UDC 519.6; 544.4 10.23947/2587-8999-2022-1-1-39-47

GENETIC ALGORITHM FOR SOLVING THE INVERSE PROBLEM OF CHEMICAL KINETICS *

G.F. Shaimardanova[1], **K.F. Koledina**[2]

[1] Ufa State Petroleum Technological University, Ufa, Russian Federation
[2] Institute of Petrochemistry and Catalysis of the Russian Academy of Sciences, Ufa, Russian Federation

✉ gulnazsajmardanova20@gmail.com, koledinakamila@mail.ru

Modeling allows to get a description of the object being modeled, in particular complications, which improves and clarifies its understanding and presents an organization of information that makes it easier to get the necessary information. Modeling of chemical kinetics problems is the solution of systems of ordinary nonlinear differential equations with determination of kinetic parameters. The inverse problem refers to incorrectly set tasks and does not have a single solution. The development of modern methods of evolutionary optimization, including the genetic algorithm, makes it possible to obtain a solution to such a high-dimensional problem in an acceptable time.

Materials and methods. The object of research is the catalytic reaction of dimethyl carbonate (DMC) with alcohols in the presence of hexacarbonyl tungsten. The solution of the direct problem is the solution of a system of ordinary nonlinear differential equations with initial data and given kinetic parameters. The solution of the inverse problem consists in determining the kinetic parameters corresponding to the minimum deviation of the calculated values of substance concentrations from the experimental data.

Results. The inverse problem of chemical kinetics is solved, which consists in calculating kinetic parameters with minimizing the functional deviation of the calculated values of component concentrations from experimental data. In the language of Python, program has been developed that implements the genetic algorithm. Numerical experiments with different numbers of iterations are performed and their impact on the accuracy of the solution, as well as the corresponding time costs, is estimated.

Discussion and conclusions. The calculated values of the component concentrations based on the kinetic parameters found correspond to the values of the experimental data.

Numerical experiments with different numbers of iterations are performed. Based on experiments, it was concluded that with an increase in the number of iterations, the accuracy of the solution increases. An increase in the number of iterations, in turn, leads to an increase in the time required for performing calculations. The time dependence on the number of iterations is analyzed.

Keywords: chemical kinetics, genetic algorithm, mathematical model, dimethyl carbonate, number of iterations.

Introduction. Modeling allows to get a description of the object being modeled, in particular complications, which improves and clarifies its understanding and presents an organization of information that makes it easier to get the necessary information. Problems of motion of a system of interacting material points, chemical kinetics, electrical circuits, and resistance of materials can be described by differential equations that can be solved numerically.

* The work was carried out within the framework of the state task of the Institute of Petrochemistry and Catalysis, Ural Federal Research Center, Russian Academy of Sciences (subject no. FMRS-2022-0078).

Modeling of chemical kinetics problems is the solution of systems of ordinary nonlinear differential equations with determination of kinetic parameters. Systems have a large dimension, depending on the number of intermediate reactants of the reaction, that is, on the detail of the scheme of chemical transformations. Accordingly, detailing also requires determining the kinetic parameters of the reaction stages – the solution of the inverse kinetic problem. The inverse problem refers to incorrectly set tasks and does not have a single solution. The dimension of the problem is determined by the number of unknown velocity constants of the stages. The development of modern methods of evolutionary optimization, including the genetic algorithm, makes it possible to obtain a solution to such a high-dimensional problem in an acceptable time.

2. Materials and methods. The object of research is the catalytic reaction of dimethyl carbonate (DMC) with alcohols in the presence of hexacarbonyl tungsten. The total reaction equation has the form:

$$ROH + (MeO)_2CO \xrightarrow{W(CO)_6} ROMe + ROCO_2Me + CO_2 + MeOH \quad (1)$$

The reaction products esters and simple esters are used in the preparation of medicinal products, and are also components of antifreeze, solvents and herbicides. For this process, it is necessary to calculate the rate constants of the stages and develop a kinetic model.

Systems of chemical kinetics equations describing reactions of particular interest usually have large dimensions, strong nonlinearities, and small singularly perturbing parameters. Their numerical study is also complicated by the fact that these systems, as a rule, hard drives, which forces us to develop special methods for approximate research.

