Method of statistical spline functions for solving problems of data approximation and prediction of objects state

Serhiy Babak[1][0000-0001-8805-1184], Vitaliy Babak[2][0000-0002-9066-4307], Artur Zaporozhets[2][0000-0002-0704-4116], Anastasia Sverdlova[2][0000-0001-8222-1357]

[1] University of Emerging Technologies, Kyiv, Ukraine
[2] Institute of Engineering Thermophysics of NAS of Ukraine, Kyiv, Ukraine

a.o.zaporozhets@nas.gov.ua

Abstract. The method of statistical spline functions is considered for problems of predicting the state of complex technical objects using the example of power transmission lines. The choice of parameters of spline fragments for building an adequate mathematical model is analyzed. Based on the experimental data, a short-term spline forecast of heating of overhead power lines has been created.

Keywords: spline functions, mathematical model, overhead power lines, prediction, technical state, diagnostics

1 Introduction

The basic idea of using the mathematical apparatus of statistical spline functions for processing the spatial characteristics of a field is that some numerical characteristics of physical processes change during the observation process (meteorological fields, airspace in the vicinity of energy facilities, navigation space, etc.) for any circumstances [1, 2, 3].

Regular observations of the history of gradual changes in these parameters over time provide an opportunity to obtain information on the trends of further changes in the studied parameters and to predict the behavior of the field at certain points [4]. For example, using the spline function algorithm, it can monitor and predict the temperature at a particular point in the airspace [5]. Such a problem arises during temperature control equipment of overhead power lines (OPL) using thermal imaging equipment installed on the unmanned aerial vehicles (UAV) [6, 7].

Using the spline functions, the problem of predict ting electrical equipment failures is also solved. In this case, the main idea of predicting failures is that some numerical characteristics of physical processes occurring in certain nodes of electrical equipment change during the occurrence or development of faults and defects, which allows to identify them [8, 9]. Observations of their gradual change in time provide information on the development trends of the defect and provide a prediction of the possible moment of failure [10, 11].

In work the forecasting of values of temperature in certain points of an arrangement of the equipment of the OPL will be considered. Due to a significant temperature increase, failures of individual blocks of OPL are possible. So, the sharp heating of metal wires of OPL can lead to sagging and short circuit on the earth. The temperature factor may contribute to the occurrence of a breakdown or overlap of insulators of OPL. It is also possible breakdown of insulators due to contamination of their surface, or aging of the materials from which they are made. In addition, due to various reasons, accumulation of microdefects can occur in the insulator material, which contributes to their breakdown [12, 13, 14].

The listed defects of equipment of OPL in the course of their operation can lead to the emergence of so-called gradual failures. Accidental failures caused by unpredictable factors (for example, the overlap of insulation of OPL by birds or animals) will not be considered.

2 Features of spline functions

2.1 Mathematical model

To predict the possible failure of a selected node, you additionally need to have statistics on the values of the monitored parameter for a certain period of time, which is called the observation interval.

Let some functional dependence be given on the segment

$$y = f(x, A),\ x \in [a, b], \tag{1}$$

where A is the deterministic vector of unknown real numeric parameters entering linearly in y and do not depend on x.

At given points $x_i \in [a, b],\ i = \overline{1, n}$, random uncorrelated values of the function y are observed, which we denote as $\{y_i, i = \overline{1, n}\}$. For definiteness, we assume that these observations are distributed according to the normal law, and

$$\begin{cases} M y_i = x_{i0} a_0 + x_{i1} a_1 + \ldots + x_{ir} a_r \\ D y_i = \sigma^2, \quad i = \overline{1, n} \end{cases}, \tag{2}$$

where $A = (a_1, a_2, \ldots, a_r)$ and σ are unknown parameters, $X = \{x_{ij},\ \ i = \overline{1, n};\ j = \overline{0, r}\}$ is a rectangular matrix of deterministic coefficients, functionally dependent from x_i and, as a rule, is called the planning matrix. It should also be noted that the dependence of these coefficients from x_i is not necessarily linear.

If we assume that Y and A are column vectors, which will be provided below, then (2) can be written in vector form: $MY = XA$, $DY = \sigma^2 I$. Here, M and D are the expectation and variance operators, and I is the n-th unit matrix order.

