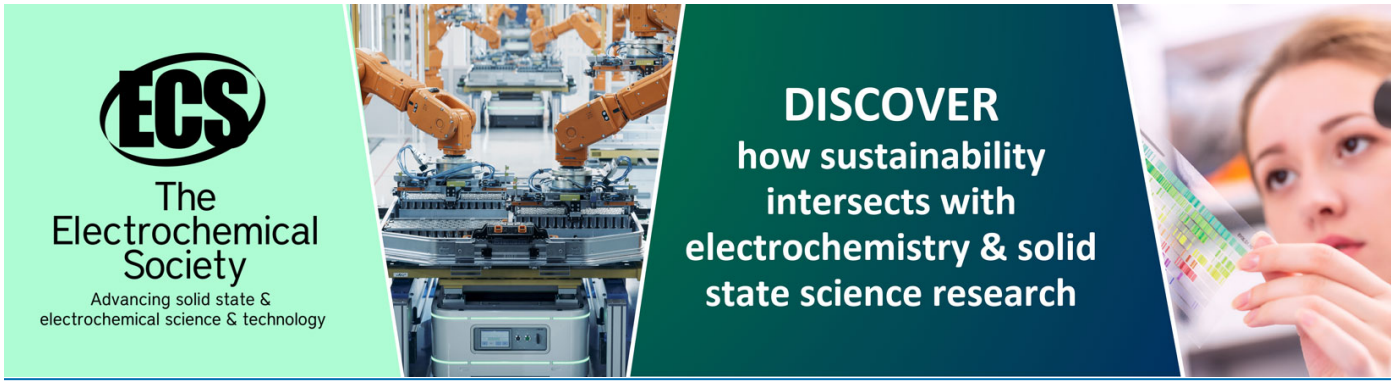
Search for the optimal temperature regime for the synthesis of benzylidenebenzylamine based on genetic algorithms



IOP Conf. Series: Journal of Physics: Conf. Series **1479** (2020) 012101 doi:10.1088/1742-6596/1479/1/012101

Search for the optimal temperature regime for the synthesis of benzylidenebenzylamine based on genetic algorithms

E V Antipina[1], S A Mustafina[1], A F Antipin[1] and E S Podval'nyi[2]

[1]Bashkir State University, Russia, Republic of Bashkortostan, 450076, Ufa, Zaki Validi st., 32
[2]The Russian Presidential Academy of National Economy and Public Administration, Russia, 119571, Moscow, Vernadsky Avenue, 82

E-mail: stepashinaev@ya.ru

Abstract. The article is devoted to the creation of an algorithm for finding the optimal temperature regime of a chemical reaction based on evolutionary calculations. The problem of optimal control for a chemical process is formulated in a general way. A genetic algorithm for solving the problem is described. A computational experiment was carried out for the synthesis reaction of benzylidenebenzylamine in order to obtain the maximum yield of the product during the reaction. The optimum temperature profile and the optimal concentration of reagents were obtained. Based on the formulated genetic algorithm, a program was developed in the visual programming environment Delphi, which allows finding the optimal temperature profile of the chemical process.

1. Introduction

The task of determining the optimal process conditions and the optimal apparatus designs is one of the most important tasks of mathematical modeling for chemical-technological processes. The application of mathematical modeling methods makes it possible to increase the productivity for the technological scheme of the process and obtain specific quantitative results. Natural and laboratory experiments are simulated using computer programs [1-5].

The search for optimal conditions for conducting the process is based on the kinetic model of the reaction. Therefore, this task has the following features [6]:

1) Nonlinearity. The right side of the differential equations (kinetic models) is an exponential function of control, as well as a nonlinear function of phase variables.

2) The presence of restrictions in the control and/or phase variables.

3) The large dimension of the problem being solved, associated with the complexity of the processes.