The mathematical model of chemical kinetics problems is a system of ordinary nonlinear differential equations with initial data, i.e., the Cauchy problem:

$$\frac{dy_i}{d\tau} = \sum_{j=1}^{J} v_{ij} w_j, i = 1,...I$$

$$w_j = k_j^0 \cdot e^{-\frac{E_j^+}{RT}} \prod_{i=1}^{I} (y_i)^{|\alpha_{ij}|} - k_{-j}^0 e^{-\frac{E_j^-}{RT}} \prod_{i=1}^{I} (y_i)^{|\beta_{ij}|} \quad (2)$$

with initial conditions: when $\tau = 0, y_i(0) = y_i^0$, where y_i are concentrations of reaction reagents, mol/l; t is the time, min; J is the number of stages; I is the quantity of substances; v_{ij} is the stoichiometric matrix; w_j is speed j-th stage, 1/min; k_j, k_{-j}– are speed constants of the stages (given), 1 / min; α_{ij} are negative elements of the matrix v_{ij}, β_{ij} *are* positive elements of the matrix v_{ij}, k_j^0,k_{-j}^0 are pre-exponential multipliers, 1 / min; E_j^+, E_j^- are activation energies of forward and reverse reactions, kcal/mol; R is the gas constant, 2 cal/(mol · K); T is the temperature, K.

The solution of the direct problem is the solution of a system of ordinary nonlinear differential equations with initial data and given kinetic parameters up to time t*.

The solution of the inverse problem consists in determining the kinetic parameters corresponding to the minimum deviation of the calculated values of substance concentrations from the experimental data.

Table 1 shows the scheme of chemical transformations of the catalytic reaction of DMC with alcohols in the presence of $W(CO)_6$, obtained in [1-3].

Table 1. Scheme of chemical transformations of the catalytic reaction of DMC with alcohols in the presence of $W(CO)_6$

№	Stages
1	$W(CO)_6$ ***[Y15]***$\rightarrow W(CO)_5^+$***[Y18]***$+ CO$***[Y13]***
2	$W(CO)_5^+$***[Y18]***$+ MeOCO_2Me$ ***[Y2]*** $\rightarrow W(CO)_5CO_2Me$***[Y16]*** $+ MeO^-$***[Y9]***
3	$W(CO)_5CO_2Me$ ***[Y16]***$+ ROH$***[Y1]*** $\rightarrow ROMe$***[Y4]*** $+ CO_2$***[Y6]***$+ HW(CO)_5$***[Y17]***
4	$W(CO)_5CO_2Me$***[Y16]*** $+ ROH$***[Y1]*** $\rightarrow ROCO_2Me$ ***[Y5]***$+ HW(CO)_5$***[Y17]***
5	$HW(CO)_5$ ***[Y17]***$+ MeO^-$***[Y9]***$\rightarrow MeOH$***[Y7]*** $+ W(CO)_5^+$***[Y18]***

For the reaction under consideration, according to Table 1, the mathematical model according to (2) has the following form:

$$\begin{cases} \frac{dy_1}{dt} = -k_3 \cdot y_{16} \cdot y_1 - k_4 \cdot y_{16} \cdot y_1 \\ \frac{dy_2}{dt} = -k_2 \cdot y_{18} \cdot y_2 \\ \frac{dy_4}{dt} = k_3 \cdot y_{16} \cdot y_1 \\ \frac{dy_5}{dt} = k_4 \cdot y_{16} \cdot y_1 \\ \frac{dy_6}{dt} = k_3 \cdot y_{16} \cdot y_1 \\ \frac{dy_7}{dt} = k_5 \cdot y_{17} \cdot y_9 \\ \frac{dy_9}{dt} = k_2 \cdot y_{18} \cdot y_2 - k_5 \cdot y_{17} \cdot y_9 \\ \frac{dy_{13}}{dt} = k_1 \cdot y_{15} \\ \frac{dy_{15}}{dt} = -k_1 \cdot y_{15} \\ \frac{dy_{16}}{dt} \dot{y}_{16} = k_2 \cdot y_{18} \cdot y_2 - k_3 \cdot y_{16} \cdot y_1 - k_4 \cdot y_{16} \cdot y_1 \\ \frac{dy_{17}}{dt} = k_3 \cdot y_{16} \cdot y_1 + k_4 \cdot y_{16} \cdot y_1 - k_5 \cdot y_{17} \cdot y_9 \\ \frac{dy_{18}}{dt} = k_1 \cdot y_{15} - k_2 \cdot y_{18} \cdot y_2 + k_5 \cdot y_{17} \cdot y_9 \end{cases} \quad (3)$$

