Information systems research

UDC 311.2 doi: 10.20998/2522-9052.2019.1.10

R. Galagan, A. Momot

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

STATISTICAL ANALYSIS OF THERMAL NONDESTRUCTIVE TESTING DATA

The features of processing of active thermal nondestructive testing results are considered. Proved the necessity of search and introduction of new informative parameters in evaluation of thermograms in order to improve the reliability of control. Task of detecting and estimating the relationships between defect parameters and optimal testing time and the maximum value of temperature signal is set. Computer simulation of active thermal testing of two samples with artificial defects with known characteristics was performed. Also obtained the sequences of thermograms and formed the sets of initial data during simulation for correlation, regression and dispersion analysis of testing results. The method of dynamic thermal tomography was used to determine the levels of maximum differential temperature signal and optimal testing time. The estimates of correlation coefficient for various informational parameters of thermal control obtained. There is a high level of relations between the optimal control time and depth of defects. A high correlation also observed between the maximum value of temperature signal and depth of defects. The nature of relationships between various informative parameters of active thermal control established by the regression analysis. A one-factor dispersion analysis of the influence of defect parameters on optimal testing time and maximum value of the temperature signal was performed. High degree of mutual influence of all informative parameters is established. The conclusion made on the necessity of developing new modern methods for analysis the data of thermal testing. Revealed patterns in relationships between data show low efficiency of traditional statistical methods in tasks of active thermal testing. Alternatively, proposed to use the artificial intelligence technologies, in particular, neural networks.

Keywords: nondestructive testing; thermal testing; correlation analysis; regression; dispersion analysis; thermograms processing.

Introduction

Thermal non-destructive testing methods used to solve a wide range of tasks due to a number of advantages. Quality of the received thermograms, which shows the distribution of thermal field in object of testing (OT), depends on many factors. Thermal images often contain high levels of noise, and the nature of defects thermal imprints does not always allow us to make unambiguous conclusions about their size, shape, position and depth. In this regard, in many cases it is impossible to analyze the technical state of OT, using only one informative parameter - temperature. It is known that temperature distribution histograms in defective and defect-free areas partially overlap each other, which reduces the reliability of decision-making. Therefore, the relevant task is to search for additional informative parameters, analysis of which will increase the reliability of testing.

An important condition for increasing the efficiency of decision-making by introducing additional informative parameters is weak correlation of these parameters among themselves [1]. For today, there are no any analyses of correlation degree of various informative features in thermal testing. As a result, there are no regression models that allow us to estimate the kind of mathematical relations between different parameters in thermal testing [2]. Another important task is to carry out a dispersion analysis of defects parameters influence on informative signals. Statistical analysis of the results of active thermal control will determine the most optimal informative parameters for solving specific tasks of defectoscopy and defectometry [3].

Task review

Dynamic thermal field of OT describes by $T(x, y, \tau)$ function. During conducting of active thermal nondestructive testing, we consider the nature of change of instantaneous temperature values in time at the points of the surface of OT. To obtain necessary data, the object of testing is heating over some time by a source of heat flow. The process of heating and subsequent cooling of OT is recording using a thermal imager. The resulting sequence of thermograms reflects the change in the temperature field on the OT surface in time [4]. Considering the temperature dynamics in separate points of thermograms (pixels) corresponding to the coordinates of the surface of OT, it is possible to construct a temperature profile - the dependence of temperature changes in time for the given area. Typically, the temperature change in defect-free areas is constant and considering to be known. It is possible to enter some reference temperature value $T_{nd}(x_{nd}, y_{nd}, \tau)$, which is taken as defect-free. In the defective zone, the regular nature of the thermal field is violated and there are local temperature differences $T_d(x, y, \tau)$, that lead to changes in temperature profile. Thus, it is possible to calculate the value of temperature difference between defective and defect-free areas:

$$\Delta T\left(x,y,\tau\right)=T_d\left(x,y,\tau\right)-T_{nd}\left(x_{nd},y_{nd},\tau_{nd}\right).$$

