Original article
DOI: https://doi.org/10.18721/JPM.15307

RECONSTRUCTING THE THERMAL PROCESS MODEL USING THE TIME-SPACE DISTRIBUTIONS OF TEMPERATURE

N. Yu. Bykov[1, 2]✉

[1] ITMO University, St. Petersburg, Russia

[2] Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

✉ nbykov2006@yandex.ru

Abstract. The method of generative model design (GMD) has been applied to reconstruct the structure and coefficients of a partial differential equation describing the target's heating and its evaporation by laser radiation. The initial synthetic data includes heating scenarios corresponding to surface energy absorption or to volume one. It was shown that reconstructing the model correctly required the use of a preprocessing technique providing the exclusion of a part of the initial data if the volume absorption took place. A modification of the method that made it possible to take into account the temperature dependence of the coefficients of the reconstructed equation was put forward. The influence of various statistical criteria used in selecting the optimal subset of elements on the accuracy of reconstructing the equation structure was discussed. The efficiency of the GMD was demonstrated for a wide range of target heating parameters and different options for setting the energy input. The possibility of model generating by noisy data was shown.

Keywords: method of generative model design, data-driven model, heat transfer equation, laser heating and evaporation

Funding: The reported study was funded by Russian Science Foundation (Grant 21-11-00296).

Citation: Bykov N. Yu., Reconstructing the thermal process model using the time-space distributions of temperature, St. Petersburg Polytechnical State University Journal. Physics and Mathematics. 15 (3) (2022) 83–99. DOI: https://doi.org/10.18721/JPM.15307

Научная статья
УДК 519.6
DOI: https://doi.org/10.18721/JPM.15307

ВОССТАНОВЛЕНИЕ МОДЕЛИ ТЕПЛОВОГО ПРОЦЕССА ПО ПРОСТРАНСТВЕННО-ВРЕМЕННЫМ РАСПРЕДЕЛЕНИЯМ ТЕМПЕРАТУРЫ

Н. Ю. Быков[1, 2]✉

[1] Национальный исследовательский университет ИТМО, Санкт-Петербург, Россия

[2] Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия

✉ nbykov2006@yandex.ru

Аннотация. Метод генеративного дизайна модели (ГДМ) применен для восстановления структуры и коэффициентов дифференциального уравнения в частных производных, описывающего процесс нагрева и испарения мишени лазерным излучением. Исходные синтетические данные включают сценарии нагрева, соответствующие поверхностному или объемному поглощению энергии. Показано, что в случае объемного поглощения для корректного восстановления модели требуется применение процедуры препроцессинга, предусматривающей исключение части исходных данных. Предложена модификация метода, позволяющая учитывать зависимость коэффициентов восстанавливаемого уравнения от температуры. Обсуждается влияние различных статистических критериев, применяемых при селекции оптимального подмножества элементов, на точность восстановления структуры уравнения. Эффективность применения метода ГДМ продемонстрирована для широкого диапазона параметров нагрева мишени и разных вариантов задания энергоподвода. Показана возможность генерации модели по зашумленным данным.

Ключевые слова: метод генеративного дизайна, управляемая данными модель, уравнение теплопроводности, лазерный нагрев

Финансирование: Исследование выполнено при финансовой поддержке Российского научного фонда (грант № 00296-11-21, https://rscf.ru/project/21-11-00296/)

Ссылка для цитирования: Быков Н. Ю. Восстановление модели теплового процесса по пространственно-временным распределениям температуры // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2022. Т. 15. № 3. С. 83–99. DOI: https://doi.org/10.18721/JPM.15307

Introduction

Data-driven models are widely used for predicting the parameters of social, political or physical processes and subsequently controlling them [1, 2]. The requirements imposed on such models generally include accuracy of the predictions and interpretability of the models themselves. If the object is characterized by quantitative predictors, then the model is typically a linear or nonlinear regression function [3]. Alternative models are formulated as ordinary differential equations (ODE) or partial differential equations (PDE), reconstructed from the available data. The data-driven models based on differential equations are expected to provide both good interpretability and accuracy of the predictions they yield [2].

Different methods for reconstructing the model of a process as a PDE based on the available data are widespread in studies of heat and mass transfer. The form of the ODE describing the majority of the thermal processes is well known: in the simplest case, it is a classical equation

of thermal conductivity. However, the equation may include a second derivative of temperature with respect to time in the general case [4, 5], and a corresponding convective term for a moving medium. If energy is released inside the object, one or more additional terms appear in the equation determining the power of internal heat sources [5]. A tool for selecting the significant terms of the equation from a large array of 'building blocks' should allow to detect the processes that occur within the internal volume of the object and cannot therefore be visualized during the experiment. Examples of such processes are phase transitions and chemical reactions. Reconstructing the convective term with the corresponding weight in the energy equation allows to assess the presence of convective processes in the object considered. What is more, data can be obtained for the qualitative changes in the thermal process (for example, transition from predominantly thermal conductivity to well-developed convection with the liquid heated non-uniformly [6]).

Problems on reconstructing the structure of PDE and determining the coefficients of the equation with respect to thermal processes can be solved by the method of generative model design (GMD) proposed in [7]. The method comprises several stages.

I. The most complete possible structure of the reconstructed equation is determined.

II. PDE elements are discretized.

III. Values are calculated for the vector elements of discretized templates based on available data on spatial-temporal temperature distributions.

IV. Statistical methods are applied to determine the optimal structure and the PDE coefficients.

Even though the core of the method has been formulated, some questions are yet to be addressed.

First of all, the effectiveness of the method clearly depends on the quality of the initial data. In addition to the above stages of the GMD algorithm, we should consider the stage when the available synthetic or experimental data are preprocessed. The quality of such data should be analyzed at this stage in order to make a decision as to whether they can be used fully or partially.

Secondly, the algorithm for constructing discretized stencils for the case of temperature-dependent parameters of the model is not discussed in detail in the literature.

Thirdly, different statistical criteria can be adopted at the stage when the optimal structure of the equation is chosen (Mallow's criterion, information criteria, etc.). The statistical criterion selected should be assessed for adequacy to determine the optimal structure of the model.

Fourthly, an important question yet to be answered definitively is whether the GMD algorithm is applicable to noisy data.

In general, the potential offered by the GMD method should explored further to accumulate experience with its practical applications.

This study concentrated on developing the GMD method proposed in [7]. Our intention was to further expand the scope of the method, trying to answer the above questions.

Generation of initial synthetic data

To better illustrate the capabilities of the GMD method, the initial data for reconstructing the thermal process model include information about the spatio-temporal evolution of temperature in the material, accounting for both the presence/absence of physical processes accompanying heating, and the temperature dependence of the thermophysical parameters of the medium.

