Estimation of aircraft power elements by areas of technical condition using clustering algorithm and statistical recognition method

Andrij Bolohin *[1 A]; Maksym Strela [2 A]

[*]**Corresponding author:** [1] The chief of research department, e-mail: bolas@ukr.net, ORCID: 0000-0003-0656-7799
[2] PhD student, e-mail: maxim.strela1991@gmail.com, ORCID: 0000-0003-4055-1600

[A] State Aviation Research Institute, Kyiv, Ukraine

Received: March 11, 2021 | Revised: March 26, 2021 | Accepted: March 31, 2021

DOI: 10.5281/zenodo.4924240

Abstract

In accordance with the concept of maintaining the serviceability of aviation of the Armed Forces of Ukraine, there was a problem of ensuring the serviceability of the fleet of aircraft and accurate classification of the technical condition of power elements of different types of aircraft for timely detection of their limit state. The purpose of the work is to find a system for classifying the technical condition – an image recognition system to solve this problem. The paper considers the issue of choosing an effective method of recognition, the main requirements for it, analyzes the chosen method of classification of technical condition (or recognition of images of technical condition), which is based on the statistical method of recognition. The method of clustering the technical condition of aircraft power elements is considered, a logical block diagram is constructed and the result of the program operation with the selected algorithm "FOREL – I" as a computer program in a graphical shell is given. The implementation of the method of classification in the graphical shell using the programming language python and auxiliary scientific and graphic libraries. The reference objects are selected, the main defining parameters that characterize the intensity of resource potential depletion are determined, which were divided into two images of the technical condition, namely "good" and "bad", and the control object. As an example, the technical condition of different chassis groups of fighter aircraft of the Armed Forces of Ukraine, which have approximately the same resource hours, and different statistics on the intensity of use during operation are analyzed.

Key words: pattern recognition, classification of technical condition, power element.

Introduction

The Military Doctrine of Ukraine (Military doctrine of Ukraine) states that the formation of national defense capabilities will be carried out by increasing combat capabilities, restoring serviceability, extending resources, and modernizing military equipment of the Armed Forces of Ukraine. The Air Force has a high priority in the development of a modern, mobile and powerful army, which in turn has led to planning to re-equip the fleet with new multi-purpose types of fighter aircraft generation 4 ++. It is planned to replace obsolete aircraft with new types as aircraft develop their resource potential. Determining the pre-destructive technical condition of any technical object outside the assigned indicators does not have a clear algorithm, and the extension of resource indicators of obsolete aircraft is performed individually. Based on this – determining the technical condition of the glider of obsolete aircraft and the responsible power elements of aircraft is an urgent task.

Material and methods

Many researchers have dealt with the problem of determining the reserves of inexhaustible resource potential of technical objects. All research in this area can be divided into 2 main types: power and statistical. Power studies of the technical condition are used based on the majority of the theory of resistance of materials. Researchers build their hypotheses around the force model of the stress-strain state, to which the loads received under operating conditions are applied, recalculate resource parameters using fatigue methods and compare with those established by the resource developer. These methods are most widely used in works such as (Strigius V. E., 2012; Vorobiev A.Z., Olkin B.I., Stebenev V.N., 1990; Brondz L.D., 1986; Serensen S.V., Kogaev V.P., Schneiderovnich R.M., 1975; Kogaev V.P., Makhutov N.A., Gusenkov A.P., 1985) and others. Statistical research usually uses the theory of classification, or the theory of pattern recognition. Based on significant statistics of the fleet of technical objects and the specific object to be diagnosed – using image recognition methods, the research results reflect the belonging of the studied object to one of the images of the technical condition, of which at least two. These methods have found application in works such as (Fomin Ya.A., 2003; Fukunaga K., 1979; Akoff R., Sasieni M., 1971; Zagoruyko N.G., 1972) and others.

The above methods have both advantages and disadvantages. Power methods provide better data on load, critical locations of objects, quantitative characteristics of the resource balance, etc. However, these methods do not take into account other operational factors (operating conditions, data of non-destructive testing, etc.) in contrast to statistical methods. To use power methods, it is necessary to have an algorithm for calculating the fatigue used by the developer, and embedded in it the calculated values of the loads that the technical object must receive during operation. Such data is usually a trade secret of the manufacturer, which complicates the calculations. For statistical methods, it is sufficient to have only the defining parameters that affect the depletion of the resource of the technical object, and at least two technical states (for example, “good” and “bad”). However, it is also necessary to have significant statistics on the operation of the facility and reference images of the technical condition, which are formed from the existing reference operating facilities. The article is devoted to the study of the technical condition using statistical methods of pattern recognition, as these data are more accessible than the algorithms for calculating the fatigue of the manufacturer. The task of the research is to analyze different methods of pattern recognition, determine the technical condition of the studied power element by the chosen method, and provide recommendations for the results.