The paper considers the problem of constructing a statistical estimate of the unknown parameters $\{a_j,\quad j = \overline{1, r}\}$ from the results of observations $\{y_i,\quad i = \overline{1, n}\}$. In this case, an arbitrary choice of elements of the planning matrix X is allowed. If nothing is assumed about the distribution of observation errors or observations

$\{y_i, \quad i=\overline{1,n}\}$, then as a result of solving this problem the largest, what can we expect, is the construction of point estimates for elements *A*. Under the assumption that the distribution law of observations is normal, we can obtain a confidence interval for the estimated parameters.

Considering that in the task set the question of introducing the planning matrix is solved in an arbitrary way, we specify this task by refining the choice of the planning matrix *X*. At the same time, we construct statistical estimates of the *A* parameters using the least squares method, which in the case of normally distributed y_i leads to the same result as maximum likelihood estimation.

Concretization in the choice of the planning matrix, first of all, is connected with the choice of the upcoming functions. The task was traditionally solved in the class of polynomials. But polynomial approximations have several disadvantages, the most significant of which is that the sequence of interpolation polynomials does not always converge to the interpolated function. Therefore, in many problems, the more natural and convenient approximation apparatus turned out to be splines, with the help of which we will solve the problem posed.

Splines are functions that are "glued together" from pieces of various functions in a specific pattern. Polynomial splines "stick together" from pieces of various polynomials in such a way as to ensure the necessary smoothness of the resulting spline. The simplest example of a polynomial spline is a broken line.

Let the grid be set on the segment $[a,b]$, $a,b\in R^+$, $a<b$ (partition):

$$\Delta_n : a = x_0 < x_1 < \ldots < x_n < x_{n+1} = b, \tag{3}$$

where $n \in N$.

Let also P_m is the set of polynomials of degree not higher than m, $m \ge 0$, and $C^{(k)} = C^{(k)}[a,b]$ is the set of functions continuous on $[a,b]$ that have a continuous *k*-th derivative, $k \in Z^+$; R^+ is the set of positive numbers; *N* is the set of natural numbers; Z^+ is the set of positive integers.

The function $S_m(x) = S_{m,k}(x,\Delta_n)$ is called a polynomial spline of degree *m* of defect *k* ($1 \le k \le m$) with nodes (3), if

- $S_m(x) \in \mathbf{P}_m$, $x \in [x_i, x_{i+1}]$, $i = \overline{0, r-1}$,
- $S_m(x) \in \mathbf{C}^{m-k}[a,b]$.

The points $\{x_i\}$ are called spline nodes, the $(m-k+1)$-th derivative of $S_m(x)$ can be discontinuous on the segment $[a,b]$. Basically, it takes *k* = 1. There is a representation (with a fixed grid Δ_n):

$$S_m(x) = S_{m,k}(x,\Delta_n) = \sum_{s=0}^{m} a_s x^s + \sum_{s=1}^{n-1} \sum_{r=m-k+1}^{m} a_{r,s}(x-x_s)_+^r, \tag{4}$$

where $(x-x_s)_+^r = [\max(0, x-x_s)]^r$ is a Peano's core.

The coefficients a_s and $a_{r,s}$ can take arbitrary values from *R*; the set $S_{m,k}(x, \Delta_n)$ with a fixed Δ_n is linear with dimension $m+1+nk$. Therefore, for an unambiguous definition of a spline, it is necessary to specify $m+1+nk$ independent conditions. For a linear spline this is $n+2$, that is, for the statistics obtained, the number of parameters that need to be estimated is found from condition $r < n+2$.

Thus, under the conditions of the formulated problem, it is necessary on the segment $[a, b]$ for given r to find a grid $\{\tilde{x}_j\}_+^r$ such that the spline $S_{k+1}(x) \in \mathbf{C}^k[a, b], k = 0,1, \ldots,$ constructed on the grid Δ_r provides the minimum (in terms of standard quadratic deviations) statistical estimate of the vector *A*, that is, we assume that

$$\sum_{j=1}^{r} x_{i1} a_1 + \ldots + x_{ir} = S_{k+1}(x_i),\ i = \overline{1, n}. \tag{5}$$

