomyces cerevisiae*) verwendet, da bei der Kultivierung dieser Organismen sowohl Änderungen in der Glucosekonzentration als auch in der Ethanolkonzentration zu beobachten sind, zwei Größen, die neben der Biomasse die Ultraschallgeschwindigkeit wesentlich beeinflussen. Das verwendete akustische Online-Sensorsystem US 2100 wurde am Institut für Automation und Kommunikation (Barleben) entwickelt und für diese Untersuchung zur Verfügung gestellt. Um den Einfluss von Gasblasen bei der Messung zu minimieren, wird die Messsonde in einem Bypass zum Reaktor betrieben und über ein Ventil ein Stop-Flow-Verfahren realisiert.

Zunächst wird die Abhängigkeit der Ultraschallgeschwindigkeit von den einzelnen Bioprozessgrößen untersucht. Dabei wird festgestellt, dass ein linearer Zusammenhang zwischen der Ultraschallgeschwindigkeit und den einzelnen Bioprozessgrößen vorliegt. Die Biomassekonzentration übt dabei den stärksten Einfluss mit einer Veränderung der Ultraschallgeschwindigkeit von 1,09 m-L/s·g aus.

Bei unterschiedlichen Kultivierungen von Hefezellen wird die Biomasse-, Glucose- und Ethanolkonzentration sowie die Ultraschallgeschwindigkeit gemessen. Um die Datenbasis zu erweitern und den Einfluss von Messfehlern zu kompensieren, wird ein Bioprozessmodell aufgestellt, welches auf einem Monod-Ansatz basiert. Die Parameter dieses Modells werden mit Hilfe der Messdaten einer Kultivierung identifiziert. Ein Vergleich der simulierten Werte mit den offline gemessenen Werten der verschiedenen Kultivierungen zeigte eine gute Übereinstimmung (Mittelwert der Validierung der mittleren Fehler der Simulation für Biomasse 0,3 g/L, für Glucose 0,7 g/L und für Ethanol 0,9 g/L).

Insgesamt kommen folgende Auswerteverfahren zum Einsatz: ein multilineares Regressionsmodell, eine Hauptkomponentenregression und ein multilineares Regressionsmodell als Messgleichung in einem kontinuierlich-diskreten erweiterten Kalman-Filter sowie neuronale Netzwerke.

Die mittleren Fehler der Vorhersage bezüglich der Messwerte betragen für Biomasse und Glucose jeweils 0,6 g/L und 1,7 g/L (berechnet mittels neuronaler Netze) und für Ethanol 1,6 g/L (berechnet mittels MLR aus simulierten Werten) Diese Fehler waren niedriger als die Fehler anderer „Software-Sensoren“. Die Ergebnisse zeigen, dass die Fehler der Vorhersage aller Verfahren die gleiche Größenordnung besitzen. Die Größenordnung dieser Fehler stimmt ebenfalls mit den Fehlern überein, die bei der Vorhersage basierend auf den tatsächlichen Messwerten erzielt wurden.

Bei der Bestimmung der Prozessgrößen mit Hilfe des multilinearen Regressionsmodells, der PCR und den neuronalen Netzwerken ist die Kenntnis der drei anderen Prozessgrößen Voraussetzung, um die vierte vorherzusagen. Somit ist für diese Verfahren zusätzliche Messtechnik erforderlich. Für das Kalman-Filter ist zwar ein Bioprozessmodell mit den entsprechenden Parameterwerten notwendig, jedoch muss während des Prozesslaufs neben der Ultraschallmessung keine weitere Messgröße erfasst werden. Somit sind für das Kalman-Filter zwar bei der Implementierung mehr Informationen als für die anderen Verfahren erforderlich, für den Betrieb sind jedoch keine weiteren Messungen notwendig.

Insgesamt zeigt diese Arbeit, dass Auswerteverfahren, die auf Ultraschallgeschwindigkeitsmessungen basieren, wertvolle quantitative Informationen über einen Bioprozess liefern. Darüber hinaus kann der Bioprozess auch qualitativ, nur basierend auf den reinen Ultraschallmesswerten, beurteilt werden, da die typische Phaseneinteilung (in Glucose-, Ethanol- und Acetatabbau) dieses komplexen Bioprozesses in den Messwerten deutlich sichtbar ist. Somit stellt das Ultraschallmesssystem eine wertvolle Ergänzung in der Bioprozessanalytik dar.

