Metaheuristic algorithms for identification of the convection velocity in the convection-diffusion transport model

A V Tsyganov[1], Yu V Tsyganova[2], A N Kuvshinova[1] and H R Tapia Garza[1]

[1] Department of Mathematics, Physics and Technology Education, Ulyanovsk State Pedagogical University named after I.N. Ulyanov, Ulyanovsk, Russian Federation
[2] Department of Mathematics, Information and Aviation Technology, Ulyanovsk State University, Ulyanovsk, Russian Federation

Abstract. This paper considers an application of metaheuristic algorithms for solving the problem of convection velocity identification in the convection-diffusion transport model. Algorithms based on numerical minimization of the parameter identification criterion for a discrete linear stochastic model using the simulated annealing and a genetic algorithm are proposed. The log-likelihood function is used as the identification criterion. Numerical experiments were conducted to compare the computational properties of the proposed algorithms.

1. Introduction and problem statement

Convection-diffusion transport models are an indispensable tool for describing various natural and anthropogenic processes [1], [2]. These models contain parameters that must be specified to uniquely determine the solution of boundary value problems, but in practice, situations often arise where some of these parameters are unknown or given approximately and they need to be determined or refined. Such problems belong to the class of inverse problems for the models of matter transfer.

In the simplest one-dimensional case, the convection-diffusion transport model can be described by equation (1) with initial condition (2) and boundary conditions (3):

$$\frac{\partial c}{\partial t} + v\frac{\partial c}{\partial x} = \alpha\frac{\partial^2 c}{\partial x^2}, a < x < b, 0 < t < +\infty, \quad (1)$$

$$c(x,0) = \varphi(x), a \le x \le b, \quad (2)$$

$$c(a,t) = f_1(t), c(b,t) = f_2(t), 0 < t < +\infty, \quad (3)$$

where $c(x,t)$ is the function of interest (for example, the concentration of the pollutant), x is the spatial coordinate, t is the time, v is the convection velocity, α is the diffusion coefficient, $\varphi(x)$, $f_1(t)$, $f_2(t)$ are given functions, a, b are boundaries of the considered segment.

When solving a wide range of problems in ecology, geophysics, seismology, and other areas, the problem of determining (identifying) the convection velocity in the convection-diffusion transport model often arises. Depending on the equation under consideration and the boundary conditions, various methods can be used to solve this problem. In [3], [4], to find the parameters of equation (1), time series analysis methods based on the method of least squares, extended Kalman filter and their combination are used.

In this paper, we propose the use of metaheuristic algorithms for numerical optimization to find the optimal estimate of a parameter of a discrete linear stochastic model with respect to a given criterion of identification quality. As in [4], this model with noisy measurements is constructed from the original model (1)–(3) using a two-layer finite-difference scheme, but unlike [4] it uses a grid not with three but with an arbitrary number of nodes on the coordinate x and the conventional, rather than the extended Kalman filter.

The choice of metaheuristic algorithms to optimize a parameter identification criterion is due to the fact that deterministic numerical methods applicability is guaranteed under the conditions of convergence theorems [5]. The convergence of numerical methods is influenced by various factors, including a good choice of the initial approximation.

If the initial approximation is not chosen correctly, the exact algorithm for finding the estimates of the parameters may diverge, which means that it is impossible to solve the identification problem. Metaheuristic algorithms can be used both directly for solving the problem of minimizing the identification criterion, and for finding a good initial approximation for exact methods. A similar approach was applied by the authors in [6–9] and proved its efficiency.

