

Analysis of Overdetermined System in Elemental Balances

A Project Report

Submitted as part of BBD852 part 2 MTech Major Project (BB5)

Submitted by:
Ayush Chachan
2016BB50005

Guided by:
Prof. Atul Narang



DEPARTMENT OF BIOCHEMICAL ENGINEERING AND BIOTECHNOLOGY
INDIAN INSTITUTE OF TECHNOLOGY DELHI

June 2022

DECLARATION

I certify that

1. The work contained in this report is original and has been done by me under the guidance of my supervisor.
2. I have followed the guideline provided by the Department in preparing the report.
3. I have conformed to the norms and the guidelines given in the Honour Code of Conduct of the Institute.
4. Whenever I have used materials (data, theoretical analysis, figures, and text) from other sources, I have given due credit to them by citing them in the text of the report and giving their details in the references. Further, I have taken permission from the copyright owners of the sources whenever necessary.

Signed by:

A handwritten signature in black ink, appearing to read 'Ayush Chachan', with a stylized flourish at the end.

Ayush Chachan

CERTIFICATE

It is certified that the work contained in this report titled “Analysis of Overdetermined System in Elemental Balances” is the original work done by Ayush Chachan and has been carried out under my supervision.

Signed by:



Prof. Atul Narang

ABSTRACT

Due to certain errors in measurements, the measurements done on biochemical reactor rarely form a consistent set of data which satisfies exactly the material and energy balances. The errors may be of random type, or they may be systematic errors. Elemental and Macroscopic balances can be used to check the consistency of the data in an overdetermined system. However, these elemental balances can also be used to identify the presence of gross errors and to locate the sources of gross errors if one has identified the gross errors in the measurements. In this project, we try to understand the complex statistical analysis of inconsistency of a material balance proposed by Madron et al. (1977)¹ for detection of possible gross and systematic errors. Finally, we will be implementing the method in python on various fermentation data sets from literatures.

¹ F. Madron, V. Veverka, and V. VanBcBk, AIChE J., 23,482 (1977).

Table of Contents

DECLARATION	2
CERTIFICATE	3
ABSTRACT.....	4
1.Introduction.....	6
2. Formulation of Elemental Balances	7
3. Best Estimate when measurement data contain only Random Errors	12
4. Classical Hypothesis Testing for Systematic Error Detection	17
5.Location of Source of Systematic Errors.....	18
6. Nomenclature.....	19
5. References.....	20
Appendix A : Analysis of Data by von Meyenburg using python	21
Appendix B: Analysis on Data by de Kok and Roels using python.....	23

1. Introduction

When carrying out the measurements necessary for the material balance of chemical reactors, one obtains data that contain certain errors, so that the measured values need not always comply with the stoichiometry of reactions taking place in the reactor. In such a case we say that the material balance data are not consistent.

The reliability of data is of great significance if they are to be used effectively in process monitoring for operational optimization, control or identification.

Measurements corrupted with undetected gross, biased or random errors result in false control, optimization or simple process identification.

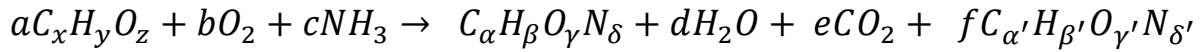
In the normal treatment of process data enhancement, it is invariably assumed that the only errors 'present in the data are normally distributed measurement errors with zero mean and known variances. In practice, the process data may also contain other types of errors, which are caused by non-random events. For instance, instrument biased may not be adequately compensated, measuring devices may malfunction. etc. We shall refer to these biases as gross errors.

In present work the systematic method for a complex statistical treatment of the material balance of a chemical reactor, proposed by Madron et al. (1977), is presented enabling us to identify data which might contain gross errors (measuring errors, effluxes of materials, occurrence of unknown side products, and the like).

It is capable of both detecting the presence of grossly biased measurement errors and locating the source of these errors in a bioreactor through statistical hypothesis testing. Equality constraints derived from material and energy balances are employed for the detection of data inconsistencies and for the subsequent identification of the suspect measurements by a process of data analysis and rectification. Maximum likelihood techniques are applied to the estimation of the states and parameters of the bioreactor after the suspect measurements have been eliminated. The level of significance is specified by the experimenter while the measurements are assumed to be randomly, normally distributed with zero mean and known variances.