9. Literaturverzeichnis

- [Asher, 1997] RC Asher: Ultrasonic Sensors for Chemical and Process Plant, Institute of Physics Publishing Bristol and Philadelphia, 1997
- [Becker et al., 2002] T. Becker, M. Mitzscherling, A. Delgado: Hybrid data model for the improvement of an ultrasonic-based gravitz measurement system, 2002, Food control 13, S.223-233
- [Bellgardt, 1984] K.-H. Bellgardt: Modellbildung des Wachstums von *Saccharomyces cerevisiae* in Rührkesselreaktoren, Dissertation, Universität Hannover, 1984
- [Bellgardt et al., 1986] K.-H. Bellgardt, W. Kuhlmann, H.-D. Meyer, K. Schügerl, M. Thoma: Application of an extended Kalman Filter for state estimation of a yeast fermentation, *IEE Proceedings*, Vol. 133, No. 5, 1986, S.226-234
- [Bergman, 1954] L. Bergmann: Der Ultraschall und seine Anwendung in Wissenschaft und Technik, Hirzel Verlag Stuttgart, 1954
- [Bogaerts, 1999] Ph. Bogaerts: A hybrid asymptotic-Kalman observer for bioprocesses, *Bioprocess Engineering* 20, 1999, S.249-255
- [Börnstein, 1967] L. Börnstein: Molekularakustik, Band 5, Vol. 5, Springer-Verlag, 1967, S.69-72
- [Brause, 1991] R. Brause: Neuronale Netze, Verlag B. G. Teubner Stuttgart, 1991
- [Cha, Hitzmann, 2004] Young-Lok Cha, Bernd Hitzmann: Ultrasonic Measurements and its Evaluation for the Monitoring of *Saccharomyces cerevisiae* Cultivation, *Bioautomation*, Vol.1, 2004, S.16-29.
- [Coppella et al., 1990] Steven J. Coppella, Prasad Dhurjati: A Mathematical Description of Recombinant Yeast, *Biotechnology and Bioengineering*, Vol.35, 1990, S.356-374.
- [Daur et al., 1998] P.-Ch. Daur, B. Henning, A. Wolf, S. Prange, K. Dierks, P. Hauptmann, A. Püttmer: In-Line-Flüssigkeitsanalysator US 2100, Tagungsband Fachtagung „Informations- und Mikrosystemtechnik“, Magdeburg, 1998, S.71-78

- [Dubach, Märkl, 1992] A.-Ch. Dubach, H. Märkl: Application of an Extended Kalman Filter Method for Monitoring High Density Cultivation of *Escherichia coli*, Journal of Fermentation and Bioengineering, Vol 73, 1992, S.396-402
- [Dymling et al., 1991] S. O. Dymling, H. W. Persson, T. G. Hertz, K. Lindström: A New Ultrasonic Method for fluid property measurements, Ultrasound in Med. & Biol. 1991, Vol. 17, No. 5, S.497-500
- [Enfors et al., 1990] S.-O. Enfors, J. Hedenberg, K. Olsson: Simulation of the dynamics in the Baker's yeast process, Bioprocess Engineering 5, 1990, S.191-198
- [Geissler et al., 2002] D. Geissler, D. Solle, E. Stärk, T. Scheper, H. Märkl, B. Hitzmann: Berechnung chemometrischer Modelle basierend auf theoretischen Modellen und 2D-Fluoreszenzspektren, Chemie Ingenieur Technik 74, 8/2002, S.1167-1170
- [Geladi, Kowalski, 1986] P. Geladi, B. R. Kowalski: Partial least-squares regression: A tutorial, Anal. Chem. Acta 185, 1986, S.1-17
- [Henning, 1990] B. Henning: Sensorsystem zur Prozessbeobachtung Fluidischer Systeme, Ultraschallmesstechnik in Forschung und Praxis in Martin-Luther Universität Halle-Wittenberg, 1990, S.66-70
- [Henning et al., 1994] B. Henning, P. Hauptmann, R. Lucklum: Ultrasonic Sensor System for Characterisation of Liquids. Sensors and Actuators, 1994 H 41-42, S.476-480
- [Henning et al., 1996] B. Henning, H. Lachmann, K.-H. Bellgardt, A. Afschar: Sensorsystem zur kontinuierlichen Beobachtung der Hefe-Fermentation, InCom 96, Düsseldorf, Tagungsband, 1996
- [Henning et al., 1998a] B. Henning, P.-Ch. Daur, S. Prange, K. Dierks: In-Line-Flüssingkeitsanalysator für verfahrenstechnische Prozesse, VDI Berichte Nr. 1397, 1998, S.165-172
- [Henning et al., 1998b] B. Henning, P.-Ch. Daur, K. Dierks, S. Prange: Schnelle Konzentrationsmessung mit Multisensorsystem, Hüthig, Chemie Technik 8, 1998, S.30-35

- [Henning et al., 1999] B. Hennig, S. Prange, P.-C. Daur, K. Dierks: Charakterisierung des akustischen Verhaltens blasenhaltiger Flüssigkeiten, *Chemie Technik*, 51 Heft 5, 1999, S.267-271
- [Henning et al., 2000] B. Hennig, P.-Ch. Daur, S. Prange, K. Dierks, P. Hauptmann: In-line concentration measurement in complex liquids using ultrasonic sensors, *Ultrasonics* 38, 2000, S.799-803
- [Hitzmann, 1996] B. Hitzmann: Optimierung von FIA-Systemen für die Bioprozesstechnik, Vieweg, 1996
- [Hitzmann et al., 1998] B. Hitzmann, T. Pekeler, C. Lindemann, S. Marose, T. Scheper: Chemometric models for the on-line estimation of Bioprocess variables from 2-D-Fluorescence spectra, *Computer Applications in Biotechnology – CAB7*, Osaka, 1998
- [Hitzmann et al., 2000] B. Hitzmann, O. Broxtermann, Y.-L. Cha, O. Sobieh, E. Stärk, T. Scheper: The control of glucose concentration during yeast fed-batch cultivations using a fast measurement complemented by an extended Kalman Filter, *Bioprocess Engineering*, 23, 2000, S.337-341
- [Karim, Rivera, 1992] M. N. Karim, S.L. Rivera: Artificial Neural Networks in Bioprocess State Estimation, *Journal of Biotechnology (Advances in Biochemical Engineering)*, Vol. 46, 1992
- [Kowalski et al., 1991] B.R. Kowalski, M.B. Seasholtz: Recent developments in Multivariate Calibration, *Journal of Chemometrics*, Vol. 5, 1991, S.129-145
- [Kruse et al., 1991] H. Kruse, R. Mangold, B. Mechler; Penger, O.: Programmierung neuronaler Netze, Addison-Wesley, Bonn, 1991
- [Lawrence, 1989] Lynnworth C. Lawrence: Ultrasonic Measurements for Process control, Academic Press INC, 1989
- [Lennox et al., 2001] B. Lennox, G.A. Montague, H.G. Hiden, G. Kornfeld, P.R. Goulding: Process Monitoring of an Industrial Fed-Batch Fermentation: *Biotechnology and Bioengineering*, Vol. 74, No.2, 2001, S.125-135

