Technical Physics Letters, 2023, Vol. 49, No. 8

03.2;01.1

Discovery of the thermal process model from noisy data

© T.A. Andreeva [1,2], N.Y. Bykov [1,2], A.K. Klimova [2], A.Ya. Lukin [1]

[1] Peter the Great Saint-Petersburg Polytechnic University, St. Petersburg, Russia
[2] ITMO University, St. Petersburg, Russia
E-mail: nbykov2006@yandex.ru

Received April 11, 2023
Revised May 25, 2023
Accepted May 25, 2023

A modification of the algorithm of the model generative design in the form of a partial differential equation for working with noisy data is proposed. Using the algorithm, the model of the heat and mass transfer process was restored from synthetic and original experimental data on heating the medium by a flooded heat source. The thermophysical parameters of the medium are determined, the possibility of using the algorithm to indicate the convection process based on data on the space-time distribution of temperature is shown.

Keywords: generative design method, data-driven model, thermal conductivity equation, convection.

DOI: 10.61011/TPL.2023.08.56679.19588

Algorithms for solving inverse problems (IP) developed by this time within the heat-exchange theory allow restoring from available data the medium thermophysical parameters (coefficient IPs), boundary conditions (boundary IPs), spatial distributions of temperature at previous time moments (retrospective IPs), and also the form and parameters of the function describing the power of internal heat sources [1,2]. A separate trend of solving inverse problems is the methods for generative model design (GDM) [3]. The GDM methods allow restoring the mathematical model structure in the form of a differential equation describing the physical process. Surely, the GDM algorithm implies also determination of the derivative coefficients and, if necessary, restoring of additional terms of the differential equation. In the context of thermal problems, the classical heat transfer equation has a well-known structure. However, in general the heat transfer equation may comprise the second time derivative as well as convective terms accounting for the heat transfer by moving medium [1,4,5]. The GDM methods are of interest in view of developing artificial intelligence techniques [6] the objective of which can be restoration and analysis of the thermal process model from experimental data; the analysis implies, for instance, revealing hidden processes in the medium, such as phase transitions and chemical reactions, indicating change of heating modes, etc.

Papers [3,7] propose a GDM algorithm for restoring the heat process model, which is based on using the procedure of best subset selection [8]. In the case of noiseless synthetic data, this approach enabled precise restoration of the model of heating a metal with a laser pulse. During the model generation, thermophysical medium parameters were determined, including temperature-dependent ones [7]. However, the proposed approach was not verified for the case of random-error data.

The goal of this work was to develop a GDM algorithm for processing noisy data, as well as its verification based on synthetic and experimental data; for this purpose, an experiment on pulsed heating of a medium with a flooded heat source was carried out, and original data on the thermal process under consideration were obtained.