The equations written in the matrix form

$$X^* X A = X^* Y \tag{6}$$

are called normal. They are obtained by minimizing the sum of squared differences between observations and their mathematical expectations:

$$\begin{gathered}(Y - X A)^*(Y - X A) = \sum_{i=1}^{n}\left[y_i - x_{i1} a_1 - \ldots - x_{ir} a_r\right]^2 = \\ = \sum_{i=1}^{n}\left[y_i - S_{k+1}(x_i)\right]^2\end{gathered} \tag{7}$$

As the estimated parameters of the spline, the a_j ordinates of the nodes of the grid Δ_r are chosen, whose estimates are now determined by solving the normal equations (6) in the matrix form

$$A = (X^* X)^{-1} X^* Y. \tag{8}$$

The planning matrix generally has the following form

$$X = \begin{vmatrix} \bar{x}_{10} & \bar{x}_{11} & 0 & 0 & \cdot & 0 & 0 \\ \bar{x}_{20} & \bar{x}_{21} & 0 & 0 & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \bar{x}_{k0} & \bar{x}_{k1} & 0 & 0 & \cdot & 0 & 0 \\ 0 & \bar{x}_{k_1 11} & \bar{x}_{k_1 12} & 0 & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \bar{x}_{k_1 k_2 1} & \bar{x}_{k_1 k_2 2} & 0 & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdot & \bar{x}_{n-k_r r-1} & \bar{x}_{n-k_r r} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdot & \bar{x}_{nr-1} & \bar{x}_{nr} \end{vmatrix} \qquad (9)$$

This matrix is rectangular, it has n rows and r+1 columns. The components of the vectors $\left\{ k_j, \quad j = \overline{1, r} \right\}$ characterize the number of observations that fall in the j-th interval, $\sum_{j=1}^{r} k_j = n$. It is assumed that $n \ge r + 2$.

Based on (9), we obtain the matrix in the form

$$C = X^{(*)} X = \begin{vmatrix} c_{00} & c_{01} & 0 & 0 & 0 & \cdot & 0 & 0 & 0 \\ c_{10} & c_{11} & c_{12} & 0 & 0 & \cdot & 0 & 0 & 0 \\ 0 & c_{21} & c_{22} & c_{23} & 0 & \cdot & 0 & 0 & 0 \\ 0 & 0 & c_{32} & c_{33} & c_{34} & \cdot & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & \cdot & c_{r-2\,r-2} & c_{r-2\,r-1} & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdot & c_{r-1\,r-2} & c_{r-1\,r-1} & c_{r-1\,r} \\ 0 & 0 & 0 & 0 & 0 & \cdot & 0 & c_{r\,r-1} & c_{rr} \end{vmatrix} \qquad (10)$$

where

$$c_{00} = \sum_{i=1}^{k_1} \bar{x}_{i0}^{2} = \sum_{i=1}^{k_1} \frac{\left(\tilde{x}_1 - x_i\right)^2}{\left(\tilde{x}_1 - \tilde{x}_0\right)^2} = \frac{\sum_{i=1}^{k_1} \left(\tilde{x}_1 - x_i\right)^2}{\left(\tilde{x}_1 - \tilde{x}_0\right)^2},$$

$$c_{jj} = \sum_{i=1+s_{j-1}}^{s_j} \bar{x}_{ij}^{2} + \sum_{i=1+s_j}^{s_{j+1}} \bar{x}_{ij}^{2} = \sum_{i=1+s_{j-1}}^{s_j} \frac{\left(\tilde{x}_{j-1} - x_i\right)^2}{\left(\tilde{x}_j - \tilde{x}_{j-1}\right)^2} + \sum_{i=1+s_j}^{s_{j+1}} \frac{\left(\tilde{x}_{j+1} - x_i\right)^2}{\left(\tilde{x}_{j+1} - \tilde{x}_j\right)^2}, \qquad (11)$$

$$c_{j+1\,j} = c_{j\,j+1} = \sum_{i=1+s_j}^{s_{j+1}} \bar{x}_{ij}\, \bar{x}_{i\,j+1} = \sum_{i=1+s_j}^{s_{j+1}} \frac{\left(\tilde{x}_{j+1} - x_i\right)\left(x_i - \tilde{x}_j\right)}{\left(\tilde{x}_{j+1} - \tilde{x}_j\right)^2}, \; j = \overline{0, r-1},$$

