Modeling the Glucose Concentration for the Recombinant *E.coli* Bioprocesses

Benas Kemesis, Renaldas Urniezius*, Vygandas Vaitkus, Arnas Survyla, Rimvydas Simutis, Povilas Babilius, Lukas Zlatkus

Kaunas University of Technology, Dept of Automation, Kaunas LT-51368, Lithuania
renaldas.urniezius@ktu.lt

This article presents the sophisticated-to-date carbon mass balance for fed-batch *E.coli* bioprocesses. The model originates from the distribution of carbon mass from glucose in biomass, off-gas, and hypothetical solutes. The suggested model complements Pirt's equation as a particular case scenario. The approach uses the linear relationship between biomass carbon content per carbon grams in glucose and average cell population age. The carbon balance brings two potential practical benefits. First, it has the potential to assess the type of cell metabolism pathway and to have a soft sensor for the concentration of dissolved products such as acetates. The measure of glucose concentration suggests another finding, assuring the reliance on off-gas information only. The paper introduces an average carbon content ratio in biomass and off-gas, with numerical values of 0.5 in growth-limiting experiments and 0.27 in nonlimiting ones, which may serve as a decision-making criterion for metabolic pathway detection in the future.
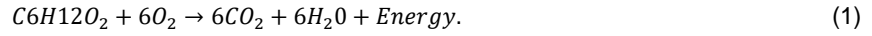
1. Introduction

Modeling glucose concentration in nonstationary fed-batch processes remains desired in research and development institutions and industrial installations. Therefore, corresponding dedicated online control techniques are unavailable. Currently, the regulators are present for bioreactor-related fed-batch setpoints (Levisauskas et al., 2019) and growth-limiting substrate feed for fed-batch (Galvanauskas et al., 2021) cultivations. Up until now, the state variables' estimation is accessible through off-gas-based analytics routines providing soft sensors for microbial biomass (Urniezius and Survyla, 2019), targeting non-soluble recombinant product (Urniezius et al., 2021), and the specific biomass growth rate (Survyla et al., 2021). All mentioned methods are noninvasive and rely on exhaust gas mixture information.

Moreover, they share a common benefit: such approaches do not suffer issues when the nutrition medium is nonhomogeneous, making them acceptable for industrial scale-up efforts. The reason is that the exhaust gas presents cumulative knowledge with the trade-off of unavoidable estimation delay due to the bioreactor outlet channel length and broth's changing volume. Therefore, the analytical optimization and resolution of crucial state parameters in fed-batch recombinant bioprocesses are still challenging, involving multiple dilutions and time constants due to asynchronous observations from different external sensors.

So far, the specific glucose consumption rate behavior was modeled offline in both growth-limiting and repeated batch microbial cultivations (Survyla et al., 2022). Glucose concentration modeling efforts returned accumulated errors due to minor deviations in the substrate concentration and flow rate (Schubert et al., 1997). Building the parametric relationship with physiological meaning is ongoing. Therefore, we share an approach for glucose concentration modeling in fed-batch aerobic bacterial cultivations. The aim is to resolve the challenges of synchronizing different state variables: glucose concentration, the carbon content in biomass, and bacteria population age linked to the off-gas biosynthesis information.
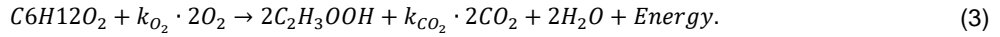
2. Metabolic Pathways

The oxidative metabolism path is the desired metabolic pathway in aerobic fed-batch cultivations of recombinant *E. coli*. The general equation for aerobic respiration (Patel et al., 2022)

$$C6H12O_2 + 6O_2 \rightarrow 6CO_2 + 6H_2O + Energy. \tag{1}$$

From a fed-batch process perspective, the latter equation expands by introducing the additional chemical compounds involved:

$$C6H12O2 + k_{O_2} \cdot 6O2 + Nutrient\ medium \rightarrow k_{CO_2} \cdot 6CO2 + 6H2O + Biomass + Solutes, \tag{2}$$
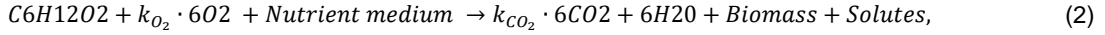
where oxygen and carbon dioxide coefficients k_{O_2} and k_{CO_2} relate to the respiratory quotient RQ. Correspondingly, the equation for the overflow metabolism (Deng et al., 2021):

$$C6H12O_2 + k_{O_2} \cdot 2O_2 \rightarrow 2C_2H_3OOH + k_{CO_2} \cdot 2CO_2 + 2H_2O + Energy. \tag{3}$$