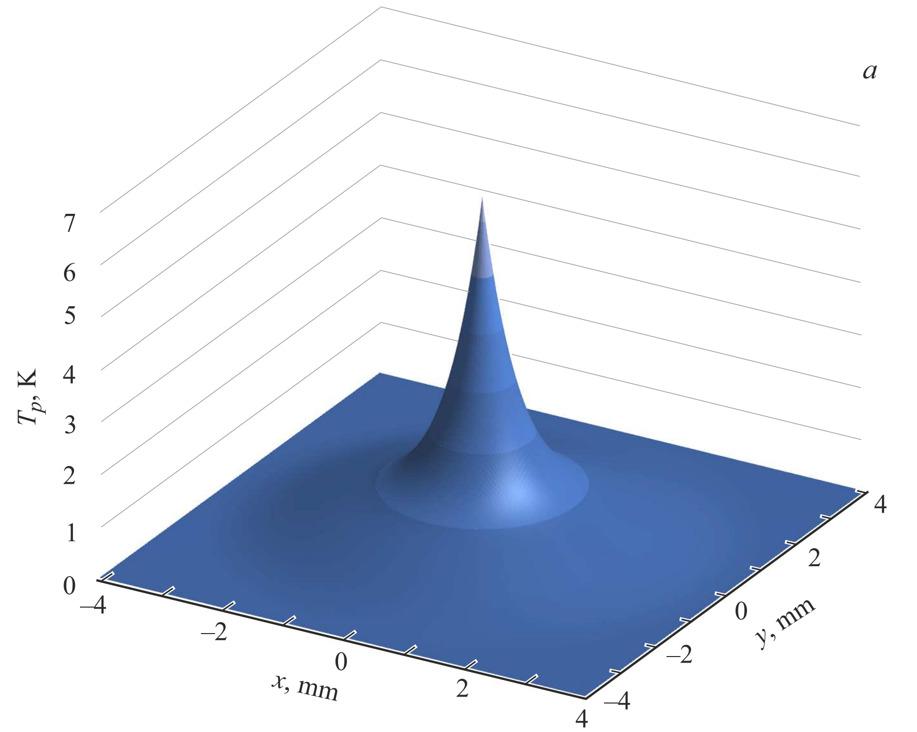
In this study, the object for the model restoration was an unsteady process of heating glycerin with a flooded constantan wire 0.1 mm in diameter which was heated by direct electric current. Experimental and synthetic data were obtained for a 30−s heating pulse; the distance to the wire centerline was varied from 0.6 to 3.1 mm with the step of 0.5 mm. The frequency of taking the thermocouple readings was 2 ms; the random error had the form of additive white Gaussian noise with the standard deviation of 0.025 K. The measurements were performed at two volumetric rates of heat release in wire: 0.38 and 1.83 W/mm^3.

The process of heating the medium is generally described by an equation of the following form [1]:

$$c\rho\left(\frac{\partial T}{\partial t}+\mathbf{v}\cdot\nabla T\right)=\nabla\cdot\left(\lambda\nabla T\right)+Q_V, \qquad (1)$$

here t is the time, T is the temperature, ρ, c, λ are the density, thermal capacity, and thermal conductivity coefficient, $\mathbf{v}$ is the medium speed, Q_V are the volumetric heat sources. In generating the synthetic data, convective motion of the medium was assumed to be absent ($\mathbf{v}=0$), and the problem was reduced to a one-dimensional one. In the case of glycerin, this assumption is valid at the initial stage of the process provided the medium overheating is insignificant [9]. In the temperature range under consideration, the thermal conductivity coefficient changes only slightly (within 1.5%). Thermophysical parameters of constantan and glycerin hbox[10–12] were assumed to be constant.

The calculations were performed by the finite-difference technique using the Crank-Nicolson scheme with the time step of $\tau=2$ ms. The medium overheating $T_p=T-T_0$