Results and discussion

In operation, due to the fact that the development of the resource potential of different classes of technical objects, as a rule, is uneven – most often there are cases of need to classify the technical condition of one specific technical object. Hence the scientific and engineering task of determining the affiliation of a particular object to one of the images of the technical state, at least to the state of “good” or “bad”. Operational practice shows that the main role in ensuring flight safety in terms of reliability is the timely detection of the pre-destructive condition of a complex technical object (or its responsible part) to take timely measures to repair or replace it. In general, it can be assumed (Kogaev V.P., Makhutov N.A., Gusenkov A.P., 1985) that the object under study can take one of two mutually exclusive states: S_1 – state with a resource reserve, or “good” state and S_2 – state “bad”, or pre-destructive. Recognition is the assignment of the observed unknown state, given by the set X_n observations on its features $X_1, X_2, \dots, X_{\mathrm{p}}$,

$$\overline{X_n} = \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ X_{p1} & X_{p2} & \cdots & X_{pn} \end{pmatrix},$$

to one of the two mutually exclusive states S_1 or S_2.

Each column

$$\overline{x_i} = \begin{pmatrix} x_{1i} \\ x_{2i} \\ \cdots \\ x_{pi} \end{pmatrix} = \left(x_{1i}\, x_{2i} \cdots x_{pi}\right)^T, i = 1, 2, \dots, n$$

matrix $\overline{X_n}$ is a p-dimensional parameter of the observed values of p features $X_1, X_2, \dots, X_p$, reflecting the most important properties for recognition.

The set of features p is usually the same for all recognizable classes S_1 and S_2. If each class S_1 and S_2 is described by its own set of features, then the recognition problem is trivial, because the unambiguous assignment of the existing set of observations to a particular class is easily detected by the set of its constituent features. The general scheme of the system for recognizing the technical condition of the object is shown in Figure 1