The process of heating a metal target with laser radiation is a convenient object to consider. Firstly, this process depends on the radiation parameters and can be accompanied by phase transitions (melting of the material and evaporation of its surface); secondly, the temperature range of the material is very wide and requires taking into account the temperature dependence of the medium parameters. Thirdly, the data can be obtained synthetically by numerical modeling [8–10].

As an example, we consider a niobium target heated by moderate-intensity laser pulses. This unsteady process can be assumed to be one-dimensional for the case when the depth to which the target is heated is much smaller depth than the diameter of the laser spot.

Data on the spatio-temporal distribution of temperature $T(x,t)$ in the target are generated to subsequently use the GMD method by numerically solving the thermal conductivity equation [5, 9, 10] taking the form

$$c\rho\left(\frac{\partial T}{\partial t}-\omega\frac{\partial T}{\partial x}\right)=\frac{\partial}{\partial x}\lambda\frac{\partial T}{\partial x}+q_V, \qquad (1)$$

where x, m, is the spatial coordinate; t, s, is the time; ρ, kg/m^3, c, J/(kg·K), λ(T), W/(m·K), are the density, heat capacity, and thermal conductivity of the material, respectively; q_v, W/m^3, is the volumetric heat source; ω, m/s, is the speed of the evaporating surface.

The velocity ω depends on the surface temperature T_S ($T_S \equiv T(x = 0)$) and is the function of time only: ω = ω(t). Eq. (1) is written in the moving coordinate system X, with the origin corresponding to the target surface. Surface evaporation does not occur for surface temperatures significantly lower than T_b (T_b = 5033 K is the boiling point of niobium at normal pressure p_b [11]), and the surface velocity is ω = 0. Surface evaporation is observed at temperatures exceeding or close to T_b. The velocity ω starts to differ from zero. The coefficient ω ≠ 0 can serve as an indicator pointing to the presence of an evaporation process. Melting of niobium is not taken into account in the given formulation (melting point of niobium T_L = 2750 K [11]).

The boundary condition on the surface is imposed as

$$-\lambda \left.\frac{\partial T}{\partial x}\right|_{x=0} = q_S - L\rho\omega, \tag{2}$$

where q_s, W/m^2, is the energy flux of laser radiation over the target surface; L, J/kg, is the latent heat of evaporation.

The initial temperature at the remote boundary $T(\infty,t) = T_0$ corresponds to $T(x,0) = T_0$ (T_0 = 300 K).

Two scenarios of laser energy absorption are considered in this paper:

(*i*) the energy flux of laser radiation through the target surface per unit time (surface flux) is given by the expression

$$q_S = (1 - R_f)W_0, \tag{3}$$

where W_0, W/m^2, is the density of the radiation flux incident on the target surface; R_f is the reflectance of the material surface. Volume absorption is absent in this case ($q_V = 0$).

(*ii*) it is assumed that the volumetric absorption $q_V \neq 0$, $q_s = 0$.

In the latter case, the heat input at a distance x from the surface is determined by the expression

$$q_V(x,t) = \alpha_a I(x,t), \tag{4}$$

$$I(x,t) = I_0(t)\exp(-\alpha_a x), \tag{5}$$

where $I(x,t)$, W/m^2, is the radiation intensity in the material at a distance x from the surface; $I_0(t)$, W/m^2, is the density of the radiant flux penetrating the material, $I_0(t) = (1 - R_f)W_0$; α_a, m^{-1}, is the absorptance.

Eq. (5) expresses the Beer–Lambert–Bouguer absorption law.

It is assumed that the radiation power does not change, i.e.,

$$W_0 = const,\ t \le t_0;\ W_0 = 0,\ t > t_0. \tag{6}$$

The required values of the optical parameters are taken from [12, 13]: $\alpha_a = 5\cdot10^7$ m^{-1} (for lasing wavelength of the order of 1 μm), R_f = 0.77. Laser pulse duration t_0 = 0.1 μs.

The generated data correspond to two temperature ranges:

I. Target temperature is substantially below the boiling point T_b (and the melting point T_L);

II. Surface temperature is above T_b.

The experimental data [14] on the thermal conductivity of niobium in the range 300 < T < 2200 K are approximated by a third degree polynomial for temperature range I:

$$\lambda = \lambda_0 + 6.687\cdot10^{-3}T + 6.652\cdot10^{-6}T^2 - 2.256\cdot10^{-9}T^3, \tag{7}$$

where λ_0 = 51.49 W/(m·K).

The density and heat capacity of niobium are assumed to be constant in this formulation, amounting to ρ = 8570 kg/m^3, c_0 = 263 J/(kg·K).

The temperature dependences given in the literature for the parameters of liquid niobium differ considerably near its melting point for Range II [14]. In this case, we assume the thermal conductivity and the heat capacity to be constant over the entire temperature range. In this case, the thermophysical parameters correspond to the parameters of liquid niobium [11, 14]:

$$\rho_L = 7580 \text{ kg/m}^3, c_{0L} = 449.9 \text{ J/(kg·K)}, \lambda_{0L} = 65 \text{ W/(m·K)}.$$

The surface moves with a velocity corresponding to Hertz's law and the assumption that 18% of the evaporated atoms are dispersed back to the surface due to collisions in the gas phase [15]. The density of saturated vapor follows the Clausius–Clapeiron law [15, 8]:

$$\omega(t) = 0.82 \frac{p_b}{\rho} \left(\frac{m}{2\pi k_B T_S} \right)^{1/2} \exp\left[\frac{Lm}{k_B} \left(\frac{1}{T_b} - \frac{1}{T_S} \right) \right], \tag{8}$$

where k_B, J/K, is the Boltzmann constant; m, kg, is the atomic mass.

The Crank-Nicolson scheme of the finite difference method is used to numerically solve Eq. (1), providing a second-order approximation with respect to both the spatial and the temporal coordinate [5]. Thermal conductivity at the point between grid nodes j and j+1 was determined by the mean temperature in the nodes:

$$\lambda_{j+1/2} = \lambda\left((T_{j+1} + T_j)/2 \right).$$

The thermal conductivities were tailored during the iterative process for each subsequent time step (n + 1), accounting for the relationship between thermal conductivity and temperature.