In connection with these features, there is a limitation for the applicability of some optimization methods. Nonlinearity does not allow the use of linear programming methods [7]. The second feature limits the use of variational calculus methods [8]. The large dimensionality of tasks makes it difficult to use dynamic programming because of the large computational costs [9]. To apply the maximum principle, an additional check of the found solution for optimality is necessary [10-12].

The use of genetic algorithms allows us to overcome these difficulties in the process of finding optimal control during the chemical-technological process. Genetic algorithms mimic the processes

that occur during evolution (heredity, variability, natural selection). As a result, the fittest individuals survive, and they give rise to a new generation of offspring [13, 14].

Genetic algorithms make it possible to find a solution of an optimization problem in an acceptable time. This solution does not depend on the initial approximation.

Genetic algorithms can be easily modified with an increase in the vector of phase variables. Therefore, genetic algorithms can be used for various processes whose kinetic models contain different amounts of reacting substances.

2. Statement of the problem

We formulate in general terms the problem of finding optimal control and its main parts:

1) A mathematical model of the process, which includes control parameters, areas of feasible solutions, initial values of the model variables.

2) Limitations of control parameters and/or phase variables.

3) The optimization criterion, which is a quantitative characteristic for the solution of the optimization problem.

Let the mathematical model of the controlled process be described by the system of ordinary differential equations [15]:

$$\frac{dx_i}{dt} = f_i(t, \boldsymbol{x}, \boldsymbol{u}), \tag{1}$$

where $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T$ is the vector of phase variables, $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)^T \in U$ is the control vector, U is the set of admissible control values, , t is time, $f_i(t, \boldsymbol{x}, \boldsymbol{u})$ are functions that are continuous together with their partial derivatives $(i = \overline{1,n})$.

Let the initial values of the phase variables be given:

$$x_i(0) = x_{i0}, \, i = \overline{1,n}. \tag{2}$$

The set of permissible processes $Q(t_0, \boldsymbol{x}_0)$ is the set of pairs $\boldsymbol{q} = (\boldsymbol{x}(t), \boldsymbol{u}(t))$, that satisfies the system (1) and the initial conditions (2).

Let the optimality criterion be a functional on the set $Q(t_0, \boldsymbol{x}_0)$:

$$G(\boldsymbol{q}) = F(\boldsymbol{x}(t_{end})), \tag{3}$$

where t_{end} is the final time the operation of the system.

It is required to find such a pair $\boldsymbol{q}^* = (\boldsymbol{x}^*(t), \boldsymbol{u}^*(t)) \in Q(t_0, \boldsymbol{x}_0)$, that $G(\boldsymbol{q}^*) = \max\limits_{q \in Q(t_0, \boldsymbol{x}_0)} G(\boldsymbol{q})$.

As the phase variables x_i $(i = \overline{1,n})$, we will consider the values of the concentrations of interacting substances. Then the system of differential equations (1) with initial conditions (2) is a kinetic model of a chemical reaction.

Temperature is one of the parameters that most strongly affects the dynamics of chemical processes. Therefore, one of the tasks of optimizing chemical processes is determining the optimal temperature regime $T = T(t)$.

The optimal temperature regime provides an extremum of the optimality criterion (3). In this case, restrictions can be imposed on the control parameter in the form of:

$$T_* \le T(t) \le T^*, \tag{4}$$

where T_* and T^* are the lower and upper limits of the permissible temperature range.

3. Genetic algorithm for finding optimal control of a chemical process

We formulate an algorithm for finding the optimal temperature regime of a chemical process based on a genetic algorithm. Let the reaction temperature $T(t)$ be a control parameter in the genetic algorithm.

The work of the genetic algorithm consists in sequent changing of the set of population vectors. Each vector is an analogue of an individual (chromosome) in a population. An individual consists of a set of numerical values called genes (traits) [16]. According to a certain rule, two individuals are selected from the population by the selection operator. Then these two individuals are crossed among themselves by the crossing operator. The resulting vector-individual is exposed to the mutation operator. As a result of mutations, one or several genes (vector coordinate values) change. The fittest individuals from the current population are moving into a new generation. The fitness of an individual is determined by calculating the optimality criterion. This procedure of changing generations of a population continues until a criterion for terminating the search is reached.