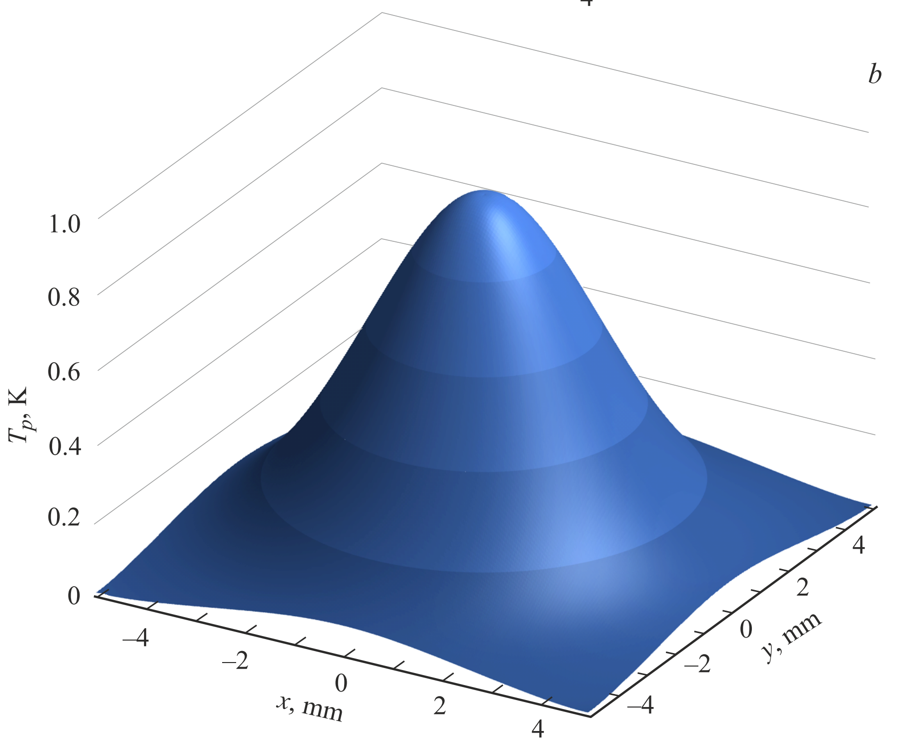


Figure 1. Temperature distributions at the end of heating (a) and 15 s after that (b). $Q_V = 0.38\,\mathrm{W/mm^3}$. The wire is arranged along the z axis.

(T_0 is the initial medium temperature) for the heat source power of $Q_V = 0.38\,\mathrm{W/mm^3}$ is illustrated in Fig. 1. In preparing the synthetic data for processing with the GDM algorithm, the calculation results were supplemented by the Gaussian noise with the standard deviation of $\sigma = 0.025\,\mathrm{K}$ and value selection range $(-3\sigma;\ +3\sigma)$. Notice that the noise level is rather high: about 1% of the maximum overheating for the location at 0.6 mm and 7% for the location at 3.1 mm.

In the experimental setup, the constantan wire 0.1 mm in diameter and 46 mm in length was placed in a cuvette filled with glycerin. The temperature was measured along