Time τ_{opt}, at which the value of $\Delta T(x, y, \tau)$ in this area of OT becomes the maximum, is called the optimal testing time:

$$\Delta T_{\max}\left(x,y,\tau\right)=\Delta T_{\max}\left(\tau_{opt}\right).$$

An important parameter influencing the shape of temperature profile is the geometric dimension of defect $h(x, y, z)$. It is known that increases of transverse dimensions of defects x and y at constant thickness z leads to an increase the heat amount needed for its heating. This process is describing by heat transfer equation. When increase the defect's size, the speed of his heating decreases, which leads to a change in shape of thermal profile. In particular, for deeper defects, the value $\Delta T_{\max}$ decreases and optimal observation time τ_{opt} increases. However, for near-surface defects, this dependence not observed [5]. In [6] argued that the size of defects significantly affects the magnitude of $\Delta T_{\max}$ signal, but almost does not affect the τ_{opt} value. Thus, it is possible to construct linear gauge dependences of optimal testing time τ_{opt} from the depth l for defects with known parameters. These dependencies are linear functions, which makes it easy to estimate the depth of defects in known optimal time τ_{opt}. This approach used in the method of dynamic thermal tomography [7].

In the simplest case, decisions about the technical state of OT are taking by one informative parameter, for example, temperature. In the presence of high levels of noise, a complex structure of OT and some other factors, the distribution of temperature values in defective and defect-free areas may be similar, which significantly reduces the probability of testing. It is possible to increase the statistical defect detection parameters by introducing additional informative parameters. This is because in the general case, defective and defect-free areas can overlap by one parameter, but be differ substantially from other parameters. For example, optimal observation time τ_{opt} increases with deeper depth, which allows separating internal defects from surface noise. Another informative parameter may be the geometric dimension h of the temperature anomaly, which allows separate the signals from scratches and internal defects with a large area.

In the general case, any area on the OT surface can be characterized by N parameters. A decision about degree of difference between the corresponding statistical distributions will be taken in the multidimensional feature space. For the effective use of additional informative parameters, the relationship between them should be weak. Of particular interest is the estimation of relations of an informative parameter with a certain characteristic of defect, for example, him size or depth [8]. Statistical analysis of such dependencies will allow us to develop methods for solving inverse tasks of thermal nondestructive testing.

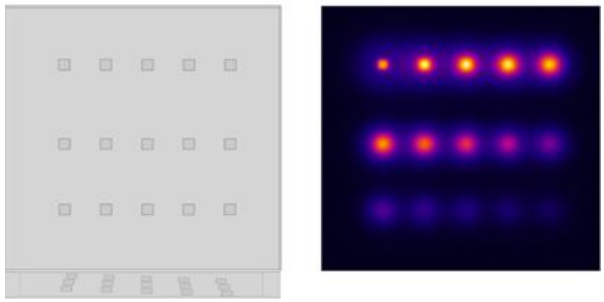
Problem statement

In the frame of this study, was perform a statistical analysis of data obtained by computer simulation of active thermal nondestructive testing of a metal plate. The aim of the work is to identify and analyze the relationships between the maximum values of differential temperature signals ΔT_{max}, the optimal observation times τ_{opt}, the depths and geometric dimensions of defects. Statistical analysis of data includes correlation analysis, building of regression models and dispersion analysis. The obtained results will allow determining the most optimal informative parameters for solving tasks of thermal defectometry and increase the reliability of testing.

Description of input data

For the purpose of statistical analysis of testing results, a simulation of OT heating and cooling process was performed by COMSOL Multiphysics software. In simulation, an artificial mathematical model of an 10 mm thick aluminum plate with 100 mm sides was considered (Fig. 1, *a*). There are artificial defects in the middle of OT, which are air cavities of square shape of various sizes and 2 mm thickness. For various experiments, the defects were located at different depths. In addition, to simulate uneven heating on the upper edge of the sample, there are linear sources of low power heat flux, not shown in the diagrams. The heating carried out by a pulse of 0.1 s duration with a heat flux of 100 kW/m^2 power density, which applied to the upper edge of the plate. Thermograms recording took place for 3 seconds. As a result, sequences of thermographs containing 50 images recorded.