Consider the temperature $T(t)$ as a population. We represent it by a set of m vectors $\boldsymbol{T}_i = (T_{i1}, T_{i2}, \ldots, T_{in-1})$ $(i = \overline{1,m})$. To calculate the fitness function (3), it is necessary to calculate the concentration vector $\boldsymbol{x}_i(t_{end}) = (x_{i1}(t_{end}), x_{i2}(t_{end}), \ldots, x_{in}(t_{end}))$, corresponding to the control vector T_i.

We formulate this algorithm.

Step 1. Generation of an initial population for temperature vectors.

Set the population size m, the number of the current population $P = 0$, the maximum number of populations $P_{\max}$, the number of split points n of the time interval $[0, t_{end}]$. Generate the initial population $\boldsymbol{T}^0 = (T_1^0, T_2^0, \ldots, T_m^0)$, $\boldsymbol{T}_i^0 = (T_{i1}^0, T_{i2}^0, \ldots, T_i^0), i = \overline{1,m}$ randomly on the interval $[T_*, T^*]$.

For each vector $\boldsymbol{T}_i^0$ solve the direct kinetic problem, that is, the system of differential equations (1) with initial conditions (2). Based on the obtained concentration values for each individual $\boldsymbol{T}_i^P$ calculate the value of the fitness function (3).

Step 2. Selection.

The selection operator selects individuals for subsequent crossbreeding. The “roulette” selection operator allows you to select individuals with the best fitness function value. The maximum value of the optimality criterion is the best value of the fitness function for the maximum problem. For each individual, the probability of selection is calculated by the formula

$$p(\boldsymbol{T}_i^P) = \frac{G(\boldsymbol{x}, \boldsymbol{T}_i^P)}{\sum_{j=1}^{m} G(\boldsymbol{x}, \boldsymbol{T}_j^P)}.$$

Two individuals are selected with the highest probabilities $\boldsymbol{a} = \boldsymbol{T}_l^P$ and $\boldsymbol{b} = \boldsymbol{T}_r^P$ where l and r are the numbers of individuals with the highest probabilities of selection.

The “panmixia” selection operator performs random equiprobable selection, while the probability of an individual participating in the selection is determined by the formula

$$p(\boldsymbol{T}_i^P) = \frac{1}{m}.$$

The “tournament selection” operator performs two tournaments. The first tournament includes a random selection of two individuals from a population. During the second tournament, an individual is randomly selected from the two individuals which have been selected in the first tournament.

Step 3. Crossing (crossover).

The cross operator generates one, two or three new descendant individuals from the parent individuals $\boldsymbol{a} = (a_1, a_2, \ldots, a_{n-1})$, $\boldsymbol{b} = (b_1, b_2, \ldots, b_{n-1})$, selected in the previous step, by exchanging some of the genes.

A discrete crossover creates one descendant $\boldsymbol{c} = (c_1, c_2, \ldots, c_{n-1})$. A parent is randomly selected. The corresponding value of the selected parent gene $(i = \overline{1, n-1})$. is assigned to the value c_i.

Arithmetic crossover generates two descendants $\boldsymbol{c} = (c_1, c_2, \ldots, c_{n-1})$ and $\boldsymbol{d} = (d_1, d_2, \ldots, d_{n-1})$ by the rule: $c_i = \lambda a_i + (1-\lambda) b_i$, $d_i = \lambda b_i + (1-\lambda) a_i$, where λ is a random number from the range (0,1), $i = \overline{1, n-1}$.