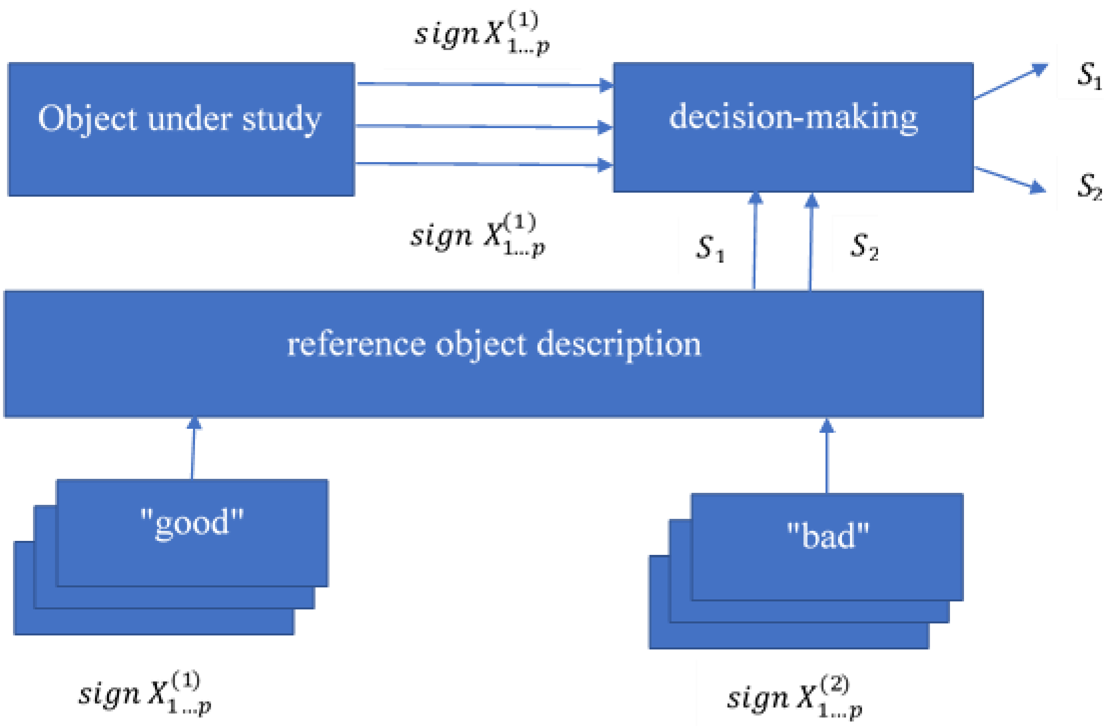

Fig. 1 – General scheme of recognition of the technical condition of the object

Thus, the considered problem of determining the technical condition of the observed object to one of the two images S_1 and S_2 should be described by the same set of features $X_1, X_2, \dots, X_p$ for all classes. In this case, the difference between the classes will be manifested only in the fact that different objects will have the same characteristics (quantitative, qualitative, etc.), and for any set of features $X_1, X_2, \dots, X_p$ it is possible to set rules according to which classes S_1 and S_2 correspond to the vector d_{12}:

$$d_{12} = \begin{vmatrix} d_1^{12} \\ \cdots \\ d_p^{12} \end{vmatrix}$$

states of p scalars, which are called between class distances and express the degree of difference in these classes of characteristics of these features.

The considered system of recognition of a technical condition has to provide the guaranteed reliability of recognition. There are several common systems and methods of recognition, such as:

deterministic (perceptron) recognition methods;

linguistic (syntactic) methods of recognition;

logical and algebraic methods of recognition;

statistical recognition methods.

Each of these methods has its advantages and disadvantages. But of all the above methods, only one can ensure the reliability of recognition – statistical methods (Fomin Ya.A., Savich A.V.,1993).

In the statistical method of recognition (Fig. 2) during training the reference description-estimations of multidimensional conditional densities of probabilities which have all information present in observations $x_1^1, \dots, x_m^1, \dots, x_1^p, \dots, x_m^p$ are formed. p and for all relationships between the features $X_1, X_2, \dots, X_p$. The estimate $\widehat{w}\left(\overline{x_1}, \dots, \overline{x_m / S_\iota}\right)$ is a random variable. The statistical theory of plausibility $\hat{L}(\bar{x}) = \hat{L}(\overline{x_1}, \dots, \overline{x_n}) = \frac{\widehat{w}(\overline{x_1}, \dots, \overline{x_m}/S_1)}{\widehat{w}(\overline{x_1}, \dots, \overline{x_m}/S_2)}$, is used to make a decision, which is a non-negative random variable obtained by the functional transformation $Z = \hat{L}(\overline{x_1}, \dots, \overline{x_n})$, which maps the points of the n-dimensional space of the samples on the real half-axis. Thus, to make a decision it is enough to use the value of one random variable – the statistics of the probability of $\hat{L}(\overline{x_1}, \dots, \overline{x_n})$, and not the value of each element of the sample $(x_1, x_2, \dots, x_n)$ separately, i.e. the likelihood ratio carries all the statistical information about the classes contained in this sample.