In order to exclude the influence of the size of defects on temperature profile, defects of the same size (4 mm) analyzed for determining the relationship between the optimal observation time and the depth of defect. Similarly, relationship between the value of temperature signal and depth of defects evaluated. The thermogram of such sample in optimal testing time showed in Fig. 1, b. Due to the effect of thermal diffusion, thermal imprints of defects on the OT surface have fuzzy boundaries.



a *b*

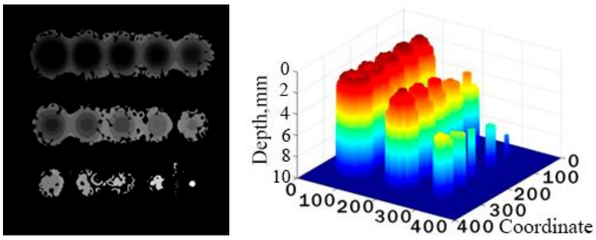
Fig. 1. Results of computer simulation of the first sample: *a* - scheme of an artificial sample; *b* – thermogram in optimal testing time

Fuzziness increases with increasing of defects depth and diminishing of their size. The high level of thermal diffusion makes it impossible to estimate the shape and size of defects. This fact greatly complicates the visual processing of thermograms, and creates considerable difficulties in digital image analysis by most existing methods.

A defect map (Fig. 2, *a*) and a thermal tomogram (Fig. 2, *b*) were obtained by the application of dynamic thermal tomography method (DTT) [9]. Information about three defects was lost using the DTT method, due to the low sensitivity and the presence of noise. The shape and size of the detected defects significantly distorted. There are external noise in the image. In this regard, during the formation of samples for statistical analysis of data, signals selected only from areas whose defective status was determined without errors. The degree of correlation between the optimal observation time and the maximum value of the temperature signal was also determined for faultlessly detected defective points of the sample.

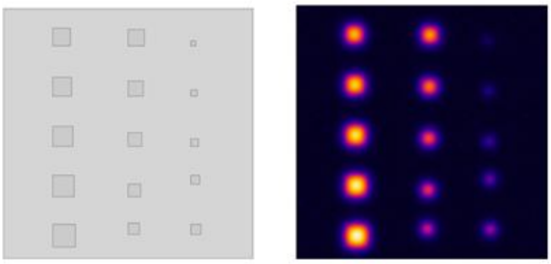
The influence of defects size on the value of temperature signal and optimal observation time was investigated on a specimen with defects located at the same depth of 2 mm. The scheme of the sample and the thermogram in optimal testing time showed in Fig. 3.

The values of ΔT_{max} and τ_{opt} measured for formation of studied examples from total set of received signal samples for each of 15 defects. Measurement carried out at random points of the thermal imprints of the defects. Samples selected only from defective points that identified correctly. Size of the samples datasets (vectors of ΔT_{max} and τ_{opt}) for each individual experiment corresponds to the total number of defects.



a *b*

Fig. 2. Results of thermograms sequence processing by the method of dynamic thermal tomography: *a* – timegram; *b* – thermal tomogram



a *б*

Fig. 3. Results of computer simulation of the second sample: *a* – scheme of an artificial sample, *b* – thermogram in optimal testing time

Correlation analysis

In order to verify the relationship between the informative parameters of active thermal testing and defect parameters, Pearson correlation coefficients for received datasets were determined [10]. Algorithms for determining the correlation coefficient and checking the significance of obtained data implemented in the MATLAB software. Since the measured values of ΔT_{max} and τ_{opt} are random variables, a check on the normality of their distribution according to the Pearson consistency criterion performed. The standard *chi2gof* function was used for this. As a result, a normal data distribution law confirmed and Pearson correlation coefficients determined. Since the number of samples is insignificant ($n = 15$), the corrected correlation coefficient was calculated:

$$\hat{r}^* = \hat{r}\left[1 + \hat{r}^2/2/(n-3)\right].$$

The significance of the correlation coefficients checked by Student and Gauss statisticians. The tabular value of Student's coefficient for the probability of 0.05 and the degree of freedom n-2 is $t_{0.05} = 1.771$. The table value of the quantile of Gaussian distribution for the probability 0.05 is $r_{\alpha 0.05} = 1.96$. In all cases, considered correlation coefficients are significant. The obtained values of corrected Pearson correlation coefficients and confidence limits for $P = 0.95$ are shown in Table 1.