The computational scheme of the problem with volume absorption assumes that the volume where the absorption occurs is present around each grid node. The volume boundaries for node j with the coordinate x_j are defined for this one-dimensional case as

$$x_j - \Delta x/2 \le x < x_j + \Delta x/2.$$

Taking into account the Booger–Lambert–Beer law (see Eq. (5)), the power density of the heat source in the volume is expressed as

$$q_V(x_j, t) = \frac{I_0(t)}{\Delta x} \left\{ \exp[-\alpha_a (x_j - \Delta x/2)] - \exp[-\alpha_a (x_j + \Delta x/2)] \right\}. \tag{9}$$

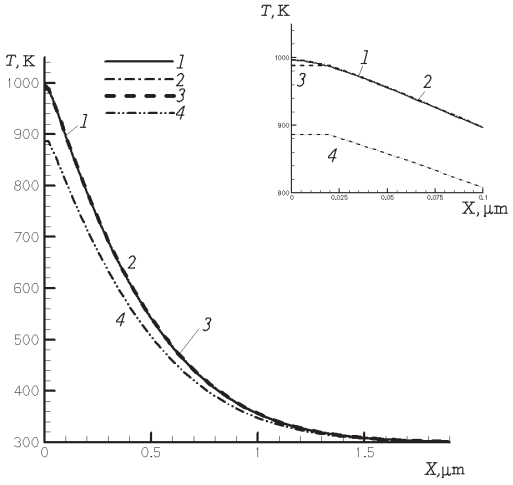

Fig. 1. Distributions of temperature T over the target depth X at time t = 10 ns, with volume absorption of laser radiation (computational case 6). The inset shows the initial region X.
Analytical solution [16] (curve *1*) is compared with the results of different computational cases: taking into account Eq. (9), with steps of 2 nm (curve *2*) and 40 nm (curve *3*); taking into account Eqs. (4), (5) and with steps of 40 nm (curve *4*)

Provided that $\Delta x \to 0$, expression (9) corresponds to expressions (4), (5). Taking into account the final value of the step Δx, the main numerical algorithm incorporates expression (9). The first node inside the body has the coordinate $x_0 = \Delta x/2$ for the problem statement with volume absorption (the target surface corresponds to the coordinate value $x = 0$). In addition to volume absorption, the boundary condition should be taken into account in form (2) at $q_s = 0$. A fictitious boundary node with the coordinate $x_{-1} = -\Delta x/2$ was included into the scheme to satisfy this boundary condition. There is no absorption of laser energy in this node,

$$T(x_{-1}, t) = T(x_0, t) = T_S(t).$$

The heating process in the target was observed for a total time of 2 μs.

The numerical algorithm intended for extracting data by GMD was verified by comparison with the analytical solutions available in monograph [16], obtained in the absence of surface evaporation ($\omega = 0$) both for the case of volume absorption (see expression (4)) and for the energy flux through the surface given by Eq. (3).

The data obtained by numerically solving the thermal conductivity equation (1) are in good agreement with the analytical solutions (Figs. 1 and 2). Using Eq. (9) and the described algorithm for volume absorption of radiation energy significantly increases the computational accuracy compared with the results obtained using expressions (4) and (5) for a large spatial step (see Fig. 1). If the surface flux is given (Fig. 2), the solution is similar to the case of volume absorption. The solutions differ only in the region of the target located directly near the surface.

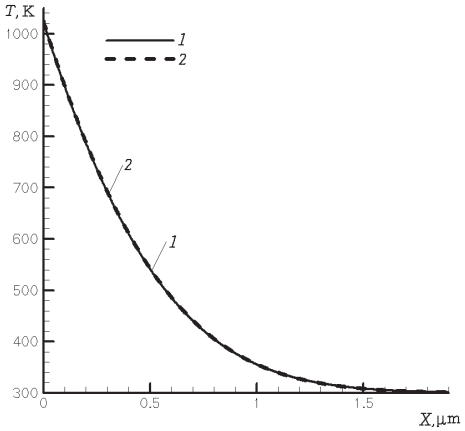

Fig. 2. Distributions of temperature T over the target depth X at time $t = 10$ ns, with radiation flux through the surface (computational case 1).
Analytical solution [16] (curve *1*) is compared with the computational result (curve *2*)

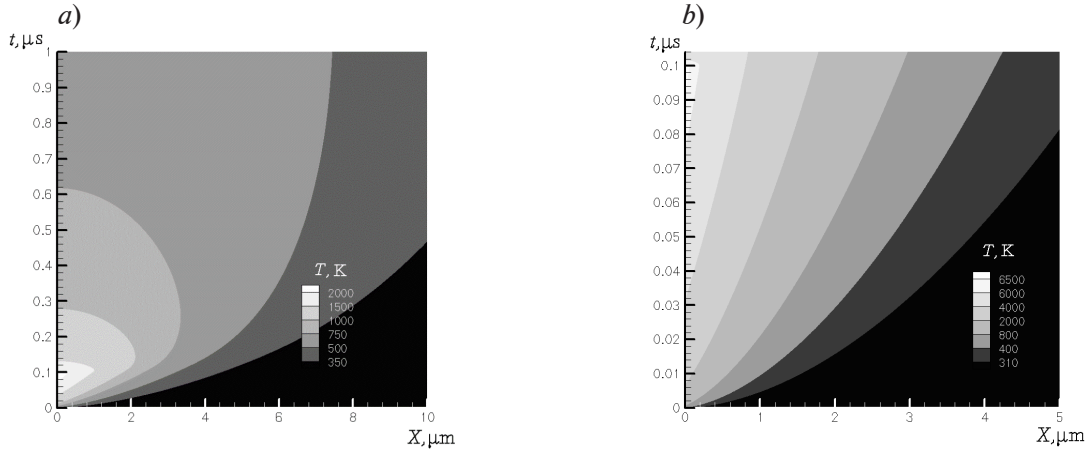

Fig. 3. Spatio-temporal temperature distributions for Cases 4 (*a*) and 11 (*b*) (see Table 1)

The cases corresponding to the generated data are given in Table 1, and example data for the temperature distributions are shown in Fig. 3.

The computational results for the second temperature range correspond to Case 11 (see Table 1 and Fig. 3,*b*). For radiation intensity of 1.2 TW/m^2 and a pulse duration of 100 ns, the surface temperature exceeds the boiling point by the time the laser pulse stops, and surface evaporation is observed. The velocity of the evaporation front is approximately 0.3 m/s by the end of the radiation time interval.

Table 1

Computational cases for generating the input data

Case	Energy input	Time sample, μs	λ	Δx, nm	N_d
1	$q_V = 0$	0.01	λ_0	2.0	973
2		1.00	λ_0		11333
3		0.01	(7)		998
4		1.00	(7)		11704
5		2.00	(7)		15962
6	$q_s = 0$	0.01	λ_0		775
7		1.00	λ_0		11333
8		0.01	(7)		801
9		1.00	(7)		11704
10		1.00	(7)	100	228
Case 11: $q_V = 0$, time sample is 0.0975 μs, $I_0(t)$ = 1.2 TW/m^2, $c = c_{0L}$, $\lambda = \lambda_{0L}$, Δx = 2 nm, N_d = 3355					

Note. $I_0(t)$= 0.3 TW/m^2, $c = c_0$ for cases 1–10.