2. Formulation of Elemental Balances

We will be using the black box model for analysing the consistency of data. In the black box, all the cellular reactions are lumped into one reaction, that of overall cell growth, in which the substrate is biologically converted to cell mass and extracellular metabolic product. It can be represented as:



where,

$C_xH_yO_z$ = substrate,

$C_\alpha H_\beta O_\gamma N_\delta$ = cell biomass,

$C_{\alpha'}H_{\beta'}O_{\gamma'}N_{\delta'}$ = extracellular metabolite product

One can write an elemental balance for above reaction as:

$$\begin{aligned} -aC_xH_yO_z - bO_2 - cNH_3 + C_\alpha H_\beta O_\gamma N_\delta \\ + dH_2O + eCO_2 + fC_{\alpha'}H_{\beta'}O_{\gamma'}N_{\delta'} = 0 \quad (1) \end{aligned}$$

If one further assumes that all the important chemical species involved in biological process (1) and their chemical formulae, including that of biomass, are known and constant, it is then possible to write the following elemental balances for C, H, O, and N:

for C:

$$xa = \alpha + e + \alpha'f \quad (2)$$

for H:

$$ya + 3c = \beta + 2d + \beta'f \quad (3)$$

for O:

$$za + 2b = \gamma + d + 2e + \gamma'f \quad (4)$$

for N:

$$c = \delta + \delta'f \quad (5)$$

One can write an energy balance also if the heat of fermentation, Q , is measured:

$$-aH_c - bH_{O_2} - cH_{NH_3} + H_b + dH_{H_2O} + eH_{CO_2} + fH_p + Q = 0 \quad (6)$$

The stoichiometric coefficients can be found by using the macroscopic balances over the bio reactor:

$$\frac{d(C_j V)}{dt} = \nu_j R V + \Phi_j \quad (7)$$

where,

C_j = concentration of component j in the system

Φ_j = net rate of input of component j to the system by transport

R = total rate of biomass production

ν_j = stoichiometric coefficient of eq. (1) for component j

V = system volume

Equation (7) can be used along with measurements of concentrations in the streams entering and leaving the system to find all the stoichiometric coefficients except that of water.

To calculate the coefficient of water (and other too, for example, coefficient of nitrogen source, if they have not been determined by experimental measurements), we use one (or two) of the elemental-energy balances (2) – (6). Since, the remaining balances must be satisfied they impose some equality constraints that must be satisfied. These constraints can be further used to test the consistency of data and to locate the erroneous measurements (if data measurement contains gross errors).

If we let \mathbf{x}'' a vector of stoichiometric coefficients, then equations (2) – (6) can be written using matrix notation as:

$$\mathbf{A}'\mathbf{x}'' = \mathbf{b}' \quad (9)$$

Where,

$$\mathbf{A}' = \begin{bmatrix} -x & 0 & 0 & \alpha & 0 & 1 & \alpha' \\ -y & 0 & -3 & \beta & 2 & 0 & \beta' \\ -z & -2 & 0 & \gamma & 1 & 2 & \gamma' \\ 0 & 0 & -1 & \delta & 0 & 0 & \delta' \\ -H_c & -H_{O_2} & -H_{NH_3} & H_b & H_{H_2O} & H_{CO_2} & H_p \end{bmatrix} \begin{matrix} \cdots C \\ \cdots H \\ \cdots O \\ \cdots N \\ \cdots \text{heat} \end{matrix}$$

$$\mathbf{x}'' = \begin{bmatrix} a \\ b \\ c \\ 1 \\ d \\ e \\ f \end{bmatrix} \quad \text{and} \quad \mathbf{b}' = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -Q \end{bmatrix}$$

We already know that all stoichiometric coefficients for equation (1) cannot be measured (for example, coefficient of water, as water is always present in excess and change in the concentration of water is very less which cannot be measured without huge errors), the unmeasured ones are then eliminated from equation (9) to yield:

$$\mathbf{Ax}' = \mathbf{b} \quad (10)$$

The vector \mathbf{x}' now contains all the measured coefficients. So, if, for example, the stoichiometric coefficient of water is not measured then \mathbf{A} , \mathbf{x}' and \mathbf{b} will be:

$$\mathbf{A} = \begin{bmatrix} -x & 0 & 0 & \alpha & 1 & \alpha' \\ -y + 2z & 4 & -3 & \beta - 2\gamma & -4 & \beta' - 2\gamma' \\ 0 & 0 & -1 & \delta & 0 & \delta' \\ \left(-z + \frac{H_c}{H_{H_2O}}\right) & \left(-2 + \frac{H_{O_2}}{H_{H_2O}}\right) & \left(\frac{H_{NH_3}}{H_{H_2O}}\right) & \left(\gamma - \frac{H_b}{H_{H_2O}}\right) & \left(2 - \frac{H_{CO_2}}{H_{H_2O}}\right) & \left(\gamma' - \frac{H_p}{H_{H_2O}}\right) \end{bmatrix} \quad (11)$$

$$\mathbf{x}' = \begin{bmatrix} a \\ b \\ c \\ 1 \\ e \\ f \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{Q}{H_{H_2O}} \end{bmatrix}$$

If heat of Fermentation Q has not been measured, then we cannot use equation (6) in the analysis so that the system of equations in (9) and (10) is homogeneous. If heat of fermentation Q is measured, which is beneficial as it contributes to extra redundancy, the system of equations in (9) and (10) is of the non-homogeneous form. In this case it is convenient to reduce equation (10) to homogeneous form by parallel translation as follows:

$$\mathbf{x} = \mathbf{x}' + \mathbf{h} \quad (12)$$

So that equation (10) reduces to:

$$\mathbf{Ax} = \mathbf{0} \quad (13)$$

Where the constant vector \mathbf{h} is some solution satisfying:

$$\mathbf{Ah} = -\mathbf{b} \quad (14)$$

Equation (13) now summarizes all the equality constraints that must be satisfied by the vector of measured variables (measured coefficients) \mathbf{x} .