Table 1 – **Pearson correlation coefficients in the confidence limits for the received data**

Parameter	τ_{opt}	ΔT_{max}
Depth l	$0.889 \le \mathbf{0.963} \le 0.989$	$-0.908 \le \mathbf{-0.793} \le -0.366$
Defects size h	$0.04 \le \mathbf{0.544} \le 0.826$	$0.995 \le \mathbf{0.998} \le 0.999$
τ_{opt}	-	$-0.835 \le \mathbf{-0.564} \le -0.072$

Regression

A regression analysis of data performed in the MATLAB for purpose of establishing the type and studying the relationship between variables [11]. Regression models that reflect these dependencies are constructed. Estimated standard and extended uncertainty. The nature of the dependence and regression equation for the studied datasets shown in Table 2. Regression graphs shown in Fig. 4.

Table 2 – **Regression for the studied data**

Y / X	τ_{opt}, c	ΔT_{max}, °K
Depth l, mm	$Y(X) = 0.084 \cdot e^{0.504X}$	$Y(X) = -0.877 \cdot e^{58.46X}$
Size h, mm	$Y(X) = 0.5 \cdot X - 2.47$	$Y(X) = 3.052 \cdot X - 3.59$
τ_{opt}, s	-	$Y(X) = 4.132 \cdot \frac{1}{X} - 3.43$

Dispersion analysis

In order to study the influence of factor features (defect depth, defect size) on the result features (ΔT_{max} and τ_{opt}), one-factor dispersion analysis was performed [12]. Four groups of observations (for l = [2 mm, 3 mm, 4 mm, 5 mm]) were formed to evaluate the influence of defect depth on informative parameters. The number of samples in each group comprised 12 to 14 elements. Similarly, for the estimation of influence of defects size on informative parameters, four groups of observations (for h = [7 mm, 6 mm, 5 mm, 4 mm]) were formed. The number of samples in each group comprised 12 to 15 elements. The significance of obtained results was checked by Fisher's criterion $F_\alpha(v_1, v_2)$ for the significance level $\alpha = 0.05$. In a result of calculations, all estimates considered significant. The obtained values of the influence of factor features on the measurement results shown in Table 3.

Discussion

In this work was established the presence of a high level of relationship ($r = 0.963$) between optimal testing time and depth of defects, which is explained by the process of distribution of heat packets from surface to the middle of specimen.

Nearly functional relations ($r = 0.998$) was found between the maximum value of temperature signal and size of defects. This explained by the fact that for homogeneous defects in the physical sense at same depth, the value of temperature signal depends linearly on the power of heating source [13]. Under conditions of uniform heating, larger defects absorb more heat, which leads to slower cooling and, consequently, higher values of differential temperature signal.

A high correlation ($r = -0.793$) is observed between the maximum value of temperature signal and depth of defects. As the depth of penetration increases, the intensity of thermal waves fades out. This leads to a decrease in amount of heat absorbed by defects of the same size. Consequently, presence of a correlation relations in this case is also confirmed by a physical substantiation [14].