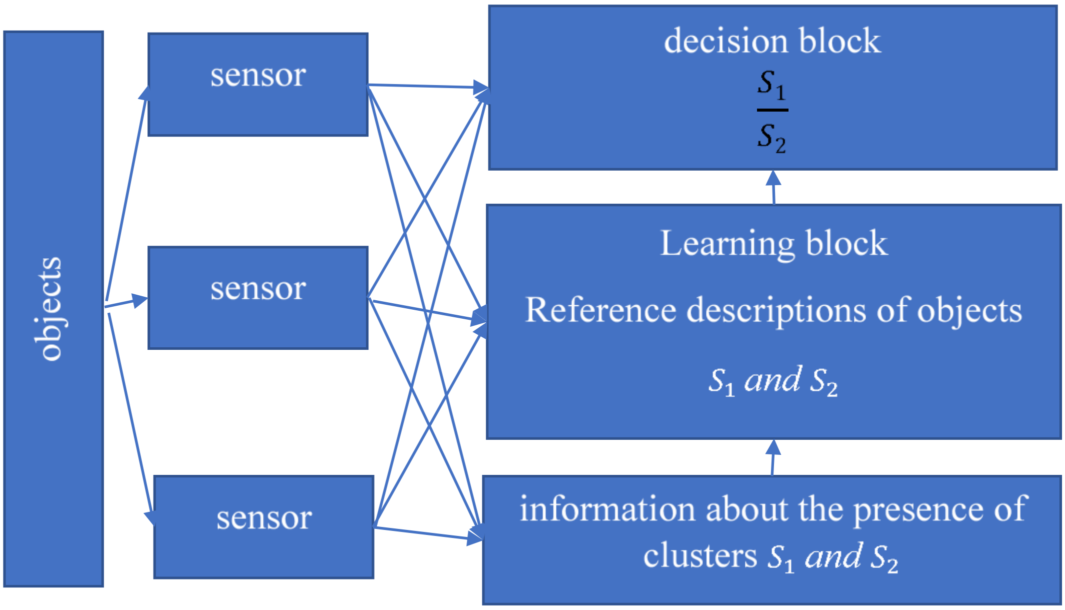

Fig. 2 – General scheme of statistical recognition

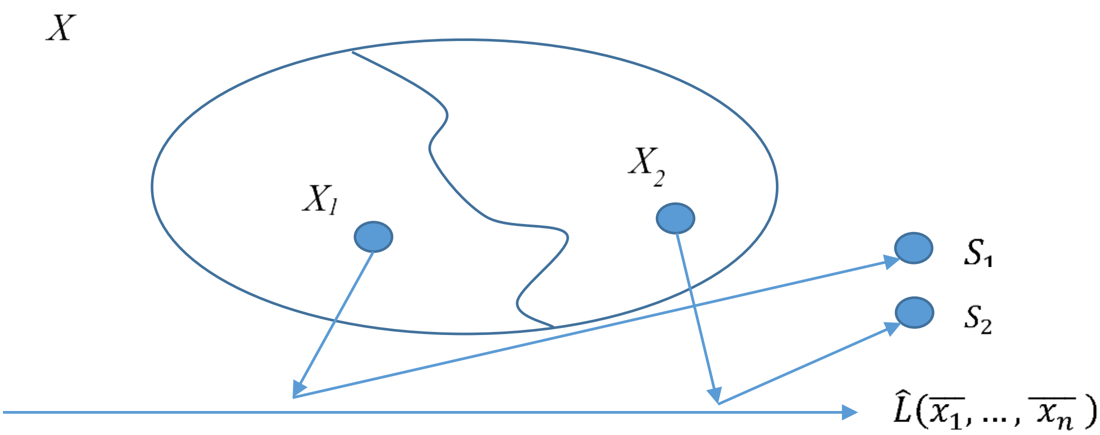

Fig. 3 – Reduction of observed data using plausibility ratio statistics

This statistic is called sufficient and leads to a reduction of the observed values: the mapping of the sample n-dimensional space X to the real positive half-axis (Fig. 3).

The surface in the n-dimensional sample space separating the space X into subspaces X_1 and X_2 is mapped to a point C on the axis $L \geq 0$. The decision at this stage consists in mapping the interval $0 < L < C$ to the point S_2 and the interval $L \geq 0$ to the point S_1.

Accordingly, the calculation algorithm is reduced to three stages.

The first stage: training.

According to the given diagnostic indicators from the educational sample estimates of vectors of means $\hat{a}_1$ and $\hat{a}_2$are calculated by formulas:

$$\hat{a}_1 = \frac{1}{m_1}\sum_{i=1}^{m_1} x_i^{(1)}$$

$$\hat{a}_2 = \frac{1}{m_2}\sum_{i=1}^{m_2} x_i^{(2)}$$

where m_1 and m_2 are the number of objects in each class, $x_i^{(1)}$ and $x_i^{(2)}$ are the values of the defining parameters of each class. Next is the calculation of the covariance matrices $\widehat{M}_1$ and $\widehat{M}_2$ and the total covariance matrix $\widehat{M}$ by the formulas:

$\widehat{M}$ за фрмулами:

$$\widehat{M}_1 = \frac{1}{m_1 - 1}\sum_{i=1}^{m_1}\left(\bar{x}_i^{(1)} - \hat{a}_1\right)\left(\bar{x}_i^{(1)} - \hat{a}_1\right)^T$$

$$\widehat{M}_2 = \frac{1}{m_2 - 1}\sum_{i=1}^{m_2}\left(\bar{x}_i^{(2)} - \hat{a}_1\right)\left(\bar{x}_i^{(2)} - \hat{a}_1\right)^T$$