2. Discrete linear stochastic model

We want to move from the model (1)–(3) to a discrete linear stochastic model, whose equations generally have the following form:

$$\begin{cases} c_{k+1} = F(\theta)c_k + B(\theta)u_k + G(\theta)w_k, \\ z_{k+1} = H(\theta)z_{k+1} + \xi_{k+1}, \\ k = 0,1,\dots \end{cases} \tag{4}$$

where $c_k \in \mathbb{R}^n$ is the system state vector, $u_k \in \mathbb{R}^r$ is the control vector, $z_k \in \mathbb{R}^m$ is the measurements vector, noises $w_k \in \mathbb{R}^q$ and $\xi_k \in \mathbb{R}^m$ form independent normally distributed sequences with zero mean and covariance matrices $Q(\theta) \geq 0$ and $R(\theta) > 0$ respectively, matrices $F(\theta) \in \mathbb{R}^{n\times n}$, $B(\theta) \in \mathbb{R}^{n\times r}$, $G(\theta) \in \mathbb{R}^{n\times q}$, $H(\theta) \in \mathbb{R}^{m\times n}$, $Q(\theta) \in \mathbb{R}^{q\times q}$, $R(\theta) \in \mathbb{R}^{m\times m}$ can depend on an unknown parameter θ.

In computational practice, for the numerical solution of non-stationary problems for convection-diffusion equations, two- and three-layer finite-difference schemes are most often used. To obtain a discrete linear stochastic model, consider in the Oxt plane a regular grid (5) with spatial step Δx and time step Δt:

$$x_i = a + i\Delta x, t_k = k\Delta t, i = 0,1,\dots,n, k = 0,1,\dots. \tag{5}$$

Let us denote $c_k^i = c(x_i, t_k)$, $\varphi^i = \varphi(x_i)$, $f_{1k} = f_1(t_k)$, $f_{2k} = f_2(t_k)$ and write down the finite-difference scheme for (1)–(3):

$$\frac{c_{k+1}^i - c_k^i}{\Delta t} + v\frac{c_k^{i+1} - c_k^{i-1}}{2\Delta x} = \alpha\frac{c_k^{i+1} - 2c_k^i + c_k^{i-1}}{\Delta x^2}, \quad i = 1,2,\dots,n-1, k = 0,1,\dots, \tag{6}$$

$$c_0^i = \varphi^i, i = 0,1,\dots,n, \tag{7}$$

$$c_k^0 = f_{1k}, c_k^n = f_{2k}, k = 0,1,\dots. \tag{8}$$

It follows from equation (6) that the value of the required function at the internal node points of the $k+1$-th time series can be found through its values at the nodal points of the k-th time series as follows:

$$c_{k+1}^i = (r_1 + r_2)c_k^{i-1} + (1 - 2r_2)c_k^i + (r_2 - r_1)c_k^{i+1} \tag{9}$$

where $r_1 = \frac{v\Delta t}{2\Delta x}$, $r_2 = \frac{\alpha\Delta t}{\Delta x^2}$. Let us write (9) in the form

$$c_{k+1}^i = a_1 c_k^{i-1} + a_2 c_k^i + a_3 c_k^{i+1} \tag{10}$$

where $a_1 = r_1 + r_2, a_2 = 1 - 2r_2, a_3 = r_2 - r_1$.

The desired discrete linear stochastic model can be represented in the following form:

$$
\begin{cases}
\begin{bmatrix} c_{k+1}^0 \\ c_{k+1}^1 \\ c_{k+1}^2 \\ \vdots \\ c_{k+1}^{n-2} \\ c_{k+1}^{n-1} \\ c_{k+1}^n \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ a_1 & a_2 & a_3 & \cdots & 0 & 0 & 0 \\ 0 & a_1 & a_2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_2 & a_3 & 0 \\ 0 & 0 & 0 & \cdots & a_1 & a_2 & a_3 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} c_k^0 \\ c_k^1 \\ c_k^2 \\ \vdots \\ c_k^{n-2} \\ c_k^{n-1} \\ c_k^n \end{bmatrix} + \begin{bmatrix} f_{1,k+1} \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ f_{2,k+1} \end{bmatrix}, \\
\begin{bmatrix} z_{k+1}^0 \\ z_{k+1}^1 \\ z_{k+1}^2 \\ \vdots \\ z_{k+1}^{n-2} \\ z_{k+1}^{n-1} \\ z_{k+1}^n \end{bmatrix} = \begin{bmatrix} c_{k+1}^0 \\ c_{k+1}^1 \\ c_{k+1}^2 \\ \vdots \\ c_{k+1}^{n-2} \\ c_{k+1}^{n-1} \\ c_{k+1}^n \end{bmatrix} + \begin{bmatrix} \xi_{k+1}^0 \\ \xi_{k+1}^1 \\ \xi_{k+1}^2 \\ \vdots \\ \xi_{k+1}^{n-2} \\ \xi_{k+1}^{n-1} \\ \xi_{k+1}^n \end{bmatrix}, \\
k = 0,1,\dots .
\end{cases}
\tag{12}
$$