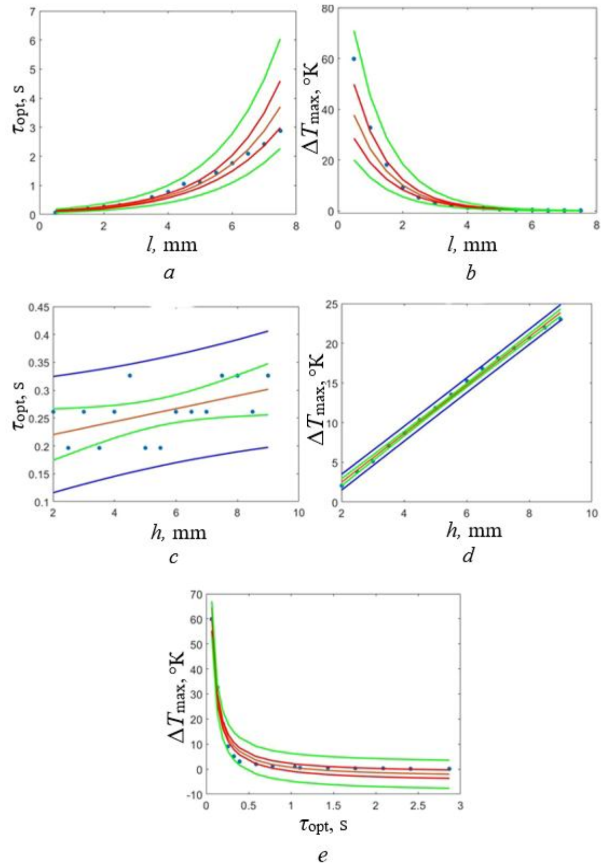


Fig. 4. Regression: *a* – exponential regression for τ_{opt} and l; *b* – exponential regression for ΔT_{max} and l; *c* – linear regression for τ_{opt} and h; *d* – linear regression for ΔT_{max} and h; *e* – hyperbolic regression for τ_{opt} and ΔT_{max}

Table 3 – **Values of the influences of depth and size of defects on informative parameters**

Parameter	τ_{opt}	ΔT_{max}
Depth l	91,89 %	99,23 %
Size h	32,05 %	96,29 %

A noticeable correlation ($r = 0.544$) is observed between optimal testing time and defect size. For the same depth and physical structure, larger defects absorb more heat. This leads to the fact that, under uniform heating conditions, defects that are more massive require more time to heat up to the maximum value of temperature signal.

There is a noticeable correlation ($r = -0.564$) and a hyperbolic relations is established between optimal testing time and maximum value of temperature signal. This calls into question the possibility of using these parameters in image recognition tasks using statistical methods. As known, to improve the accuracy of recognition, a weak correlation between the parameters that form the set of signs is required.

Parameters related to the defects depth have a complex nonlinear relationship. This may be due to the nature of extinction of thermal radiation when propagated deep into OT, which described by exponential dependences [15]. The obtained dependences can be used in determining of optimal heating parameters.

A one-factor dispersion analysis of the influence of defect parameters on optimal testing time and maximum value of temperature signal was carried out. The high degree of influence of defect depth on both informative parameters ($\eta = 91.89$ % on τ_{opt} and $\eta = 99.23$ % on ΔT_{max}) is established. High influence (η = 96.29 %) of defect size on the maximum value of temperature signal was also detected. Obtained results may be explained by the presence of a significant correlation between these parameters, which caused by peculiarities of physical processes occurring in OT during active thermal testing. Determined high levels of influence allow linking the considered parameters of signal with the defect parameters and use this dependence data in defectometry tasks. Less significant effect (η = 32.05 %) has the size of defect at optimal testing time. However, the value of influence does not allow neglecting it in data analysis.

Conclusions

Obtained results testify to the complexity of unambiguous interpretation of thermal testing data due to a large number of relationship parameters. The nature of the relations is mostly complex and nonlinear, which complicates the data processing by traditional statistical methods.

The presence of a significant correlation between considered informative parameters calls into question the possibility of their use for forming feature space in tasks of pattern recognition. At the same time, obtained numerical and mathematical dependences can be used in the development of new methods of thermal testing or analysis of temperature data. The optimal testing time depends strongly on the defects depth and, to a lesser extent, on the size of the defects, which allows using this parameter to construct the gauge dependencies in the tasks of thermal tomography. The maximum value of temperature signal significantly related to both depth and the size of the defects. This proves the need for further analysis and introduction of additional informative parameters to improve the reliability of testing. Obtained results also should be taken into account during selection of experiment parameters in active thermal testing.