The simplest crossover creates two descendants $\boldsymbol{c} = (c_1, c_2, \ldots, c_{n-1})$ and $\boldsymbol{d} = (d_1, d_2, \ldots, d_{n-1})$ by the rule: $\boldsymbol{c} = (a_1, \ldots, a_s, b_{s+1}, \ldots, b_{n-1})$, $\boldsymbol{d} = (b_1, \ldots, b_s, a_{s+1}, \ldots, a_{n-1})$, where s is a random number from the range $[1, n-2]$.

A linear crossover generates three descendants $\boldsymbol{c} = (c_1, c_2, \ldots, c_{n-1})$, $\boldsymbol{d} = (d_1, d_2, \ldots, d_{n-1})$, $\boldsymbol{e} = (e_1, e_2, \ldots, e_{n-1})$ according to the rule:

$$c_i = \frac{a_i + b_i}{2}, \quad d_i = \frac{3a_i - b_i}{2}, \quad e_i = \frac{-a_i + 3b_i}{2}, \quad i = \overline{1, n-1}.$$

Step 4. Mutation.

The mutation operator is designed to overcome the entry of a population into a local extremum. At this step, the genes of the descendants are transformed. For this, the descendant gene is randomly selected and replaced with a random value from the range of acceptable values $[T_*, T^*]$. The fitness function value for the altered offspring vectors is calculated by solving the direct kinetic problem.

Step 5. Updating the population.

The individual $\boldsymbol{T}_i^P$ with the worst value of the fitness function is selected from the current population of temperature vectors $\boldsymbol{T}^P = (T_1^P, T_2^P, \ldots, T_m^P)$. This vector T_i^P is replaced by one randomly selected descendant *c* or *d*.

Step 6. Verification of the end of the search for a solution.

If $P \le P_{\max}$, you go to step 2, otherwise stop the search. Choose the individual with the best fitness function from the last population $\boldsymbol{T}^{P_{\max}} = (T_1^{P_{\max}}, T_2^{P_{\max}}, \ldots, T_m^{P_{\max}})$. The selected vector $\boldsymbol{T}_i^{P_{\max}}$ represents the optimal temperature profile $T(t)$ of the chemical process.

This algorithm is implemented in the Delphi visual programming environment.

4. Computing experiment

We will find the optimal temperature profile for the synthesis reaction of benzylidenebenzylamine based on the genetic algorithm. Benzylidenebenzylamine is the starting material for the synthesis of a number of heterocycles.

The synthesis of benzylidenebenzylamine under the action of $FeCl_3 \cdot 6H_2O$ is shown in Figure 1.

The mechanism of the chemical reaction for the synthesis of benzylidenebenzylamine in the presence of a catalyst $FeCl_3 \cdot 6H_2O$ is represented by a set of stages [17]:

$$\begin{aligned} &X_1 + X_2 \rightarrow X_3 + X_4, \\ &X_3 \rightarrow X_5 + X_6, \\ &X_5 + X_1 \rightarrow X_7 + X_8, \\ &X_8 + X_6 \rightarrow X_9, \end{aligned} \tag{5}$$

where X_1 – benzylamine (C_7H_9N), X_2 – carbon tetrachloride (CCl_4), X_3 – chlorobenzylamine (C_7H_8NCl), X_4 – chloroform ($CHCl_3$), X_5 – 1-phenylmethanimine (C_7H_7N), X_6 – hydrogen chloride (HCl), X_7 – benzylidenebenzylamine ($C_{14}H_{13}N$), X_8 – ammonium (NH_3), X_9 – ammonium chloride (NH_4Cl).