Note that the first equation in model (12) is deterministic, the initial value of the state vector is (7), and the boundary conditions (8) act as control parameters.

Suppose that the diffusion coefficient α and the characteristics of the noise in the measurer are known, and the steps of the space-time grid Δx and Δt are given, then the unknown parameter of model (12) to be determined is the convection velocity v, on which coefficients a_1 and a_3 of equation (10) and, consequently, matrix F depend. Briefly, the model (12) can be written in the form:

$$
\begin{cases}
c_{k+1} = F(\theta)c_k + u_{k+1}, \\
z_{k+1} = c_{k+1} + \xi_{k+1}, \\
k = 0,1,\dots
\end{cases}
\tag{13}
$$

where $\theta = v$.

3. Metaheuristic algorithms for parameter identification

Let us consider the problem of parameter identification of model (13) from noisy measurements data for estimating an unknown parameter. The parameter identification problem consists in finding an unknown parameter θ from known input signals $U_0^{N-1} = \{u_0, u_1, \dots, u_{N-1}\}$ and the output observation data $Z_1^N = \{z_1, z_2, \dots, z_N\}$ in accordance with the chosen identification criterion $J(\theta; Z_1^N, U_0^{N-1})$. In this case, the problem of estimating an unknown parameter θ requires solving the nonlinear programming problem with constraints

$$
\hat{\theta}_{min} = \operatorname{argmin} J(\theta; Z_1^N, U_0^{N-1}) \tag{14}
$$

where $\theta \in D(\theta)$ (the domain of θ).

In this paper, we use metaheuristic algorithms to solve problem (14). Metaheuristic is a high-level search strategy for finding solutions, applicable to a wide range of optimization tasks. Metaheuristics have the following properties: they are based on fairly simple ideas, for example, imitating biological or physical processes, they are problem-independent, practically all of them are nondeterministic.

Most metaheuristic optimization algorithms can be divided into two large groups according to the method of obtaining a solution: trajectory and population-based ones. In trajectory algorithms, the process of finding a solution can be considered as a movement between individual solutions of a problem while in population-based algorithms a group of solutions called population changes in the process of finding the solution.

One of the most popular trajectory algorithms used in solving global optimization problems is the simulated annealing (SA) method. A key feature of the method is the use of a control parameter– a temperature, which allows controlling the nondeterministic process of solution search. As a rule, the temperature decreases during the operation of the algorithm according to a certain law, starting with some initial value. At each iteration of the algorithm, the randomly generated new solution from the neighborhood of the current solution is taken with probability 1 if it improves it, and with probability less than 1, if worsens, and the probability of making the worst decision decreases with a decreasing temperature. The quality of decisions is estimated using a cost function (integer or real).

Simulated Annealing (SA)

```
Solution ← InitialSolution()
BestSolution ← Solution
BestCost ← Cost(Solution)
T ← InitialTemperature()
n ← 0
while not StopCondition() do
    NewSolution ← ChooseRandomOf(Neighborhood(Solution))
    NewCost ← Cost(NewSolution)
    if NewCost < BestCost then
        BestSolution ← NewSolution
        BestCost ← NewCost
    end if
    Solution ← AcceptWithProbability(Solution, NewSolution, T)
    n ← n + 1
    T ← UpdateTemperature(T, n)
end while
Output: BestSolution
```

The genetic algorithm (GA) is a popular version of the so-called evolutionary optimization algorithms based on the simulation of natural selection processes. In evolutionary algorithms, the quality of solutions is estimated using the fitness function, and the main idea of algorithms is that solutions with the best values of a given function “survive” in the course of evolution.