For $t=0$: $y_1(0)=y_1^0, y_2(0)=y_2^0$, $y_{15}(0)=y_{15}^0$, $y_i(0)=0$, $i=3,4,5,6,7,9,13,14$; where $y_1, y_2, y_4, y_5, y_6, y_7, y_9, y_{13}, y_{15}, y_{16}, y_{17}, y_{18}$ are reagent concentrations, mol/l; k_1, k_2, k_3, k_4, k_5 are rate constants of stages 1-5 (shown), 1 / min, respectively (Table 1).

To calculate the velocity constants of the stages, the inverse kinetic problem was set. Which consists in minimizing the functional deviation of the calculated values of the concentrations of components according to (3) from the experimental data, in the form (4).

$$E = \sum_i^I \sum_l^L \left| y_{il}^{calc} - y_{il}^{\exp} \right| \longrightarrow \min \quad (4)$$

where I is the quantity of substances, L is the number of time-based measurements.

When solving the inverse problem of chemical kinetics for determining kinetic parameters, corresponding to the minimum deviation of the calculated values of substance concentrations from the experimental data (4), the method of genetic algorithms was used.

Genetic algorithms are adaptive search methods that are quite often used today in solving a wide variety of problems. They are based on the genetic processes of biological organisms: biological populations evolve over several generations, following the laws of natural selection and following the principle of «survival of the fittest» [4].

The genetic algorithm consists of the following components:

1) The gene. In our case, the genome is understood as one of the coefficients of the desired solution.

2) A chromosome consisting of genes. It is one of the solutions to this problem.

3) The initial population (set) of chromosomes.

4) A set of operators for generating new solutions from the previous population. These include crossbreeding and mutation methods.

5) Objective function for evaluating the fitness of solutions.

Genetic algorithms work with a collection of «individuals» – a population, each of which represents a possible solution to a given problem. Each individual is evaluated by the measure of its «fitness» according to how «good» the corresponding solution of the problem is [4].

Selection rule. At each iteration, two individuals from the population are randomly selected. They cross according to the crossbreeding rule. The resulting individual is compared with the parent individuals. If a new individual is better than any parent, it is added to the population, and the worst parent is removed from the population. Gradually, the population improves, but there is a certain probability of «degeneration» of the population, when individuals become either the same or almost similar to each other and further improvement is impossible. In this case, a local minimum is reached [8]. In order to avoid this, the mutation method is used, when a random gene is randomly changed according to the mutation rule.

3. Research results. In the software system Python [9] developed a program with the implementation of a genetic algorithm, according to Figure 1. An individual is understood as a certain solution that is randomly generated in the initial population [7]. Each individual has five parameters, which are velocity constants in the system of differential equations. When solving the direct problem of chemical kinetics, these coefficients are known. When solving the inverse problem of chemical kinetics, these coefficients are unknown. In the initial population, the coefficients are randomly generated over specified intervals.

Table 2. The generated population before applying the genetic algorithm

Individual	k_1	k_2	k_3	k_4	k_5	*Deviation, E*
1	93.605	0.103	0.840	0.052	264.753	32.204
2	95.384	0.216	1.199	0.091	342.385	15.201
3	70.333	0.047	0.933	0.060	397.131	51.280
4	74.774	0.252	0.894	0.096	207.481	18.285
5	79.345	0.170	1.129	0.047	302.302	19.620

Table 3. Population after applying the genetic algorithm

Individual	k_1	k_2	k_3	k_4	k_5	*Deviation, E*
1	98.273	0.252	1.199	0.096	342.385	13.521
2	98.273	0.252	1.199	0.096	342.385	13.521
3	98.273	0.252	1.199	0.096	342.385	13.521
4	98.273	0.252	1.199	0.096	342.385	13.521
5	98.273	0.252	1.199	0.096	342.385	13.521

Tables 1 and 2 show the population values before applying the genetic algorithm and after applying the genetic algorithm.