In the second stage: decision making. At this stage, the estimate of the logarithm of the likelihood ratio $\ln \hat{L}$ by the formula

$$\ln \hat{L} = \frac{n}{2}(\hat{a}_1 - \hat{a}_2)^T \widehat{M}^{-1}\left[\frac{2}{n}\sum_{i=1}^{n}\bar{x}_i - (\hat{a}_1 + \hat{a}_2)\right] \langle comparison \rangle\, 0$$

where n is the number of studied objects, $\bar{x}_i$ – parameters of the studied objects. At the third stage: assessment of the reliability of the diagnosis. At the last stage, the recognition errors of the first and second kind α and β are calculated by the formula

$$\alpha = \beta = F\left(\frac{d}{\sigma_1}\right)F\left(-\frac{d}{\sigma_2}\right) + F\left(-\frac{d}{\sigma_1}\right)F\left(\frac{d}{\sigma_2}\right) + \left[\frac{1}{\sqrt{2\pi}}\frac{\sigma_1\sigma_2}{d(\sigma_1^2 - \sigma_2^2)}\right] \cdot \left[\sigma_2 exp^{-\frac{d^2}{2\sigma_1^2}}\right] \cdot \left[F\left(\frac{d}{\sigma_2}\right) - F\left(-\frac{d}{\sigma_2}\right)\right] - \left(\sigma_1 exp^{-\frac{d^2}{2\sigma_2^2}}\right) \cdot \left(F\left(\frac{d}{\sigma_1}\right) - F\left(-\frac{d}{\sigma_1}\right)\right)$$

where d is the distance of Mahalanobis,

$$d^2 = (\hat{a}_1 - \hat{a}_2)^T \widehat{M}^{-1}(\hat{a}_1 - \hat{a}_2)$$

$$\sigma_1^2 = \frac{1}{m_1} + \frac{1}{m_2}$$

$$\sigma_2^2 = \frac{1}{m_1} + \frac{1}{m_2} + \frac{4}{n}$$

$F(X)$ is the Laplace integral.

To ensure further research by the chosen method, it is necessary to form a training sample, which will create at least two technical images (clusters). To do this, you can use clustering methods.

There are many clustering methods. It is necessary to decide which method is better than others. Analyzing (Borisova I.A., Zagoruyko N.G., 2008) it is clear that the best recognition system is a person. Orienting in the external environment, a person continuously classifies, recognizes, identifies important features, predicts and more. But most of all a person uses the same universal psychophysiological function. This function consists of determining the degree of similarity between objects and phenomena. The algorithms of the FOREL family (Rormal Element) have a fairly high correlation with the human mechanism of perception of similarities and differences.

In (Zagoruyko N.G., 1972) the algorithm “ROREL – I” is considered, in which hyperspheres are used as decisive functions. The algorithm at the input takes the coordinates of the points and the specified radius of the hypersphere. At the output – clusters with the corresponding coordinates of the points included in them.

Algorithms of the FOREL family have certain advantages:

accuracy of minimization of quality functionality;

clarity of clustering visualization;

convergence of the algorithm;

the possibility of operations on the centers of clusters – they are known in the process of the algorithm.

They also have disadvantages:

poor applicability of the algorithm to the clustered sample;

instability of the algorithm (depending on the choice of the initial object, with a bad breakdown of the sample);

the need for a priori knowledge of the approximate width (diameter) of clusters.

The chassis riser was chosen as the object of study as one of the main and responsible power elements.

Applying the algorithm "ROREL – I" to the chassis risers – its shortcomings are almost completely eliminated, as the partitioning of the sample is quite successful for this algorithm and the need for a priori knowledge of the approximate width (diameter) of clusters can be neglected.

Analyzing the above data, a logical block diagram of the FOREL-I clustering algorithm is constructed (Fig. 4).