In GA, at each iteration of the evolutionary process, a new population is obtained from the current population using one or more genetic operators successively. The most common genetic operators are the crossover (recombination) which is used to generate the descendant solutions from the parent solutions and the mutation – an accidental change in the solution.

Genetic Algorithm (GA)

```
Population ← InitialPopulation()
for all p_i ∈ Population do
    EvaluateFitness(p_i)
end for
while not StopCondition() do
    Parents ← SelectParents(Population)
    Offspring ← Crossover(Parents)
    Offspring ← Mutation(Offspring)
    for all p_i ∈ Offspring do
        EvaluateFitness(p_i)
    end for
    Population ← UpdatePopulation(Population ∪ Offspring)
end while
Solution ← ChooseBestOf(Population)
Output: Solution.
```

The algorithms SA and GA are discussed in more detail, for example, in [10].

Since (13) is a discrete linear stochastic model with Gaussian noise, it is advisable to select the identification criterion (14) in the form of a negative logarithmic likelihood function as the cost/fitness function for implementing metaheuristic algorithms [11]

$$J_{MLF}(\theta; Z_1^N, U_0^{N-1}) = -\frac{Nm}{2}\ln(2\pi) - \frac{1}{2}\sum_{k=1}^{N}\left\{\ln[\det(B_k)] + e_k^T R_{e,k} e_k\right\} \tag{15}$$

where the residual vector e_k and its covariance matrix $R_{e,k}$ for a given value of the parameter θ in (4) are calculated from the known Kalman filter equations [12]:

A. Time update.

For $k = 0,1, \ldots$ the Kalman filter computes extrapolated estimates $\hat{c}_{k+1}^-$ for c_{k+1}. They are obtained through the temporal update from k to $k+1$ as

$$\hat{c}_{k+1}^- = F(\theta)\hat{c}_k^+ + B(\theta)u_k$$

with $\hat{c}_0^+ \coloneqq \bar{c}_0 \triangleq E\{c_0\}$ and the covariance matrices

$$P_{k+1}^- = Q(\theta) + F(\theta)P_k^+ F^T(\theta)$$

where $P_0^+ \coloneqq P_0 \triangleq E\{[c_0 - \bar{c}_0][c_0 - \bar{c}_0]^T\}$.

B. Measurement update.

For $k = 0,1, \ldots$ the Kalman filter computes the so-called filtered (i.e., measurement updated) estimates $\hat{c}_k^+$. They are obtained through the measurement update using z_k with noise covariance $R(\theta) > 0$, as

$$\hat{c}_k^+ = \hat{c}_k^- + K_k^f e_k, \quad e_k = z_k - H(\theta)\hat{c}_k^-,$$

with filter gain

$$K_k^f = P_k^- H^T(\theta) R_{e,k}^{-1}, \quad R_{e,k} = R(\theta) + H(\theta)P_k^- H^T(\theta)$$

and filtered estimates covariance matrices

$$P_k^+ = P_k^- - K_k^f H(\theta) P_k^-.$$

At present, when solving practical problems with the use of a computer, it is preferable to apply square-root and UD-implementations that are numerically stable against machine round-off errors instead of the conventional form of the Kalman algorithm [12, Chapter 6].

The maximum likelihood method consists in optimizing criterion (15) with respect to the system parameter θ. It is often used in practice to solve parameter identification problems of discrete linear stochastic systems [12, 13].