3. Best Estimate when measurement data contain only Random Errors

Bioreactor measurements data always contains random error, and in some cases, there may even be systematic errors. As a result of such error equation (13) generally not satisfied. There will be some residuals different from zero when the measured stoichiometric coefficients are multiplied with the matrix \mathbf{A} .

This is better expressed by recognizing that the vector of true values (stoichiometric coefficients better known as yield coefficients) \mathbf{x} equals the sum of the measured vector $\bar{\mathbf{x}}$ and its corrupting general measurement error $\boldsymbol{\delta}$. That is,

$$\mathbf{x} = \bar{\mathbf{x}} + \boldsymbol{\delta} \quad (15)$$

If we assume that the error vector is distributed normally with a mean value of zero and a variance-covariance matrix $\boldsymbol{\psi}$, then we get:

$$\mathbf{E}(\boldsymbol{\delta}) = \mathbf{0}$$

$$\begin{aligned} \boldsymbol{\psi} &\equiv \text{Covar}(\bar{\mathbf{x}}) = \mathbf{E}[(\bar{\mathbf{x}} - \mathbf{E}(\mathbf{x})) (\bar{\mathbf{x}} - \mathbf{E}(\mathbf{x}))^T] \\ &= \mathbf{E}[(\bar{\mathbf{x}} - \mathbf{x}) (\bar{\mathbf{x}} - \mathbf{x})^T] \\ &= \mathbf{E}[\boldsymbol{\delta}\boldsymbol{\delta}^T] \end{aligned}$$

If each measurement error is assumed to be independent of other, that is, $\delta_1, \delta_2, \delta_3, \dots, \delta_n$ are assumed completely independent. The presence of the random errors $\boldsymbol{\delta}$ in the measurements will produce a vector of residuals $\boldsymbol{\epsilon}$ in the balance equations. It is easy to show that $\boldsymbol{\epsilon}$ is related to $\boldsymbol{\delta}$ by:

$$\epsilon \equiv \mathbf{A}\delta = -\mathbf{A}\bar{\mathbf{x}} \quad (18)$$

$$E[\epsilon] = \mathbf{A}E[\delta] = \mathbf{0} \quad (19)$$

$$\phi \equiv E[\epsilon\epsilon^T] = \mathbf{A}E[\delta\delta^T]\mathbf{A}^T = \mathbf{A}\psi\mathbf{A}^T \quad (20)$$

We have

$$\mathbf{x} = \bar{\mathbf{x}} + \delta$$

$$\mathbf{A}\mathbf{x} = \mathbf{A}\bar{\mathbf{x}} + \mathbf{A}\delta$$

$$\Rightarrow \quad \mathbf{0} = \mathbf{A}\bar{\mathbf{x}} + \epsilon$$

$$\epsilon = -\mathbf{A}\bar{\mathbf{x}}$$

Where,

\mathbf{x} = vector of true values of coefficients

$\bar{\mathbf{x}}$ = vector of measured values of coefficients

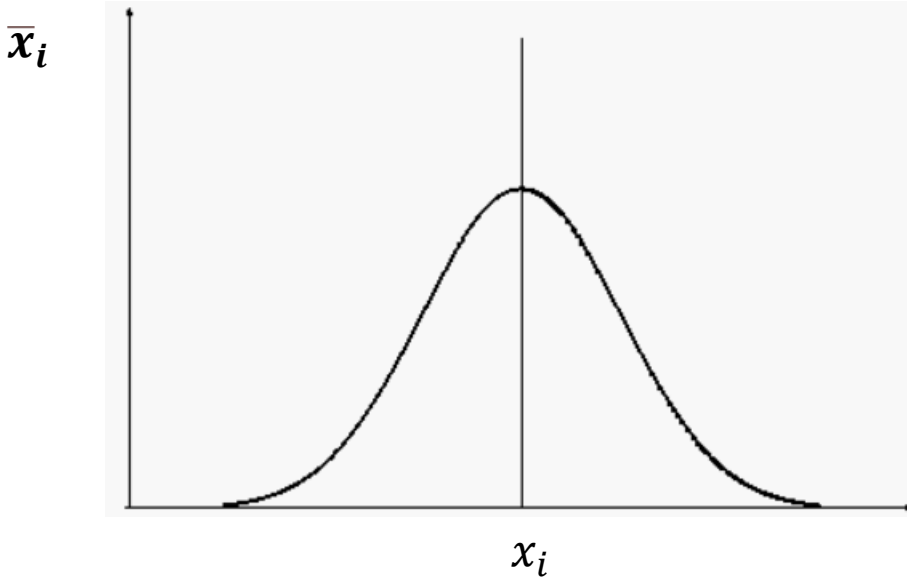
δ = vector of measurement errors

ϵ = vector of residuals

Also,

$$\mathbf{E}(\bar{\mathbf{x}}) = \mathbf{x} \quad \Rightarrow \quad \mathbf{E}(\delta) = \mathbf{0}$$

