

Multiparameter assessment of accumulated fatigue damages level

Iskander Kuzeev

Department of Technological machines and equipment
Ufa State Petroleum Technological University
Ufa, Russia

Evgeniy Naumkin

Department of Technological machines and equipment
Ufa State Petroleum Technological University
Ufa, Russia

Yulia Shkel

Department of Technological machines and equipment
Ufa State Petroleum Technological University
Ufa, Russia

Abstract— The performed studies show that application of only one parameter for assessing structure service life does not allow accurately determining time of limit state occurrence. That is why emergencies, leading to the structure damage, right after assessment of technical state of the facility are present. The present paper proposes assessing the accumulated damage level by studying its properties divided into three groups: magnetic, electrical, mechanical. It is proposed to solve the problem of multiparameter assessment of the accumulated fatigue damage level by numerical methods based on data obtained during laboratory testing of metal samples.

Keywords— service life, structure, fatigue damages, material properties.

I. INTRODUCTION

Although many ways of improvement of oil and gas equipment safety were proposed, accidents are present, which indicates to the relevance of searching new methods for assessing level of accumulated damages in the metal of the structure. The case under discussion is oil storage tanks. Accidents with partial or complete damage of structures occur every year (figure 1). One of the reasons of such accidents is operation in excess of the service life period, as currently 60-80 % of the tank battery in Russia are worn out. According to the Federal Service for Environmental, Technological, and Nuclear Supervision (Rostekhnadzor) data, last year (2016) there were 5 accidents in tank batteries, including due to the operation of the tank with expired service life.

Literature contains description of different ways of classifying reasons of tank failure. Galeev V.B. divides reasons into the following 6 groups: insufficient quality control of construction and erection works; natural disasters; manufacturing defects; defects of installation; violation of operating instructions; objective factors. Rosenstein proposes division into direct (brittle and viscous fractures, welded joints defects, combined grading of rolled products, uneven precipitations) and indirect (design errors, violation of installation procedure, poor control, low quality materials) damages. .

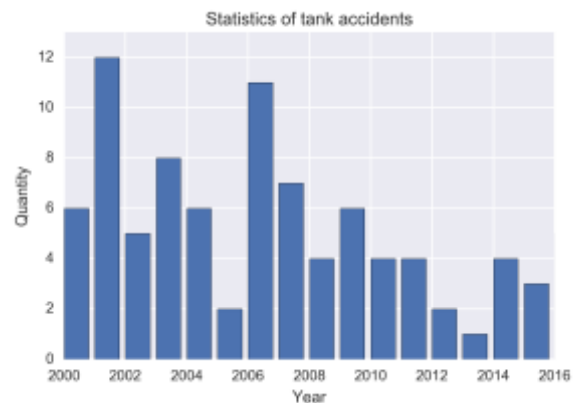


Fig. 1. Statistics of tank accidents in 2000-2016.

Tarasenko A.A. widens the classification and considers defects of vessels' structure [1].

It is possible to point out two main sources of the reasons of vessels failure: human factor (faults of design, installation, operation, maintenance), physical wear. On the basis of the abovementioned information and due to the fact that great number of vessels are operated in excess of the service life period (over 40 years), the aim of accurate determination of the remaining service life period is still urgent.

Assessment of remaining service life of vessels takes into account corrosive wear, low-cycle fatigue and fracture resistance. Assessment is carried out separately by elements: walls, bottom, cover. Values, characterizing mechanical properties of metal, which are determined by reference data or experimentally (yield point, break point and etc.), are taken into account when calculating the remaining service life period of the equipment. However, study of mechanical properties is permitted within the framework of complete examination of the vessel, which should be carried out at least once every ten years. This could lead to reassessment of the remaining service life period. Another factor impairing accuracy of calculated values is application of coefficients

obtained experimentally (corrosion factor, service factor and etc.). Calculations also include average, maximum values of parameters and do not take into account change of these values over the time (average thickness of the vessel element, petroleum product density). This affects the accuracy of assessment of the remaining service life period that leads to the necessity of developing a method without the abovementioned deficiencies. It could be noted that for this facility, the key factor in remaining service life period assessment is consideration of the material state of the vessel.