The revealed and estimated patterns confirm the relevance of issue of choosing or developing new modern methods for analyzing the thermal testing data. Promising statistical method of temperature data processing is a principal components analysis method, which allows reducing the dimension of initial dataset by excluding from the consideration of interrelated and uninformative parameters. Another effective means of thermograms processing are artificial neural networks, which have the ability to work with complex nonlinear dependencies. Application of artificial intelligence in the tasks of thermal testing is of high interest for further study.

References

1. Grinzato, E.G., Vavilov, V.P., Bison, P.G. and Marinetti, S. (2005), "Methodology of processing experimental data in transient thermal nondestructive testing (NDT)", *Proc.: Thermosense XVII Int. Conf. on TSIDA*, Vol. 2473. pp. 12–18.

2. Balageas, D.L., Roche J.M. and Leroy F.H. (2017) “Comparative Assessment of Thermal NDT Data Processing Techniques for Carbon Fiber Reinforced Polymers”, *Materials Evaluation*, Vol. 75(8), pp. 1019–1031.
3. Vavilov, V.P., Chulkov A.O., Derusova D.A. and Pan Y. (2016), “Thermal NDT research at Tomsk Polytechnic University”, *Quantitative InfraRed Thermography Journal*, Vol. 13(2) pp. 128–143.
4. Galagan, R.M. and Momot, A.S. (2018), “Analysis of the digital processing of thermograms”, *Visnyk Natsionalnoho Tekhnichnoho Universytetu Ukrayiny «KPI». Seriya pryladobuduvannya*, Kyiv, No. 55(1), pp. 108–117.
5. Vavilov V.P, Plesovskikh A.V. and Chulkov A.O. (2015), “A complex approach to the development of the method and equipment for thermal nondestructive testing of CFRP cylindrical parts”, *Composites Part B: Engineering*, Vol. 68, pp. 375–384.
6. Vavilov V.P. and Burleigh D.D. (2016), “Review of pulsed thermal NDT: Physical principles, theory and data processing”, *Ndt & E International*, Vol. 73, pp. 28–52.
7. Vavilov V.P. and Burleigh D.D. (2015), “Pulsed thermal NDT in tables, figures, and formulas”, *In Thermosense: Thermal Infrared Applications XXXVII International Society for Optics and Photonics*, Vol. 9485, [94850Q], pp. 1–15.
8. Vavilov, V.P. (2015), “Dynamic thermal tomography: Recent improvements and applications”, *NDT & E Int.*, Vol. 71, pp. 23–32.
9. Maldague, X. and Ziadi, A. (2002), “Advances in pulsed phase thermography”, *Infrared Physics & Techn.*, Vol. 43, pp. 175–181.
10. Dutta, T., Santra, D., Peng-Lim, C., Sil, J. and Chottopadhyay, P. (2018), “Statistical Feature Analysis of Thermal Images from Electrical Equipment”, *Decision Science in Action*, Springer, Singapore, Vol. 1, pp. 127–137.
11. Addepalli, S., Zhao, Y., Roy, R., Galhenege, W. and Colle M. (2018), “Non-destructive evaluation of localised heat damage occurring in carbon composites using thermography and thermal diffusivity measurement”, *Measurement*, Vol. 131. pp. 706–713.
12. Khodayar, F., Sojasi, S. and Maldague, X. (2018), “Infrared thermography and NDT: 2050 horizon”, *Quantitative InfraRed Thermography Journal*, Vol. 13 (2), pp. 210–231.
13. Vavilov, V.P. (2013), *Infrared thermography and heat control*, Spectrum ID, Moscow, 544 p.
14. Vavilov, V.P. (2006), “Dynamic thermal tomography”, *Plant Laboratory*, No. 3, pp. 26–36.
15. Venegas, P., Perán, J., Usamentiaga, R. and de Ocáriz I.S. (2018), “Projected thermal diffusivity analysis for thermographic nondestructive inspections”, *Int. Journal of Thermal Sciences*, Vol. 124, pp. 251–262.