$$\begin{aligned} \psi \equiv \text{Covar}(\bar{\mathbf{x}}) &= \mathbf{E}[(\bar{\mathbf{x}} - \mathbf{E}(\bar{\mathbf{x}}))(\bar{\mathbf{x}} - \mathbf{E}(\bar{\mathbf{x}}))^T] \\ &= \mathbf{E}[(\bar{\mathbf{x}} - \mathbf{x})(\bar{\mathbf{x}} - \mathbf{x})^T] \\ &= \mathbf{E}[\delta\delta^T] \end{aligned}$$



If $\delta_1, \delta_2, \dots, \delta_n$ are assumed to be independent then, ψ will be diagonal. Since actual values of ψ are not known, simple variances will suffice. Each \bar{x}_i is a normal random variable with mean x_i and some variance σ_i^2 . we can write joint density function for n-dimensional random variable $\bar{\mathbf{x}}$:

$$f(\bar{\mathbf{x}}) = \frac{1}{(2\pi)^{n/2} \sqrt{|\psi|}} \exp\left\{-\frac{1}{2}(\bar{\mathbf{x}} - \mathbf{x})^T \psi^{-1}(\bar{\mathbf{x}} - \mathbf{x})\right\}$$

likelihood function of $\bar{\mathbf{x}}$,

$$L(\bar{\mathbf{x}} | \mathbf{x}) = \frac{1}{(2\pi)^{n/2} \sqrt{|\psi|}} \exp\left\{-\frac{1}{2}(\bar{\mathbf{x}} - \mathbf{x})^T \psi^{-1}(\bar{\mathbf{x}} - \mathbf{x})\right\}$$

We want to find the value of x which maximize the *likelihood function of \bar{x}* , that is, we need to solve the following optimization problem:

$$\underset{x}{\text{Maximize}} L(\bar{x} | x) \equiv \underset{x}{\text{Minimize}} (\bar{x} - x)^T \psi^{-1} (\bar{x} - x) \equiv \underset{\delta}{\text{Minimize}} \delta \psi^{-1} \delta^T$$

Now, our optimization problem reduces to

$$\underset{\delta}{\text{Minimize}} J = \delta \psi^{-1} \delta^T \quad \text{where } \delta \text{ must satisfy the constraint } A\delta = \epsilon$$

To solve the above problem, the **method of Lagrange multipliers** is used:

$$\underset{\delta}{\text{Minimize}} J = \delta \psi^{-1} \delta^T \quad \text{where } \delta \text{ must satisfy the constraint } A\delta = \epsilon$$

The Lagrange function is given by

$$\mathcal{L}(\delta, \lambda) = \delta \psi^{-1} \delta^T + \lambda^T A \delta$$

The optimality conditions are

$$\nabla_{\delta} \mathcal{L} = 0$$

Solving $\nabla_{\delta}\mathcal{L} = \mathbf{0}$

$$\Rightarrow \nabla_{\delta}\mathcal{L} = \nabla_{\delta}(\delta\psi^{-1}\delta^T + \lambda^T A\delta) = \mathbf{0} \quad (1)$$

$$\Rightarrow 2\psi^{-1}\delta + (\lambda^T A)^T = \mathbf{0} \text{ (using } \nabla_x \mathbf{b}^T x = \mathbf{0} \text{ for } \mathbf{b} \in \mathbb{R}^n \text{)} \quad (2)$$

$$\Rightarrow 2\psi^{-1}\delta + A^T\lambda = \mathbf{0} \quad (3)$$

$$\Rightarrow \delta = \frac{-\psi A^T \lambda}{2} \quad \text{Now, Putting in } A\delta = \epsilon, \text{ we get} \quad (4)$$

$$\Rightarrow \frac{-A\psi A^T \lambda}{2} = \epsilon \quad (5)$$

$$\Rightarrow \lambda = -2(A\psi A^T)^{-1}\epsilon \quad (6)$$

From (4) and (5) we get,

$$\hat{\delta} = \psi A^T (A\psi A^T)^{-1} \epsilon$$

$$\Rightarrow \hat{\delta} = \psi A^T \phi^{-1} \epsilon$$

The better estimate for \mathbf{x} is given by

$$\hat{\mathbf{x}} = \bar{\mathbf{x}} + \hat{\delta}$$

4. Classical Hypothesis Testing for Systematic Error Detection

Null Hypothesis:

H_0 : $\delta_1, \delta_2, \dots, \delta_n$ are independent of one another and uncorrelated.