A lot of articles are devoted to the issues of the safety operation life period of equipment, including works of N.A. Makhutov, A.P. Gusenkov, V.Z. Patron et al. [2 - 5]. These studies point out necessity of searching new methods for assessment of limit state of technological facilities. It is possible to assess the accumulated damages level by studying its properties divided into three groups: magnetic, electrical, mechanical.

The parameter of the material state, which could be connected with the time period of equipment operation, is the level of accumulated fatigue damages (N_i/N_p) — a ratio of the number of load cycles as of the moment of parameter measurement to the number of cycles till damage. Zero refers to the metal state at the moment of its delivery, one – to the limit state (damage).

II. DATA ANALYSIS

As there are no analytical dependences uniting all groups of material properties, the problem of multiparameter assessment of the accumulated fatigue damages level could be solved by numerical methods on the basis of data obtained during laboratory testing of metal samples. In the present work, the studied material was steel 09G2s, as the material is often used in oil and gas industry.

The following metal characteristics were examined:

- range of gradient of magnetic field intensity,
- change of magnetic field intensity,
- magnetoelastic hysteresis loop area,
- change of latent crystallinity parameter,
- ultrasound velocity,
- actual stress,
- average stress,
- fracture resistance factor $K1c$.

Magnetic field changes its properties near the defect: the value and the direction of the intensity vector are changed. In papers by Naumkin E.A., there are graphs of dependence of the accumulated fatigue damages level on various magnetic characteristics (figures 2-4). Change of magnetic properties can be initiated not only by defects, but also by surface roughness, structural heterogeneity of metal, mechanical processing and other effects. That is why it is advisable to examine not only magnetic field intensity, but also a gradient of magnetic field intensity. The value of the gradient over the

defect is significantly greater than its value over heterogeneities of metal of the detail. Difference between steady state maximum and minimum values of magnetic parameters is the range of the gradient of magnetic field intensity. 2 zones can be observed in figure 3, depending on the range value: in the first one, there is a sharp drop and then its gradual decrease. It was found out that its decrease is observed while accumulating fatigue damages of the material.

The material structure can adapt to the changing ambient conditions. The change of the mechanism of material adaptation to the external effects allows following stages of its damage. The method for adaptive properties determination is multifractal parameterization of metal structures, subject to different types of impact. Crystallinity, being one of the most informative parameters, was chosen as the quantitative characteristic. It characterizes a degree of symmetry disturbance of the studied structure. Naumkin E.A. in his paper presents dependence of latent crystallinity parameter change on the level of accumulated fatigue damages of the material (figure 5).

One of the most dangerous processes of structure damage is brittle failure resulted from fracture initiation in the material.

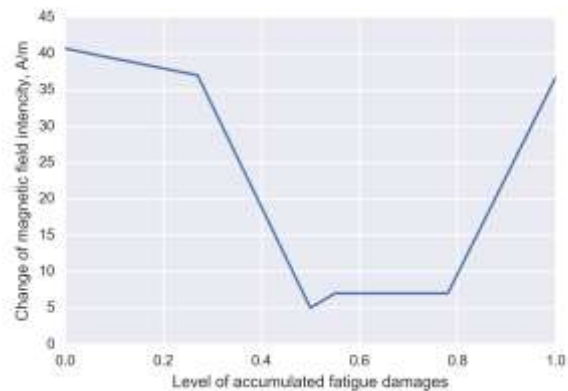


Fig. 2. Change of magnetic field intensity

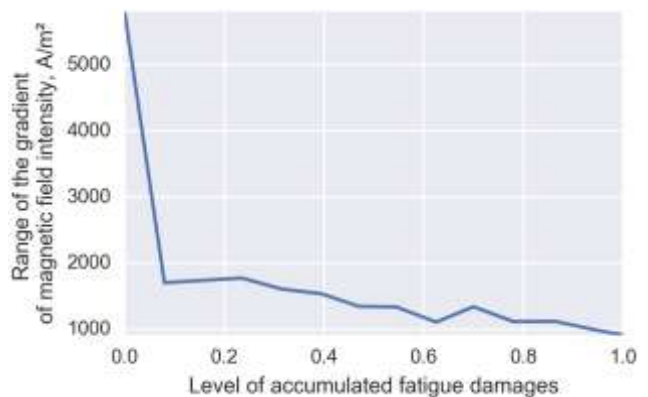


Fig. 3. Range of the gradient of magnetic field intensity

Fracture-like defects can form during installation and operation of technological equipment. Fracture development occurs under the influence of various loads, which can lead to emergency situations.