Notations: q_V is the bulk absorption of laser radiation; q_s is the energy flux through the target surface; $I_0(t)$ is the density of the radiant flux penetrating the material; c is the heat capacity, c_0 = 263 J/(kg·K), c_{0L} = 449.9 J/(kg·K); λ is the coefficient of thermal conductivity, λ_0 = 51.49 W/(m·K), λ_{0L} = 65 W/(m·K), (7) is the number of the polynomial formula; N_d is the number of degrees of freedom.

Method of generative design

The first stage of the GMD algorithm consists of determining the total number of possible elements in the reconstructed equation (depending on the type of problem). The full stencil of the PDE for the problem considered includes a convective term and is written as follows:

$$-\frac{\partial T}{\partial t}+\frac{1}{c\rho}\frac{\partial}{\partial x}\lambda\frac{\partial T}{\partial x}+\omega\frac{\partial T}{\partial x}+\frac{q_v}{c\rho}=0, \tag{10}$$

where the coefficients λ and ω are assumed to be unknown, while the coefficients ρ, s and the heat source power q_v are known.

The temperature dependence of thermal conductivity λ is assumed to be known:

$$\lambda=\beta_0+\beta_1 T+\beta_2 T^2+\beta_3 T^3, \tag{11}$$

where β_p are the unknown coefficients.

The second stage of the GMD algorithm involves discretization of the equation elements by finite difference (FD) [7] or finite element (FE) methods [17] and calculating the values of discretized stencils based on the available data. We applied an FD method in this study.

The difference template of the second derivative in space for the regular mesh takes the form

$$\frac{\partial}{\partial x}\left(\lambda\frac{\partial T}{\partial x}\right)\approx\frac{1}{\Delta x^2}\left[\lambda_{j+1/2}T_{j+1}-(\lambda_{j+1/2}+\lambda_{j-1/2})T_j+\lambda_{j-1/2}T_j\right]. \tag{12}$$

In view of dependence (11), expression (12) takes the form

$$\frac{\partial}{\partial x}\left(\lambda\frac{\partial T}{\partial x}\right)\approx\sum_{s=0}^{3}K_s^{i,l}\beta_s, \tag{13}$$

$$K_s^{i,l}=0.5K_s^{i,n}+0.5K_s^{i,n+1}, \tag{14}$$

$$K_s^{i,n}=\frac{1}{\Delta x^2}\left[\left(\frac{T_{j+1}^n+T_j^n}{2}\right)^s(T_{j+1}^n-T_j^n)+\left(\frac{T_{j-1}^n+T_j^n}{2}\right)^s(T_{j-1}^n-T_j^n)\right]. \tag{15}$$

The superscript index i in Eqs. (13)–(15) and below corresponds to a spatial slice comprising three nodes: $j-1$, j and $j+1$; the superscript l corresponds to a time slice comprising two time layers: n and $n+1$. The coefficients $K_s^{i,n+1}$ are determined by Eqs. (15) substituting the index n by the index $n+1$.

The discretized equation (10) for the spatial slice i and the time slice l takes the form

$$\sum_{p=1}^{P_t}a_p^{il}\alpha_p=0, \tag{16}$$

where $P_t=7$ for near-surface nodes during the pulse with volume absorption, and $P_t=6$ in other cases; $\alpha_1=-1$, $\alpha_p=\beta_{p-2}/(c\rho)$ for $2\le p\le 5$; $\alpha_6=\omega$, $\alpha_7=q_V/(c\rho)$. Parameters α_2–α_6 are unknown and should be determined.

The coefficients a_p^{il} can be found from the available synthetic (or experimental) data:

$$a_1^{il}=\frac{T_j^{n+1}-T_j^n}{\Delta t}, \tag{17}$$

$$a_s^{il}=K_{s-2}^{i,l}\ \ for\ 2\le s\le 5, \tag{18}$$

$$a_6^{il}=0.5\frac{T_{j+1}^{n+1}-T_{j-1}^{n+1}}{2\Delta x}+0.5\frac{T_{j+1}^n-T_{j-1}^n}{2\Delta x}, \tag{19}$$

$$a_7^{il}=1. \tag{20}$$

In vector form, expression (16) can be represented as

$$\sum_{p=1}^{P_t}\mathbf{V}_p\alpha_p=\mathbf{Z}_0, \tag{21}$$

where $\mathbf{Z}_0$ is a zero vector; the vector $\mathbf{V}_p$ consists of elements a_p^{il}, where the superscript i varies from 2 to $N-1$ (the spatial index j varies from 1 to N (N is the number of spatial nodes)); the superscript l varies from 1 to L (L is the number of time slices) $L<n$ (n is the number of time steps):

$$\begin{gathered}\mathbf{V}_p^T=(v^{p1}\ v^{p2}\ v^{p3}\ \ldots v^{p\ (N-2)\times L})=\\=(a_p^{21}\ a_p^{31}\ a_p^{41}\ \ldots a_p^{N-1\,1}\ a_p^{22}\ a_p^{32}\ a_p^{42}\ \ldots a_p^{N-1\,2}\ \ldots a_p^{2L}\ a_p^{3L}\ a_p^{4L}\ \ldots a_p^{N-1\ L}).\end{gathered} \tag{22}$$

Expression (21) can be represented as

$$\mathbf{Y} = \alpha_0 + \sum_{p=2}^{P_t} \mathbf{V}_p \alpha_p, \tag{23}$$

where $\mathbf{Y} = -\alpha_1 \mathbf{V_1}$, $\alpha_1 = -1$, $\alpha_0 = 0$.

Estimates of coefficients α_p can be obtained by the ordinary least squares (OLS) method.

The presence of the initial data allows us to calculate the component values of the vectors $\mathbf{V}_p$.

The third stage of the GMD algorithm involves applying statistical learning methods to expression (23). These methods rely on algorithms for selecting the optimal subset of elements and statistical criteria for selecting a single 'correct' combination of elements [3, 18]. The Bayesian information criterion (BIC) or C_p (Mallow's criterion) can be used for this purpose [3, 19, 20]. The respective criteria are calculated as

$$\mathrm{BIC} = n \ln \frac{\mathrm{RSS}}{n} + k \ln n, \tag{24}$$

where $n = (N - 2)\ L$ is the number of observations; RSS is the residual square sum; $k = p_e + 2$, p_e – is the number of elements included in the sum in the right-hand side of expression (23) (the maximum possible number of elements: $P = P_t - 1$),

$$\tilde{N}_p = \frac{\mathrm{RSS}}{S^2} - n + 2(p_e + 1), \tag{25}$$

where S^2 is mean squared residual after regression over the entire set of predictors.