4. Numerical experiments

Consider the following problem:

$$\frac{\partial c}{\partial t} + v\frac{\partial c}{\partial x} = \alpha\frac{\partial^2 c}{\partial x^2}, 0 < x < \pi, 0 < t < +\infty, \tag{16}$$

$$c(x,0) = \exp\left(\frac{vx}{2\alpha}\right)\sin x, 0 \le x \le \pi, \tag{17}$$

$$c(0,t) = 0, c(\pi,t) = 0, 0 < t < +\infty \tag{18}$$

where $c(x,t)$ is the concentration of the pollutant in one-dimensional flow, (17) is the initial distribution of the pollutant and boundary conditions (18) correspond to the case of two absorbing walls.

The exact solution of (16)–(18) has the form:

$$c(x,t) = \exp\left[\frac{v}{2\alpha}\left(x - \frac{vt}{2}\right)\right]\sin x \exp(-\alpha t). \tag{19}$$

Suppose that the value of parameter α in equation (16) is known and it is required to determine the value of parameter v, provided that noisy measurements from sensors located at nodes of some regular grid are available. Let a grid with 10 nodes ($\Delta x = \frac{\pi}{9}$) be given on the Ox axis, and the time step $\Delta t = \frac{\Delta x^2}{4\alpha}$.

Let $\alpha = 1$, $v = 2$. The plot of the exact solution for this case is shown in figure 1.

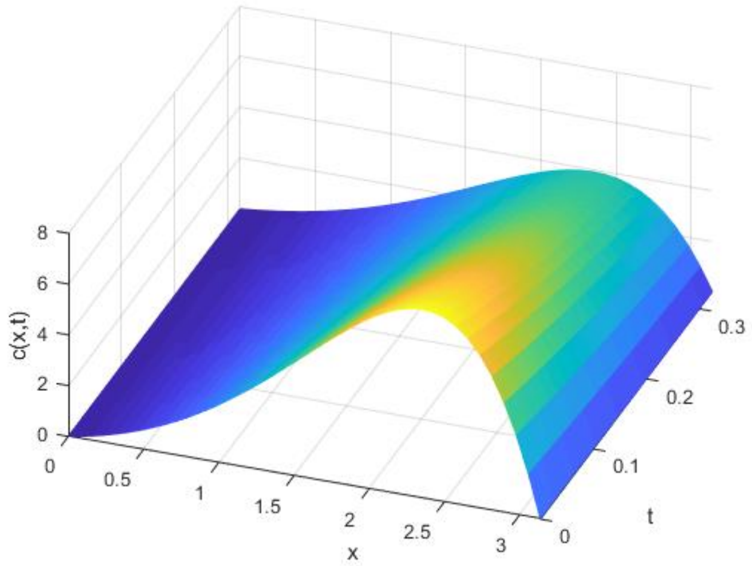

Figure 1. The plot of the exact solution.

To simulate noisy measurements we add random errors (white noise) to the values of the exact solution (19) at the grid nodes. The covariance matrix of the noise is $R = \sigma^2 I_{n\times n}$, where I is the identity matrix and σ^2 is the known variance. An example of a noisy solution is shown in figure 2.

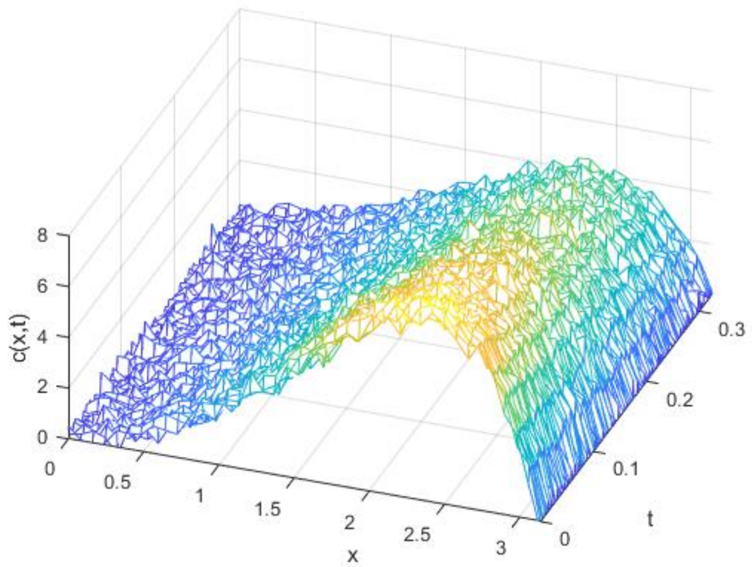

Figure 2. The plot of the noised solution.