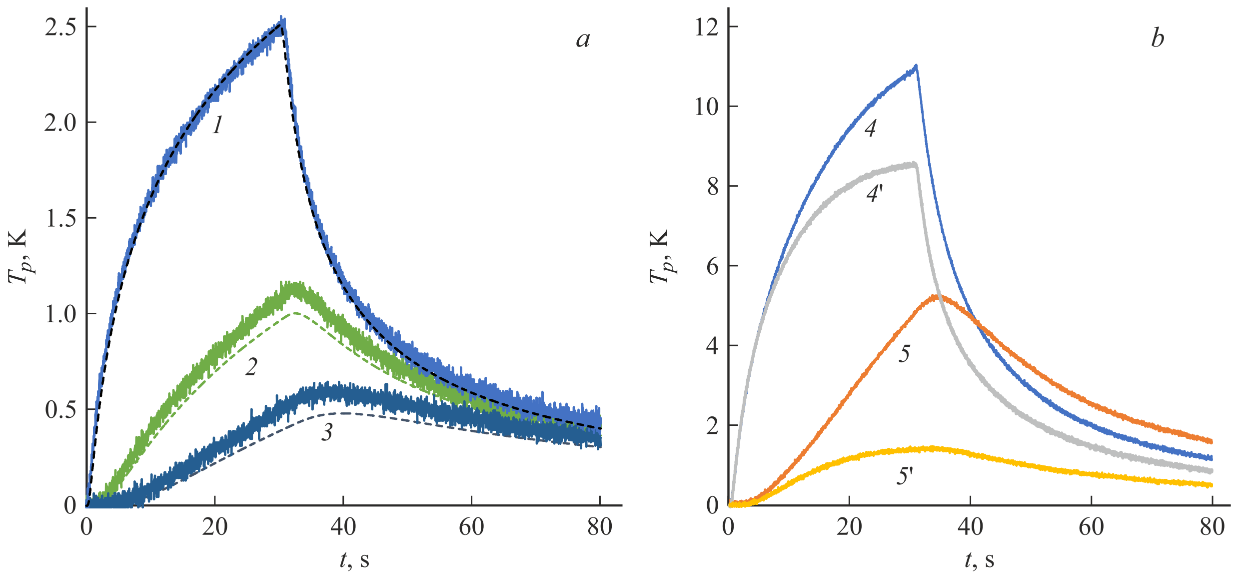


Figure 2. Experimental time dependences of temperature at points $r = 0.6$ (*1*), 1.6 (*2*), 2.6 mm (*3*) above the wire for $Q_V = 0.38\,\mathrm{W/mm^3}$ (*a*) and $r = 0.6$ (*4*, *4'*) and 2.6 mm (*5*, *5'*) for $Q_V = 1.83\,\mathrm{W/mm^3}$ above (*4, 5*) and below (*4'*, *5'*) the wire (*b*). Dashed lines represent the calculations obtaind ignoring the convection, $t = 0$ is the beginning of heating.

the vertical line passing through the wire center by using a thermocouple mounted on the positionerś support. The thermocouple signal was read using a special analog-digital converter ADS1220 produced by Texas Instruments and transmitted to the computer with the polling period of 2 ms.

The experimental data are presented in Fig. 2. The maximum overheating at the 0.6 mm point is about 2.5 K for 0.38 W/mm^3 and 11 K for 1.83 W/mm^3. In the first case, convection is weak, which confirms the conclusions made in [9]. In the second case, a significant difference between the medium temperatures measured below and above the wire is observed; this evidences for a developed convection.

The GDM algorithm [3,7] begins restoring an unknown equation structure based on available data from considering the complete possible template of the required equation. In the one-dimensional problem of heating a medium with constant thermophysical parameters, the energy equation template in the cylindrical frame of reference gets the following form [1]:

$$a_0\frac{\partial T}{\partial t} + a_1\frac{\partial^2 T}{\partial t^2} + a_2\frac{\partial^2 T}{\partial r^2} + a_3\frac{1}{r}\frac{\partial T}{\partial r} = 0. \qquad (2)$$

As compared with (1), this template contains the second time derivative of temperature which may appear during high-intensity unsteady processes. Coefficients a_i in the equation may be identified with the following terms of equation (1): $a_0 = -1$, $a_1 = -\tau_{rx}$ (τ_{rx} is the relaxation time), $a_2 = a_t = \lambda/(c\rho)$ (a_t is the thermal conductivity coefficient) in the absence of convection ($a_2 = a_t$). In the presence of convective motion of the medium, the problem cannot be anymore regarded as a one-dimensional one. Due to the symmetry along the vertical line for which the experimental data has been obtained, the convective term has only one radial component $a_3(r, v_r) = a_t - rv_r$ (v_r is the radial component of the medium speed). The transverse thermal conductivity exists, however, its possible contribution to equation (2) $\frac{a_t}{r^2}\frac{\partial^2 T}{\partial \varphi^2}$ (here φ is the azimuthal angle) is essentially lower than that of other terms. The considered template does not include the internal heat source since the data used in restoring does not include the wire location. The presence of noise in the data, substitution of derivatives in (2) with finite differences, as well as convection occurring in the real experiment, result in appearance in the equation (2) right-hand part of extra constant term α playing the role of an integral discrepancy.

Notice that the goal of this study is to restore the equation of the process from a minimal number of data. In the case of availability of complete data on the medium temperature spatial distribution $T(r, \varphi)$ and of the speed field for the case of prevailing convection, the proposed approach may be used to restore the complete set of the mediumś motion equations, including the momentum-conservation equation accounting for the presence of the gravity force.

The proposed approach implies that, after creating the possible potential equation template, vectors containing finite-difference templates of the equation (2) terms will be calculated based on the synthetic or experimental data. After that, a procedure of best subset selection (2) terms is applied, which results in eliminating negligible terms and determining the required coefficients. The optimal model will be selected by using the Bayesian information criterion (BIC) [8,13].