$$c_{j+m\,j} = c_{j\,j+m} = 0 \text{ at } m \ge 2, \quad j = \overline{0, r-2},$$

$$c_{rr} = \sum_{i=1+n-k_r}^{n} \bar{x}_{ir}^2 = \sum_{i=1+n-k_r}^{n} \frac{\left(\tilde{x}_{r-1} - x_i\right)^2}{\left(\tilde{x}_r - \tilde{x}_{r-1}\right)^2},$$

and x_i are the elements of the source data (experimental points), $S_j = \sum_{m=1}^{j} k_m$, $j = \overline{1, r}$, $s_0 = 0$, $s_r = n$. The matrix *C* is rectangular $(r+1) \times (r+1)$.

The above relations are obtained by directly multiplying the matrices $X^{(*)}$ and *X* with (9) taken into account.

Let us briefly discuss some properties of the matrix *C*.

1. The sum of the diagonal elements of the matrix *C* is equal to the sum of the squares of all the elements of the planning matrix *X*, that is,

$$\sum_{j=0}^{r} c_{jj} = \sum_{i=1}^{n} \sum_{j=1}^{r} \bar{x}_{ij}^2 = \sum_{i=1}^{n} \left\{ \left(\frac{\tilde{x}_1 - x_{i1}}{\tilde{x}_1 - \tilde{x}_0} \right)^2 I_{x_{i1}}\left(\Delta x_1\right) + \right.$$

$$+ \left(\frac{x_{r-1} - x_{ir}}{x_r - x_{r-1}} \right)^2 I_{x_{iir}}\left(\Delta x_r\right) + \qquad (12)$$

$$\left. + \sum_{j=1}^{r-1} \left[\frac{x_{j+1} - x_{ij+1}}{x_{j+1} - x_j} I_{x_{ij+1}}\left(\Delta x_{j+1}\right) \frac{x_{j+1} - x_{ij}}{x_j - x_{j-1}} I_{x_{ij}}\left(\Delta x_j\right) \right]^2 \right\}.$$

2. The rectangular matrix C has all sums of elements in the rows equal to 1.
3. The matrix *C* (and C^{-1}) is only a function of the sums $\sum x_i$ and $\sum x_i^2$, which are taken over all intervals and do not change if the individual x_i change their values in such a way that these sums remain unchanged.

The determinant of the matrix *C* is denoted by $|C| = |C|_{r+1}$, where the index $r+1$ is equal to the number of rows or columns of the matrix *C*.

If the determinants $|C|_{r-1}$ and $|C|_r$ are calculated, then the determinant $|C|_{r+1}$ is found by the formula

$$|C|_{r+1} = c_{rr} |C|_r - c_{rr}^2 |C|_{r-1} \text{ at } r = 2, 3, \ldots \qquad (13)$$

The recurrence formula (13) can be obtained by direct calculation, by decomposing the *C* row (or column) in $r+1$ minors.

Thus, if for generalization we denote $|C|_1 = c_{00}$, then to calculate the determinant of the matrix *C*, we obtain the following recurrence relations

$$\begin{cases} |C|_0 = 1; \quad |C|_1 = c_{00}; \\ |C|_{j+1} = c_{jj} |C|_j - c_{jj-1}^2 |C|_{j-1}; \end{cases} \quad j = \overline{1, r}. \qquad (14)$$

It can construct recurrent formulas for finding the elements of the inverse matrix *C*. Let us briefly discuss the construction of these formulas. The determinant $\left|C\right|_{r+1}$ is calculated by the recurrent formula (14).

The algebraic complement of the element $c_{\min}$ which located on the main diagonal is equal to the multiplication of two determinants similar in structure to the determinant of the matrix and calculated by recurrent formulas (14), which in this case take the form

$$\begin{cases} \left|C\right|_0 = 1; \quad \left|C\right|_1 = c_{00}; \\ \left|C\right|_{j+1} = c_{jj}\left|C\right|_j - c_{jj-1}^2\left|C\right|_{j-1}; \end{cases} \quad j = \overline{1, m-1};$$

$$\begin{cases} \left|\overline{C}\right|_0 = 1; \quad \left|\overline{C}\right|_1 = c_{rr}; \\ \left|\overline{C}\right|_{j+1} = c_{r-jr-j}\left|\overline{C}\right|_j - c_{r-jr-j+1}^2\left|\overline{C}\right|_{j-1}; \end{cases} \quad j = \overline{1, r-m-1}, \tag{15}$$

where $\left|\overline{C}\right|_j$ is the determinant of the matrix obtained with *C* by crossing out the first $r - j + 1$ g rows and first $r - j + 1$ columns in it.