In most cases, the respiratory quotient helps to determine the current cell metabolic pathway (Heyman et al., 2020). This paper proposes that the respiratory quotient numeric value is close to one for all metabolic pathways in *E.coli* fed-batch bioprocesses

$$RQ = \frac{6CO_2}{6O_2} = \frac{k_{O_2}}{k_{CO_2}} \cong 1. \tag{4}$$

Experimental proof confirms this proposition when measuring the off-gas carbon production rate CPR and oxygen uptake rate OUR no matter the chosen overflow mode, i.e., glucose feed can be limiting or non-limiting the biomass growth.

2.1 Carbon Mass Balance

Looking at Eq(2) through the perspective of the carbon element, the total carbon balance emerges

$$TC_{C6H12O2} = \sum_{i=1}^{m} TC_i, \tag{5}$$

where total carbon $TC_{C6H12O2}$ in fed glucose and nutrient medium is equal to the sum of total carbon TC_i in each of the products *m,* produced from the aerobic process

$$\sum_{i=1}^{m} TC_i = TC_X + TC_{sol} + TC_{CPR}. \tag{6}$$

The sum of total carbon TC involves the carbon content in biomass TC_X, soluble byproducts TC_{sol} (e.g., acetates) and produced off-gas TC_{CPR}. Given this, Eq(2) turns into the mass balance equation for carbon

$$m_{C6H12O2} \cdot C_s = C_X \cdot \Delta X_{total} + TC_{sol} + C_{CPR} \cdot \int CPR \cdot dt, \tag{7}$$

where the first term $m_{C6H12O2}$ represents the total mass of carbon in glucose, $C_s = 0.03$ and $C_X = 0.04$ coefficients are the carbon content in moles per gram of glucose *s* and total biomass X_{total}, respectively (Xu et al., 1999). The constant $C_{CPR} = 0.272$ relates to the weight ratio of carbon and oxygen in carbon dioxide.
Data analysis showed that the coefficient (C_X) is not a process constant. It instead represents a linear relationship with the average age of the cell population $\overline{Age}$, which is a function of biomass concentration *X* (Urniezius et al., 2021)

$$\overline{Age}(t) \cong \frac{\sum_{j=0}^{i}(t_i - t_j)\Delta X(t_j)}{X_i}. \tag{8}$$

The variable $m_{C6H12O2}$ consists of carbon masses in feeding glucose $m_{fed,C6H12O2}$, initial glucose m_0 and glucose concentration m_s as follows

$$m_{C6H12O2} = m_{fed,C6H12O2} + m_0 - m_s. \tag{9}$$

It leads to an equivalent form of Eq(7), solved for boundary conditions

$$\left(\frac{(F(t) - F_0) \cdot S_f}{S_d} + \left(s_0 \cdot W_0 - s(t) \cdot W(t)\right)\right) \cdot C_s$$
$$= C_X(t) \cdot (X_{total}(t) - X_0) + TC_{sol,total}(t) + C_{CPR} \cdot \int_{t_0}^{t} CPR(t^*) \cdot dt^*, \tag{10}$$

In Eq(10), the variable F is the total fed glucose, and W stands for the total volume of the cultivation medium. The initial values of the state variables after inoculation into the bioreactor are as follows

$$\begin{aligned} F_0 &\equiv F(t = 0), \\ W_0 &\equiv W(t = 0), \\ X_0 &\equiv X_{total}(t = 0), \\ s_0 &\equiv s(t = 0). \end{aligned} \tag{11}$$

The coefficient S_f is the concentration of glucose in the feed, while the coefficient (S_d) is the glucose density which ensures the balance is in grams.