Test function for Hypothesis Testing:

$$\mathbf{h} = \boldsymbol{\epsilon}^T \boldsymbol{\phi}^{-1} \boldsymbol{\epsilon} \sim \chi^2(m)$$

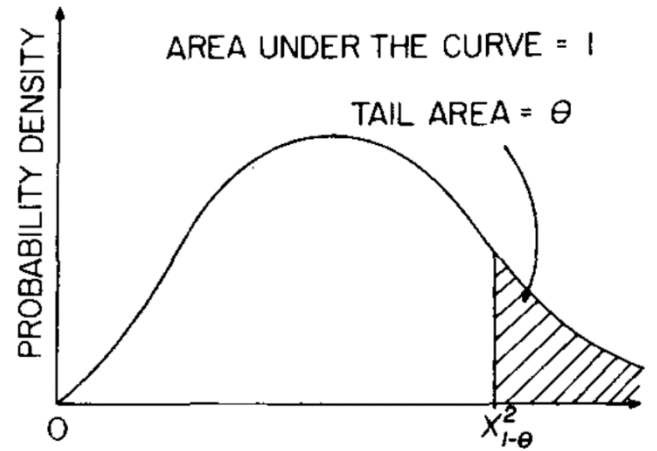


Fig: $\chi^2_{1-\theta}$ probability distribution with m degrees of freedom

It turns out that if the error vectors $\delta_1, \delta_2, \dots, \delta_n$ are independent of one another and uncorrelated then the test function h will follow chi-square distribution with n degrees of freedom.

Our judgment on error detection is strongly dependent on the outcome of the hypothesis testing performed on the test function.

If $\mathbf{h} > \chi^2_{1-\theta}(\mathbf{m})$ the test fails and we reject the hypothesis with a confidence level of $1 - \theta$.

If, however, $\mathbf{h} \leq \chi^2_{1-\theta}(\mathbf{m})$ the test passes and we fail to reject the null hypothesis.

5. Location of Source of Systematic Errors

If the source of a gross error remains unidentified, a single large correction, which is actually needed only on the erroneous measurement, is distributed among all other statistically correct measurements which need little adjustment. To locate the measurements with gross errors when the original complete set of data fail the hypothesis test, we delete one measurement at a time, and calculate the performance index after the deletion of each measurement.

Due to the numerical identity of the test function \mathbf{h} , the calculation of performance indices can be carried out before the actual adjustment is made. These performance indices obtained from each deletion of one measurement are then compared with one another and also with the χ^2 criteria, with the degree of freedom appropriately decreased by one to account for the loss of one constraint equation due to the deletion of one measurement.

In short:

1. We want to find the measurements with gross errors when the original set of data fail the hypothesis test.
2. We delete one measurement at a time and calculate the test function $\mathbf{h} = \boldsymbol{\epsilon}^T \boldsymbol{\phi}^{-1} \boldsymbol{\epsilon}$ after the deletion of each measurement.
3. \mathbf{h} obtained from each deletion of one measurement are then compared with the χ^2 criteria with the degree of freedom appropriately decreased by 1.

6. Nomenclature

a, b, c, d, e, f	stoichiometric coefficients in eq. (1); a is in mol substrate/C mol biomass; b is in mol O ₂ /C mol biomass; c is in mol NH ₃ , /C mol biomass; d is in mol H ₂ O/C mol biomass; e is in C ₀₂ /C mol biomass; f is in mol product/C mol biomass
Q	heat evolution in eq. (6) (kcal/C mol biomass)
x, y, z	composition of substrate in eq. (1)
$\alpha, \beta, \gamma, \delta$	composition of biomass in eq. (1)
$\alpha', \beta', \gamma', \delta'$	composition of metabolic product in eq. (1)
C_j	concentration of component j in the system
Φ_j	net rate of input of component j to the system by transport
R	total rate of biomass production
ν_j	stoichiometric coefficient of eq. (1) for component j
V	system volume
\mathbf{x}	true parameter vector for measured variables (dimension = n)
$\bar{\mathbf{x}}$	measurement vector (dimension = n)
$\boldsymbol{\delta}$	measurement error vector (dimension = n)
$\boldsymbol{\epsilon}$	balance equation residual vector (dimension = m)

5. References

- i. Metabolic Engineering: Principles and Methodologies
- ii. Wang and Stephanopoulos (1983). Application of Macroscopic Balances to the identification of Gross Measurement Errors.
- iii. H. E. de Kok and J. A. Roels, Biotechnol. Bioeng., 22, 1097 (1980).
- iv. Python Data Science Handbook Essential Tools for Working with Data by Jake VanderPlas.
- v. F. Madron, V. Veverka, and V. VanBcBk, AIChE J., 23,482 (1977).
- vi. Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython, Wes McKinney
- vii. Bijkerk and Hall (1977)
- viii. Nielsen and Villadsen (1994)

