Possibilities of R programming language in simulating microbiological synthesis processes

Marina A. Nikitina, Irina M. Chernukha
V.M. Gorbatov Federal Research Center for Food Systems of RAS, Moscow, 26, str. Talalikhina, 109316, Russia
nikitinama@yandex.ru

Abstract. Information technologies of biotechnological processes are based on the use of mathematical models to describe microbiological synthesis. Application of digital technologies in analysis of microbial growth patterns is mainly determined by the ability of modern programming languages to numerically integrate systems of differential equations describing the development of the microbial process in time. In Jupyter Notebook environment in the R programming language, the solution of the kinetic growth model of the E.coli microbial population was shown. Two solution methods were used - the one-step Runge-Kutta method of the fourth order of accuracy and the universal solver ODE (General Solver for Ordinary Differential Equations). Initial data of the problem in question:

\[ \frac{K_s}{S_0} = 2 \quad (K_s \text{ is substrate affinity constant for the biomass (microorganism)}, S_0 \text{ is initial concentration of substrate}); \]

repopulating cells \( m_{a0} = 0.01 \); total number of cells \( m_0 = 0.05 \); stoichiometric ratio \( Y_s = 0.5 \); various ratios

1. \( \frac{\mu}{\mu_m} = 0.0357 \); 2. \( \frac{\mu}{\mu_m} = 0.0714 \); 3. \( \frac{\mu}{\mu_m} = 0.1071 \); 4. \( \frac{\mu}{\mu_m} = 0.1428 \); 5. \( \frac{\mu}{\mu_m} = 0.2142 \) (\( \lambda \) is specific growth rate of dividing cells, \( \mu_m \) is inactivation rate constant).

As a result, the simulation and verification of microbial biomass growth process - its visual representation in the form of tabular and graphical data were carried out. In the process of simulation of E.coli growth the following peculiarity was revealed. In addition to cell division, a fairly intensive loss of their ability to divide occurs. This process is supposedly determinant in population development and limits the growth and ultimate density of the culture. Thus, information technology will help the researcher not only in studying the process, establishing patterns and predicting results, but also in making reasoned decisions.

1. Introduction

Information technology for biotechnological processes is based on the use of mathematical models to describe microbial synthesis. Currently, a great number of mathematical models such as Monod, Moser, Haldane, Teissier models, etc. [1-8] are used to describe kinetics of microbial synthesis, which usually refers to dynamics of substrate consumption, biomass and metabolite production.

Simulation is of particular importance in scientific study of biotechnological systems, both isolated and interconnected. With virtual process simulation, one can study behavior of a complex system with a given set of characteristics when internal and/or external parameters change.

Metadata from scientific literature show that growth phases of microbial populations have been well studied and described. Kinetic parameters such as specific bacterial cell growth and die-off rates, average generation time, etc. are used to quantitatively characterize physiological activity [8].

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Published under licence by IOP Publishing Ltd
industrial production of cell biomass, mathematical simulation method is widely used to establish regularities of growth processes and to determine influence of environmental factors [9, 10]. This approach is of particular interest when studying dynamics of producer growth in the inducible promoter-controlled expression system of E. coli [11].

Our paper focuses on the detailed analysis of the kinetic model of microbial population growth with inactivation of replicative capacity of cells in R software environment.

2. Methods
In order to solve differential equations we used the one-step Runge-Kutta method of the fourth order of accuracy and the universal solver ODE (General Solver for Ordinary Differential Equations).

Numerical integration was performed in Jupyter Notebooks software environment in the R programming language [12, 13].

3. Results
The change in viable cell concentration $M_a$ can be represented as

$$\frac{dM_a}{dt} = \mu_a M_a - \lambda M_a$$

where $\mu_a$ is specific growth rate of dividing cells, which depends on the concentration of limiting substrate, which in turn is a function of time; $\lambda$ is inactivation rate constant, the frequency of "failures" leading to the loss of cells to divide

The accumulation of "incompetent" cells incapable to divide $M_i$ in the system is described by the equation

$$\frac{dM_i}{dt} = \lambda M_a$$

Total number of cells ($M$) in the system is defined by the sum

$$M = M_a + M_i$$

Accordingly

$$\frac{dM}{dt} = \mu_a M_a$$

The system of equations (1)-(4) describes basic patterns of population development. Unfortunately, analytical integration of this system of equations is not feasible. However, it is possible to obtain sufficiently detailed information using numerical integration.

For this purpose it is necessary to transform the system of equations (1)-(4) to dimensionless variables. As a result of this transformation we got

$$\frac{dm_a}{dt} = \frac{m_a s}{K_s S_0 + s} - \frac{\lambda}{\mu_m} m_a$$

$$\frac{dm_i}{dt} = \frac{\lambda}{\mu_m} m_a$$

$$s = \frac{S}{S_0}; m = \frac{M}{M_\infty - N_\infty}; p = \frac{p}{P_\infty}; \tau = \mu_m t$$

where $s$ is the substrate concentration in dimensionless quantities; $m$ is total number of cells; $m_a$ is replicating cells; $S$ is the concentration of substrate in physical (dimensional) quantities; $M$ is the concentration of biomass (microorganisms) in physical (dimensional) quantities; $S_0$ is the initial substrate concentration; $M_0$ is the initial concentration of biomass (microorganisms); $K_s$ is the constant of affinity of substrate to biomass (to microorganism); $M_\infty$ is the limiting quantity of biomass (microorganisms) and product formed at infinite great time of the process. $M_\infty$ is related to initial conditions: substrate and biomass (microorganisms) concentrations.
Numerical values of parameters given in [14] (Table 1) were used in simulation.