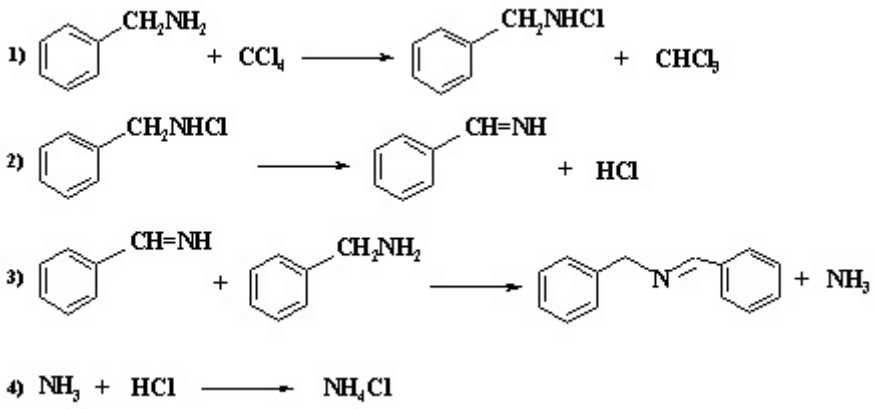


Figure 1. Reaction for the synthesis of benzylidenebenzylamine.

The speed of each stage is determined according to the law of mass action:

$$\begin{aligned} \omega_1 &= k_1 x_1 x_2, \\ \omega_2 &= k_2 x_3, \\ \omega_3 &= k_3 x_1 x_5, \\ \omega_4 &= k_4 x_6 x_8, \end{aligned} \tag{6}$$

where x_i – concentration of the i-th substance $(i = \overline{1,9})$ (mole fraction), k_j – rate constant of the j-th reaction stage $(j = \overline{1,4})$, calculated according to the Arrhenius equation

$$k_j(T) = k_{0j} \exp\left(-\frac{E_j}{RT}\right),$$

where E_j – energy of activation of the j-th stage (kJ/mol), T – reaction temperature (K), R – universal gas constant (8,31 J/(mol·K)).

The matrix of stoichiometric coefficients of substances (γ_{ij}) $(i = \overline{1,4},\ j = \overline{1,9})$ is given in Table 1.

Table 1. Matrix of stoichiometric coefficients of the reaction of synthesis of benzylidenebenzylamine.

	ω_1	ω_2	ω_3	ω_4
X_1	-1	0	-1	0
X_2	-1	0	0	0
X_3	1	-1	0	0
X_4	1	0	0	0
X_5	0	1	-1	0
X_6	0	1	0	-1
X_7	0	0	1	0
X_8	0	0	1	-1
X_9	0	0	0	1
δ	0	1	0	-1

Since there are nonzero elements among the elements of the last line δ, the reaction proceeds with a change in the reaction volume [18]. Then the kinetic model of the reaction is a system of differential equations:

$$\begin{cases} \dfrac{dx_i}{dt} = \dfrac{F_i(\boldsymbol{x},T) - x_i \cdot F_n(\boldsymbol{x},T)}{N}, \ i = \overline{1,9}, \\ \dfrac{dN}{dt} = F_n(\boldsymbol{x},T), \end{cases} \tag{7}$$

with initial conditions:

$$x_i(0) = x_i^0, \ i = \overline{1,9}, \ \ N(0) = 1, \tag{8}$$

where N – variable reaction volume, $F_i = \sum_{j=1}^{4} \gamma_{ij} W_j, \ i = \overline{1,9}, \ \ F_n = \sum_{j=1}^{4} W_j \sum_{i=1}^{9} \gamma_{ij}, \ \ W_j = \dfrac{\omega_j}{C_0}$ – values of rate of reaction stages ($j = \overline{1,4}$) (h^{-1}), C_0 – initial total concentration of substances (mol/l).