To regularize the algorithm, it is proposed to increase the numerical-differentiation time step. According to [2], the necessary value of the time step may be approximately estimated for the cases of first and second time derivatives

Results of the model restoration from noisy data

Version	τ_{reg}, s	$\alpha \cdot 10^3$, K/s	$a_2 \cdot 10^2$, mm^2/s	δa_2, %	$a_3 \cdot 10^2$, mm^2/s
Theory					
		0	9.263	0	9.263
Results of the model restoration from synthetic data					
1	0.002	−6.025	3.457	62.6	–
2	0.02	−6.568	3.584	61.3	–
3	0.2	−1.542	6.886	25.7	6.188
4	1	−1.467	7.057	23.8	6.446
5	2	−1.849	6.799	26.9	6.0523
6	5	−1.451	8.171	11.8	8.072
7	10	−5.373	6.880	25.7	–
Results of the model restoration from experimental data (0.38 W/mm^3)					
8	0.002	−6.283	4.547	50.9	–
9	0.02	−6.700	4.590	50.4	–
10	0.2	−2.095	7.131	23.0	4.608
11	1	−2.190	6.832	26.2	4.374
12	2	−2.312	7.000	24.4	4.502
13	5	−1.863	7.261	21.6	4.809
14	10	1.568	8.984	3.0	7.674
Results of the model restoration from experimental data (1.83 W/mm^3)					
15	0.002	−43.47	8.054	13.0	–
16	0.02	−42.89	8.005	13.6	–
17	0.2	−40.29	8.818	4.8	1.354
18	1	−39.90	8.813	4.9	1.425
19	2	−44.59	8.074	12.8	–
20	5	−44.14	7.951	14.2	–
21	10	−47.00	7.846	15.3	–

Note. The coefficient at the second time derivative is in all the cases $a_1 = 0$ and is not presented in the Table.

as

$$\tau_{reg,1} = \frac{2\sigma}{\varepsilon_1}, \quad \tau_{reg,2} = \left(\frac{4\sigma}{\varepsilon_2}\right)^{1/2}. \tag{3}$$

Here ε_1 and ε_2 are the absolute errors in the first and second time derivatives of temperature. It is assumed that the step should be larger than $\tau_{reg,1}$ and $\tau_{reg,2}$ but, at the same time, much less than the typical process duration. In addition, it is necessary to take into account the increase in the numerical differentiation error with increasing step.

Calculations of the first and second time derivatives for points 0.6 and 2.1 mm show that the maximum value of derivative $D_{1,T} = (dT/dt)_{\max}$ is 0.33 K/s. Assuming that $\varepsilon_1 = \delta_1 D_{1,T}$ and selecting relative error $\delta_1 = 0.1$, obtain from (3) $\tau_{reg,1} \approx 1.5$ s at $\sigma = 0.025$ K $\tau_{reg,1} \approx 1.5$ s. The maximum value of derivative $D_{2,T} = (d^2T/dt^2)_{\max}$ is 0.85 K/s^2. Assuming that $\varepsilon_2 = \delta_2 D_{2,T}$ and $\delta_2 = 0.1$, obtain $\tau_{reg,2} \approx 1.1$ s. Lower errors ε_1 and ε_2 lead to higher values of τ_{reg}.

Regularization enables significant improvement of the algorithm operation (see the table)). Beginning from the time step of 0.2 s, restoration of the equation structure which does not include the second time derivative of temperature but includes the first derivative with respect to radial component becomes correct. The optimal time step for the synthetic data is 5 s. Thermal conductivity coefficient a_t is restorable accurately to 12%.

For the experimental data obtained at $Q_V = 0.38$ W/mm^3, the time step increase also provides correct restoration of the model. When the step is $\tau_{reg} = 5$ s, the restoration error a_t is 20%. In the experiment with $Q_V = 1.83$ W/mm^3, overheating is 4 times higher, and relative error in temperature is 4 times lower; therefore, restoration of a_t is more accurate. The minimal error of 5% corresponds to the numerical-differentiation time steps of 0.2 and 1 s.