The recurrence formulas (15) allow to find the corresponding determinants of the matrices $\left|\overline{C}\right|_{r-m}$ moving in the direction of the main diagonal from the periphery inward towards the center of the matrix. These formulas can also be rewritten in another form, using the movement from the middle to the periphery, namely, taking into account (14) we have

$$\begin{cases} \left|C\right|_0 = 1; \quad \left|C\right|_1 = c_{m-1m-1}; \\ \left|C\right|_{j+1} = c_{m-1-jm-1-j}\left|C\right|_j - c_{m-1-j\,m-j}^2\left|C\right|_{j-1}; \end{cases} \quad j = \overline{1, m-1};$$

$$\begin{cases} \left|\overline{C}\right|_0 = 1; \quad \left|\overline{C}\right|_1 = c_{m+1m+1}; \\ \left|\overline{C}\right|_{j+1} = c_{m+1+jm+1+j}\left|\overline{C}\right|_j - c_{m+1+jm+j}^2\left|\overline{C}\right|_{j-1}; \end{cases} \quad j = \overline{1, r-m-1}, \tag{16}$$

and now $\left|C\right|_j$ is the determinant of the matrix obtained by crossing out the first rows $j-1$ and columns in the *C* matrix and $r - m + 1$ last rows and columns, and $\left|\overline{C}\right|_j$ is the determinant of the matrix obtained from the matrix *C* by crossing out $m+1$ first rows and columns and $r - j + 1$ last rows and columns, $\left|C\right|_j$ and $\left|\overline{C}\right|_j$ are the main minors of the matrix.

It should be noted that the final results of calculations by the recurrence relation (15) and (16) for a fixed matrix are the same, and the intermediate ones may differ.

Thus, the algebraic complement of the element of the matrix *C*, standing on the main diagonal, is determined by the ratio

$$C_{jj} = \left|C\right|_j \left|\overline{C}\right|_{r-1}. \tag{17}$$

The ratio for calculating the algebraic complements of the elements of the matrix *C* that are not on the main diagonal is calculated by the ratio

$$\begin{cases} c_{ji} = (-1)^{(i-j)} \left\{ 1 + sign(i-j) \left[\prod_{p=j}^{i-2} C_{p\,p+1} - 1 \right] \right\} |C|_j \, |C|_{r-1}, \; j \le i = \overline{0, r}; \\ c_{ji} = c_{ij}, \qquad\qquad\qquad\qquad r \ge 1. \end{cases} \tag{18}$$

Denote by H the matrix, which is determined by the expression

$$H = X^{(*)} Y. \tag{19}$$

where $X^{(*)}$ is the matrix transposed with respect to the planning matrix, and Y (matrix-column) is the output of the observations.

The elements of the matrix H are defined as follows.

$$h_0 = \sum_{i=1}^{k_1} \bar{x}_{i0} y_i = \sum_{i=1}^{k_1} \frac{(\tilde{x}_1 - x_i) y_i}{\tilde{x}_1 - \tilde{x}_0};$$

$$h_j = \sum_{i=1+s_{j-1}}^{s_{j+1}} \bar{x}_{ij} y_i = \sum_{i=1+s_{j-1}}^{s_j} \frac{(\tilde{x}_{j-1} - x_i) y_i}{\tilde{x}_j - \tilde{x}_{j-1}} + \sum_{i=1+s_{j-1}}^{s_{j+1}} \frac{(\tilde{x}_{j+1} - x_i) y_i}{\tilde{x}_{j+1} - \tilde{x}_j}; \tag{20}$$

$$h_j = \sum_{i=1+n-k_r}^{n} \bar{x}_{ir} y_i = \sum_{i=1+n-k_r}^{n} \frac{(\tilde{x}_{r-1} - x_i) y_i}{\tilde{x}_r - \tilde{x}_{r-1}}; \quad j = \overline{1, r-1}.$$