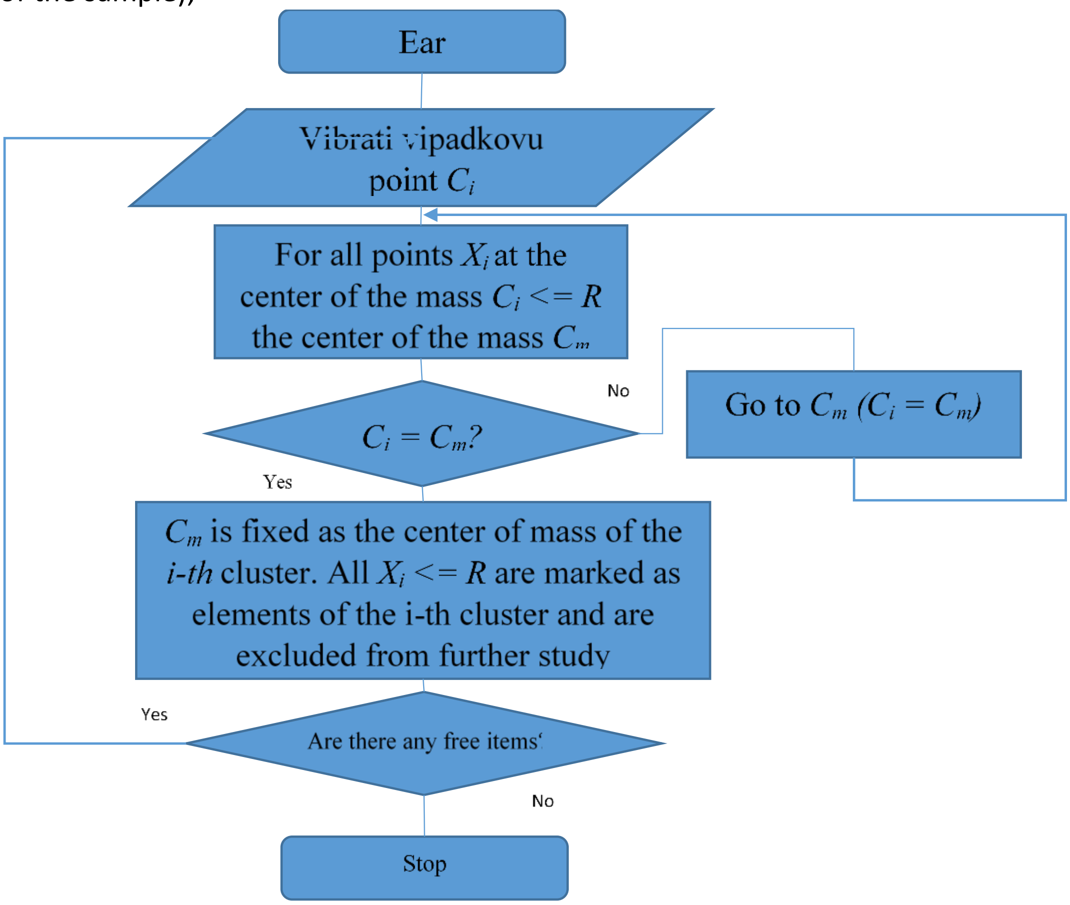

Fig. 4 – is a logical block diagram of the FOREL-I clustering algorithm

To implement the algorithm as a software product and write a logical part of the program, the Python programming language was chosen. The PySimpleGUI graphics library was selected to implement the graphical part of the program, the capacities of standard Python libraries were used to implement logical operations, and the "numpy" research library was added. The matplotlib library of scientific graphs and mappings and its "Pyplot" event handler were used to plot and visualize the results of the study.

The program receives the number of studied objects, the coordinates of the studied objects in the two-dimensional plane and the radius of the hyperspheres (in the scale of coordinates).

After processing the input data, the program returns the result of the research as a graph with points (objects under study) and the centers of mass, which are in the hyperspheres. Each hypersphere is a separate cluster, and the points in the hypersphere are objects that belong to the latter. The log window also displays the coordinates of objects that have reached a specific cluster.

9 aircraft (as groups of landing gear risers) were selected for research. The №5 aircraft is a control aircraft in which the landing gear riser was destroyed during operation. As the determining parameters, according to previous studies, the standard deviation of the overload during landing σ_{n_y} and the mathematical expectation of the landing mass $M[m_{\text{пос}}]$ were chosen (Table 1).

Table 1 – defining parameters of the selected aircraft

aircraft number	σ_{n_y}	$M[m_{\text{пос}}]$
1	0,196	18443
2	0,168	19250
3	0,075	18269
4	0,118	18303
5	0,11	19584
6	0,32	19927
7	0,12	19927
8	0,13	18154
9	0,12	18313

For better calculation and clarity, the method of transition to dimensionless values is used (Table 2).

Table 2 – defining parameters of the selected aircraft in dimensionless values

aircraft number	σ_{n_y} (dimensionless)	$M[m_{\text{пос}}]$ (dimensionless)
1	0,495652174	0,206810215
2	0,414492754	0,610916375
3	0,144927536	0,119679519
4	0,269565217	0,136705058
5	0,855072464	0,778167251
6	0,275362319	0,949924887
7	0,304347826	0,949924887
8	0,275362319	0,06209314
9	0,246376812	0,141712569

After entering data into the program, you must select the radius of the hypersphere. According to the algorithm – the radius of the hypersphere is selected by the method of successive approximations. The radius R is selected 0.47 dimensionless units.

After performing the calculations, the program returned the result in the form of a graphic image (Fig. 5) with clusters and points in them, in the log window more detailed information.