Appendix A : Analysis of Data by von Meyenburg² using python

Original Data:

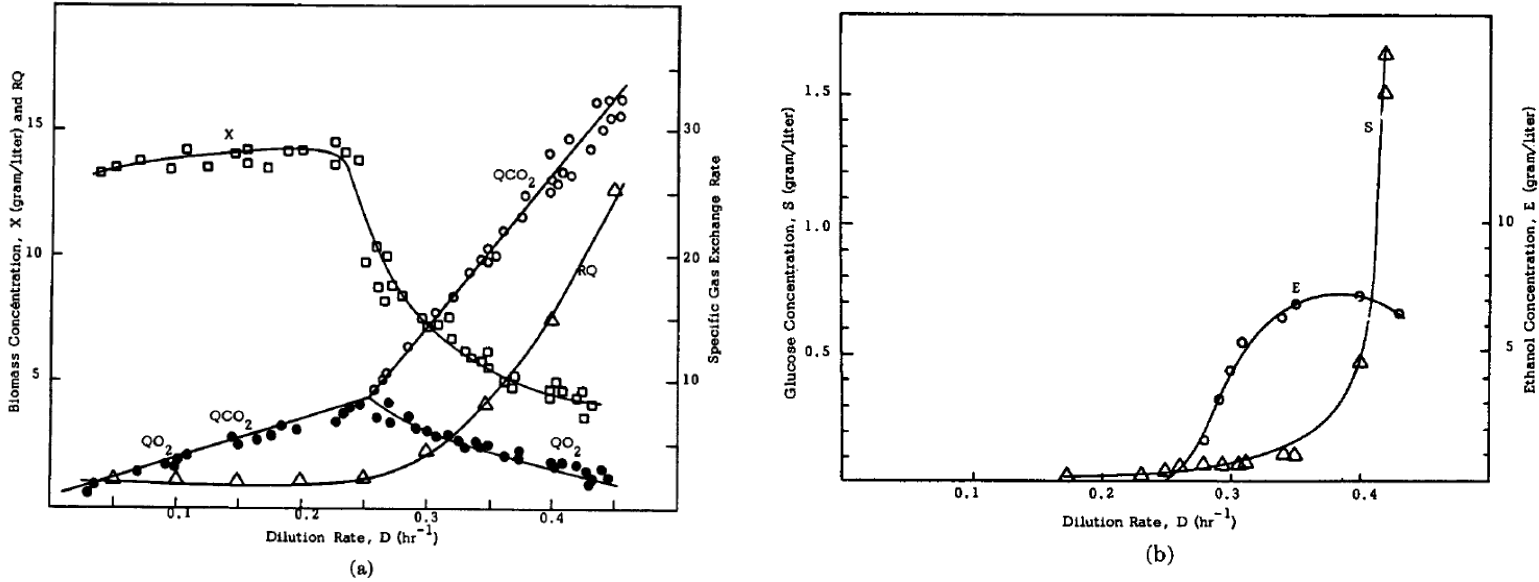


Fig: Chemostat culture of *S. cerevisiae*. (a) biomass conc., x (gDW/L) \square ; specific O_2 uptake, r_{O_2} (mmoles / gDW.h) \bullet ; specific CO_2 production (r_{CO_2} (mmoles / gDW.h) \circ ; (b) glucose conc., s (g/L) Δ ; ethanol conc., p (g/L) \circ .

Black Box Model for Chemostat culture of *S. cerevisiae*:



- The feed was 28 (g/L) sterile glucose solution.
- At low dilution rate no ethanol is produced.

² Bijkerk, Hall, VOL. XIX, PAGES 267-296 (1977)

Out[2]:

	Dilution Rate (1/h)	biomass concentration (g/L)	substrate concentration (g/L)	ethanol concentration (g/L)	specific CO2 production rate (mmoles / gDW.h)	specific O2 consumption rate (mmoles / gDW.h)
0	0.15	14.0	0.00	0.0	4.5	4.50
1	0.30	7.1	0.10	4.2	14.4	6.57
2	0.40	4.4	0.46	8.3	26.0	3.66

Table 1 : Measured Experimental Data by von Meyenburg

Out[4]:

	D	q_x (C mole / L.h)	$-q_s$ (C mole / L.h)	q_p (C mole / L.h)	q_c (C mole / L.h)	q_o (mole / L.h)
0	0.15	0.076758	0.1400	0.000000	0.06300	0.063000
1	0.30	0.077855	0.2790	0.054783	0.10224	0.046647
2	0.40	0.064331	0.3672	0.144348	0.11440	0.016104