Each individual can be qualitatively estimated by calculating the value of the residual functional according to (4). The lower the value, the «closer» the obtained solution is to the experimental data [5].

Tables 2 and 3 show that the genetic algorithm converges to a single value regardless of the initial values of the populations. After applying the genetic algorithm, the deviation decreases, i.e. it becomes closer to the experimental data. This can be clearly seen in the graphs (Figs. 1, 2 and 3).

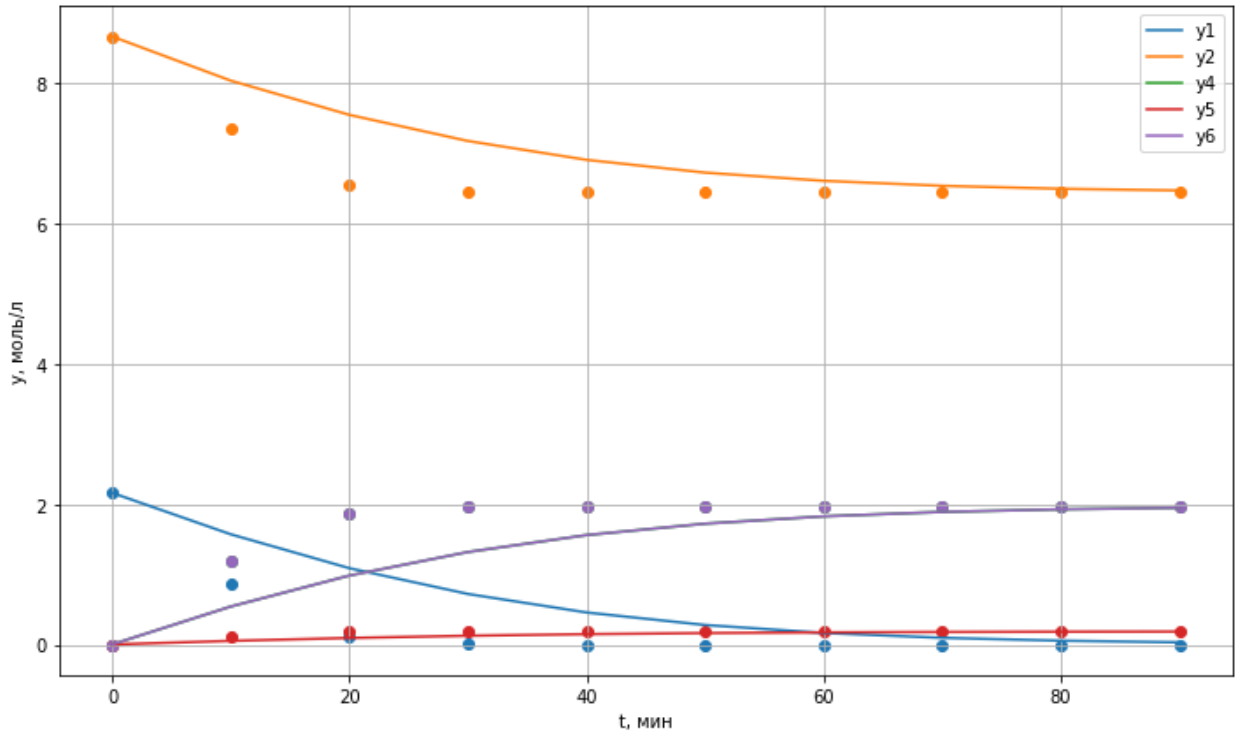

Fig. 1. Changes in component concentration's catalytic reaction of dimethyl carbonate with alcohols with a 10-iteration genetic algorithm solution.

Figures 1, 2 and 3 show graphs of functions of concentrations of substances that change over time. The dots represent graphs of functions based on experimental data. Curves represent graphs of functions constructed by solving the differential equation system of the best individual from the population for 10, 100, and 1000 iterations, respectively [6]. As the number of iterations increases, the graphs of functions represented by lines become closer to the graphs of functions represented by points. Thus, increasing iterations improves the description of experimental data.