- [Linko, Zhu, 1991] Pekka Linko, Yihong Zhu: Neural network programming in bioprocess variable estimation and state prediction, *Journal of Biotechnology*, Vol.21, 1991, S. 253-270
- [Locher et al., 1993] G. Locher, U. Hahmemann, B. Sonnleitner, A. Fiechter: Automatic bioprocess control. 4. A prototype batch of *Saccharomyces cerevisiae*, *Journal of Biotechnology*, 29, 1993, S.57-74
- [Löhmer et al., 1998] H. Löhmer, R. Säuberlich, S. Heppner: Messung der Ultraschallgeschwindigkeit in blasenhaltigen Flüssigkeiten und Suspensionen, *Technisches Messen* 65, 12/1998, S.427-433
- [Löhmer, et al., 1999] H. Löhmer, R. Säuberlich: Ultraschall-Konzentrationsmessung in fluiden hochviskosen Medien bei der Anwesenheit von Gasblasen, *Chemie Technik*, Vol.51, Heft 5, 1999, S.267-271
- [Marose et al., 1998] S. Marose, C. Lindemann, T. Scheper: Two-Dimensional Fluorescence Spectroscopy: A New Tool for On-Line Bioprocess Monitoring, *Biotechnology Progress* Vol.14, No.1, 1998, S.63-74
- [Martens, Næs, 1993] H. Martens, T. Næs: *Multivariate Calibration*, 2nd Edition, John Wiley & Sons, Chichester, 1993
- [Massart et al., 1987] D.-L. Massart, R.-G. Brereton, R.-E. Dessy, P.K. Hpké, C.H. Spiegelman, W. Wegscheider, *Chemometrics and Intelligent Laboratory Systems*, 2, 1997, S.37-52
- [Náhlík, Burianec, 1988] J. Náhlík, Z. Burianec: On-line parameter and state estimation of continuous cultivation by extended Kalman Filter, *Applied Microbiology and Biotechnology*, Vol.28, 1988, S.128-134
- [Niepmann, 1983] Rolf Niepmann: Messung der Ultraschallgeschwindigkeit in Flüssigkeiten über einen größeren Temperatur- und Druckbereich, Dissertation, Ruhr-Universität Bochum, 1983
- [Otto, 1997] M. Otto: *Chemometrie: Statistik und Computereinsatz in der Analytik*, VCH, Weinheim [u.a.], 1997
- [Petrova et al., 1997] M. Petrova, P. Koprinkova, T. Patarinska: Neural network modelling of fermentation processes: Microorganisms cultivation model, *Bioprocess Engineering* 16, 1997, S.145-149

- [Rehm, Reed, 1991] H.-J. Rehm, G. Reed: Biotechnology, 2th edition, Vol. 4, Kap. 9 Cell Models, 1991, S.267-298
- [Ritter et al., 1990] H. Ritter, T. Martinetz, K. Schulten: Neuronale Netze – Eine Einführung in die Neuroinformatik selbstorganisierender Netzwerke, Addison-Wesley, 1990
- [Ritzka et al., 2002] A. Ritzka, R. Ulber, B. Hitzmann, T. Scheper: Regelung von Bioprocessen mit neuronalen Netzen, Chemie Ingenieur Technik 74, 1-2/2002, S.149-152
- [Safi, 1996] Raid Safi: Ultraschalluntersuchungen zur Konzentrationsbestimmung an flüssigen Mehrkomponentensystemen unter Nutzung des Prinzips der Parameterabhängigkeit von Stoffgrößen am Beispiel der Schallgeschwindigkeit und der Dichte, Dissertation, Martin-Luther-Universität Halle-Wittenberg, 1996
- [Schatzman, 1975] H. Schatzmann: Anaerobes Wachstum von *Saccharomyces cerevisiae*, Dissertation, ETH Zürich No. 5504, 1975
- [Schügerl et al., 1993] K. Schügerl, K.-H. Bellgardt, G. Kretzmer, B. Hitzmann, T. Scheper: In-situ- und On-line-Überwachung und Regelung biotechnologischer Prozesse, Chemie Ingenieur Technik 65, Nr. 12, 1993, S.1447-1456
- [Schügerl, Bellgardt, 2000] K. Schügerl, K.-H. Bellgardt: Bioreaction Engineering – Modeling and Control, Springer-Verlag, 2000
- [Haykin, 1994] Simon Haykin: Neural Networks – a comprehensive foundation, McMaster University, Macmillan College Publishing Company, 1994
- [Svrcek et al., 1974] W.Y. Svrcek, R.F. Elliott, J.E. Zajic: The Extended Kalman Filter Applied to a Continuous Culture Model, Biotechnology and Bioengineering, 1974, S.827-846
- [Roland et al., 2001] U. Roland, B. Hitzmann, T. Scheper: Innovative Bioprocessanalytik – Neue Wege zum Besseren Verständnis biotechnologische Prozesse, Chemie Ingenieur Technik (73), 2001, S.19-26