Figure 3 demonstrates the averaged plot of criterion (15), obtained from the results of 100 experiments. To minimize it, we have used functions simulannealbnd() and ga() from the Global Optimization Toolbox of the MATLAB system. Numerical experiments were conducted on a hardware-software platform: Intel Core 2 Quad Q6600 @ 2.40 GHz, 4 Gb RAM, Microsoft Windows 10 Pro x64, MATLAB R2017a.

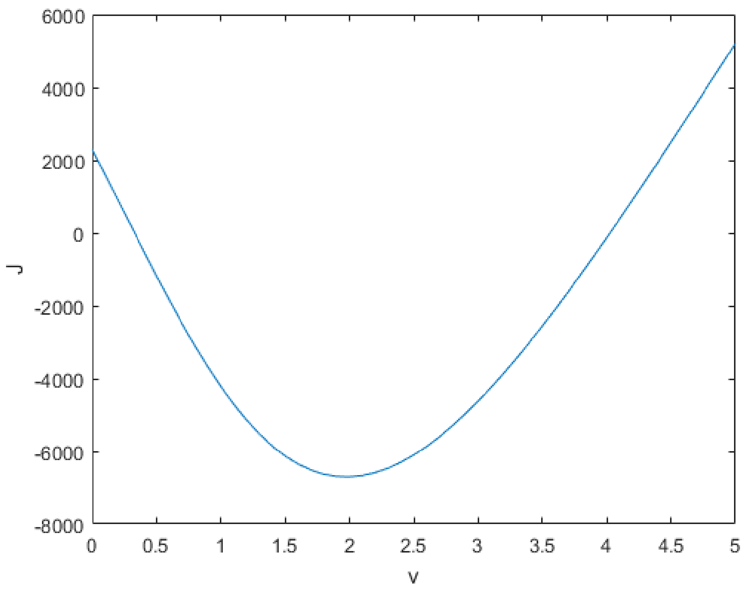

Figure 3. The plot of the identification criterion.

For each of the algorithms, the corresponding cost/fitness and output functions were written. Basic settings of the algorithms are given in table 1. As a stopping criterion for both algorithms, a time limit of 5 seconds was used. The search for solutions was carried out on the interval [0; 5], the number of steps in time (i.e., the number of time series) is $m = 500$.

Table 1. SA and GA settings.

SA		GA	
TimeLimit	5	TimeLimit	5
MaxIter	Inf	Generations	Inf
MaxFunEvals	Inf	StallGenLimit	Inf
StallIterLimit	Inf	PopulationSize	20
ReannealInterval	100	PopInitRange	0..5

Table 2 shows the results of computational experiments for different values of the noise variance σ^2. For each value of σ^2, a series of 100 experiments was conducted and for each series, the following values were calculated: mean value of the identified parameter (Mean), mean absolute percentage error (MAPE) and root mean squared error (RMSE). The obtained results show that, with the selected settings, both algorithms allow identifying the convection velocity with an acceptable accuracy. At high noise values, the quality of parameter identification is approximately the same for both algorithms, with a low noise level, the simulated annealing method yields smaller MAPE and RMSE errors.

Table 2. Experiment results.