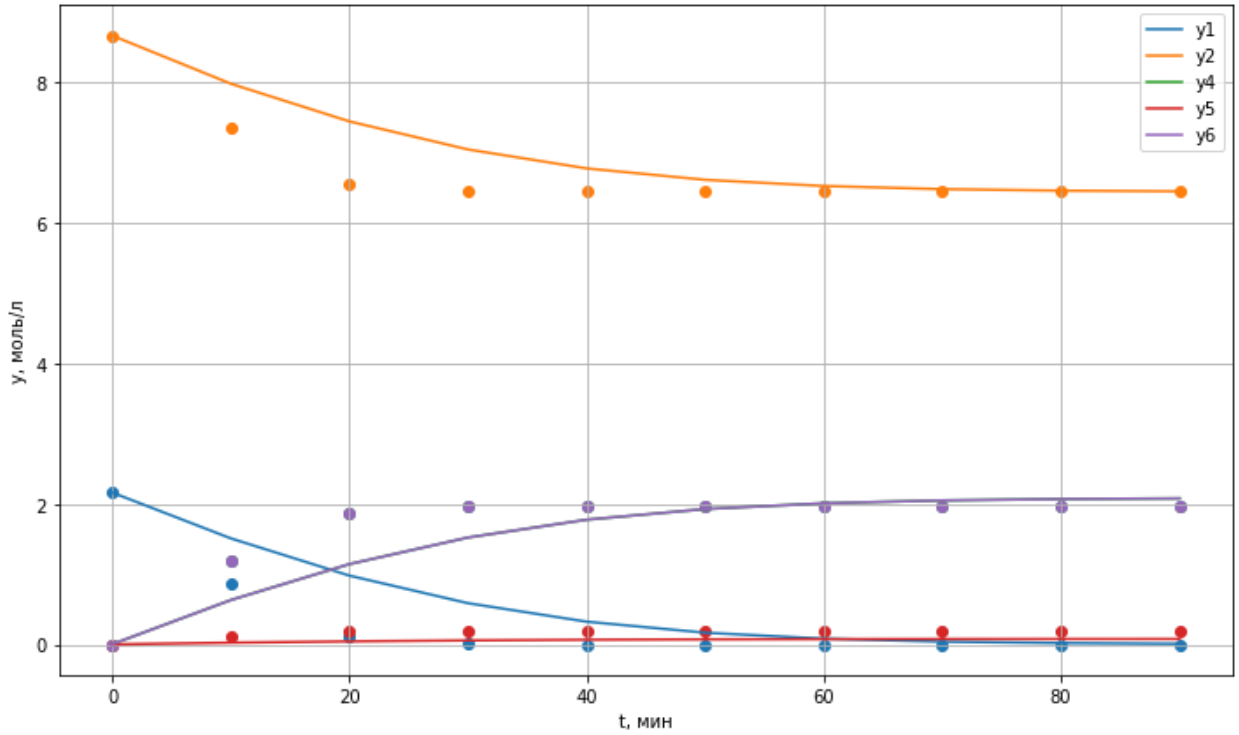

Fig. 2. Changes in component concentration's catalytic reaction of dimethyl carbonate with alcohols with a genetic algorithm solution with 100 iterations