Property of the materials to prevent fracture development under mechanical and other effects is fracture resistance; its quantitative characteristic is critical stress intensity factor K_{Ic} . Its value corresponds to the critical value of external load under which fracture initiation to the material's body takes place. Dependence of the level of accumulated fatigue damages on this characteristic is provided in the paper by Kuzeev M.I (figure 6).

One of the methods to control equipment's material state is ultrasound diagnostics. Heterogeneity of the material's structure, stressed state affects the uniformity of acoustic properties. Naumkin E.A. in his paper shows the nature of interrelation between the rate of ultrasound velocity change and the level of accumulated fatigue damages (figure 7).

During the process of equipment operation, it accumulates stress. Stress raisers in the material can lead to deformations and damage of the facility. Bikbulatov T.R. determined dependences of actual and average stress values on the level of accumulated fatigue damages. The obtained results are presented in figure 8.

Performance of data analysis can be divided into following stages:

- Selection of characteristics (properties of material).
- Selection of mathematical models, choosing its parameters and training.
- Quality assessment of models.

To reduce expenses for examination and to minimize time expenditures, it is reasonable to determine informative characteristics showing the actual state of material. Moreover, selection of characteristics helps to improve quality and to simplify models, to reduce the risk of retraining and to solve the problem of data multicollinearity. At this stage, careful attention is paid to the degree of interconnection of dependent variables to each other and their contribution to the dependent variable.

After visualization of characteristics, certain peculiarities can be observed:

- Part of properties possesses high correlation dependence (figures 2, 5, 8).
- Presence of low informative characteristics (figure 3).

