A systematic approach for the processing of experimental data from anaerobic syngas fermentations

Benalcázar, Eduardo Almeida; Noorman, Henk; Filho, Rubens Maciel; Posada, John

DOI
10.1016/B978-0-323-95879-0.50218-6

Publication date
2022

Document Version
Final published version

Published in
Computer Aided Chemical Engineering

PROCEEDINGS OF THE 32nd European Symposium on Computer Aided Process Engineering (ESCAPE32), June 12-15, 2022, Toulouse, France
L. Montastruc, S. Negny (Editors)

http://dx.doi.org/10.1016/B978-0-323-95879-0.50218-6

A systematic approach for the processing of experimental data from anaerobic syngas fermentations

Eduardo Almeida Benalcázar [a,b], Henk Noorman [b,c], Rubens Maciel Filho [a], John Posada [b]

[a] *Department of Product and Process Development, Faculty of Chemical Engineering, State University of Campinas, Av. Albert Einstein 500, 13083-852, Campinas – SP, Brazil.*
[b] *Department of Biotechnology, Faculty of Applied Sciences, Delft University of Technology, Van der Maasweg 9, 2629 HZ, Delft, the Netherlands.*
[c] *DSM Biotechnology Center, A. Fleminglaan 1, 2613 AX, Delft, the Netherlands.*

Abstract

This study describes a methodological framework designed for the systematic processing of experimental syngas fermentation data for its use by metabolic models at pseudo-steady state and at transient state. The developed approach allows the use of not only own experimental data but also from experiments reported in literature which employ a wide range of gas feed compositions (from pure CO to a mixture between H_2 and CO_2), different pH values, two different bacterial strains and bioreactor configurations (stirred tanks and bubble columns).

The developed data processing framework includes *i)* the smoothing of time-dependent concentrations data (using moving averages and statistical methods that reduce the relevance of outliers), *ii)* the reconciliation of net conversion rates such that mass balances are satisfied from a black-box perspective (using minimizations), and *iii)* the estimation of dissolved concentrations of the syngas components (CO, H_2 and CO_2) in the fermentation broth (using mass transfer models). Special care has been given such that the framework allows the estimation of missing or unreported net conversion data and metabolite concentrations at the intra or extracellular spaces (considering that there is availability of at least two replicate experiments) through the use of approximative kinetic equations.

Keywords: Syngas fermentation, experimental data processing, fermentation data reconstruction, data reconciliation.

1. Introduction

The development of a new fermentation technology starts at the laboratory where the chemostat is one of the main tools for studying microbial behavior. The chemostat is a fermentation setup where the concentrations of substances and cells are maintained nearly constant, and therefore a pseudo-steady state is reached at the intra and extracellular spaces (Noorman et al., 1996; Villadsen et al., 2011). The chemostat serves for gaining insights on the steady-state regulations of, on one hand, the whole metabolism such as the growth rate dependencies on substrates and products concentrations or the substrate requirements for maintenance, and on the other hand, the intracellular reaction rates that

allow cells to exhibit their behavior expressed as net conversion rates, or q-rates, as well as through the rates of growth and decay.
The measurements of substance concentrations is the main source of information about the state of a fermentation process. The net conversion rates are estimated from these measurements. The substances that are generally measured are the carbon and the electron sources, the excreted products, cells, oxygen (in aerobic fermentations), CO_2, H^+ ions and the nitrogen source, (Stephanopoulos and Tsiveriotis, 1989). Yet, measurements are commonly prone to errors, both random and systematic (van der Heijden et al., 1994b). The reconciliation process is an adjustment of the measured rates aimed at improving their accuracy, such that they fulfill constraints formulated by, for instance a black-box description, which contains the energetic as well as the elemental balances of carbon, hydrogen, oxygen, nitrogen and charge (Noorman et al., 1996, 1991; van der Heijden et al., 1994a). In underdetermined systems, the reconciliation may allow the estimation of unmeasured quantities, whereas in overdetermined systems, the reconciliation allows to assess the consistency of the collected data and the identification of gross measurement errors (Stephanopoulos and Tsiveriotis, 1989; van der Heijden et al., 1994b, 1994a).
The development and up-scaling of the anaerobic syngas fermentor requires the design of mathematical models that are able to reproduce the *i)* characteristics of mass transfer from the gas to the liquid and the *ii)* microbial metabolic responses to the stimulus provided by the extracellular environment. The models on the side of the microorganism are much nurtured by the adequate collection and processing of experimental information. In this document we describe a systematic methodological framework we developed to process the experimental metabolite concentrations and conversion rates from syngas fermentation experiments. We have previously used the proposed framework for assessing the performance of the reported and own fermentation experiments in terms of the distribution of carbon and electrons among the fermentation products, as well as for the parameterization of metabolic models at pseudo-steady state and transient state.
The following pages describe the methodological framework formed by: *i)* data curation and smoothing, *ii)* reconciliation of net conversion rates and *iii)* reconstruction of missing concentrations of the dissolved gases.