Illustration of GMD method

In general, the reconstruction accuracy for a model in the form of a PDE with derivative coefficients depending on the solution is determined by a quantitative parameter that is the number of degrees of freedom (DoF), as well as by the quality of the data. The number of DoF in the given problem is the number of nodes (points) with a known temperature value. This number depends on the grid spacing and the instant in time that the data correspond to. Heat penetrates deep into the material over time, with an increase in the heated area and, accordingly, the number of nodes for the case of a uniform grid. Data quality refers to several factors: the temperature range covered by the data; possible correlation of the data, as well as the presence and type of energy supply.

If the heat flux through the surface is given as condition (2) and expression (3), the thermal balance of a unit volume associated with the grid node j is determined only by the process of thermal conductivity. It is expected that data in all nodes of the area under consideration can be used to generate a model.

In the case of volume absorption, this process occurs in the near-surface layer of the target with a characteristic scale of about $\delta \approx 1/\alpha_a = 0.2 \cdot 10^{-7}$ m (20 nm). The contribution of the term q_V to equation (1) for near-surface nodes is significant with respect to the thermal process. The quantity Os_L, which is the ratio of the energy supplied to the unit volume corresponding to the node j with the boundaries

$$x_j - \Delta x / 2 \le x < x_j + \Delta x / 2$$

for absorption of radiation to the energy supplied/discharged from the unit volume due to thermal conductivity, can be defined as

$$\mathrm{Os}_L = \left| \frac{q_V(x_j, t)\Delta x}{q_{-\Delta x/2} - q_{\Delta x/2}} \right|, \tag{26}$$

$$q_{-\Delta x/2} = -\lambda \frac{\partial T}{\partial x}\bigg|_{x_j - \Delta x/2}, \quad q_{\Delta x/2} = -\lambda \frac{\partial T}{\partial x}\bigg|_{x_j + \Delta x/2}.$$

The boundary of 'influence' of energy absorption on the heat balance depends on the distribution of the modulus of the local Ostrogradsky number $|\mathrm{Os}_L|$. Fig. 4 shows an example for the variation in $|\mathrm{Os}_L|$ for the case of volume absorption and the time $t = 0.1$ μs. $|\mathrm{Os}_L| >> 1$ for points in the near-surface layer around the coordinate $x = 7 \cdot 10^{-8}$ and the thermal conductivity process also plays a secondary role. The latter circumstance affects the reconstruction accuracy for the thermophysical parameters and the quality of generation for the equation model.

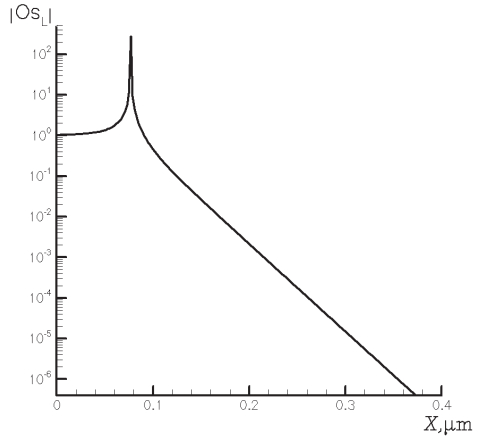

Fig. 4. Modulus distribution of the local Ostrogradsky number along the coordinate X

A total of 200 near-surface nodes were excluded from the data intended for generative design for the irradiation stage and the case of volume absorption, while the sum in expression (23) included the total number of predictors $P = P_t - 1 = 5$ ($P_t = 6$). The reconstruction error of the coefficients exceeded 50% accounting for either the near-surface nodes in the initial data or only the near-surface nodes for which expression (23) also included the term associated with volume energy release ($P = 6$, $P_t = 7$), while the reconstructed model included an 'extra' convective term for the first range of temperatures below the melting point T_L. One near-boundary node was excluded from the data for the heat diffusion stage after the pulse stopped to account for the specifics of how finite-difference stencils are set.

Cases with the energy flux through the surface were not sensitive to the presence of information about near-surface nodes in the data under consideration. If the surface energy flux was set regardless of the process stage, one node immediately adjacent to the target surface was excluded from the data.

The target temperature varies in a wide range during heating: $T_0 = 300 \le \mathrm{T} < 2200$ K for the first temperature range and $T_0 = 300 \le \mathrm{T} < 7000$ K for the second one. Internal tests indicate that it is ineffective to use data for points deep within the target with a temperature below $1.003 T_0 = 301$ K. These nodes were also excluded from consideration.

Data generated for different time slices correlate with each other [3]. The model of the equation is reconstructed from the data corresponding to a single time slice consisting of two time layers. We considered the data corresponding to the time slices at both the irradiation stage ($t = 10$ and 100 ns) and after the laser pulse stopped (1 and 2 μs).

The number of DoF used to generate the model ranged from 220,000 to 16,000 and depended on the case (see Table 1).

In the first stage of the study, the GMD method was applied to the data corresponding to Cases 1, 2, 6 and 7, assuming constant thermal conductivity λ_0 and a temperature range not exceeding the melting point (see Table 1). The results of the algorithm for selecting the optimal subset of elements for Case 1 are summarized in Table 2. The number in the first column corresponds to the number of elements p_e included under the summation sign in Eq. (23). According to the results obtained, the minimum values of the BIC and C_p criteria correspond to an equation with one term with the coefficient α_2 (aside from the term with the coefficient $\alpha_1 = -1$) for Case 1. The statistical procedure of the package *R* [18] for all cases determines the value of the coefficient α_2 as equal to $2.2845 \cdot 10^{-5}$ (Table 3), which corresponds to the value of thermal conductivity

$$\alpha_2 = \lambda_0 / (\rho c_0) = 2.2845 \cdot 10^{-5} \ \mathrm{m^2/s}$$

(see expression (7)), which is used to generate data.

Notably, in addition to exact reconstruction of the coefficient, the structure of the equation, which should not include the convective term with the coefficient α_6, is also reconstructed correctly. Thus, the results obtained for generating an equation with constant coefficients can be assumed to be satisfactory.