- [Wenz, 1992] Roland Wenz: Anwendung neuronaler Netzwerke zur Zustandserkennung bei biotechnologischen Prozessen, Diplomarbeit, 1992, Universität Hannover.
- [Zell, 1995] Andreas Zell: *SNNS Stuttgart Neural Network Simulator*, Institut for Parallel and Distributed High Performance Systems, 1995.
- [Zell, 1997] Andreas Zell: *Simulation neuronaler Netze*, R. Oldenbourg Verlag, 1997
- [Zhang et al., 1997] Z. Zhang, J.M. Scharer, M. Moo-Young: Mathematical model for aerobic culture of a recombinant yeast, *Bioprocess Engineering* 17, 1997, S.235-240.
- [Zips, 1993] Alf Zips: Beitrag zum Einsatz von Ultraschallsensoren zur Konzentrationsmessung in Biosuspensionen, Dissertation, Universität Stuttgart, 1993

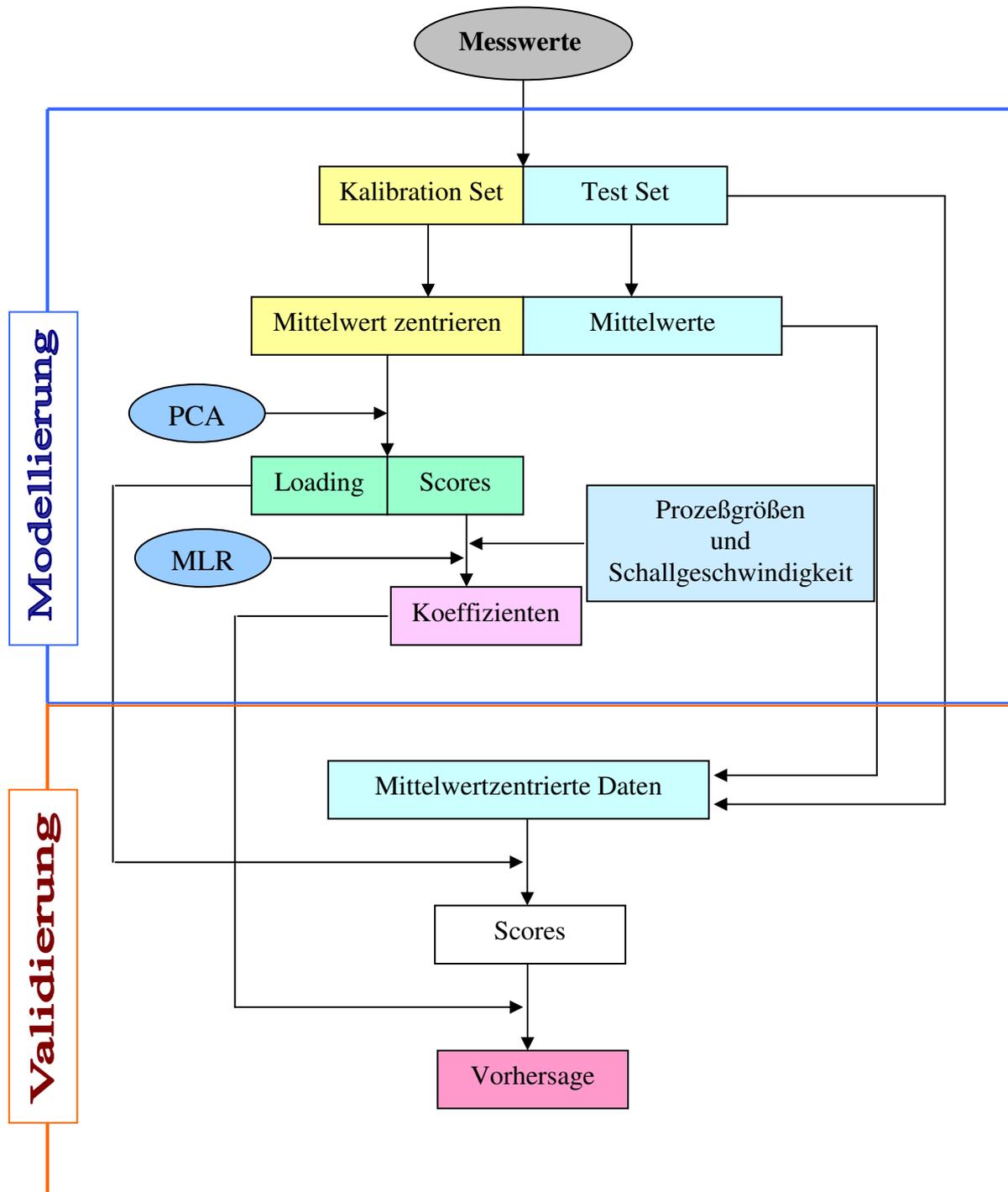
10. Anhang

10.1. Abkürzungs- und Symbolverzeichnis

a	Aktivität	K	Kalmanverstärkungsmatrix
A	Empfangsamplitude	K	Koeffizient
Abb.	Abbildung	K_S	Limitierungskonstant für Glucose
b	Koeffizientenvektor	K_E	Limitierungskonstant für Ethanol
B	Regressionskoeffizient	L	Liter
B	Koeffizient	m	Meter
B	Echoamplitude	M	Matrix der unabhängigen Variablen
BTM	Biotrockenmasse	M	Messwerte
C	Konzentration	MEP	mean error of prediction
C	Schallgeschwindigkeit	min	Minute
CPR	Kohlendioxidbildungsrate	MLR	Multi-Lineare-Regression
E	Ethanolkonzentration	n	Anzahl der Messungen
E	Residual-Matrix	NIPALS	Nonlinear Iterative Partial Least Square
E_p	Quadratischer Fehler des p -ten Lernmusters	o	Ausgangswerte des Neurons
EKF	Erweitertes Kalman-Filter	OD	Optische Dichte
f_{sig}	Sigmoide Aktivierungsfunktion	OUR	Sauerstoffaufnahme
f	Prozessvariablengleichung	P'	Loadingmatrix
F	Jakobimatrix von f	P	Schätzfehlerkovarianzmatrix
g	Gramm	P	Schallwechseldruck
h	Stunde	PCA	Principal Component Analysis
h	Wahrer Wert von x	PCR	Principal Component Regression
H	Jakobimatrix von h	pO_2	Gelöstsauerstoffgehalt
i	Laufvariable für die Messung	Q	Prozessrauschleistungsmatrix
i	Index für Ausgabeschicht	R	Messrauschkovarianzmatrix
IFAK	Institut für Automation und Kommunikation (in Barleben)	RISP	Realtime Integrating Software Plattform
J	Laufvariable für die Messgröße		