σ^2	SA			GA		
	Mean	MAPE	RMSE	Mean	MAPE	RMSE
0.001	1.9710	2.2439	0.0566	1.9739	4.7195	0.1316
0.005	1.9786	2.0756	0.0532	1.9937	4.9573	0.1343
0.01	1.9665	2.3950	0.0660	1.9813	4.8876	0.1274
0.05	1.9762	2.6874	0.0708	1.9710	4.8828	0.1315
0.1	1.9662	2.7181	0.0725	1.9985	4.9098	0.1273
0.5	1.9642	5.0677	0.1290	1.9953	6.4925	0.1869
1	1.9643	6.9710	0.1739	1.9635	7.0476	0.1750

5. Conclusion

This paper demonstrates the practical applicability of metaheuristic algorithms for solving the problem of parameter identification in the model of convection-diffusion transport. The convection velocity was considered as an unknown parameter of the model. The problem solution was obtained with the use of the maximum likelihood method, in which the simulated annealing method and the genetic algorithm were used to numerically minimize the negative logarithmic likelihood function.

Numerical experiments were conducted in MATLAB. The results obtained make it possible to conclude that the application of metaheuristic algorithms is expedient since it allows to obtain acceptable estimates of the parameter for different levels of noisy measurements.

Further research will focus on the construction and software implementation of new exact and hybrid algorithms for parameter identification of convection-diffusion transport models and their application to real life problems.

6. References

[1] Leontiev A I 1997 *The Theory of Heat and Mass Transfer* ed 2 (Moscow: Bauman Moscow State Technical University) p 683

[2] Farlow S J 1982 *Partial Differential Equations for Scientists and Engineers* (New York: Wiley) p 402

[3] Matveev M G, Sirota E A, Semenov M E and Kopytin A V 2017 Verification of the Convective Diffusion Process Based on the Analysis of Multidimensional Time Series *Proc. of the DAMDID/ RCDL'2017. XIX International Conference on Data Analytics and Management in Data Intensive Domains* (Moscow: Federal research center "Informatics and Control" of the Russian Academy of Sciences) pp 433-437

[4] Matveev M G, Kopytin A V and Sirota E A 2018 Combined method for identifying the parameters of a distributed dynamic model *Proc. of the ITNT-2018. IV International Conference on Information Technology and Nanotechnology* (SAMARA: New Technique) pp 1651-1656

[5] Vasilyev V P 1982 *Numerical Methods for Solving Extreme Problems* (Moscow: Mir) p 372

[6] Tsyganova Yu V and Tsyganov A V 2010 Simulation normalization in the problem of identification of parameters of stochastic linear system *Stochastic Optimization in Informatics* **6** 1 pp 147-159

[7] Tsyganov A V, Bulychov O I and Tsyganova Yu V 2011 Parallel hybrid algorithms for the problem of parameter identification in stochastic linear systems *Vector of Sciences. Togliatti State University* 3(17) pp 45-49

[8] Tsyganova Yu V and Tsyganov A V 2013 Program for parameter identification in stochastic linear systems ISLSP v.1.1 *The Certificate of State Registration of Computer Programs* no 2013612686

[9] Tsyganov A V, Semushin I V, Tsyganova Yu V, Golubkov A V and Vinokurov S D 2017 Metaheuristic algorithms in the issue of identification of the moving object mathematical model parameters *Automation of Control Processes* 1 pp 16-23

[10] Hromkovič J. 2002 *Algorithmics for Hard Problems: Introduction to Combinatorial Optimization, Randomization, Approximation and Heuristics* ed 2 (Springer-Verlag Berlin, Heidelberg) p 557
[11] Åström K J 1980 Maximum Likelihood and Prediction Error Methods *Automatica* **16** 5 pp 551-574
[12] Grewal M S and Andrews A P 2001 *Kalman Filtering: Theory and Practice Using MATLAB* ed 2 (New Jersey: Prentice Hall) p 401
[13] Gibbs V P 2011 *Advanced Kalman Filtering, Least-squares and Modeling: Practical Handbook* (Hoboken, New Jersey: John Wiley & Sons, Inc.) p 632