The „one-dimensional" model generated based on the data on the temperature distribution along the vertical line accommodating the wire centerline (symmetry line) may be used to indicate the presence of convection in the medium. For the synthetic data, coefficients a_2 and a_3 differ only slightly, and integral discrepancy α is low. In the experiment with $Q_V = 0.38$ W/mm^3, weak convection takes place, restored coefficients a_2 and a_3 differ by 35%, and α differs slightly from that without convection. When $Q_V = 1.83$ W/mm^3, a_2 and a_3 differ by 6 times, while the integral discrepancy increases by an order of magnitude. It is possible to state that convection plays an important role in this case. Notice that conclusions about the convection existence and intensity may be made without measuring the speed field of the medium but by applying the GDM method to the data on the spatial temperature distribution.

Thus, the previously proposed GDM algorithm has been expanded for processing noisy data. In the considered modification of the algorithm, the time and spatial numerical-differentiation steps were selected depending on the level of the initial data noise.To test the algorithm, experiments with heating the medium with a flooded heat source were carried out. Taking as an example the case of processing noisy synthetic and experimental data, there has been demonstrated the efficiency of the GDM algorithm in restoring the model structure in the form of a partial differential equation and in determining the medium thermophysical parameters. The possibility of using the GDM method to indicate the presence of special heat-exchange modes has been shown.

Financial support

The study was supported by the Russian Science Foundation (project № 21-11-00296). (https://rscf.ru/project/21-11-00296/).

Conflict of interests

The authors declare that they have no conflict of interests.

References

[1] A.A. Samarskii, P.N. Vabishevich, *Computational heat transfer* (Willey, Chichester, 1995).
[2] A.O. Vatul'yan, *Koeffitsientnye obratnye zadachi mekhaniki* (Fizmatlit, M., 2019). (in Russian)
[3] N.Y. Bykov, A.A. Hvatov, A.V. Kalyuzhnaya, A.V. Boukhanovsky, Tech. Phys. Lett., **48** (15), 50 (2022). DOI: 10.21883/TPL.2022.15.55281.18967.
[4] A.V. Lykov, *Teoriya teploprovodnosti*(Vyssh. shk., M., 1967). (in Russian)
[5] A.I. Zhmakin, Tech. Phys., **66** (1), 1 (2021). DOI: 10.1134/S1063784221010242.
[6] F. Hutter, L. Kotthoff, J. Vanschoren, *Automated machine learning. Methods, systems, challenges* (Springer, Cham, Switzerland, 2019). DOI: 10.1007/978-3-030-05318-5
[7] N.Y. Bykov, Nauch.-tekhn. vedomosti SPbGPU. Fiz.-mat. nauki, **15** (3), 83 (2022). DOI: 10.18721/JPM.15307 (in Russian)
[8] G. James, D. Witten, T. Hastie, R. Tibshirani, *An introduction to statistical learning: with applications in R* (Springer, N.Y., 2013). DOI: 10.1007/978-1-4614-7138-7.
[9] B.G. Manukhin, M.E. Gusev, D.A. Kucher, S.A. Chivilikhin, O.V. Andreeva, Opt. Spectrosc., **119** (3), 392 (2015). DOI: 10.1134/S0030400X15090180.
[10] N.B. Vargaftik, *Handbook of physical properties of liquids and gases* (Springer, Berlin–Heidelberg, 1975).
[11] V.A. Rabinovich, Z.Ya. Khavin, *Kratkiy khimicheskiy spravochnik* (Khimiya, L., 1972). (in Russian)
[12] *Fizicheskie velichiny: spravochnik*, pod red. I.S. Grigor'eva, E.Z. Meylikhova (Energoatomizdat, M., 1991). (in Russian)
[13] M.B. Priestley, *Spectral analysis and time series* (Academic Press, London, 1981).

Translated by Solonitsyna Anna