Expressing biomass growth $\Delta X_G = (X_{total}(t) - X_0)$ and dividing Eq(10) by total biomass $X_{total}(t)$ and $\Delta t \equiv t - t_0$ produces specific rate terms concerning total biomass and time

$$\frac{\left(\frac{(F(t) - F_0) \cdot S_f}{S_d} + \left(s_0 \cdot W_0 - s(t) \cdot W(t)\right)\right)}{X_{total}(t) \cdot \Delta t} \cdot C_s = C_X(t) \cdot \frac{\Delta X_G}{\Delta t} \cdot \frac{1}{X_{total}(t)} +$$
$$\frac{TC_{sol,total}(t)}{X_{total}(t) \cdot \Delta t} + C_{CPR} \cdot \frac{\int_{t_0}^{t} CPR(t^*) \cdot dt^*}{X_{total}(t) \cdot \Delta t}, \tag{12}$$

where the biomass growth term becomes the average specific growth rate $\bar{\mu}(t)$

$$\bar{\mu}(t) = \frac{\Delta X_G}{\Delta t} \cdot \frac{1}{X_{total}(t)}. \tag{13}$$

Extending the idea of specific rates, Xu (1999) presented specific rates of glucose consumption for the oxidative pathway $qS_{ox,an}$ and cell maintenance q_m. Taking this into account, Eq(12) takes a different form

$$(qS_{ox,an} + q_m) \cdot C_s = C_X(t) \cdot \bar{\mu}(t) + \frac{TC_{sol,total}(t)}{X_{total}(t) \cdot \Delta t} + C_{CPR} \cdot \frac{\int_{t_0}^{t} CPR(t^*) \cdot dt^*}{X_{total}(t) \cdot \Delta t}$$
$$\rightarrow C_X(t) \cdot \bar{\mu}(t) + \overline{TC_{sol_X}}(t) + C_{CPR} \cdot \overline{CPR_X}(t). \tag{14}$$

Rearranging and dividing Eq(14) by the average specific growth rate $\bar{\mu}(t)$ and the coefficient C_s yields

$$\frac{C_X(t)}{C_s} = \frac{(qS_{ox,an} + q_m)}{\bar{\mu}(t)} + \frac{\overline{TC_{sol_X}}(t) + C_{CPR} \cdot \overline{CPR_X}(t)}{C_s \cdot \bar{\mu}(t)}, \tag{15}$$

where the summand relating to the total carbon content in off-gas and solutes turns into

$$\frac{\overline{TC_{sol_X}}(t) + C_{CPR} \cdot \overline{CPR_X}(t)}{C_s \cdot \bar{\mu}(t)} = \frac{TC_{sol}(t)}{X(t) \cdot \Delta t \cdot \frac{\Delta X_G}{\Delta t} \cdot \frac{1}{X(t)} \cdot C_s} + \frac{C_{CPR} \cdot \int_{t_0}^{t} CPR(t^*) \cdot dt^*}{X(t) \cdot \Delta t \cdot \frac{\Delta X_G}{\Delta t} \cdot \frac{1}{X(t)} \cdot C_s}$$
$$\rightarrow \frac{TC_{sol}(t)}{\Delta X_G \cdot C_s} + \frac{C_{CPR} \cdot \int_{t_0}^{t} CPR(t^*) \cdot dt^*}{\Delta X_G \cdot C_s}. \tag{16}$$

Combining Eq(15) and Eq(16) results in:

$$\frac{C_X(t)}{C_s} = \frac{(qS_{ox,an} + q_m)}{\bar{\mu}(t)} + \frac{TC_{sol}(t)}{\Delta X_G \cdot C_s} + \frac{C_{CPR} \cdot \int_{t_0}^{t} CPR(t^*) \cdot dt^*}{\Delta X_G \cdot C_s}. \tag{17}$$