Received (Надійшла) 14.11.2018
Accepted for publication (Прийнята до друку) 06.02.2019

Статистичний аналіз даних теплового неруйнівного контролю

Р. М. Галаган, А. С. Момот

Розглянуто особливості обробки результатів активного теплового неруйнівного контролю. Доведено необхідність пошуку та введення нових інформативних параметрів при оцінці термограм з метою покращення достовірності контролю. Поставлена задача виявлення та оцінки взаємозв’язків між параметрами дефектів та оптимальним часом контролю і максимальним значенням температурного сигналу. Проведено комп’ютерне моделювання процесу активного теплового контролю двох зразків зі штучними дефектами з відомими характеристиками. Отримано послідовності термограм та сформовано набори початкових даних для проведення кореляційного, регресійного та дисперсійного аналізів результатів контролю. Для визначення рівнів максимального диференціального температурного сигналу та оптимального часу спостереження використовувався метод динамічної теплової томографії. Отримано оцінки коефіцієнта кореляції для різних інформативних параметрів теплового контролю. Встановлена наявність високого рівня зв’язку між оптимальним часом контролю та глибиною залягання дефектів. Висока кореляція також спостерігається між максимальним значенням температурного сигналу та глибиною залягання дефектів. В результаті проведення регресійного аналізу встановлено характер взаємозв’язків між різними інформативними параметрами активного теплового контролю. Проведено однофакторний дисперсійний аналіз впливу параметрів дефектів на оптимальний час контролю та максимальне значення температурного сигналу. Встановлено високий ступінь взаємовпливу всіх інформативних параметрів. Зроблено висновки про необхідність розробки нових сучасних методів аналізу даних теплового контролю. Виявлені закономірності у взаємозв’язках між даними свідчать про низьку ефективність традиційних статистичних методів в задачах активного теплового контролю. В якості альтернативи пропонується застосування технологій штучного інтелекту, зокрема, нейронних мереж.

Ключові слова: неруйнівний контроль; тепловий контроль; кореляційний аналіз; регресія; дисперсійний аналіз; обробка термограм.

Статистический анализ данных теплового неразрушающего контроля

Р. М. Галаган, А. С. Момот

Рассмотрены особенности обработки результатов активного теплового неразрушающего контроля. Доказана необходимость поиска и введения новых информативных параметров при оценке термограмм с целью улучшения достоверности контроля. Поставленная задача выявления и оценки взаимосвязей между параметрами дефектов, оптимальным временем контроля и максимальным значением температурного сигнала. Проведено компьютерное моделирование процесса активного теплового контроля двух образцов с искусственными дефектами с известными характеристиками. Получены последовательности термограмм и сформированы наборы исходных данных для проведения корреляционного, регрессионного и дисперсионного анализов результатов контроля. Для определения уровней максимального дифференциального температурного сигнала и оптимального времени наблюдения использовался метод динамической тепловой томографии. Получены оценки коэффициента корреляции для различных информативных параметров теплового контроля. В результате проведения регрессионного анализа установлено характер взаимосвязей между различными информативными параметрами активного теплового контроля. Проведен однофакторный дисперсионный анализ влияния параметров дефектов на оптимальное время контроля и максимальное значение температурного сигнала. Установлена высокая степень взаимосвязей всех рассмотренных информативных параметров. Сделаны выводы о необходимости разработки новых современных методов анализа данных теплового контроля. Обнаруженные закономерности во взаимосвязях между данными свидетельствуют о низкой эффективности традиционных статистических методов в задачах активного теплового контроля. В качестве альтернативы предлагается применение технологий искусственного интеллекта, в частности, нейронных сетей.

Ключевые слова: неразрушающий контроль; тепловой контроль; корреляционный анализ; регрессия; дисперсионный анализ; обработка термограмм.