j	Index für innere Schicht		
k	Index für Eingabeschicht		

RQ	respiratorischer Quotient	x	Zustandsvektor
s	Sekunde	x	Unabhängige Variable
s	Messrauschen	\bar{x}	Mittelwert der Variablen x
S	Glucosekonzentration	\hat{x}	Schätzwert des Zustandsvektors
S	Schallweg	X	Biomassekonzentration
S	simulierter Wert	X	Matrix der unabhängigen Variablen (Messwerte)
SNNS	Stuttgart Neural Network Simulator	X^{-1}	Inverse der Matrix X
SSE	Sum of Squared Errors	X'	Transponierte der Matrix X
t	Zeit	y	Unabhängige Variable
t_1	Laufzeit	Y	Matrix der abhängigen Variablen (Konzentrationen)
t_2	Laufzeit	Y_{XS}	Ausbeutekoeffizient für Biomasse aus Glucose
T	Scorematrix	Y_{XE}	Ausbeutekoeffizient für Biomasse aus Ethanol
TCI	Institut für Technische Chemie, Universität Hannover	Y_{EX}	Ausbeutekoeffizient für Ethanol aus Glucose
UpM	Umdrehungen pro Minute (Rührerdrehzahl)	z	Prozessrauschen
US	Ultraschallgeschwindigkeit	Z	Schallkennimpedanz
V_{US}	Berechneter Wert der Ultraschallgeschwindigkeit	α	Lernrate
V_m	Vorhersage aus Messwerten	β_{ad}	Adiabatische Kompressibilität
V_s	Vorhersage aus simulierten Werten	η	Momentum
V_s -M	Fehler zwischen Messwerten und Vorhersage basierend auf simulierten Werten	Θ	Bias
V_m -M	Fehler zwischen Messwerten und Vorhersage basierend auf Messwerten	μ	Spezifische Wachstumsrate
V_s -S	Fehler zwischen simulierten Werten und Vorhersage basierend auf simulierten Werten	μ_S^{\max}	Maximale Wachstumsrate auf Glucose
w	Verbindungsgewicht des Netzes	μ_E^{\max}	Maximale Wachstumsrate auf Ethanol
		μL	Mikroliter
		σ	Standardabweichung
		τ_a	Eigenwert

10.2. Blockdiagramm einer PCR mit Testset-Validierung

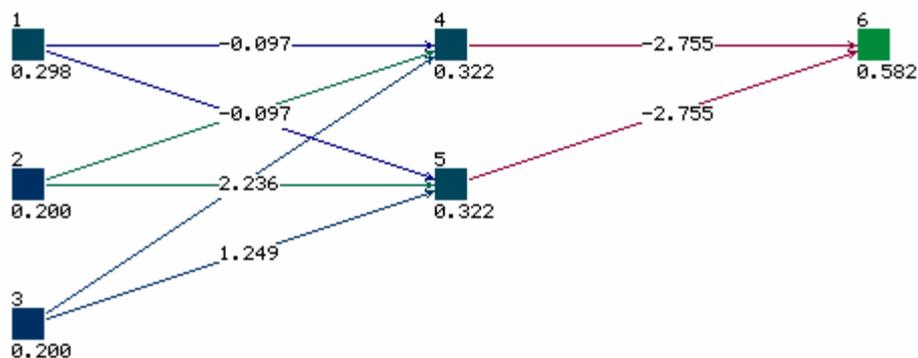


10.3. Schatzmann-Medium

	Substanz	Massenkonzentration [g/L]
Komponenten	Glucose	30,0
	Na-Citrat	5,8
	(NH ₄) ₂ SO ₄	4,5
	(NH ₄) ₂ HPO ₄	1,9
	KCl	0,9
	MgSO ₄ 7H ₂ O	0,34
	CaCl ₂ 2H ₂ O	0,42
	Mineralsalzlösung	FeCl ₃ 6H ₂ O
ZnSO ₄ 7H ₂ O		0,9·10 ⁻²
MnSO ₄ 2H ₂ O		1,05·10 ⁻²
CuSO ₄ 5H ₂ O		0,24·10 ⁻²
Vitaminlösung		m-Inosit
	Ca-Pantothenat	3,0·10 ⁻²
	Thiamin HCl	0,6·10 ⁻²
	Pyridoxol HCl	0,15·10 ⁻²
	Biotin	0,3·10 ⁻⁴