Similarly, the GMD method was applied to reconstructing the equation from the data implying a temperature dependence for the thermal conductivity of the parameters in the form (7). Table 2 presents an example of applying the procedure for selecting the optimal subset of elements for Case 3 (surface energy flux and the time t = 10 ns) corresponding to the irradiation mode. The procedure correctly reproduces the stencil for the equation without the convective term corresponding to the coefficient α_6 for this case. Cases 4 and 5 (see Table 1) correspond to later times of the process: t = 1 and 2 μs, respectively. The temperature of the target decreases, and its variation range is 300–550 K by the time of 2 μs. The number of DoF increases to 16,000. The reconstruction quality of the model structure remains virtually unchanged for later times, and the convective term is not reproduced. The structure of the required equation is also reconstructed correctly for the cases with volume absorption (Cases 8 and 9). Varying the computational parameters, i.e., increasing the spatial step by 50 times and decreasing the number of DoF in Case 10 proportionally (relative to Case 9) does not affect the reconstruction quality of the structure.

The results obtained by running the GMD algorithm for Cases 1 – 10 are given in Table 3. The normalized coefficients corresponding to expression (7) (in bold) and the coefficient α_6 associated with the surface evaporation rate are given here. The rest of the rows show the values of the reconstructed coefficients for the given cases, the values of the normalized coefficients in the dependence $\lambda(T)$ are reproduced with good accuracy. The reconstruction error of the coefficients only weakly depends on the time slice corresponding to the data and on the type of energy input, provided that some of the data for the near-surface layer are excluded for the case of volume energy absorption.

The maximum reconstruction error for the coefficients of polynomial dependence of thermal conductivity is 0.01%. The error is determined as

$$\varepsilon = \left| \alpha_m / \alpha_t - 1 \right| \cdot 100\%,$$

where α_m is the coefficient for which the maximum discrepancy with the theoretical value α_t is observed.

The error for the total thermal conductivity is even less and does not exceed 0.002%. The error is the same for Case 9 and 10, even though the number of DoF differs by 50 times.

The initial data for Case 11 assumes an evaporation process on the target surface, since the surface temperature for the corresponding high radiation intensity (see Table 1) and the pulse time exceeds the boiling point of niobium. The GMD method allows to correctly reconstruct the structure of the model, which includes a convective term for this case (see Table 3). The reconstructed coefficient value α_6 for Case 11 is different from 0.

Table 2 shows the values of criteria C_p and BIC for two cases: 1 and 3. The minimum criteria correspond to the same set of elements for these cases. However, unlike BIC, the criterion C_p predicts an incorrect structure of the model for several cases. Moreover, using BIC produces a sparser model with fewer elements.

We can therefore conclude that the BIC criterion is preferable for the given class of problems.

It is extremely important for real-life applications that the GMD method can work with noisy data. Below, we consider an additional computational case to illustrate this capability, corresponding to the time slice t = 0.0975 ms; its initial parameters are similar to Cases 1 and 2 from Table 1. The same as above (Cases 1 and 2), the synthetic data are generated on a grid with a step Δx = 2 nm. The effect of noise was achieved by simulating the temperature values by the law

$$T_\delta(x,t) = T(x,t)\left(1 + \theta\delta_r\right), \tag{27}$$

where $T(x,t)$ is the temperature found from numerical solution (1); θ is a random variable with a uniform distribution from the interval [–1, 1], δ_r is the relative error.

Table 2

Procedure for selecting the optimal element subset for Cases 1 and 3

Number of elements p_e	α_2	α_3	α_4	α_5	α_6	BIC	C_p
Case 1							
1	*	–	–	–	–	**–29,196.57**	**2.59**
2	*	–	–	–	*	– 29,190.57	3.71
3	*	*	–	–	*	– 29,183.85	5.54
4	*	*	*	*	–	– 29,178.95	5.56
5	*	*	*	*	*	– 29,173.63	6.00
Case 3							
1	*	–	–	–	–	– 3,628.85	$2.44 \cdot 10^{14}$
2	*	–	–	–	*	– 1,098.50	$1.60 \cdot 10^{11}$
3	*	*	–	–	*	– 12,940.78	$2.26 \cdot 10^{10}$
4	*	*	*	*	–	– **29,934.64**	**4.53**
5	*	*	*	*	*	– 29,928.26	6.00

Notations: α_2–α_6 are the coefficients from Eq. (23); BIC, C_p are statistical criteria. Notes. 1. The number of elements p_e included under the summation sign in Eq. (23) is given in the left column. 2. The selected results are highlighted in bold.

Table 3

Illustration of GMD method

Case	α_0	α_2 ($\times 10^5$)	α_3 ($\times 10^9$)	α_4 ($\times 10^{12}$)	α_5 ($\times 10^{15}$)	α_6
Comparison of theoretical and reconstructed values						
1, 2, 6, 7	**0**	**2.2845**	**0**			**0**
1	$-7.1669 \cdot 10^{-5}$	2.2845	0			0
2	$1.0482 \cdot 10^{-7}$		0			
6	$-3.6375 \cdot 10^{-5}$		0			
7	$-8.8505 \cdot 10^{-8}$		0			
Comparison of theoretical and reconstructed values						
3–5, 8–10	**0**	**2.2845**	**2.9668**	**2.9513**	**–1.0009**	**0**
3	$2.2415 \cdot 10^{-5}$	2.2845	2.9669	2.9513	–1.0009	0
4	$1.2132 \cdot 10^{-7}$		2.9668			
5	$-3.1960 \cdot 10^{-8}$			2.9514		
8	$8.7904 \cdot 10^{-6}$		2.9669	2.9511	–1.0008	
9	$4.1593 \cdot 10^{-8}$		2.9668	2.9513	–1.0010	
10	$-3.8343 \cdot 10^{-7}$			2.9515		
Comparison of theoretical and reconstructed values						
11	**0**	**1.9060**	**0**			–
11	$1.4656 \cdot 10^{-5}$	1.9060	0			0,3

Notes. 1. The values of normalized coefficients corresponding to expression (7) (theoretical) are highlighted in bold. 2. For all cases, $\alpha_1 = -1$.

Noisy data were generated with the value of δ_r varying from 10^{-5} to 10^{-2} (see Table 4).

The GMD method does not allow to correctly reconstruct both the structure of the PDE and the coefficients for the derivatives for the given noise level. For this reason, an additional regularization procedure had to be applied for noisy data, accounting for the function assignment error generated when the grid step is selected [21].

As noted above, synthetic data are generated with a small step with respect to the spatial coordinate, allowing to tailor a larger step, a multiple of Δx. Using the procedure for determining the step (described in [21]) for the relative errors of the initial data given in Table 4, we obtained the value of $\Delta x_{reg} = 0.1$ μm. With the given step value, the GMD correctly reproduces the structure of the equation without the convective term ($\alpha_6 = 0$) for all values of δ_r considered, and the reconstruction error of the thermal conductivity coefficient

$$\varepsilon = \left|\alpha_2 / \alpha_t - 1\right| \cdot 100\%$$

($\alpha_t = 2.2845 \cdot 10^{-5}$ m^2/s is the theoretical value of the coefficient) varies from 0.04% for the relative error of the initial data amounting to 10^{-5} to 47% for $\delta_r = 10^{-2}$ (see Table 4). The relative error $\delta_r = 10^{-2}$ corresponds to the absolute error of determining the temperature determination at about 25 K near the target surface and 3 K near the remote boundary.