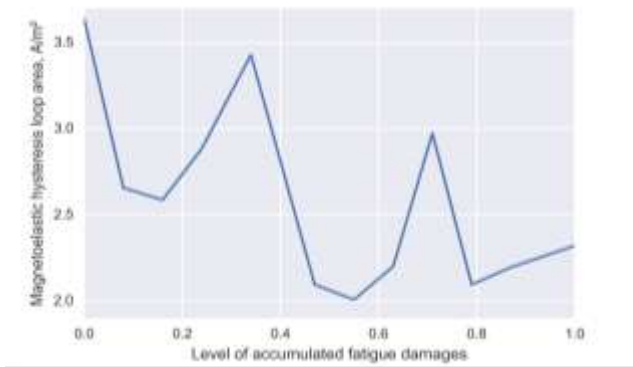


Fig. 4. Magnetoelastic hysteresis loop area

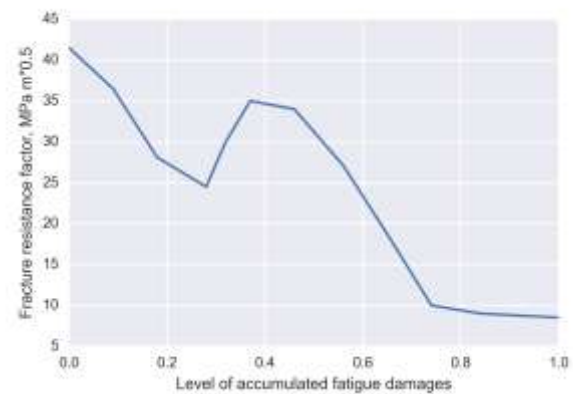


Fig. 6. Critical stress intensity factor

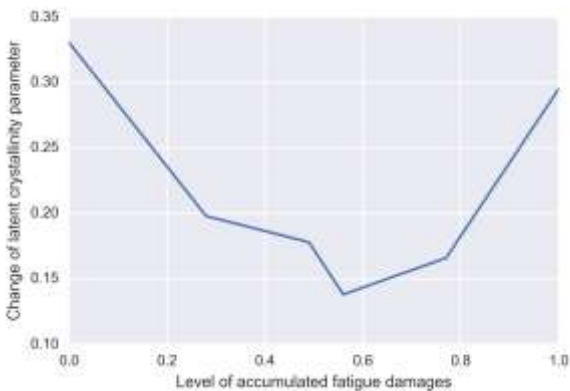


Fig. 5. Change of latent crystallinity parameter

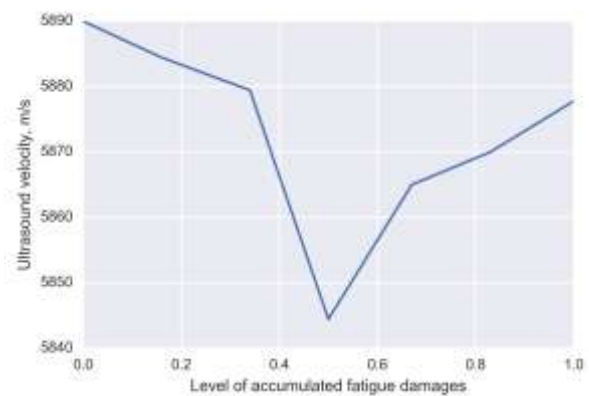
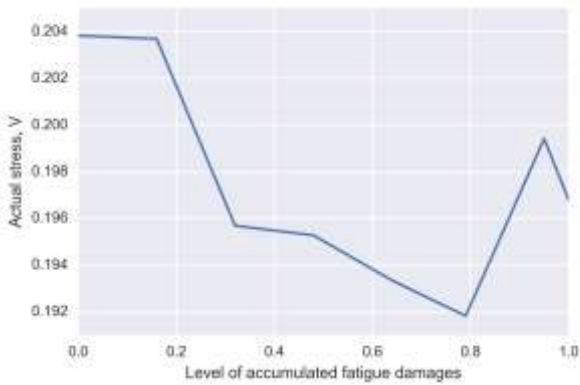
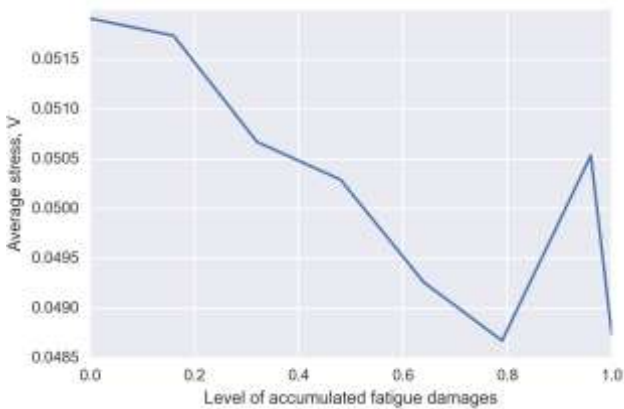


Fig. 7. Ultrasound velocity

b.



a) Actual stress



b) Average stress

Fig. 8. Dependence of actual (a) and average (b) stress on the level of accumulated fatigue damages

The authors will exclude low informative characteristics from further examination. Only one of correlated characteristics should be taken into account: with the greatest correlation relation with the level of accumulated fatigue damages.

First of all, one should scale up all values because many models are sensitive to such data (for example support vectors machine, nearest neighbors algorithm). This could be performed by deducting sample mean and dividing by sample standard deviation.

Selection of the method for correlation coefficient calculation depends on the type of scale the variables refer to (interval, ordinal) and on the type of data distribution. Moreover, there is the problem of multiple comparisons, which occur if it is necessary to make a family of statistical conclusions. Several approaches were developed (Bonferroni correction, Holm procedure, Benjamini–Hochberg procedure) to eliminate this affect. Taking into account small sample size and not making suggestion on the normality of data distribution, Spearman method and Benjamini–Hochberg corrections [6] were chosen as the method of correlation analysis. Final coefficients are illustrated in figure 9.

Based on obtained values, the following characteristics remained for further analysis: the magnetoelastic hysteresis loop area, the average stress value, fracture resistance factor K1c.

Performance of data analysis will be carried out within the framework of classification issue: the stage of material damage will be identified by previously determined characteristics.

Depending on the value of the accumulated fatigue damages level, let us point out 3 classes:

- 1) *Ni/Np up to 0.3 or «high level of equipment operation safety».*
- 2) *Ni/Np from 0.3 up to 0.7 or «medium level».*
- 3) *Ni/Np more than 0.7 or «low level».*

Taking into account proportional data distribution in relation to the characteristic of interest, classes are balanced out and there is no need to perform any additional operations. Unbalance of classes leads to the fact that models have tendency to refer new observations to the classes, presented by a greater number of training samples.