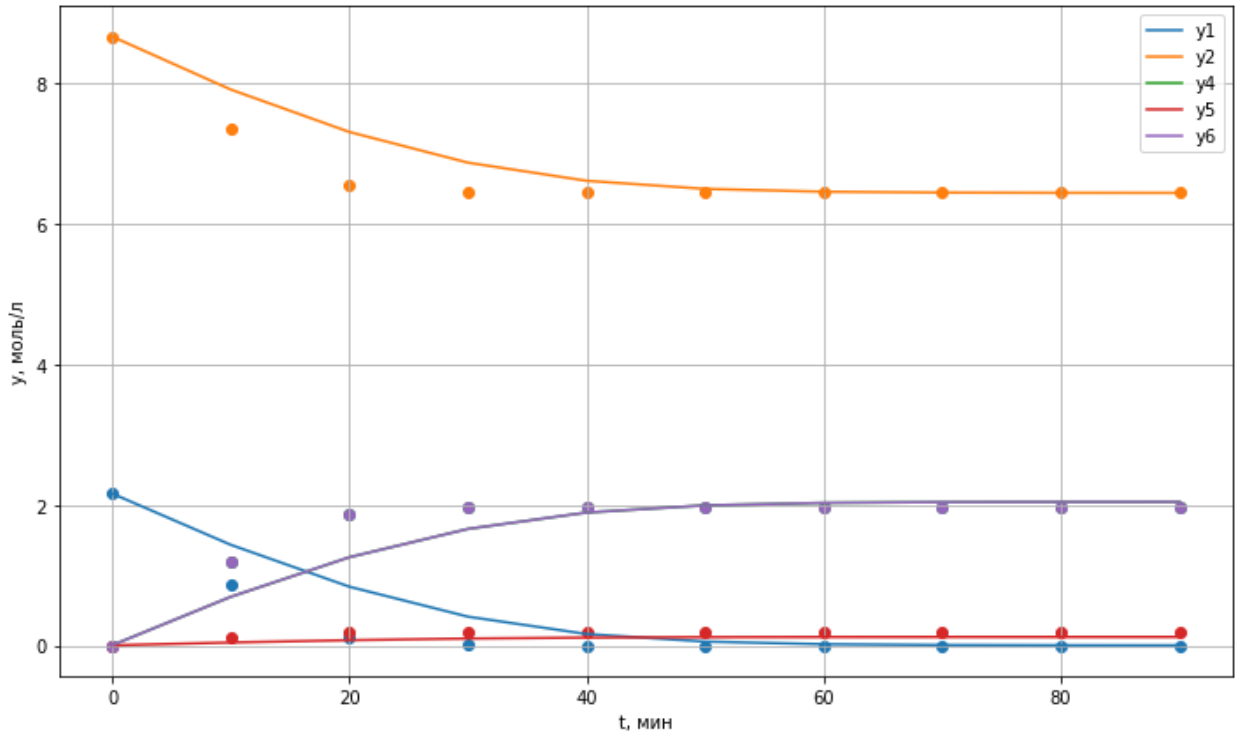

Fig. 3. Changes in component concentration's catalytic reaction of dimethyl carbonate with alcohols with a genetic algorithm solution with 1000 iterations

3. Evaluation of the program execution time. Increasing the number of iterations significantly increases the execution time of the program [10]. It has been suggested that increasing the number of iterations in N once the population size is maintained, it increases the time spent on calculations by approximately N once. This assumption was confirmed by numerical experiments. Fig. 4 shows the results of one of these experiments. There is an almost linear increase in time with an increase in the number of iterations.

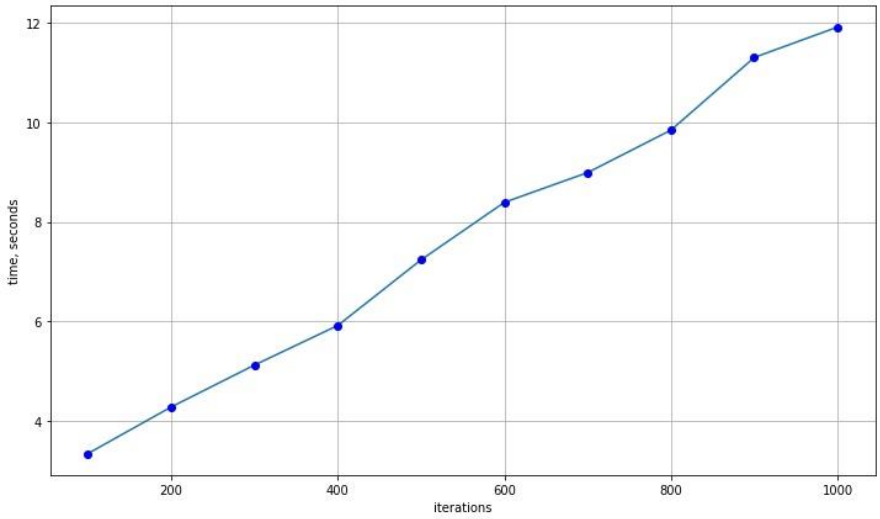

Fig. 4. Dependence of the program execution time on the number of iterations

Increasing the number of iterations while maintaining the population size with some error directly proportionally increases the time spent on calculations. This relationship allows you to estimate the time required to perform a sufficiently large number of iterations when searching for the best solution.

In the course of further numerical experiments, it was found that increasing the population size while maintaining the number of iterations also increases the time that must be spent on calculations. But in this case, the relationship is not linear. Fig. 5 shows the calculation execution time increases, but the dependence is not as obvious as in fig. 4.