If the step with respect to the spatial coordinate cannot be varied for synthetic or experimental data, an alternative regularization procedure can be used, based on temperature interpolation with spline functions [21].

Table 4

Application of GMD method to noisy data

Case	δ_r	α_2	ε, %
s1	10^{-5}	$2.2854 \cdot 10^{-5}$	0.04
s2	10^{-4}	$2.2912 \cdot 10^{-5}$	0.29
s3	10^{-3}	$2.5112 \cdot 10^{-5}$	9.92
s4	10^{-2}	$3.3634 \cdot 10^{-5}$	47.22

Note. $\alpha_6 = 0$ for cases s1–s4.

Notations: δ_r is the variable relative error; α_2, α_6 are the coefficients from Eq. (23); ε is the error, $\varepsilon = |\alpha_2/\alpha_t - 1| \cdot 100\%$ (α_t is the theoretical value).

Conclusions

The GMD method holds much promise for reconstructing the PDE describing thermal processes. GMD can be used both to directly construct a mathematical model of a complex phenomenon from the available data and to visualize the accompanying processes, such as chemical reactions or phase transformations, as well as to refine the thermophysical parameters of materials. However, little experience has been accumulated so far with applying the GMD in practice, so the method itself needs to be developed further.

This paper continues the efforts on optimizing the algorithm for generative model design as applied to reconstructing a model of a thermal process, which may generally include a convective term, with the material parameters assumed to depend on temperature.

We report on the reconstruction of a partial differential equation describing heating and evaporation of a metal target by a laser pulse.

The initial data for subsequent application of the GMD method were generated by numerically solving the thermal conductivity equation an unsteady process for different approaches to describing the laser energy. We propose a method for improving the computational accuracy for the case of volume absorption and limited computational resources, which require a large step with respect to the spatial coordinate. The computational algorithm is verified by comparison with existing analytical solutions.

Our findings indicate that the GMD method is sensitive to the type of data used. The proportion of data that can be used to generate the model has to be additionally estimated for the case where the initial data for reconstructing the model correspond to volume absorption of the material and the stage when the target is irradiated. Excluding the data on temperature in near-surface nodes allows to reconstruct the model structure with better quality and minimum error in finding the temperature-dependent coefficients of the PDE generated. On the other hand, there is no problem with excluding near-surface nodes and pre-processing of the data if the data corresponds to the stage of heat diffusion after the laser pulse stops for volume absorption or any stage of the process with the heat flux through the target surface.

If the model is generated as a PDE with variable coefficients, a significantly larger number of predictors have to be taken into account to apply statistical learning methods in the search for the optimal structure and coefficients of the equation. Despite this circumstance, we have confirmed that applying generative model design to the available data yields good results in reconstructing the structure of the model. The structure of the equation does not include a convective term at temperatures substantially lower than the boiling point of the material. If the surface temperature of the target exceeds the boiling point, the convective term associated with the surface evaporation process is reconstructed. The reconstruction error for the temperature-dependent coefficient of thermal conductivity is less than 0.002% for the given number of degrees of freedom (more than 200).

We have established that the GMD method can be used to reconstruct a model from noisy data. In this case, additional regularization procedures are to be introduced to obtain the correct structure of the equation and the coefficient values of the derivatives.

As a direction for future research, we intend to validate the method of generative design with real experimental data.

REFERENCES

1. **Hutter F., Kotthoff L., Vanschoren J.** (Eds.), Automated machine learning. Methods, systems, challenges, (The Springer Series on Challenges in Machine Learning), Springer Nature, Cham, Switzerland, 2019.

2. **Maslyaev M., Hvatov A., Kalyuzhnaya A. V.,** Partial differential equations discovery with EPDE framework: Application for real and synthetic data, J. Comp. Sci. 53 (July) (2021) 101345.

3. **James G., Witten D., Hastie T., Tibshirani R.,** An introduction to statistical learning: with applications in R, Springer, New York, 2013.

4. **Lykov A. V.**, **Mikhailov Y. A.,** Theory of heat and mass transfer, Institute for the Promotion of Teaching Science and Technology (IPST), Thailand, 1963.

5. **Samarskii A. A., Vabishchevich P. N.,** Computational heat transfer, in 2 Vols., Willey, Chichester, 1995.

6. **Manukhin B. G., Kucher D. A., Chivilikhin S. A., et al.,** Optical diagnostics of the process of free liquid convection, Optics and Spectroscopy. 119 (3) (2015) 392–397.

7. **Bykov N. Yu., Khvatov A. A., Kalyuzhnaya A. V., Bukhanovskiy A. V.,** A method for reconstructing models of heat and mass transfer from the spatio-temporal distribution of parameters, Technical Physics Letters. 47 (24) (2021) 9–12 (in Russian).

8. **Bulgakov A. V., Bulgakova N. M., Burakov I. M., Bykov N. Y., et al.,** Nanosized material synthesis by action of high-power energy fluxes on matter, Institute of Thermophysics SB RAS, Novosibirsk, 2009 (in Russian).

9. **Bykov N. Y., Bulgakova N. M., Bulgakov A. V., Loukianov G. A.,** Pulsed laser ablation of metals in vacuum: DSMC study versus experiment, Appl. Phys. A. 79 (4–6) (2004) 1097–1100.

10. **Bykov N. Y., Lukyanov G. A.,** The direct simulation Monte Carlo of cluster formation processes in laser plume, Proc. 25-th Int. Symp. Rarefied Gas Dynamics, Publishing House of SB RAS, Novosibirsk (2007) 645–650.

11. Niobiy [Niobium], In Book: Khimicheskaya entsiklopediya, T. 3. Pod red. N. S. Zefirova i dr. [Encyclopedia of chemistry, Vol. 3, Edited by N. S. Zefirov et al.], Soviet Encyclopedia Publishing, Moscow, 1992. P. 249 (in Russian).

12. **Zolotarev V. M., Morozov V. N., Smirnova E. V.,** Opticheskiye postoyannyye prirodnykh i tekhnicheskikh sred [Handbook of optical constants of natural and technical media], Khimiya Publishing, Leningrad, 1984 (in Russian).

13. **Grigoriev I.S., Melikhov E. Z.** (Eds.), Handbook of physical quantities, CRC Press, Boca Raton, Florida, USA, 1997.