10.4. Die trainierten Netze mit „3-2-1“ Struktur

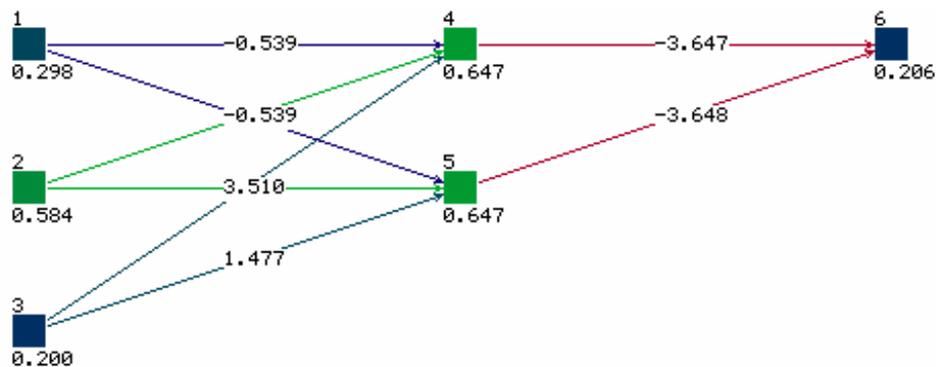
Zum Training des Netzes für die Vorhersage der Biomasse (Ausgabeneuron, Neuron 6) werden die drei Eingabeneuronen Ultraschallgeschwindigkeit (Neuron 1) und Glucose (Neuron 2) sowie Ethanol (Neuron 3) verwendet. Bias-Werte sind unter dem Neuronen angegeben:



Gewichtsfaktoren für die Vorhersage der Biomasse:

Source	Target		
	4	5	6
1	-0,09655	-0,0966	
2	2,23555	2,23564	
3	1,24909	1,249	
4			-2,7546
5			-2,75472

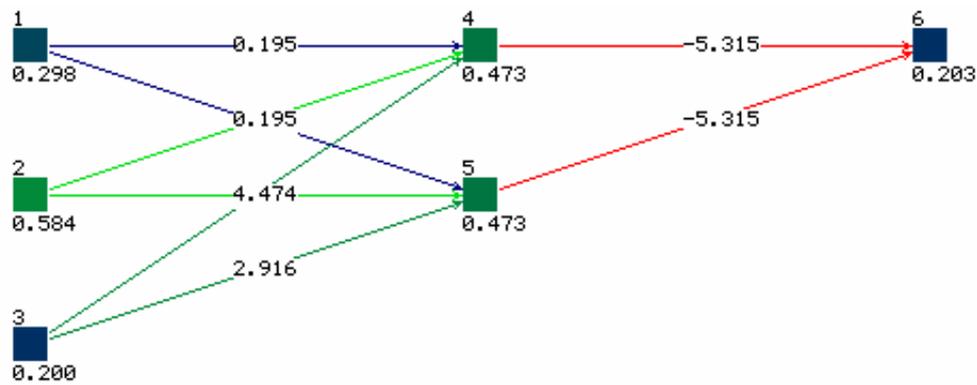
Zum Training des Netzes für die Vorhersage der Glucose (Ausgabeneuron, Neuron 6) werden die drei Eingabeneuronen Ultraschallgeschwindigkeit (Neuron 1) und Biomasse (Neuron 2) sowie Ethanol (Neuron 3) verwendet. Bias-Werte sind unter dem Neuronen angegeben:



Gewichtsfaktoren für die Vorhersage der Glucose:

Source	Target		
	4	5	6
1	-0,53893	-0,53909	
2	3,50933	3,5097	
3	1,47683	1,47681	
4			-3,64717
5			-3,64758

Zum Training des Netzes für die Vorhersage der Ethanol (Ausgabeneuron, Neuron 6) werden die drei Eingabeneuronen Ultraschallgeschwindigkeit (Neuron 1) und Biomasse (Neuron 2) sowie Glucose (Neuron 3) verwendet. Bias-Werte sind unter dem Neuronen angegeben:

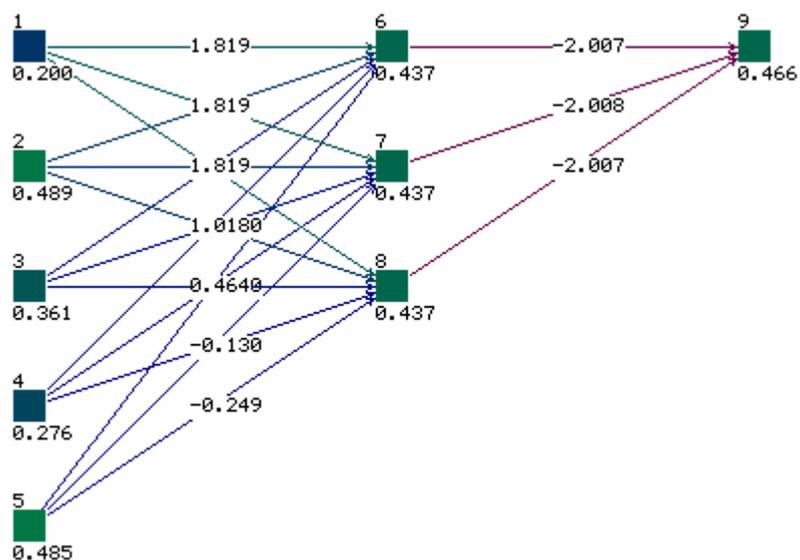


Gewichtsfaktoren für die Vorhersage der Ethanol:

Source	Target		
	4	5	6
1	0,19509	0,19503	
2	4,47354	4,47377	
3	2,91539	2,91556	
4			-5,31517
5			-5,31548