Then the estimates of the ordinates of the points of connection of the linear parts of the spline (the elements of the A matrix-column) are determined by the formula

$$a_i = \sum \left\{ C^{-1} \right\}_{ji} h_j, \quad i = \overline{0, r}. \tag{21}$$

The accuracy of approximation of the desired dependence using the selected spline is estimated by the sum of the squares of the deviations of the ordinates of the observation points from the found dependence

$$d = \sum_{j=1}^{r} \sum_{u=1+s_{j-1}}^{s_j} \left[\frac{a_j (x_u - \tilde{x}_{j-1}) + a_{j-1} (\tilde{x}_j - x_u)}{\tilde{x}_j - \tilde{x}_{j-1}} - y_u \right]^2. \tag{22}$$

The confidence interval for the found estimates of a_j is calculated by the formula

$$I_\beta^{(j)} = \left[a_j \mp \gamma_\beta \sqrt{\left\{ C^{-1} \right\}_{jj} \frac{d}{n-r-1}} \right]. \tag{23}$$

where γ_β is a value that satisfies the relation $\mathsf{P}\left\{ |t_{n-r-1}| \le \gamma_\beta \right\} = 1 - \beta$, if the random variable t_{n-r-1} is distributed according to Student's law with $n-r-1$ degrees of freedom; d is the sum of the squares of the deviations of the observations y_i s from the values of the resulting spline at the corresponding points.

The prediction of the confidence interval for the values of the function $x(t)$ at the point $\tilde{x}_{r+1}$ is carried out by enumerating all the splines and choosing one that minimizes this interval at the prediction point. In this case, the spline itself, in the mean-square sense, is closer to the points, and is observed experimentally.

The relation (23) allows to build confidence intervals in each node of the spline, and on the whole interval - a confidence corridor.

To obtain a forecast using a statistical spline, an additional node is introduced to the set of nodes of the spline, the abscissa of which corresponds to the forecast interval. Using a computer search method, a grid is selected that satisfies equation (8) and at the same time minimizes it on the set of possible non-uniform grids and the width of the confidence interval in the forecast node.

As a result, the expected value and the confidence interval for the selected controlled parameter at the end of the forecast interval will be obtained. The boundaries of the confidence corridor are formed by linear interpolation of the upper and lower boundaries of the confidence intervals in all nodes of the resulting spline, including the forecast node.

2.2 Experimental results

Based on the considered spline forecast method, algorithms are built and a computer program is developed that implements this method. It is a computer program that is the main part of the proposed method, which makes it possible to carry out a short-term forecast for determining the time interval in which the temperature of the OPL can reach critical values. All the basic information about this program, practical issues of its application are set out in the work, as well as in special documentation.

The application of the proposed method for constructing a forecast will be considered on a specific example.

For the implementation of such a forecast, it was necessary, first of all, to obtain experimental statistical data on the temperature state of the wires of OPL. In addition, the initial data for the program are:

- the number of temperature measurements $n = 11$ (in our case);
- the number of hours of the forecast $n_{for} = 3$;
- the number of intervals, interpolates spline $r = 5$;
- the number of intervals of the time interval, on which the temperature is measured (the breakdown is performed in order to find the optimal interpolation spline), $n_{break} = 10$;
- confidence probability of estimating the prediction of the length of the time interval to reach the critical temperature of the power lines, $P = 0.95$.

As a result of the calculation of the program we obtain the data that are presented in Table 1 and in Fig. 1.

Table 1. Experimental data

Hours	9.0	9.5	10.0	10.5	11.0	11.5	12.0	12.5	13.0	13.5	14.0
t°, C	40.5	39.5	40.5	45	47.5	47.5	50.0	49.0	51.0	55.0	57.0

In the top line of the Table 1 shows the hours in which the temperature measurement of power lines was carried out. The bottom line contains the corresponding temperature values.

The graph presented in Fig. 1 is a spline approximation in the form of a channel consisting of six sections. The spline docking points at the upper boundary of the channel are designated 1u - 7u (these points are denoted by the symbol Δ), and at the lower boundary 1d - 7d (the symbol is used to designate these points). In the main line of approximation, the notation □ is selected for the designation of docking points.