2. Data curation and smoothing

In general, the curation and smoothing of the experimental data is useful for the design of kinetic expressions and the early assessment of trends. The use of curated and smoothed data is however not recommended for the parameterization of models because such procedure may introduce errors or may lead to the negligence of exceptional phenomena producing outliers in the data trends.
Curation is here referred to the process of filling missing concentration data in time points where samples have been withdrawn and for any particular reason the concentration of one or more substances have not been able to be measured or the measurements fell below the equipment detection limits. In either of the two cases the empty data points may be filled using the overall trends of the data as reference. For instance, if the missing data correspond to points where low concentrations are expected, as the low points in the oscillatory data collected by (Mahamkali et al., 2020), then the missing data can be replaced by values somewhere around the detection limit of the measuring equipment. If the missing data is instead expected higher than the detection limits of the measuring equipment, their estimation may be supported on the overall trend, where spline interpolation may be far more useful than linear interpolation. The maintenance of the trends in the time derivatives may also be used as a strategy for selecting the interpolation

method. Time derivatives of experimental concentrations (C_j) may be estimated using forward, central and backward differentiation for the first, middle and last points in a time series, respectively. Table 1 shows the equations applicable to the three mentioned differentiation methods.

Table 1 Equations used for the estimation of the concentration gradients with the experimental data

Type of differentiation	Equation	nr.
Central	$\frac{\mathrm{d}C_j}{\mathrm{d}t} \approx \frac{\Delta C_j}{\Delta t} = \frac{C_{t+1} - C_{t-1}}{2 \cdot \Delta t}$	(1)
Forward	$\frac{\mathrm{d}C_j}{\mathrm{d}t} \approx \frac{\Delta C_j}{\Delta t} = \frac{C_{t+1} - C_t}{\Delta t}$	(2)
Backward	$\frac{\mathrm{d}C_j}{\mathrm{d}t} \approx \frac{\Delta C_j}{\Delta t} = \frac{C_t - C_{t-1}}{\Delta t}$	(3)

Smoothing is here referred to the reduction of the noise in the trends of dynamic experiments. The earliest smoothing method to be tried may be moving averages. However, the higher and lower points in oscillatory data, as in (Mahamkali et al., 2020), may be neglected by moving averages. As more refined alternatives, the smoothing may be done using local regression methods that use weighted linear least squares or 1st degree polynomial models that assign lower weight to outliers in the regression. The last two methods may be found coded into MatLab's 'smooth' function as the algorithms 'lowess' and 'rlowess', respectively. In any case, the selection of smoothing algorithm may be, as described for the curation of data, supported on the trends of the time derivatives and on trends from a second experiment, provided it is available.

3. Reconciliation of net conversion rates

In general terms, the reconciliation process consists on the minimization of the mismatch between the reconciled data and the raw experimental data while closing the carbon, hydrogen, oxygen, nitrogen elemental balances as well as the charge balances. Prior to describing how the minimization problem is created, it is necessary to provide an introduction on how the mass balances are structured for the fermentation of syngas.