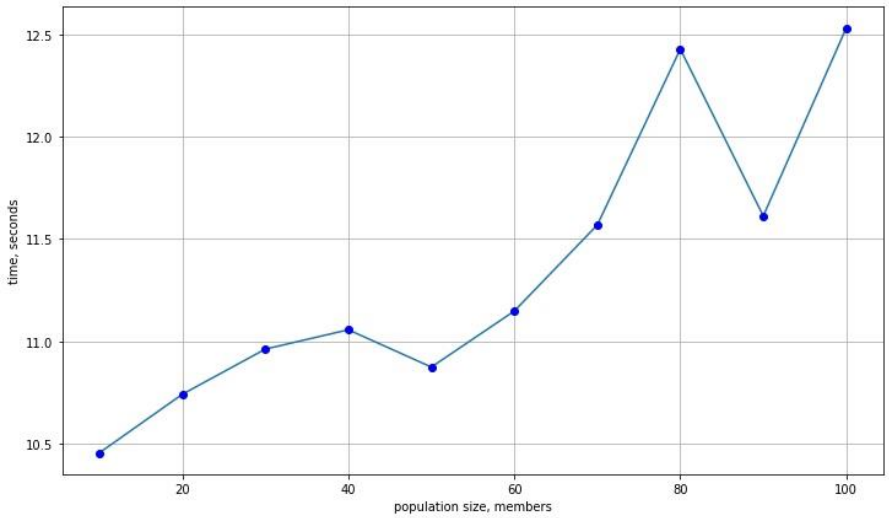

Fig. 5. Dependence of program execution time on population size

Conclusion. In this paper, we consider the inverse problem of chemical kinetics, which consists in calculating kinetic parameters with minimizing the functional deviation of the calculated values of component concentrations from experimental data. These calculated values were obtained using a genetic algorithm.

To find the optimal solution for the considered inverse problem of chemical kinetics, a genetic algorithm was used, the main stages of which are presented in the form of a block diagram. In the language of Python, program was developed that implements the considered genetic algorithm.

Numerical experiments with different numbers of iterations are performed. Based on experiments, it was concluded that with an increase in the number of iterations, the accuracy of the solution increases.

An increase in the number of iterations, in turn, leads to an increase in the time required for performing calculations. The time dependence on the number of iterations is analyzed.

References

1. Koledina, K.F. Kinetics and mechanism of catalytic reaction of alcohols with dimethyl carbonate / K.F. Koledina, S.N. Koledin, N.A. Shchadneva, I.M. Gubaidullin // Journal of Physical Chemistry. 2017. Vol. 91. No. 3. pp. 422-428.

2. Koledina, K.F. Automated system for identifying conditions for homogeneous and heterogeneous reactions in multi-purpose optimization problems / K.F. Koledina, S.N. Koledin, I.M. Gubaidullin // Siberian Journal of Computational Mathematics. 2019. Vol. 22. No. 2. pp. 137-151.

3. Koledin, S.N. Information system for assessing the relationship of objective functions and the study of optimal conditions for complex catalytic reaction by multi-purpose optimization methods / S.N. Koledin, K.F. Koledin, A.P. Karpenko, I.M. Gubaidullin // Electrical and information complexes and systems. 2017. Vol. 13. No. 4. pp. 71-80.

4. Genetic algorithms / L. A. Gladkov, V. V. Kureychik, V. M. Kureychik // M.: Fizmatlit. 2006. 87 p.

5. Genetic algorithms for solving extreme problems / D.I. Batishchev // Nizhny Novgorod State University. Nizhniy Novgorod. 1995. 62 p.

6. Automatic programming / N.I.Polikarpova, A. A. Shalyto // M.: Peter. 2010. 167 p.

7. Genetic algorithms, artificial neural networks and problems of virtual reality / G. K. Voronovsky, K. V. Makhotilo, S. N. Petrashev, S. A. Sergeev // Kharkiv, OSNOVA. 1997. 112 p.

8. Tikhonova, M.V. Parallel conditional global optimization in mathematical modeling of the kinetics of chemical reactions / M.V. Tikhonova, V.V. Ryabov, S.I. Spivak, I.M. Gubaidullin // Computational methods and programming: new computing technologies. 2013. Vol. 14. No. 1. pp. 262-268.