Final space dimension allows performing visual data analysis. Figure 10 shows the type of characteristics interdependence: red color refers to Ni/Np value more than 0,7, yellow – Ni/Np from 0,3 up to 0,7, green – Ni/Np up to 0,3. Quality of initial data depends on the accuracy of measuring devices and the initial state of the material. To take into account these values, let us add some noise in the initial curve area (figure 10 b).

The figure shows that dependence is not linear and is of complex nature. In the “red” zone, the main data is contained in interrelation of average stress and the magnetoelastic hysteresis loop area; in “yellow” zone – all characteristics, in “green” zone – interrelation of the fracture resistance factor and the magnetoelastic hysteresis loop area.

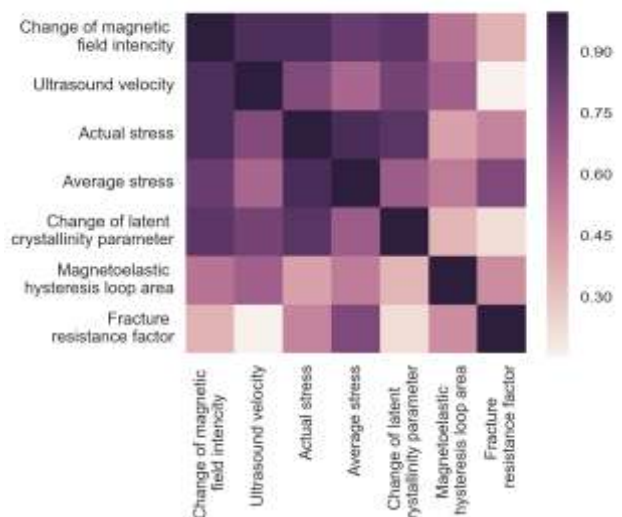
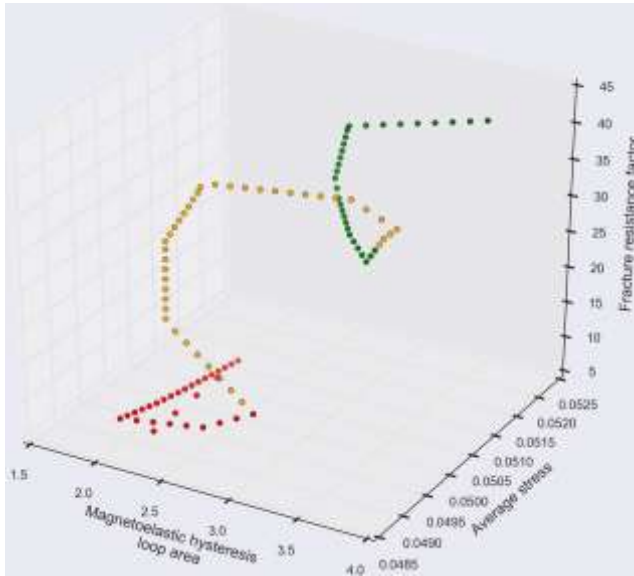


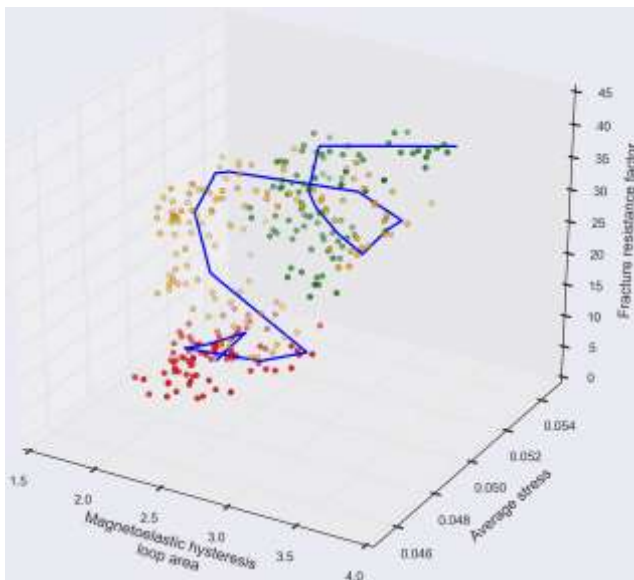
Fig. 9. Characteristic correlation degree

Based on the fact that real data possess some deviations from the curve illustrated in figure 10 (a), the authors will solve the problem from the point of view of classification issue.