The fermentation of syngas employs the ability of diverse types of bacteria and archaea, often called as acetogens, to harvest the electrons from CO and H_2 and the carbon from CO and CO_2. CO, H_2 and CO_2 can be the sole sources of energy and carbon for these microorganisms. Acetic acid, ethanol and 2,3-butanediol are the most common native products of acetogenic microorganisms. The first step on the processing of experimental data is the establishment of stoichiometric relations between electron and energy sources, the products and cells. Equations 4 – 9 in Table 2, show the stoichiometries of the reactions leading from CO and H_2 towards acetic acid, ethanol and 2,3-butanediol. In equations 4 – 9, the amount of electron donor required to generate each product is calculated by balancing the degree of reduction (or the amount of electron available for redox exchange (Heijnen, 2002)) between the donor and the catabolic product; CO_2 is added to balance the carbon, H_2O balances the oxygen atoms, and H^+ ions balance the hydrogen atoms. If the balancing is made correctly, in the mentioned order of steps, the charge should also be balanced with no additional modifications to the stoichiometry.

For simplicity, here we show only apparent species. CO_2, acetic acid and potentially the nitrogen source have different forms depending on the pH. More thorough descriptions can include these species and their equilibrium relations to improve accuracy.

The complexity of the sequence of reactions leading to the production of cells, or anabolism, may also be summarized using two equations, one for each electron donor (see equations 10 and 11 in Table 1). The derivation of the stoichiometry of the anabolic reaction follows the same steps as described previously.
In theory, the catabolic reactions of any syngas fermentation process where CO and H_2 are consumed and where acetate, ethanol and 2,3-butanediol are the main products should be described by a combination between equations 4 – 9. The same is applied to the anabolic reaction, which is derived from a combination between equations 10 and 11. The resulting catabolic and anabolic reactions may be joined to form one metabolic reaction using the biomass yield, which can be estimated using thermodynamics (Heijnen and van Dijken, 1992).

Table 2. Common reactions in syngas fermentations

Part of metabolism	Reaction	nr.
Catabolism from CO as the electron donor	$-4CO - 2H_2O + C_2H_3O_2^- + 2CO_2 + H^+$	(4)
	$-6CO - 3H_2O + C_2H_6O + 4CO_2$	(5)
	$-11CO - 5H_2O + C_4H_{10}O_2 + 7CO_2$	(6)
Catabolism from H_2 as the electron donor	$-4H_2 - 2CO_2 + C_2H_3O_2^- + 2H_2O + H^+$	(7)
	$-6H_2 - 2CO_2 + C_2H_6O + 3H_2O$	(8)
	$-11H_2 - 4CO_2 + C_4H_{10}O_2 + 6H_2O$	(9)
Anabolism	$-2CO - 0.25NH_4^+ - 0.5H_2O + CH_{1.75}O_{0.5}N_{0.25} + CO_2 + 0.25H^+$	(10)
	$-2H_2 - 0.25NH_4^+ - CO_2 + CH_{1.75}O_{0.5}N_{0.25} + 1.5H_2O + 0.25H^+$	(11)

Going back to the structuring of the minimization problem for reconciling the net conversion rates, the mismatch between the reconciled and the experimental rates is the objective function; the mismatch may be quantified by the sum of the squared differences divided by the standard deviation (Villadsen et al., 2011). The reconciled net conversion rates for CO, H_2, CO_2, acetate, ethanol, 2,3-butanediol, biomass growth, water, nitrogen source and H^+ ions may be used as decision variables. The elemental balances may be used as constraints. Additional constraints may also be applied for improving the precision; for instance, the mass balances of the different species in the two phases of the bioreactor, the liquid and the gas phases. The experimental preference for CO and H_2 uptake and for product generation may also be used as additional constraints.
It is standard procedure to fix the allowed variations of the experimental q-rates according to the error in the experimental data (known using data from multiple experiments at the same conditions); however, if that error is not reported in the source of experimental data, one may assume a percent variation of maximum 10 – 20 % for each reconciled rate. In this case, the use of additional constraints become very helpful to guide the reconciliation to values that are consistent not only with elemental balances but also with time dependent or independent balances of species.