In Eq(17), the carbon term for growth G of biomass has a different expression when maintenance energy is zero (Pirt, 1965). It relates to the aerobic yield coefficient $\left(Y_{\frac{X}{S},ox}\right)$ exclusive of maintenance, which is assumed constant throughout the process (Xu et al., 1999)

$$qS_{ox,an} \over \bar{\mu}(t)} = {\Delta TS_G \over \Delta X_G} = {1 \over Y_{\frac{X}{S},ox}}, \qquad (18)$$

where ΔTS is the total fed glucose.
Joining Eq(17) and Eq(18) yields

$${C_X(t) \over C_S} = {q_m \over \bar{\mu}(t)} + {1 \over Y_{\frac{X}{S},ox}} + {TC_{sol}(t) \over \Delta X_G \cdot C_S} + {C_{CPR} \cdot \int_{t_0}^{t} CPR(t^*) \cdot dt^* \over \Delta X_G \cdot C_S}. \qquad (19)$$

Assuming that the observed growth yield of cells $Y(t)$ is inversely proportional to carbon balance terms

$${1 \over Y(t)} = {\Delta TS \over \Delta X} \cong {C_X(t) \over C_S} \qquad (20)$$

concludes that the Pirt equation is a particular case of Eq(20) that does not account for the carbon content in solubles and off-gas carbon dioxide (Pirt, 1965):

$${1 \over Y(t)} = {q_m \over \bar{\mu}(t)} + {1 \over Y_{\frac{X}{S},ox}}. \qquad (21)$$

A simplified illustration of proposed model:

Figure 1: Simplified proposition of the paper – Pirt equation is a special case of the full carbon mass balance

2.2 Glucose Estimation

The glucose concentration expression solved from Eq(10):

$$s(t) = {\left(s_0 \cdot W_0 + {(F(t) - F_0) \cdot S_f \over S_d} - \left({C_X(t) \cdot (X(t) - X_0) + C_{CPR} \cdot \int_{t_0}^{t} CPR(t^*) \cdot dt^* \over C_S} \right) \right) \over W_i}. \qquad (22)$$

Numerical integration of Eq(22) leads to intermediate values of glucose concentration $\hat{s}_i^*$:

$$\hat{s}_i^* = {\left(s_0 \cdot W_0 + {(F_i - F_0) \cdot S_f \over S_d} - \left({C_{X,i} \cdot (X_i - X_0) + C_{CPR} \cdot \sum_{j=1}^{i} CPR_i \cdot \Delta tj \over C_S} \right) \right) \over W_i}. \qquad (23)$$

The final estimated form $(\hat{s}_i)$ is then dependent on a fuzzy rule

$$\hat{s}_i = \begin{cases} \hat{s}_i = 0, & if\ \hat{s}_i^* \le 0 \\ \hat{s}_i = \hat{s}_i^*, & otherwise \end{cases}, \qquad (24)$$

3. where the logical condition assures the validity of the boundary condition, i.e., the computed glucose concentration is always non-negative.Experimental Verification

The mean absolute error (MAE) operated as evaluation criteria to compare glucose estimation results

$$MAE = \frac{\sum_{i=1}^{n}|\widehat{s_i} - s_i^*|}{n}, \tag{25}$$

where the glucose concentration (s_i^*) had discrete values observed every 15-60 minutes during the cultivation process. The total number of samples $n = 171$ covered 5 experiments, of which 3 had growth-limiting feeding. The experimental data came from one site with Escherichia coli BL21(DE3) pETM-11+EGFP strain.

Figure 2 provides examples of linear regression models of the coefficient C_X for each experiment.
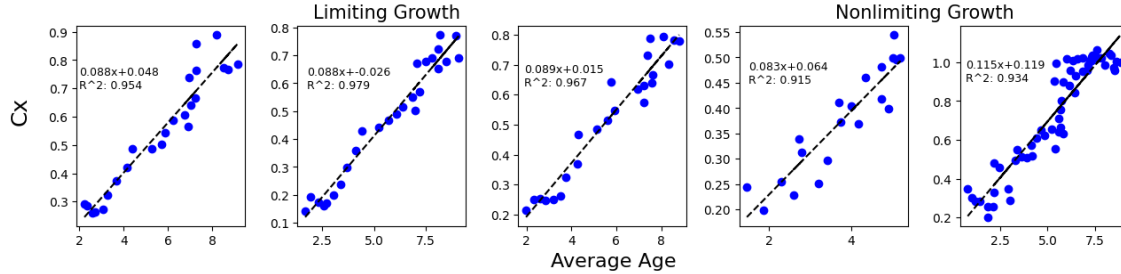

Figure 2: Carbon content in grams per gram of biomass C_X dependence on average cell population age presenting slope, offset and R^2 numerical values.