It is necessary to determine models which will be used for classification performance. This task is not formalized because one and the same problem could be solved with the same accuracy by different methods. Rational choice of the type of the model could be justified, taking into account some criteria: invariability, capacity for generalization, interpretability. In this regard, interpretability provides the possibility of model understanding by the analyst, i.e. provision of conceptual meaning by the model coefficient.



a) Initial data



b) Data with noise

Fig. 10. Characteristics' space: a) initial data, b) data with noise

Capacity for generalization means that the algorithm can recover values of the goal variable on the basis of data not available at the moment of training; and invariability of classification methods provides the possibility of these methods application and operation under condition of noises and emissions presence in the data set.

Nine algorithms for classification performance were chosen for examination: nearest neighbors algorithm, linear support vectors machines, support vectors machines with radial-basic kernel, decision tree, random forest, naive Bayes classifier, linear discriminant analysis, quadratic discriminant analysis, AdaBoost meta-algorithm. [7]

Classification by the method of the k-nearest neighbors algorithm consists in reference of the object to the class, to which nearest objects of training sample refer. The distance criterion when all characteristics are objective is Euclidean. Nevertheless, "curse of dimensionality" should be taken into account, i.e. when in the high dimensionality space all objects are approximately equally faraway from each other, selection of the nearest neighbors becomes almost user-defined. In this case, the dimensionality is not high, that is why this method is applicable. This method requires data scaling. Typical tasks of this method are tasks of low dimensionality by the number of classes and variables.

The idea of Support Vector Machine consists in transduction of initial data into the space of higher dimensionality and search of separating hyperplane with maximum margin to the nearest data point. The disadvantage of this method is that not all samples are used for classification, but only those which are on the margins; the advantage of this method is that a quite small data set is enough. [8]

Linear SVM can solve only linearly divisible tasks. In case of linearly indivisible data, the nonlinear core, for example radially basic, can be used.

Naive Bayes classifier is based on the principle of the maximum of posterior probability. Likelihood functions are determined for the classified object and, based on the obtained results, posterior probabilities of the classes are defined. Object refers to the class for which posterior probability is maximum. In this task, densities of classes distribution are unknown and they will be determined by the training sample. Characteristics for this method should be independent. The disadvantage of this method is that the less the size of the sample is, the greater probability of retraining effect occurrence. The advantage is simplicity of realization and low calculation efforts.

Linear discriminant analysis is applied for determination of linear combinations of characteristics providing the best division of object classes. Conditions of discriminant analysis applicability are: characteristics are independent from each other and continuous; intragroup covariations are approximately equal; characteristics distribution - multivariate normal. In case any conditions are not met, good quality is provided by a less strict variant of the method - quadratic discriminant analysis. [9]

A solution tree is the most well-known algorithm in the theory of equipment reliability. The process of making decision, regarding reference of the object to some class, is presented in the form of the tree-like structure of principles. Leaves of the tree are classes to which the object can be referred. Limitations in relation to the tree depth are set to minimize the retraining effect. Another algorithm operates based on this one algorithm - Random forest.

The Random forest method includes construction of several independent solution trees. Determination of the object class is performed by means of voting: each tree refers the classified object to one of the classes and the class, which has earned the most votes, wins. [10]

Adaptive boosting is the meta-algorithm developing composition of algorithms. Each next classifier is based on the objects which can poorly be classified by previous classifiers. Updating the objects balances occurs at each iteration of the algorithm. The disadvantages are high sensitivity to noise and emissions; it requires quite large training samples. The main advantage is good generalization capacity. [11]

Quality of the classification is assessed by control by separate objects (leave-one-out, special case of cross-validation). The principle of the method lies in multiple divisions of the initial sample into two: training and test, where the latter consists of only one element. Algorithm adjustment by training subsample is carried out for each division, and then the algorithm average error is assessed by the control object. Final assessment is an average error value of all divisions.