Table 2 : Calculated Data; q_x = biomass production rate; $-q_s$ = substrate consumption rate; q_p = ethanol production rate; q_c = CO₂ production rate; q_o = O₂ consumption rate;

	D	DW	h (no deletion)	h after deleting CH ₂ O	h after deleting O ₂	h after deleting CH _{1.83} O _{0.56} N _{0.17}	h after deleting CO ₂	h after deleting CH ₃ O _{0.5}
0	0.15	14.0	0.22	0.19	0.00	0.18	0.12	0.11
1	0.30	7.1	20.93	4.61	6.05	3.52	18.73	0.02
2	0.40	4.4	11.57	3.57	3.67	2.70	9.59	0.00

Table 3 : Statistical analysis performed using python; h = hypothesis test function

☐ No systematic error at $D = 0.15$ (1/h)

☐ Error in ethanol measurement at $D = 0.3$ (1/h) and $D = 0.4$ (1/h)

Note: $\chi^2(2)$ at 90.0 % confidence level = 4.60;
 $\chi^2(1)$ at 90.0 % confidence level = 2.70

Appendix B: Analysis on Data by de Kok and Roels³ using python

The method explained above was applied to full data set of de kok and Roels.

Saccharomyces cerevisiae CBS 426 was cultivated in a medium in which glucose is utilized as the carbon source and ammonium sulphate as the nitrogen source.

The steady-state flows of the substrate, oxygen, biomass, and carbon dioxide were measured for a range of dilution rates. $CH_{1.83}O_{0.56}N_{0.17}$ used as the elemental composition for the biomass in correcting the experimental data to conform to the elemental balances. No extracellular metabolic product is assumed to be present.

Thus, reaction for this problem is:



$$A'x' = 0$$

$$\begin{array}{l} \text{C} \quad \dots \\ \text{H} \quad \dots \\ \text{O} \quad \dots \\ \text{N} \quad \dots \end{array} \begin{bmatrix} -6 & 0 & 0 & 1.00 & 0 & 1 \\ -12 & 0 & -3 & 1.83 & 2 & 0 \\ -6 & -2 & 0 & 0.56 & 1 & 2 \\ 0 & 0 & -1 & 0.17 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ B \\ d \\ e \end{bmatrix} = 0$$

³ H. E. de Kok and J. A. Roels, *Biotechnol. Bioeng.*, 22, 1097 (1980).

TABLE I
Dry Weight (DW) and Glucose Consumption Rate, Oxygen
Consumption Rate, and Carbon Dioxide Production Rate per Unit of
Biomass Produced of a Glucose-Limited Continuous Culture of *S.*
cerevisiae CBS 426

D (hr ⁻¹)	DW (g/liter)	ϕ_{glucose} (C mol/C mol biomass)	ϕ_{O_2} (mol/C mol biomass)	ϕ_{CO_2} (mol/C mol biomass)	
0.008	3.6	2.0	1.1	1.4	a
		2.18	1.09	1.18	b
0.008	3.8	2.1	3.8	1.4	a
		2.39	1.34	1.39	b
0.017	3.2	2.3	1.5	1.6	a
		2.50	1.43	1.49	b
0.033	4.0	1.9	1.04	1.1	a
		2.06	0.99	1.06	b
0.047	3.17	2.48	1.2	1.2	a
		2.42	1.19	1.30	b
0.052	4.2	1.8	0.82	0.86	a
		1.87	0.81	0.86	b
0.072	4.41	1.74	0.54	0.73	a
		1.69	0.63	0.69	b
0.076	4.0	1.93	0.67	0.73	a
		1.73	0.68	0.73	b
0.092	4.4	1.8	0.7	0.7	a
		1.73	0.67	0.73	b
0.092	3.8	1.9	0.7	0.8	a
		1.80	0.74	0.80	b
0.102	4.4	1.8	0.62	0.85	a
		1.70	0.65	0.71	b
0.112	3.6	2.1	0.78	0.92	a
		1.94	0.88	0.94	b
0.113	4.3	1.8	0.8	1.0	a
		1.88	0.83	0.89	b
0.118	4.5	1.71	0.57	0.67	a
		1.66	0.60	0.66	b

Data from H. E. de Kok and J. A. Roels, Biotechnol. Bioeng.,
22, 1097 (1980)

Out[2]:

	Dilution Rate (1/h)	Dry weight (g/L)	Glucose Consumption Rate (C mol/C mol biomass)	Oxygen Consumption Rate (mol O2/C mol biomass)	Carbon Dioxide Production Rate (mol CO2/C mol biomass)
0	0.008	3.60	2.00	1.10	1.40
1	0.008	3.80	2.10	3.80	1.40
2	0.017	3.20	2.30	1.50	1.60
3	0.033	4.00	1.90	1.04	1.10
4	0.047	3.17	2.48	1.20	1.20
5	0.052	4.20	1.80	0.82	0.86
6	0.072	4.41	1.74	0.54	0.73
7	0.076	4.00	1.93	0.67	0.73
8	0.092	4.40	1.80	0.70	0.70
9	0.092	3.80	1.90	0.70	0.80
10	0.102	4.40	1.80	0.62	0.85
11	0.112	3.60	2.10	0.78	0.92
12	0.113	4.30	1.80	0.80	1.00
13	0.118	4.50	1.71	0.57	0.67