14. **Zinoviev V. E.,** Teplofizicheskiye svoystva metallov pri vysokikh temperaturakh [Thermal properties of metals at high temperatures], Handbook, Metallurgia, Moscow, 1989 (in Russian).

15. **Anisimov S. I., Imas Ya. I., Romanov G. S., Khodyko Yu. V.,** Deystviye izlucheniya bolshoy moshchnosti na metally [Effect of high-power radiation on metals], Nauka Publishing, Moscow, 1970 (in Russian).

16. **Carslaw H. S., Jaeger J. C.,** Conduction of heat in solids, 2-nd ed., Oxford University Press, Oxford, UK, 1959.

17. **Bykov N., Hvatov A., Kalyuzhnaya A., Boukhanovsky A.,** A method of generative model design based on irregular data in application to heat transfer problems, J. Phys. Conf. Ser. 1959 (2021) 012012.

18. R Core Team. R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, 2020. Available online at https://www.R-project.org/.

19. **Priestley M. B.,** Spectral analysis and time series (Probability and Mathematical Statistics), Academic Press, Cambridge, UK (1981).

20. **Mallows C. L.,** Some comments on C_p, Technometrics. 15 (4) (1973) 661–675.

21. **Vatulyan A. O.,** Koeffitsiyentnyye obratnyye zadachi mekhaniki [Inverse coefficient problems in mechanics], Fizmatlit, Moscow, 2019 (in Russian).

СПИСОК ЛИТЕРАТУРЫ

1. **Hutter F., Kotthoff L., Vanschoren J.** (Eds.) Automated machine learning. Methods, systems, challenges. (The Springer Series on Challenges in Machine Learning). Cham, Switzerland: Springer Nature, 2019. 220 p.

2. **Maslyaev M., Hvatov A., Kalyuzhnaya A. V.** Partial differential equations discovery with EPDE framework: Application for real and synthetic data // Journal of Computer Science. 2021. Vol. 53. July. P. 101345.

3. **Гарет Д., Уиттон Д., Хасти Т., Тибширани Р.** Введение в статистическое обучение с примерами на языке R. Пер. с англ. С. Э. Мастицкого. Изд. 2-е, испр. М.: ДМК Пресс, 2017. 456 с.

4. **Лыков А. В. Теория теплопроводности.** М.: Высшая школа, 1967. 600 с.

5. **Самарский А. А., Вабищевич П. Н.** Вычислительная теплопередача. М.: Едиториал УРСС, 2003. 784 с.

6. **Манухин Б. Г., Гусев М. Е., Кучер Д. А., Чивилихин С. А., Андреева О. В.** Оптическая диагностика процесса свободной конвекции жидкости // Оптика и спектроскопия. 2015. Т. 3 № .119. С. 423–418.

7. **Быков Н. Ю., Хватов А. А., Калюжная А. В., Бухановский А. В.** Метод восстановления моделей тепломассопереноса по пространственно-временным распределениям параметров // Письма в Журнал технической физики. 2021. Т. 47. № 24. С. 9–12.

8. **Булгаков А. В., Булгакова Н. М., Бураков И. М., Быков Н. Ю. и др.** Синтез наноразмерных материалов при воздействии мощных потоков энергии на вещество. Новосибирск: Институт теплофизики СО РАН, 2009. 462 с.

9. **Bykov N. Y., Bulgakova N. M., Bulgakov A. V., Loukianov G. A.** Pulsed laser ablation of metals in vacuum: DSMC study versus experiment // Applied Physics A. 2004. Vol. 79. No. 4–6. Pp. 1097–1100.

10. **Bykov N. Y., Lukyanov G. A.** The direct simulation Monte Carlo of cluster formation processes in laser plume // Proceedings of the 25-th International Symposium on Rarefied Gas Dynamics. Novosibirsk: Publishing House of SB RAS, 2007. Pp. 645–650.

11. Ниобий // Химическая энциклопедия. В 5 тт. (1998–1988). Т. 3. Мед–Пол. Под ред. Н. С. Зефирова, И. Л. Кнунянца, Н. Н. Кулова. М.: Советская энциклопедия, 1992. С. 249.

12. **Золотарев В. М., Морозов В. Н., Смирнова Е. В.** Оптические постоянные природных и технических сред. Ленинград: Химия, 1984. 216 с.

13. Физические величины. Справочник. Под ред. И. С. Григорьева и Е. З. Мелихова. М.: Энергоатомиздат, 1991. 1230 с.

14. Зиновьев В. Е. Теплофизические свойства металлов при высоких температурах. М.: Металлургия, 384 .1989 с.

15. Анисимов С. И., Имас **Я. И., Романов Г. С.,** Ходыко **Ю. В.** Действие излучения большой мощности на металлы. М.: Наука, 272 .1970 с.

16. **Карслоу Г., Егер Д.** Теплопроводность твердых тел. Пер. с англ. М.: Наука, 1964. 488 с.

17. **Bykov N., Hvatov A., Kalyuzhnaya A., Boukhanovsky A.** A method of generative model design based on irregular data in application to heat transfer problems // Journal of Physics: Conference Series. 2021. Vol. 1959. P. 012012.

18. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2020. Available online at https://www.R-project.org/.

19. **Priestley M. B.** Spectral analysis and time series (Probability and Mathematical Statistics). Cambridge, UK: Academic Press, 1981. 706 p.

20. **Mallows C. L.** Some comments on C_p // Technometrics. 1973. Vol. 15. No. 4. Pp. 661–675.

21. **Ватульян А. О.** Коэффициентные обратные задачи механики. М.: Физматлит, 2019. 272 с.

THE AUTHOR

BYKOV Nikolay Yu.
ITMO University, St. Petersburg,
Peter the Great St. Petersburg Polytechnic University
49A, Kronverksky Ave., St. Petersburg, 197101, Russia
nbykov2006@yandex.ru
ORCID: 0000-0003-0041-9971

СВЕДЕНИЯ ОБ АВТОРЕ

БЫКОВ Николай Юрьевич – *доктор физико-математических наук, старший научный сотрудник факультета цифровых трансформаций Научно-исследовательского университета ИТМО; ведущий научный сотрудник, профессор кафедры физики Санкт-Петербургского политехнического университета Петра Великого.*
195251, Россия, г. Санкт-Петербург, Кронверкский пр., 49А
nbykov2006@yandex.ru
ORCID: 0000-0003-0041-9971

Received 08.04.2022. Approved after reviewing 06.06.2022. Accepted 06.06.2022.
Статья поступила в редакцию 08.04.2022. Одобрена после рецензирования 06.06.2022. Принята 06.06.2022.