The following values can be calculated for assessing quality of models: accuracy, precision, recall, F-score. Accuracy is the ratio between the number of correct positive answers and the sum of correct and false positive answers, completeness – the ratio between correct positive answers and the sum of correct positive and false negative answers, and F-measure is the mean-harmonic between accuracy and completeness. As the leave-one-out method is used, all these quality metrics give one and the same answer. [12]

III. RESULTS

Results of calculations in the python language with the application of the sklearn library showed that the highest accuracy was provided by algorithms based on trees. This is due to the fact that densities of distribution are unknown and the sample size is small. Interpretability of models plays an important role in the field of equipment operation safety. Chosen models possess this parameter. They also require preparation (standardization) of data. The algorithm of the decision tree has a tendency of retraining; that is why, random forest was chosen as the final algorithm. This algorithm possesses such advantage as assessment of significance of separate parameters of the model. The value of variable importance could be used for making decision in relation to the improvement of the method of value measurement. To improve the quality of classification, it is necessary to adjust parameters of the model by searching optimal values in the grid: by The best classifier informativity of parameters shows

that all parameters make contribution to the obtainment of a final solution.

IV. CONCLUSION

All efforts of predicting the fracture process based on linear dependencies lead to inexact results, because dynamics of the process is not taken into account and dynamics of the process is connected with the change of the mechanism of material adaptation to the external loads. When the material state comes close to the limit state, the nature of informative characteristics interconnection starts looking like a loop. Some linearity is observed only in certain periods of time. The obtained result explains those cases when accidents occur upon a short period of time after the last examination of the technical state of facility and assessment of remaining life time. This problem could be solved by means of numerical methods, using accumulated statistical data.

Consequently, the performed analysis shows that for determination of the level of equipment operation safety:

- it is required to control all three groups of material properties;
- one characteristic from each group of properties is sufficient for making decision regarding the remaining life time;
- random forest is the optimal classification algorithm for the considered case.

To improve the method, a wider property set of material should be considered. This approach could be applied for the structure as a whole, made of various materials, but it requires the corresponding dataset.

References

- [1] A.A. Tarasenko, P.V. Chepur, P.V. Gruchenkova, "Evaluation of technical condition of tanks with geometrical imperfections form wall", Oil Industry, vol. 6, pp. 118-121, 2017.
- [2] N.A. Makhutov, "Strenght, service life, and safety of machine systems", Journal of Machinery Manufacture and Reliability, vol. 43, No. 3, pp. 217-232, 2014.
- [3] N.A. Makhutov, "A criterion base for assessment of strength, lifetime, reliability, survivability, and security of machines and man-machine systems", Journal of Machinery Manufacture and Reliability, vol. 42, No 5, pp. 364-373, 2013.
- [4] Yu. S. Kovshova, I. R. Kuzeev, E. A. Naumkin, N. A. Makhutov, M. M. Gadenin, "The influence of quasi-static loading regimes on the strength of pressure vessels", Inorganic Materials, vol. 51, No. 15, pp. 1502–1507, December 2015.
- [5] Yu.S. Kovshova, I.R. Kuzeev, E.A. Naumkin, I.G. Fattakhov, "The process of initiation of potential fracture zones in the material of shell structures under static load", International Journal of Applied Engineering Research, Vol. 11, No 3, pp. 1630-1636, 2016.
- [6] G.W. Corder, D.I. Foreman, Nonparametric Statistics: A Step-by-Step Approach, 2nd ed., Wiley, 2014, pp.122-136.
- [7] S. Shalev-Shwartz, S. Ben-David, Understanding Machine Learning: From Theory to Algorithms, Cambridge University Press, 2014, pp. 202-265.

- [8] N. Cristianini, J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*, Cambridge University Press, 2003, pp.93-122.
- [9] G. McLachlan, *Discriminant Analysis and Statistical Pattern Recognition*, Wiley, 2004, pp.129-144.
- [10] L. Breiman, "Random forests", *Machine Learning*, vol. 45, No. 1, pp. 5-32, 2001.
- [11] Y. Freund, R. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting", *Journal of Computer and System Sciences*, vol. 55, pp. 119-139, 1997.
- [12] P. Flach, *Machine Learning: The Art and Science of Algorithms that Make Sense of Data*, Cambridge University Press, 2012, pp.81-104.