The final model of glucose concentration used a shared set of fitted gain and slope parameters

$$C_X(t) = 0.0813 \cdot \overline{Age}(t) - 0.0803. \tag{26}$$

Experimental data had no information about the concentrations of secreted compounds in the culture medium. Therefore, an average ratio of carbon content cross-verified the hypothesis by using

$$(s_n^* - \widehat{s_n}) \propto \left(\frac{(CPR \cdot C_{CPR}) \cdot \mathrm{d}t}{\mathrm{d}(X_{total} \cdot C_X)} \right). \tag{27}$$

A decrease in the average ratio between the exhaust carbon and the carbon in the biomass itself implies that the remaining carbon accumulates in the growth medium. The terms in Eq(27) support this assumption by resulting in a correlation coefficient being 0.949. Hence, the estimation errors in Figure 3 represent accumulated byproducts with their total carbon content.

Table 1: The results of model fitting on experimental data

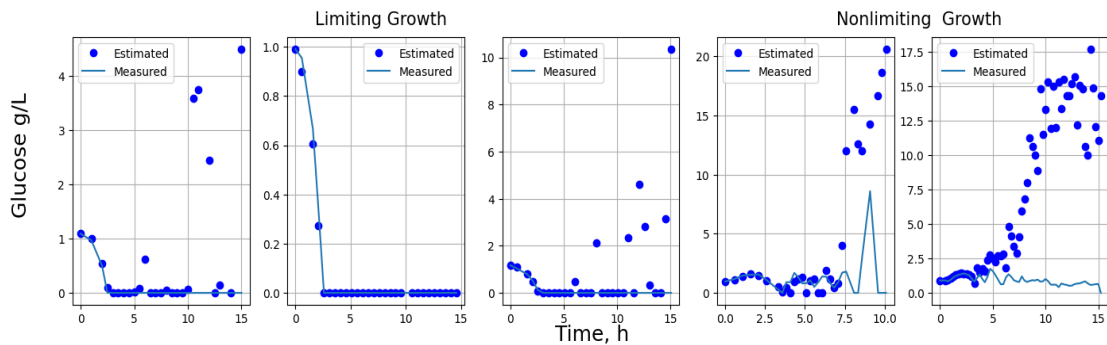| Experiment No.	MAE, g/L	n	Growth Type	$\overline{Ratio_{CPR/X}}$	$(s_n^* - \widehat{s_n})$, g/L
1	0.621	25		1.765	-0.622
2	1.434	27	Limiting	2.725	-0.06
3	0.927	29		2.376	-0.927
4	3.967	30	Nonlimiting	1.824	-3.967
5	6.642	60		0.917	-6.642

Figure 3: Estimated and measured glucose concentrations for all five experiments.

4. Conclusions

The article presents a development of the Pirt equation considering the dynamical carbon balance in a fed-batch bioprocess. The newly introduced terms of carbon contents in off-gas and byproducts allowed us to advance to a complete mathematical model that models the glucose concentration during the bioprocess. Experimental data of five fed-batch bioprocesses, three growth-limited, verified the model, with glucose concentration estimation errors ranging from 0.6 to 6.7 g/L. The article also presents the discovered linear relationship between carbon content in grams per gram of biomass and the average age of the cell population, which reduces the model parameter count. A presented model verification criterion can also indicate the cellular metabolic pathway, thus allowing bioprocess operators to make appropriate decisions. The glucose estimation errors indicate that the model's current state requires further development. The lack of produced solubles measurements also affects the model's prediction accuracy, especially in non-limiting growth experiments. Overall, the described carbon balance hypothesis further expands the potential of off-gas for non-invasive estimation of bioprocess state variables. It is especially relevant when monitoring glucose concentration, where an adequate model eliminates or reduces the need for excessive sampling.