The functions $F_n(\boldsymbol{x},T)$, $F_i(\boldsymbol{x},T)$, ($i = \overline{1,9}$) have the form:

$$\begin{aligned} F_1(\boldsymbol{x},T) &= -W_1(\boldsymbol{x},T) - W_3(\boldsymbol{x},T), \\ F_2(\boldsymbol{x},T) &= -W_1(\boldsymbol{x},T), \\ F_3(\boldsymbol{x},T) &= W_1(\boldsymbol{x},T) - W_2(\boldsymbol{x},T), \\ F_4(\boldsymbol{x},T) &= W_1(\boldsymbol{x},T), \\ F_5(\boldsymbol{x},T) &= W_2(\boldsymbol{x},T) - W_3(\boldsymbol{x},T), \\ F_6(\boldsymbol{x},T) &= W_2(\boldsymbol{x},T) - W_4(\boldsymbol{x},T), \\ F_7(\boldsymbol{x},T) &= W_3(\boldsymbol{x},T), \\ F_8(\boldsymbol{x},T) &= W_3(\boldsymbol{x},T) - W_4(\boldsymbol{x},T), \\ F_9(\boldsymbol{x},T) &= W_4(\boldsymbol{x},T), \\ F_n(\boldsymbol{x},T) &= W_2(\boldsymbol{x},T) - W_4(\boldsymbol{x},T). \end{aligned} \tag{9}$$

Let the temperature in the reactor be a control parameter. It is necessary to determine the optimal temperature regime $T = T(t)$ of the chemical process (5), which is described by a system of differential equations (7) with initial conditions (8). So the maximum yield of the reaction target product (benzylidenebenzylamine) is achieved.

$$G(\boldsymbol{x},T) = x_7(t_{end}) \to \max. \tag{10}$$

Allowable temperature values are given by the inequality:

$$285\,\text{K} \le T(t) \le 373\,\text{K}. \tag{11}$$

The problem (5)-(11) has been solved using the developed genetic algorithm with the following parameters: the population size is 60, the maximum number of populations is 3000, the number of genes in the chromosome is 100, a selection operator is a tournament selection, a crossover operator is an arithmetic crossover. The reaction time is 8 hours. The direct kinetic problem has been solved by the forecast and correction method. The initial conditions (8) are given by the following values [17]:

$$x_1(0) = 0.513,$$
$$x_2(0) = 0.487,$$
$$x_i(0) = 0,\ i = \overline{3,9},$$
$$N(0) = 1.$$

5. Results and discussion

As a result of a computational experiment, the optimal concentrations of substances (Figure 2) and the optimal temperature regime (Figure 3) were determined for the synthesis reaction of benzylidenebenzylamine.

The calculation of the optimal reaction temperature (5) showed that it is necessary to keep the temperature at the maximum acceptable level of 373 K for 3 hours. Then, the process should be carried out at a temperature of 330 K, gradually increasing it during 5 hours to 356 K. The maximum yield of the target product X_7 will be 0.328 molar fractions.

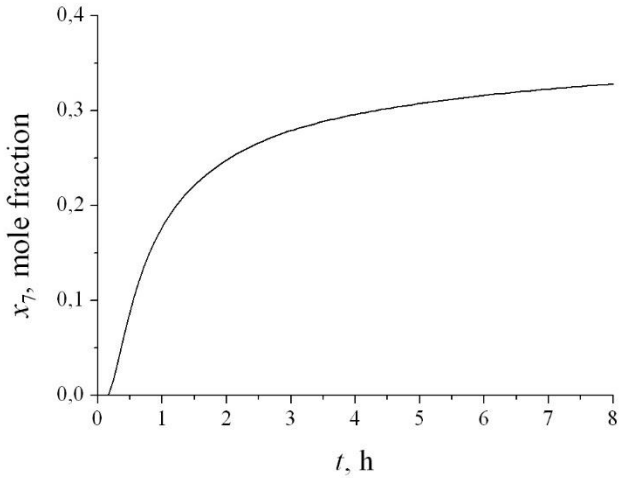


Figure 2. The optimal concentration of benzylidenebenzylamine (X_7).