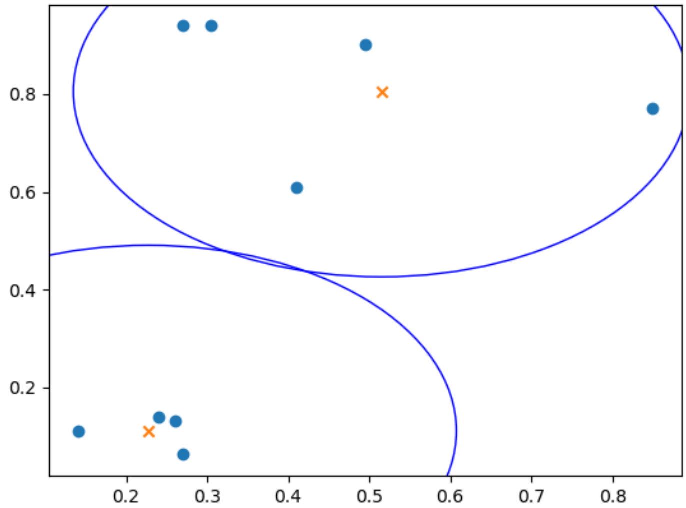

Fig. 5 – clustering of the studied objects

Figure 5 on the X axis are the values of SLE on the vertical overload σ_{n_y}, and on the Y axis are the values of the mathematical expectation of the planting mass $M[m_{пос}]$.

Analyzing the graph data, it is clear that aircraft (or groups of risers) are divided into two clusters of technical condition “good” and “bad”. The technical condition "good" included the following aircraft: 3, 4, 8, 9. The technical condition “bad” included aircraft 1, 2, 5, 6, 7. From this we can conclude that even with the same values of operating time (landing)) chassis risers accumulate damage in different ways, and their technical condition can vary significantly. The resulting cluster of risers “bad” characterizes the risers, which during operation received fatigue damage, which is close to the control riser. These risers need more thorough inspections and application of advanced methods of non-destructive testing, as well as to extend their service life or resource performance is not recommended. It also follows from the results that ensuring sustainable and economical depletion of resource indicators depends not only on the actions of engineering and technical staff, but also on the training of pilots.

Thus, having received a training and control sample of the selected power elements, it is possible to conduct further research by the method of recognizing images of the technical condition.

To use the method of technical condition recognition as a software product, it was decided to use the Python programming language. The PySimpleGUI graphics library was selected to implement the graphical part of the program, the capacities of standard “Python” libraries (math package) and the addition of the “Numpy” research library were used to implement logical operations.

The program receives a training sample as two classes S_1 and S_2 with certain defining features $x_1^1, \dots, x_m^1, \dots, x_1^p, \dots, x_m^p$ for each object X in each class. An object with certain features is also obtained. After the calculations, the program returns the result to the log window in the form of a solution, which indicates that the object under study belongs to a certain class with calculated reliability.

Class S_1 includes four aircraft (groups of risers) from the cluster “good”. Class S_2 includes five aircraft (groups of risers) from the cluster “bad”. The test aircraft (chassis group) was also selected. The determining parameters were: standard deviation of vertical overload during landing σ_{n_y}, mathematical expectation of landing mass $M[m_{пос}]$ and the number of landings from the beginning of operation $N_{ЗПЕ}$ (Table 1).

Table 1 – Data entered into the calculation program

Signs	S_1 "good"				S_2 "bad"					The investigated power element $\overline{X}$
	$X_1^{(1)}$	$X_2^{(1)}$	$X_3^{(1)}$	$X_4^{(1)}$	$X_1^{(2)}$	$X_2^{(2)}$	$X_3^{(2)}$	$X_4^{(2)}$	$X_5^{(2)}$	
$N_{ЗПЕ}$	997	691	1102	1379	1149	901	1470	1630	1420	1168
$M[m_{пос}]$	18269	18303	18154	18313	19835	19250	19584	19927	19914	18130
σ_{n_y}	0,075	0,118	0,12	0,11	0,196	0,168	0,32	0,12	0,13	0,16

A general view of the user interface with the entered data and the received calculations is shown in Figure 6.