The width of the channel is determined by the adopted confidence probability P=0.95. Sections 1 - 6 of the graph characterize the observation interval ΔT_{observ}, and sections 6 - 7 characterize the forecast interval $\Delta T_{forecast}$. The forecast was carried out at 3:00 ahead regarding to14:00 hours, the last point from the observation interval. As can be seen from the graph, with an increase in the time interval of the forecast, its upper and lower limits expand, that is, the probability of the forecast decreases with increasing time.

The statistical spline (Fig. 1) is based on the results of observation of the values of the temperature of wires in OPL, which is in operation. Temperature measurement was carried out on the time interval $\Delta T_{observ} \in 9 \ldots 14$ every 0.5 hours using a thermal imager mounted on the UAV. The quantitative values of the measured temperature, given in the bottom line of the table, are shown on the graph as black dots located on the observation interval ΔT_{observ}.

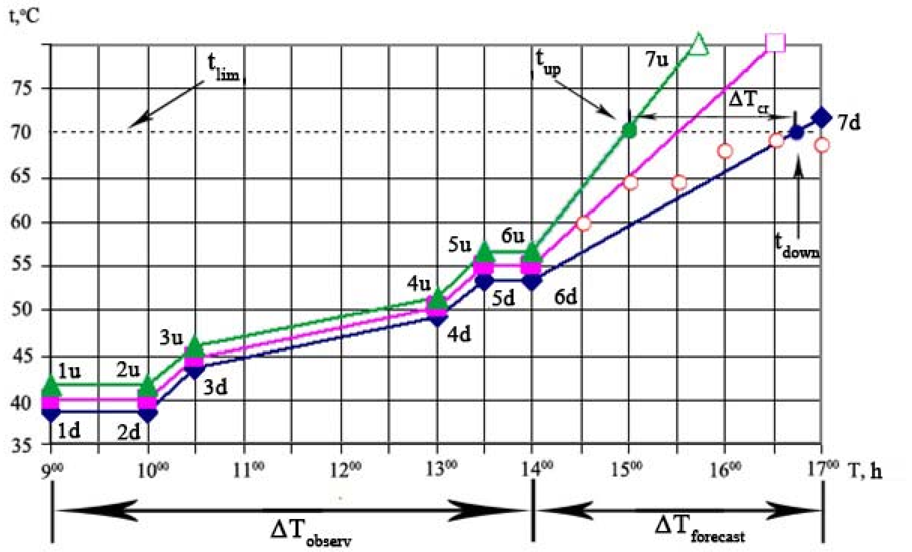

Fig. 1. Prediction using a statistical spline based on observations of temperature values of OPL

A spline was built (Fig. 1), divided into two sections, namely, the section where interpolation of the data on the measured temperature of the power transmission lines over the time interval $\Delta T_{observ} \in 9 \ldots 14$ and section $\Delta T_{forecast} \in 14 \ldots 17$ was carried out with the given probability (P = 0.95). The predicted value of the time interval ΔT_{cr} where the temperature of OPL can reach a critical level is determined by the length of the time interval between the points of intersection of the temperature limit line with the upper t_{up} and lower t_d boundaries of the constructed spline. The threshold temperature value line t_{lim}=70°C is constructed parallel to the x-axis in accordance with the existing regulatory documents on the operation of OPL. In fig. 1 the predicted length ΔT_{cr} is indicated by a black line with arrows.

In order to verify the performance of the proposed method for predicting the possible values of the temperature of power lines, an experimental measurement of the temperature of these wires in the time period was carried out using the UAV, which corresponds to the forecast interval $\Delta T_{forecast}$. The data of the results of these measurements (6 points denoted by ○) is plotted on a specified interval $\Delta T_{forecast}$ of the graph. Almost all (except for one point, the temperature value, which was measured at $T=17^{00}$), experimental data were obtained that did not exceed the upper and lower limits of the constructed spline forecast. This confirms the efficiency of the proposed method for predicting the values of the time interval in which the wires of OPL can reach unacceptable temperature values.

Based on the constructed spline forecast, the time interval when the power transmission lines can reach the maximum permissible temperature value of 70 °C is between 3:00 pm and 4:45 pm with a confidence level of P = 0.95. It should also be emphasized once again that the achievement of such a temperature can occur only under constant conditions of operation of OPL.

During constructing a short-term spline forecast, it was assumed that the heating of the wires was caused by the random nature of changes in the load of OPL and the invariance of other (primarily meteorological) conditions during the entire observation interval.