9. Koledin, S.N. Optimal control and optimum sensitivity in problems of chemical kinetics / S.N. Koledin, K.F. Koledin // Journal of the Middle Volga Mathematical Society. 2016. Vol. 18. No. 3. pp. 137-144.

10. Gubaidullin, I.M. Sequential-parallel determination of kinetic parameters / Gubaidullin I.M., Koledina K.F., Spivak S.I. // Journal of the Middle Volga Mathematical Society. 2009. Vol. 11. No. 2. pp. 14-25.

Authors:

Gulnaz F. Shaymardanova, student, Ufa State Petroleum Technological University (Kosmonavtov St., Ufa, Russian Federation), nastena1425@gmail.com

Kamila F. Koledina, Doct.Sci. (Phys.-Math.), associate professor, researcher, associate professor of the Mathematics Department, Institute of Petrochemistry and Catalysis of RAS, Ufa State Petroleum Technological University (Kosmonavtov St., Ufa, Russian Federation), koledinakamila @mail.ru

УДК 519.6; 544.4 10.23947/2587-8999-2022-1-1-39-47

ГЕНЕТИЧЕСКИЙ АЛГОРИТМ ПРИ РЕШЕНИИ ОБРАТНОЙ ЗАДАЧИ ХИМИЧЕСКОЙ КИНЕТИКИ*

Г.Ф. Шаймарданова[1], К.Ф. Коледина[2]

[1]Уфимский государственный нефтяной технический университет, г. Уфа, Российская Федерация
[2]Институт нефтехимии и катализа УФИЦ РАН, г. Уфа, Российская Федерация

✉ gulnazsajmardanova20@gmail.com, koledinakamila@mail.ru

Моделирование позволяет получить описание моделируемого объекта, в частности осложнения, улучшающее и уточняющее его понимание и представляющее такую организацию информации, которая обеспечивает более легкое получение необходимой информации. Моделирование задач химической кинетики представляет собой решение систем обыкновенных нелинейных дифференциальных уравнений с определением кинетических параметров. Обратная задача относится к некорректно поставленным задачам и не имеет единственного решения. Развитие современных методов эволюционной оптимизации, в том числе генетического алгоритма, позволяет получить решение такой высокоразмерной задачи за приемлемое время.

Материалы и методы. Объектом исследования является каталитическая реакция диметилкарбоната (ДМК) со спиртами в присутствии гексакарбонил вольфрама. Решение прямой задачи представляет собой решение системы обыкновенных нелинейных дифференциальных уравнений с начальными данными и заданными кинетическими параметрами. Решение обратной задачи заключается в определении кинетических параметров, соответствующих минимальному отклонению расчётных значений концентраций веществ от экспериментальных данных.

Результаты исследования. Решена обратная задача химической кинетики, которая заключается в расчете кинетических параметров с минимизацией функционала отклонения расчетных значений концентраций компонент от экспериментальных данных. На языке Python разработана программа, реализующая генетический алгоритм. Проведены численные эксперименты с разным количеством итераций и оценено их влияние на точность решения, а также соответствующие временные затраты.

Обсуждение и заключения. Рассчитанные значения концентраций компонент по найденным кинетическим параметрам соответствуют значениям экспериментальных данных. Проведены численные эксперименты с разным количеством итераций. На основе экспериментов был сделан вывод, что с увеличением количества итераций увеличивается точность решения. Увеличение количества итераций в свою очередь приводит к увеличению времени, которое необходимо для проведения вычислений. Проведен анализ зависимости времени от количества итераций.

Ключевые слова: химическая кинетика, генетический алгоритм, математическая модель, диметилкарбонат, число итераций.

Авторы:

Гульназ Шаймарданова, студентка Уфимского государственного нефтяного технологического университета (ул. Космонавтов, 1, Уфа, Российская Федерация)

Камила Коледина, доктор физико-математических наук, доцент кафедры математики Уфимского государственного нефтяного технологического университета (ул. Космонавтов, 1, Уфа, Российская Федерация)

* Работа выполнена в рамках государственного задания Института нефтехимии и катализа УФИЦ РАН (тема № FMRS-2022-0078).