4. Calculation of the dissolved gas concentrations

Techniques for in-line measurement of the dissolved concentrations of CO, H_2 and CO_2 have been developed (Mann et al., 2021; Mislov et al., 2015). However, it is rare to find reports that describe the use of such techniques in syngas fermentations. The calculation of the dissolved concentrations of CO, H_2 and CO_2 is therefore often necessary. This

calculation can be done using two approaches that are different if the experimental data can be assumed at steady state or at transient state.
If the experimental data is likely at steady-state, the time gradients are assumed equal to zero and the mass balances of CO and H_2 in the gas and the liquid phases, one overall CO_2 mass balance and one summation for the composition of the off-gas (see equations 12 - 17) may be enough to estimate the unknown concentrations of CO, H_2, CO_2 in the liquid phase, as well as the molar fractions of CO and H_2 in the off-gas and their flow rates. If CO_2 is a product of the fermentation, it can be assumed that it is saturated in the liquid, thus Henry's law may be applied to find its fraction in the gas phase.

CO in the liquid phase	$0 = q_{CO} \cdot C_x + k_L a_{CO} \cdot (C^*_{CO} - C_{CO}) - C_{CO} \cdot D$	(12)
H_2 in the liquid phase	$0 = q_{H_2} \cdot C_x + k_L a_{H_2} \cdot \left(C^*_{H_2} - C_{H_2}\right) - C_{H_2} \cdot D$	(13)
CO in the gas phase	$0 = F_{G,in} \cdot y_{CO,in} - F_{G,off} \cdot y_{CO,off} - k_L a_{CO} \cdot (C^*_{CO} - C_{CO}) - C_{CO} \cdot D$	(14)
H_2 in the gas phase	$0 = F_{G,in} \cdot y_{H_2,in} - F_{G,off} \cdot y_{H_2,off} - k_L a_{H_2} \cdot \left(C^*_{H_2} - C_{H_2}\right) - C_{H_2} \cdot D$	(15)
CO_2 overall	$0 = F_{G,in} \cdot y_{CO_2,in} - F_{G,off} \cdot y_{CO_2,off} + q_{CO} \cdot C_x \cdot V_L - C_{CO_2} \cdot D \cdot V_L$	(16)
Summation of gas phase composition	$1 = y_{CO,off} + y_{H_2,off} + y_{CO_2,off}$	(17)

The saturation concentrations of the gases (C^*_{CO}, $C^*_{H_2}$ and $C^*_{CO_2}$) may be estimated using Henry's equation (a list of coefficients is given by (Sander, 2015)). Moreover, the estimation of the mass transfer coefficients ($k_L a$) requires making several assumptions about the fermentation broth, the dimensions of the bioreactor vessel and the dimensions of the stirrer; unfortunately, these details are commonly not reported. The calculation of $k_L a$ may be based on the power input by the stirrer (P_S in equation 6) and by the gas sparging (P_{SG} in equation 7). Note that equation 18 and 19 were developed for the transfer of oxygen to pure water at 20 °C; therefore $k_L a$ may be further corrected for the process temperature and the specific gas through their the film diffusivity (Đ) of CO and H_2 in pure water compared to that of O_2 (see equation 20). Equation 20 also contains a correction factor for the mass transfer coefficient ($f_{k_L a}$), which may account for other differences in the experimental set-up and liquid phase composition compared to the ideal case for which equation 18 - 20 were initially constructed and parametrized.

$$P_S = N_P \cdot \rho \cdot N^3 \cdot {D_S}^5 \tag{18}$$

$$P_{SG} = P_S \cdot \left[0.1 \cdot \left(\frac{N \cdot V_L}{\dot{V}_G}\right)^{\frac{1}{4}} \cdot \left(\frac{g \cdot H_S \cdot {V_L}^{\frac{2}{3}}}{N^2 \cdot {D_S}^4}\right)^{\frac{1}{5}}\right] \tag{19}$$

$$k_L a_j = f_{k_L a} \cdot \left[1.022^{(T-293.15)}\right] \cdot \left[0.002 \cdot \left(\frac{P_{sG}}{V_L}\right)^{0.7} \cdot {v^c_{Gs}}^{0.2}\right] \cdot \frac{Đ_{O_2}}{Đ_j} \tag{20}$$

If the experimental data is instead expected to be at transient state, the concentration gradients represent one more unknown variable per each gas in the balances. The additional equations needed may be formulated using kinetic equations linking the net conversion rates of CO, H_2 and CO_2 to their dissolved concentrations. Mechanistic or

approximative formats may be used (Heijnen, 2005; Rizzi et al., 1997). The combination between both formats has been useful in our work.
To improve the accuracy of the reconciliation of the ten conversion rates, the calculation of dissolved gas concentrations may be done simultaneously by including the unknown C_{CO}, C_{H_2} and C_{CO_2} , $y_{CO,off}$, $y_{H_2,off}$ and $F_{G,off}$ among the decision variables list.