Acknowledgments

References

Deng Y., Beahm D. R., Ionov S., Sarpeshkar R., "Measuring and modeling energy and power consumption in living microbial cells with a synthetic ATP reporter," BMC Biol, vol. 19, no. 1, p. 101, Dec. 2021, doi: 10.1186/s12915-021-01023-2.

Galvanauskas V., Simutis R., Levisauskas D., Urniezius R., Vaitkus V., Tekorius T., Butkus M., Survyla A. Patent LT6861B, 2021b-11-10. Method for design and adaptation of feeding profiles for recombinant E.coli fed-batch cultivation processes. [Accessed 2022-01-28]

Heyman B., Tulke H., Putri S. P., Fukusaki E., and Büchs J., "Online monitoring of the respiratory quotient reveals metabolic phases during microaerobic 2,3-butanediol production with Bacillus licheniformis," Eng. Life Sci., vol. 20, no. 3–4, pp. 133–144, Mar. 2020, doi: 10.1002/elsc.201900121.

Levisauskas D., Simutis R., Galvanauskas V., Urniezius R., 2019, Simple Control Systems for Set-point Control of Dissolved Oxygen Concentration in Batch Fermentation Processes, Chemical Engineering Transactions, 74, 127-132.

Patel H., Kerndt C. C., and Bhardwaj A., "Physiology, Respiratory Quotient," in StatPearls, Treasure Island (FL): StatPearls Publishing, 2022. Accessed: Feb. 12, 2023. [Online]. Available: http://www.ncbi.nlm.nih.gov/books/NBK531494/

Pirt S. J. (1965). The Maintenance Energy of Bacteria in Growing Cultures. Proceedings of the Royal Society of London. Series B, Biological Sciences, 163(991), 224–231. doi:10.2307/75569

Schubert, J., Simutis, R., Dors, M., Havlik, I., & Lübbert, A. (1994). Bioprocess optimization and control: Application of hybrid modeling. In Journal of Biotechnology (Vol. 35, Issue 1, pp. 51–68). Elsevier BV. https://doi.org/10.1016/0168-1656(94)90189-9

Survyla, A., Levisauskas, D., Urniezius, R., & Simutis, R. (2021). An oxygen-uptake-rate-based estimator of the specific growth rate in Escherichia coli BL21 strains cultivation processes. In Computational and Structural Biotechnology Journal (Vol. 19, pp. 5856–5863). Elsevier BV. https://doi.org/10.1016/j.csbj.2021.10.015.

Survyla, A., Urniezius, R., Kemesis, B., Zlatkus, L., Masaitis, D., & Galvanauskas, V. (2022). Modeling the Specific Glucose Consumption Rate for the Recombinant E. coli Bioprocesses Based on Aging-specific Growth Rate. Chemical Engineering Transactions, 93, 265-270.

Urniezius, R., & Survyla, A. (2019). Identification of Functional Bioprocess Model for Recombinant E. Coli Cultivation Process. In Entropy (Vol. 21, Issue 12, p. 1221). MDPI AG. https://doi.org/10.3390/e21121221

Urniezius, R., Kemesis, B., & Simutis, R. (2021). Bridging Offline Functional Model Carrying Aging-Specific Growth Rate Information and Recombinant Protein Expression: Entropic Extension of Akaike Information Criterion. In Entropy (Vol. 23, Issue 8, p. 1057). MDPI AG. https://doi.org/10.3390/e23081057

van Bodegom P., "Microbial Maintenance: A Critical Review on Its Quantification," Microb Ecol, vol. 53, no. 4, pp. 513–523, May 2007, doi: 10.1007/s00248-006-9049-5.

Xu B., Jahic M., and Enfors S.-O., "Modeling of Overflow Metabolism in Batch and Fed-Batch Cultures of Escherichia coli," Biotechnol. Prog., vol. 15, no. 1, pp. 81–90, Feb. 1999, doi: 10.1021/bp9801087.