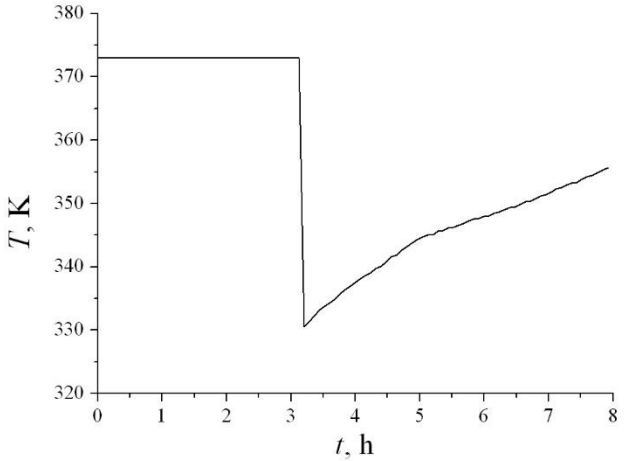


Figure 3. Optimal temperature regime.

Table 2 shows the results of calculating the concentration of the target substance X_7 at some constant permissible temperature values at a finite point in time.

Figure 4 shows the dynamics of the concentration of benzylidenebenzylamine at some constant temperature values.

Table 2 and Figure 4 show that the optimal concentration of the target substance X_7 (0.328 mole fractions) is higher than the calculated values at arbitrary constant temperature values. Therefore, we can conclude that the developed genetic algorithm works correctly.

Table 2. The value of the concentration of benzylidenebenzylamine.

№	*T*, K	The concentration of benzylidenebenzylamine $x_7(t_{end})$, mole fraction
1	285	0.013
2	295	0.024
3	305	0.040
4	315	0.061
5	325	0.084
6	335	0.108
7	345	0.130
8	355	0.149
9	365	0.167
10	373	0.178

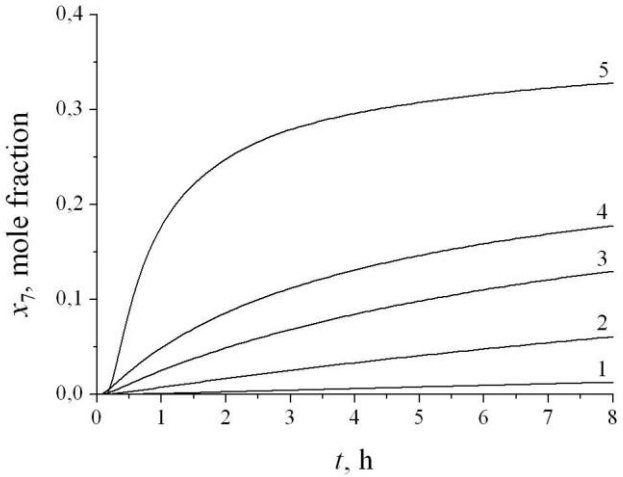


Figure 4. The dynamics of the concentration of benzylidenebenzylamine at various temperature regimes (1 – *T*=285 K, 2 – *T*=315 K, 3 – *T*=345 K, 4 – *T*=373 K, 5 – calculated optimal temperature regime).

6. Conclusion

Thus, the developed genetic algorithm for finding the optimal temperature regime can be used to solve the problems of optimization for chemical processes. Moreover, the solution found for the optimal problem does not depend on the choice of the starting point for finding a solution. The algorithm can be applied to the problems of finding optimal control for other processes, which are described by nonlinear systems of differential equations. There are restrictions on the control parameter. Based on

this algorithm, a program has been developed in the Delphi environment. The program allows you to calculate the optimal temperature regime that delivers the maximum to the criterion of optimality.

A computational experiment was carried out for the synthesis reaction of benzylidenebenzylamine. The optimum temperature regime was determined, which provides the maximum yield of the reaction product (benzylidenebenzylamine) at a finite time point. The optimal concentration of reagents was calculated. It was shown that the concentration of benzylidenebenzylamine which was calculated at the optimum temperature regime, exceeds the concentration which was calculated at other temperatures.