Table 1. Numerical values of parameters

| $\frac{K_s}{S_0}$ | $m_{a0}$ | $m_0$ | $Y_s$ | $\lambda$ |
|-------------------|---------|-------|-------|-----------|
| 2                 | 0.01    | 0.05  | 0.5   | 0.0357    |
|                   |         |       |       | 0.0714    |
|                   |         |       |       | 0.1071    |
|                   |         |       |       | 0.1428    |
|                   |         |       |       | 0.2142    |

Solution in the R programming language in Jupyter Notebooks environment. A fragment of listing is shown in Figure 1.

```r
f1 <- function (t, y, parms){
    with(as.list(y), {
        dma.dt <- ma * s / (ks*s) - mu*m
        dmi.dt <- mu * m
        ds.dt <- -ma * s / (ks*s)*2
        list(c(dma.dt, dmi.dt, ds.dt))
    })

t0 <- seq (0.01, 0.05, 0.01)
mum <- 0.0357
ks <- 2

y0 <- c(ma=0.01, m1=0.05, s=1)
out1 <- ode (y=y0, t=t0, func=f1, parms=NULL, method="ode45")
```

Figure 1. Software code listing

Integrated equations (5) and (6) with initial conditions gave the following results (Table 2).

Table 2. Fragment of the solution (calculation) result

| time | ma   | mi   | s    |
|------|------|------|------|
| 0.00 | 0.01 | 0.05 | 1.00 |
| 0.01 | 0.01 | 0.05 | 0.99 |
| 0.02 | 0.01 | 0.05 | 0.99 |
| 0.03 | 0.01 | 0.05 | 0.99 |
| 0.04 | 0.01 | 0.05 | 0.99 |
| 0.05 | 0.01 | 0.05 | 0.99 |

Along with tabulated values, the dependencies of microbial population growth with inactivation of cell replicating ability were obtained.
Comparison and analysis of graphical dependencies obtained (Figure 2) indicate that maximum concentration of dividing cells decreased as the rate constant of replicative capacity inactivation increased, while this value increased as the maximum rate increased. Consequently, the considered growth of *E. coli* is characterized by a very important feature. Besides cell division, cells lose their ability to divide quite intensively. This process is decisive in population development and limits the growth and ultimate density of the culture.

4. Conclusion

Use of modern information technology and modern programming languages, which undoubtedly include the R programming language, is an apparent support in studying the process of microbiological synthesis, establishing regularities and predicting results. There are systems of equations describing basic regularities of population development that cannot be solved analytically. But, using the capabilities of computers, software environments, and mathematical methods, it is possible to perform numerical integration with any accuracy. The solution of such system of differential equations is shown in the paper using Jupyter Notebooks software environment and the R programming language.

5. Acknowledgment

This article is published as part of scientific research theme No. 0585-2019-0008 under the state assignment of the federal state budgetary scientific institution ‘V.M. Gorbatov Federal Research Centre for Food Systems’ of RAS.

References

[1] Vasil’ev N N, Ambrosov V A and Skladnev A A 1975 Modeling the processes of microbiological synthesis (Moscow: Lesnaya promyshlennost') p 338
[2] Kafarov V V, Vinarov A Yu and Gordeev L S 1979 Modeling biochemical reactors (Moscow: Lesnaya promyshlennost') p 342
[3] Monod J 1949 Annual Review of Microbial. 111 371-394
[4] Henson M A and Seborg D E 1992 Chem. Eng. Sci. 47 821-835
[5] McLain R B, Kurtz M J and Henson M A 1996 Ind. Eng. Chem. Res. 35 4067-4077
[6] Ruan L and Chen X D 1996 Biotechnol. Prog. 12 286-288.
[7] Saha P, Partwardhan S C and Ramachandra R V S 1999 *Bioprocess Engineering*. 20 15-21
[8] Pechurkin Ya S 1978 *Population microbiology* Ed. I I Gitel’zon (Novosibirsk: Nauka) p 278
[9] Pinchuk R J, Brown W A, Hoghes S M and Cooper D G 2000 *Biotechnology and Bioengineering* 67(1) 19-24
[10] Berkholz R, Rohling D and Guthke R 2000 *Enzyme and Microb. Technol.* 27 784-788
[11] Cubarsi R, Corchero J L, Vila P and Villaverde A 1998 *J. Math. Appl. Med. Biol.* 15(3) 257-278
[12] Wickham H and Grolemund G 2016 *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data* (O'Reilly Media, Inc., Sebastopol) p 520
[13] Kabacoff R I 2011 *R in Action: Data analysis and graphics with R* (Manning, Shelter Island) p 474
[14] Varfolomeev S D and Kalyuzhnyj S V 1990 *Biotechnology: Kinetic Foundations of Microbiological Processes* (Moscow: Vysshaya shkola) p 296