Out[4]:

	D	DW	C6H12O6	O2	NH3	CH1.83O0.56N0.17	H2O	CO2
0	0.008	3.60	-0.333333	-1.10	None	1.0	None	1.40
1	0.008	3.80	-0.350000	-3.80	None	1.0	None	1.40
2	0.017	3.20	-0.383333	-1.50	None	1.0	None	1.60
3	0.033	4.00	-0.316667	-1.04	None	1.0	None	1.10
4	0.047	3.17	-0.413333	-1.20	None	1.0	None	1.20
5	0.052	4.20	-0.300000	-0.82	None	1.0	None	0.86
6	0.072	4.41	-0.290000	-0.54	None	1.0	None	0.73
7	0.076	4.00	-0.321667	-0.67	None	1.0	None	0.73
8	0.092	4.40	-0.300000	-0.70	None	1.0	None	0.70
9	0.092	3.80	-0.316667	-0.70	None	1.0	None	0.80
10	0.102	4.40	-0.300000	-0.62	None	1.0	None	0.85
11	0.112	3.60	-0.350000	-0.78	None	1.0	None	0.92
12	0.113	4.30	-0.300000	-0.80	None	1.0	None	1.00
13	0.118	4.50	-0.285000	-0.57	None	1.0	None	0.67

Table 4 : Calculated yield coefficients; all are in $\left(\frac{\text{C mole}}{\text{C mol substrate}}\right) D =$
dilution rate $\left(\frac{1}{h}\right)$, *DW* = *biomass concentration in* $\frac{g}{L}$.

	D	DW	h (no deletion)	h after deleting C6H12O6	h after deleting O2	h after deleting CH1.83O0.56N0.17	h after deleting CO2
0	0.008	3.60	3.91	1.53	3.90	1.69	0.67
1	0.008	3.80	35.07	27.06	2.12	26.43	34.96
2	0.017	3.20	2.07	0.04	1.70	0.06	1.19
3	0.033	4.00	1.65	0.00	1.31	0.01	1.18
4	0.047	3.17	1.99	0.07	1.85	0.10	1.18
5	0.052	4.20	0.23	0.01	0.15	0.00	0.21
6	0.072	4.41	2.42	1.86	0.01	1.73	1.27
7	0.076	4.00	2.20	0.01	1.78	0.00	1.98
8	0.092	4.40	0.53	0.20	0.50	0.23	0.12
9	0.092	3.80	1.01	0.17	0.43	0.13	1.00
10	0.102	4.40	2.50	2.29	0.11	2.20	0.86
11	0.112	3.60	2.71	0.43	1.12	0.33	2.70
12	0.113	4.30	1.73	1.07	1.51	1.14	0.11
13	0.118	4.50	0.54	0.25	0.09	0.22	0.46

Table 5 : Statistical analysis performed using python; *h* = hypothesis test function

Out[9]:

	D	DW	C6H12O6	O2	NH3	CH1.83O0.56N0.17	H2O	CO2	RQ
0	0.008	3.60	-0.333333	-1.10	None	1.0	None	1.40	1.272727
1	0.008	3.80	-0.350000	-3.80	None	1.0	None	1.40	0.368421
2	0.017	3.20	-0.383333	-1.50	None	1.0	None	1.60	1.066667
3	0.033	4.00	-0.316667	-1.04	None	1.0	None	1.10	1.057692
4	0.047	3.17	-0.413333	-1.20	None	1.0	None	1.20	1.000000
5	0.052	4.20	-0.300000	-0.82	None	1.0	None	0.86	1.048780
6	0.072	4.41	-0.290000	-0.54	None	1.0	None	0.73	1.351852
7	0.076	4.00	-0.321667	-0.67	None	1.0	None	0.73	1.089552
8	0.092	4.40	-0.300000	-0.70	None	1.0	None	0.70	1.000000
9	0.092	3.80	-0.316667	-0.70	None	1.0	None	0.80	1.142857
10	0.102	4.40	-0.300000	-0.62	None	1.0	None	0.85	1.370968
11	0.112	3.60	-0.350000	-0.78	None	1.0	None	0.92	1.179487
12	0.113	4.30	-0.300000	-0.80	None	1.0	None	1.00	1.250000
13	0.118	4.50	-0.285000	-0.57	None	1.0	None	0.67	1.175439

Table 6 : Calculated yield coefficients; Respiratory Coefficient (RQ).