10.5. Die trainierten Netze mit „5-3-1“ Struktur

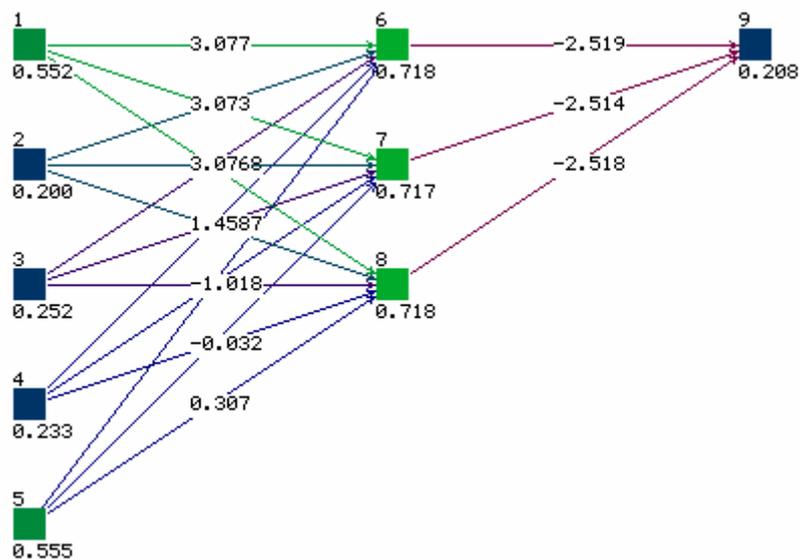
Zum Training des Netzes für die Vorhersage der Biomasse (Ausgabeneuron, Neuron 9) werden die fünf Eingabeneuronen Ultraschallgeschwindigkeit (Neuron 1), Glucose (Neuron 2), Ethanol (Neuron 3), CO₂- (Neuron 4), und O₂ Abgasmesswerte (Neuron 5) verwendet. Bias-Werte sind unter dem Neuronen angegeben:



Gewichtsfaktoren für die Vorhersage der Biomasse:

Source	Target			
	6	7	8	9
1	1,81912	1,81922	1,81917	
2	1,01844	1,01838	1,01844	
3	0,46385	0,46382	0,46382	
4	-0,12965	-0,12981	-0,12972	
5	-0,24914	-0,24923	-0,24918	
6				-2,00737
7				-2,00755
8				-2,00746

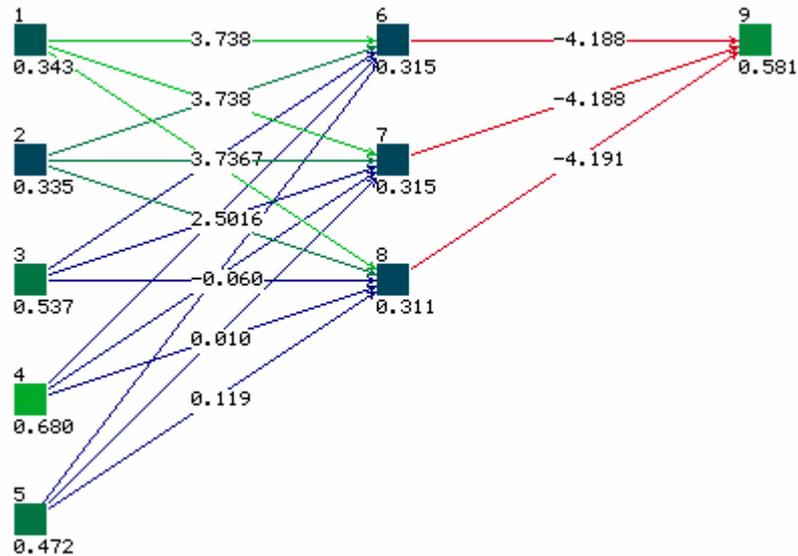
Zum Training des Netzes für die Vorhersage der Glucose (Ausgabeneuron, Neuron 9) werden die fünf Eingabeneuronen Ultraschallgeschwindigkeit (Neuron 1), Biomasse (Neuron 2), Ethanol (Neuron 3), CO₂- (Neuron 4) und O₂ Abgasmesswerte (Neuron 5) verwendet. Bias-Werte sind unter dem Neuronen angegeben:



Gewichtsfaktoren für die Vorhersage der Glucose:

Source	Target			
	6	7	8	9
1	3,07742	3,07338	3,07624	
2	1,45823	1,45649	1,45773	
3	-1,01814	-1,0166	-1,01771	
4	-0,03257	-0,03188	-0,03237	
5	0,30759	0,30717	0,30748	
6				-2,5194
7				-2,51444
8				-2,518

Zum Training des Netzes für die Vorhersage der Ethanol (Ausgabeneuron, Neuron 9) werden die fünf Eingabeneuronen Ultraschallgeschwindigkeit (Neuron 1), Biomasse (Neuron 2), Glucose (Neuron 3), CO₂- (Neuron 4) und O₂ Abgasmesswerte (Neuron 5) verwendet. Bias-Werte sind unter dem Neuronen angegeben:

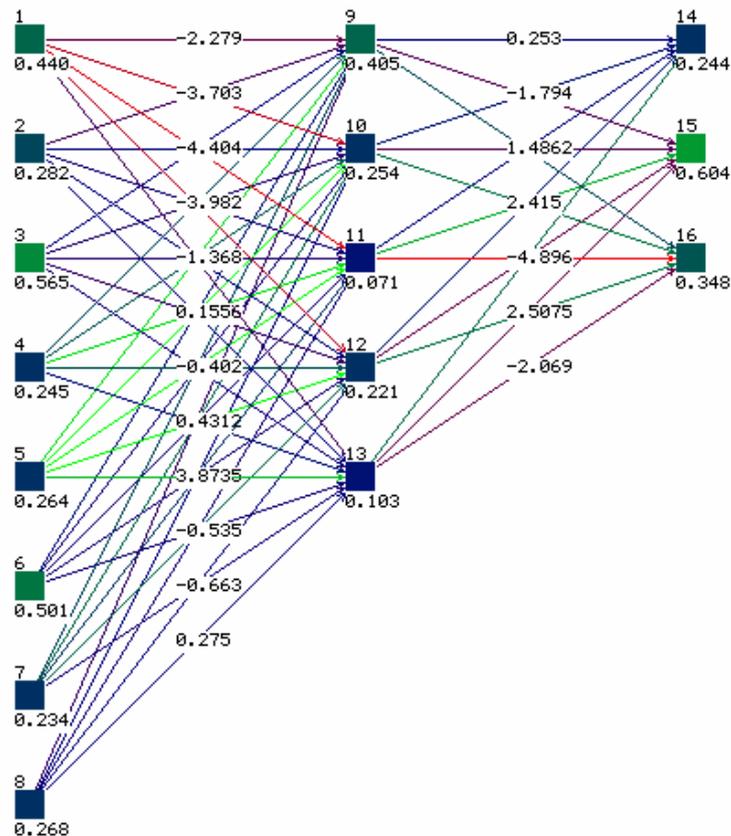


Gewichtsfaktoren für die Vorhersage der Ethanol:

Source	Target			
	6	7	8	9
1	3,73759	3,73771	3,73599	
2	2,49544	2,49471	2,50128	
3	-0,05685	-0,0565	-0,05989	
4	0,02418	0,02589	0,00954	
5	0,11994	0,12009	0,11885	
6				-4,18801
7				-4,18765
8				-4,1911

10.6. Die trainierten Netze mit „8-5-3“ Struktur

Als Eingabeneuronen werden die Ultraschallgeschwindigkeitsmesswerte (Neuron 1) und die online gemessenen Bioprozessgröße CO_2 - (Neuron 2) und O_2 Abgas (Neuron 3), Säure (Neuron 4), Base (Neuron 5), RQ (Neuron 6), OTR (Neuron 7), CPR (Neuron 8) verwendet. Die drei Ausgangsneuronen sind Biomasse (Neuron 14), Glucose (Neuron 15) und Ethanol (Neuron 16). Bias-Werte sind unter dem Neuronen angegeben:



Gewichtsfaktoren für die Vorhersage der Biomasse, Glucose, Ethanol:

Source	Target				
	9	10	11	12	13
1	-2,2791	-3,70329	-4,4043	-3,98208	-1,36826
2	-1,41502	-0,11998	-0,66897	-0,0738	0,15466
3	-0,12084	-0,94095	-0,989	-1,05647	-0,40245
4	1,60455	1,88405	4,37734	1,98102	0,43056
5	4,67818	6,7047	5,6858	7,06436	3,87337
6	0,67525	-0,38963	-0,65165	-0,57504	-0,53501
7	1,48425	2,05651	1,02164	2,18755	-0,6629
8	-1,14656	0,0939	-0,6263	0,14307	0,27469
Source	Target				
	14	15	16		
9	0,25332	-1,79359	1,48584		
10	0,54004	-1,80235	2,41515		
11	-0,33055	3,42685	-4,89576		
12	0,44688	-2,12249	2,50718		
13	2,0587	-2,12521	-2,06864		

Wissenschaftlicher Werdegang

Familienname : Cha
Vorname : Young-Lok
Geburtsdatum : 03. 10. 1966
Geburtsort : Kwang-Ju, Südkorea
Geschlecht : männlich
Familienstand : verheiratet
Staatsangehörigkeit : koreanisch

Schulausbildung

03. 1973 - 01.1979 : Elementary School in Na-Ju Korea
03. 1979 – 01.1982 : Middle School in Kwang-Ju Korea
03. 1982 – 01. 1985 : High School in Kwang-Ju Korea

Wehrdienst

03.1992 – 08.1993

Studium

03.1986 – 02. 1990 : Bachelor of engineering im Fachbereich “fine chemical technology“ an der Universität Chonnam in Kwang-Ju Korea
03.1990 – 02.1992 : Master of engineering im Fachbereich “Chemical Technology” an der Universität Chonnam in Kwang-Ju Korea
04.1996 – 02.1997 : Sprachkurs für Deutsch als Fremdsprache (PNdS-Kurs) an der Universität Hannover
04.1997 – 10.1998 : Besuch der Vorlesung für Chemie an der Universität Hannover
11.1998 – 11. 2005 : Promotion zum Doktor der Naturwissenschaften am Institut für Technische Chemie der Universität Hannover unter der Leitung von Prof. Dr. Bernd Hitzmann

Publikationen

- Young-Lok Cha, Bernd Hitzmann : Ultrasonic Measurements and its Evaluation for the Monitoring of *Saccharomyces cerevisiae* Cultivation, *Bioautomation* (2004) vol.1, pp.16-29
- B. Hitzmann, O.broxtermann, Y.-L. Cha, O.Sobieh, E. Staerk, T. Scheper: The control of glucose concentration during yeast fed-batch cultivation using a fast measurement complemented by an extended Kalman filter, *Bioprocess Engineering* 23 (2000) 337-341