3 Conclusions

It has been experimentally confirmed that using the method of statistical spline functions in combination with new information technologies implemented using UAV and maintaining unchanged meteorological conditions, allows a short-term forecasting of the time interval during which the OPL can reach critical temperatures with a given probability.

References

1. Velimirovic, L., Peric, Z., Stankovic, M.: Approximation of the optimal compressor function combining different spline functions in segments. 2013 11th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services (TELSIKS), 16-19 October, 2013, Nis, Serbia. doi: 10.1109/TELSKS.2013.6704436
2. Kabore, P., Wang, H.: A B-spline neural network based actuator fault diagnosis in nonlinear systems. Proceedings of the 2001 American Control Conference. (Cat No.01CH37148) 25-27 June, 2001, Arlington, VA, USA. doi: 10.1109/ACC.2001.945873
3. Zaporozhets, A.A., Eremenko, V.S., Serhiienko, R.V., Ivanov, S.A.: Development of an Intelligent System for Diagnosing the Technical Condition of the Heat Power Equipment.

2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT), 11-14 September, 2018, Lviv, Ukraine. doi: 10.1109/STC-CSIT.2018.8526742

4. Zaporozhets, A., Redko, O., Babak, V., Eremenko, V., Mokiychuk, V.: Method of indirect measurement of oxygen concentration in the air. *Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu*, 2018, №5, pp. 105-114. doi: 10.29202/nvngu/2018-5/14
5. Yao, L., Guan, Y.: Minimum entropy fault diagnosis and fault tolerant control for the non-Gaussian stochastic system. 2016 American Control Conference (ACC). 6-8 July, 2016, Boston, MA, USA. doi: 10.1109/ACC.2016.7526753
6. Peng, X., Qian, G., Li, X., Gao, F.: Intelligent trip-out fault diagnosis of overhead transmission line. High Voltage Engineering, 2012, Vol. 38, Issue 8, pp. 1965-1972.
7. Li, P., Li, Q., Cao, M., Gao, S., Huang, H.: Time series prediction for icing process of overhead power transmissions line based on BP neural networks. Proceedings of the 30th Chinese Control Conference. 22-24 July, 2011, Yantai, China.
8. Kuts, Y., Eremenko, V., Monchenko, E., Protasov, A.: Ultrasound method of multi-layer material thickness measurement. *AIP Conference Proceedings 1096*, 2009, Vol. 1115. doi: 10.1063/1.3114079
9. Zaporozhets, A., Eremenko, V., Serhiienko, R., Ivanov, S.: Methods and Hardware for Dianosing Thermal Power Equipment Based on Smart Grid Technology, *Advances in Intelligent Systems and Computing III*, 2019, Vol. 871, pp. 476-492. doi: 10.1007/978-3-030-01069-0_34
10. Babak., V., Mokiychuk, V., Zaporozhets, A., Redko, O.: Improving the efficiency of fuel combustion with regard to the uncertainty of measuring oxygen concentration. *Eastern-European Journal of Enterprise Technologies*, 2016, Vol. 6, №8, pp. 54-59. doi: 10.15587/1729-4061.2016.85408
11. Klevtsov, S.: Identification of the state of technical objects based on analyzing a limited set of parameters. 2016 International Siberian Conference on Control and Communications (SIBCON), 12-14 May, 2016, Moscow, Russia. doi: 10.1109/SIBCON.2016.7491752
12. Manusov, V.Z., Orlov, D.V., Frolova, V.V.: Diagnostic of Technical State of Modern Transformer Equipment Using the Analytical Hierarchy Process. 2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe), 12-15 June, 2018, Palermo, Italy. doi: 10.1109/EEEIC.2018.8493904
13. Ming, J., Zhang, L., Sun, J., Zhang, Y.: Analysis models of technical and economic data of mining enterprises based on big data analysis. 2018 IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), 20-22 April, 2018, Chengdu, China. doi: 10.1109/ICCCBDA.2018.8386516
14. Dmitriev, S., Manusov, V., Ahyoev, J.: Diagnosing of the current technical condition of electric equipment on the basis of expert models with fuzzy logic. 2016 57th International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON), 13-14 October, 2016, Riga, Latvia. doi: 10.1109/RTUCON.2016.7763126