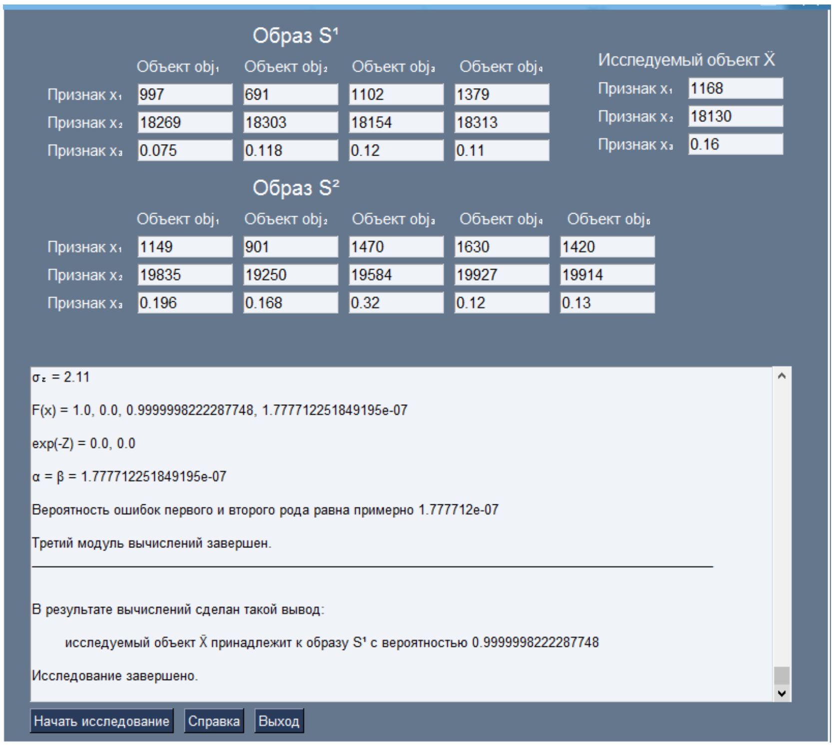

Fig. 6 – Working interface with entered data and received calculations

After performing the calculations, the logarithm of plausibility has the value $\ln\hat{L} = 35.65$ and according to the decisive rule $\ln\hat{L} = 35{,}65 > 0$ the investigated aircraft (chassis group) belongs to the class S_1 "good". The probability of reliability of the test aircraft (chassis group) to a certain class or image is calculated equal to 0.999.

Analyzing this distribution by a group of experts formed of five specialists of the institute, she agreed with the distribution and did not make any adjustments.

The obtained image of the power element in the following studies on the classification of technical condition is included in the general base of images "good" and "bad", which already make up 10 power elements, which in turn improves the quality of recognition in future studies.

Conclusions

1. From the analyzed methods of pattern recognition it is concluded that the statistical method allows to calculate the required reliability of recognition.

2. When applying the method to aircraft (chassis groups), it is determined that the method performs its functions well and provides correct results.

3. Therefore, based on the research, the options for the investigated aircraft (chassis group) may be as follows: to continue the assigned resource is allowed within the military unit, subject to strict control by promising structuroscopy in critical locations and prevent the operation of these power elements pilots 2-th grade and below; to continue the assigned resource is allowed in the framework of overhaul with complete defect of the riser, especially in critical places and allowed to operate without restrictions, subject to strict control by promising means of structuroscopy in critical places.

References

Akoff R., Sasieni M. Fundamentals of operations research. Moscow: Mir, 1971.

Borisova I.A., Zagoruyko N.G. SDX problem – the limit case of recognition problems of the combined type. *Problems of control and informatics*. 2008, №2

Brondz L.D. Aircraft resource technology. Moscow: Mashinostroenie, 1986, 184 p.

Fomin Ya.A. Diagnosis of the crisis state of the enterprise: Moscow: UNITI-DANA, 2003, 349 p.

Fomin Ya.A., Savich A.V. Optimization of recognition systems. Moscow, *Mechanical Engineering*, 1993.

Fukunaga K. Introduction to the statistical theory of pattern recognition. Moscow: *Nauka*, 1979.

Kogaev V.P., Makhutov N.A., Gusenkov A.P. Calculations of machine parts and structures for strength and durability. Moscow.: Mashinostroenie, 1985, 224 s.

Military doctrine of Ukraine: as of September 2, 2015, / President of Ukraine. Kyiv: President. type, 2015, 27 p. Library of official publications.

Proceedings of TsAGI. 1981, Issue. 2117.

Serensen S.V., Kogaev V.P., Schneiderovnich R.M. Load-bearing capacity and calculations of machine parts for strength. Moscow: Mashinostroenie, 1975, 488 s.

Strigius V. E. Methods for calculating the fatigue life of aircraft structural elements. Moscow: Mashinostroenie, 2012, 272 p.

Vorobiev A.Z., Olkin B.I., Stebenev V.N., etc. Fatigue resistance of structural elements. Moscow: Mashinostroenie, 1990, 240 s.

Zagoruyko N.G. Methods of recognition and their application. *Soviet Radio*, Moscow. 1972.

Zagoruyko N.G. Recognition methods and their application. Moscow: *Soviet Radio*, 1972, 208 p.