References

Heijnen, J.J., 2005. Approximative kinetic formats used in metabolic network modeling. Biotechnology and Bioengineering 91, 534–545.

Heijnen, J.J., 2002. Bioenergetics of Microbial Growth, in: Flickinger, M.C., Drew, S.W. (Eds.), Encyclopedia of Bioprocess Technology. John Wiley & Sons, Inc., Hoboken, NJ, USA, pp. 265–291.

Heijnen, J.J., van Dijken, J.P., 1992. In search of a thermodynamic description of biomass yields for the chemotrophic growth of microorganisms. Biotechnology and Bioengineering, 39, 833–858.

Mahamkali, V., Valgepea, K., de Souza Pinto Lemgruber, R., Plan, M., Tappel, R., Köpke, M., Simpson, S.D., Nielsen, L.K., Marcellin, E., 2020. Redox controls metabolic robustness in the gas-fermenting acetogen *Clostridium autoethanogenum*. Proc Natl Acad Sci USA 117, 13168–13175.

Mann, M., Miebach, K., Büchs, J., 2021. Online measurement of dissolved carbon monoxide concentrations reveals critical operating conditions in gas fermentation experiments. Biotechnology and Bioengineering, 118, 253–264.

Mislov, D., Cifrek, M., Krois, I., Dzapo, H., 2015. Measurement of dissolved hydrogen concentration with clark electrode, in: 2015 IEEE Sensors Applications Symposium (SAS). Presented at the 2015 IEEE Sensors Applications Symposium (SAS), IEEE, Zadar, Croatia, pp. 1–5.

Noorman, H.J., Heijnen, J.J., Ch. A. M. Luyben, K., 1991. Linear relations in microbial reaction systems: A general overview of their origin, form, and use. Biotechnology and Bioengineering, 38, 603–618.

Noorman, H.J., Romein, B., Luyben, K.Ch.A.M., Heijnen, J.J., 1996. Classification, error detection, and reconciliation of process information in complex biochemical systems. Biotechnology and Bioengineering 49, 364–376.

Rizzi, M., Baltes, M., Theobald, U., Reuss, M., 1997. In vivo analysis of metabolic dynamics in Saccharomyces cerevisiae: II. Mathematical model. Biotechnology and Bioengineering 55, 592–608.

Sander, R., 2015. Compilation of Henry's law constants (version 4.0) for water as solvent. Atmospheric Chemistry and Physics 15, 4399–4981.

Stephanopoulos, G., Tsiveriotis, C., 1989. Toward a Systematic Method for the Generalization of Fermentation Data, in: Fish, N.M., Fox, R.I., Thornhill, N.F. (Eds.), Computer Applications in Fermentation Technology: Modelling and Control of Biotechnological Processes. Springer Netherlands, Dordrecht, pp. 169–178.

van der Heijden, R.T.J.M., Heijnen, J.J., Hellinga, C., Romein, B., Luyben, K.Ch.A.M., 1994a. Linear constraint relations in biochemical reaction systems: I. Classification of the calculability and the balanceability of conversion rates. Biotechnology and Bioengineering, 43, 3–10.

van der Heijden, R.T.J.M., Romein, B., Heijnen, J.J., Hellinga, C., Luyben, K.Ch.A.M., 1994b. Linear constraint relations in biochemical reaction systems: II. Diagnosis and estimation of gross errors. Biotechnology and Bioengineering, 43, 11–20.

Villadsen, J., Nielsen, J.H., Lidén, G., Nielsen, J.H., 2011. Bioreaction engineering principles, 3rd ed. Springer, New York.