The developed genetic algorithm is universal. It can be easily modified when the chemical reaction mechanism is changed. For this, it is necessary to change only the block for solving the direct kinetic problem in the program. It is also easy to modify the optimization criterion. Therefore, the genetic algorithm can be also used to solve other problems for optimization of chemical-technological processes.

Acknowledgments
The reported study was funded by RFBR according to the research project № 19-37-51001.

References

[1] Ostrovsky G M, Ziyatdinov N N and Lapteva T V 2013 Optimization problem with normally distributed uncertain parametrses *AIChE Journal* **59** (7) 2471–2484
[2] Grossmann I, Aguirre P and Barttfeld M 2005 Optimal synthesis of complex distillation columns using rigorousmodels *Comput. Chem. Eng.* **29** 1203
[3] Barttfeld M, Aguirre P and Grossmann I 2004 A decomposition method for synthesizing complex column configurations using tray-by-tray GDP models *Comput. Chem. Eng.* **28** 2165
[4] Ostrovsky G M, Ziyatdinov N N, Lapteva T V and Zaitsev I V 2012 Optimization of chemical processes with dependent uncertain parametres *Chemical Engineering Science* **83** 119–127
[5] Ostrovsky G, Zyatdinov N and Lapteva T 2011 One stage optimization problem with chance constraints *Chemical Engineering Transactions* **25** 225–230
[6] Antipin A F 2013 A computer-aided system for designing multidimensional logic controllers with variables representing a set of binary logic arguments *Automation and Remote Control* **74** (9) 1573–1581
[7] Ostrovskii G M, Ziyatdinov N N and Emel'yanov I I 2015 Synthesis of optimal systems of simple distillation columns with heat recovery *Doklady Chemistry* **461** (1) 89–92
[8] Biegler L T 2010 Nonlinear Programming: Concepts, Algoritms and Applications to Chemical Processes (Philadelphia: SIAM) p. 399
[9] Biegler L T 2017 Integrated Optimization Strategies for Dynamic Process Operations *Theor. Found. Chem. Eng.* **51** (6) 910–927
[10] Sedov V A and Sedova N A 2014 Modelling collision avoidance actions in closest approach zones by means of neural networks *Asia-Pacific Journal of Marine Science & Education* **4** (2) 104–111
[11] Antipin A F 2015 Improving response time of real time control systems based on multidimensional interval-logical controllers *Automation and Remote Control* **76** (3) 480–486
[12] Sedova N A, Sedov V A and Bazhenov R I 2018 Analysis of emergency level at sea using fuzzy logic approaches *Advances in Intelligent Systems and Computing* **658** 314–322
[13] Holland J N 1975 Adaptation in Natural and Artificial Systems (Ann Arbor, Michigan: Univ. Michigan Press) p. 232
[14] Herrera F, Lozano M and Verdegay J L 1998 Tackling real-coded genetic algorithms: operators and tools for the behavior analysis *Artific. Intelligence Rev.* **12** (4) 265–319

[15] Panteleev A V and Metlitskaya D V 2011 An application of genetic algorithms with binary and real coding for approximate synthesis of suboptimal control in deterministic systems *Automation and Remote Control* **72** (11) 2328–2338
[16] Wright A 1991 Genetic algorithms for real parameter optimization *Foundat. Genet. Algorithm.* **1** 205–218
[17] Mazitov A A, Osipova A G, Akhmetov I V and Gubaidullin I M 2016 Solution of inverse problem of chemical kinetics of synthesis reaction benzylidenebenzylamine *Middle Volga Mathematical Society Journal* **18** (3) 145–152
[18] Mustafina S A, Vaytiev V A and Stepashina E V 2014 Search the kinetic parameters of the reduced scheme of α-methylstyrene dimerization reaction *ARPN Journal of Engineering and Applied Sciences* **9** (7) 1118–1120