ISSN 0040-6015, Thermal Engineering, 2021, Vol. 68, No. 10, pp. 763–769. © Pleiades Publishing, Inc., 2021.
Russian Text © The Author(s), 2021, published in Teploenergetika.

STEAM-TURBINE, GAS-TURBINE, AND COMBINED-CYCLE POWER PLANTS AND THEIR AUXILIARY EQUIPMENT

Model of Emergency Conditions' Early Detection in Power Plant Equipment Based on the Least Potentials Method

A. A. Korshikova[a, *] **and O. M. Idzon**[a]
[a] *OOO Incontrol, Moscow, 115280 Russia*
**e-mail: korshikova@inctrl.ru*

Received December 2, 2020; revised January 12, 2021; accepted January 20, 2021

Abstract—A method is considered for detecting and predicting the abnormality in operation of power unit equipment by an example of a gas-turbine unit (GTU). A problem of detecting abnormality in operation is formulated as the mathematical problem of modeling an abnormality criterion taking the values from 0 to 1. It has been assumed that the predictive analytics methods can be effective for predicting the future state of process equipment based on the existing scope of measurements without any increase. It is assumed that, even when each individual measurement is within the range taken as the range of normal functioning, their cumulative dynamics enables us to judge a developing defect, i.e., about the transition of the diagnosed process equipment (DPE) to the zone of abnormal operation. To solve this problem, an approach is proposed based on calculating the value of the "abnormality indicator," which can be interpreted as a conditional potential created by points in a multidimensional space of indicators that characterize the state of equipment at the given time. By learning the model against the indicators that set the regions of states (the state of normal operation and the state for various kinds of fixed defects), one can then apply the trained model to determine the type of state: the closer the value of the abnormality indicator to the values inherent in a particular region of functioning, the greater the probability that the state of DPE corresponds to this region. It is shown that, due to certain objective circumstances, there is no practical possibility of training the model against the data obtained during abnormal operation with specific types of defects in DPE. This reduces the problem to adapting the method to the case when we have only the region of normal operation for learning the model. The proposed model was trained and tested during normal operation of the equipment. The test results indicate that the proposed method is consistent (i.e., it does not yield false positive response).

Keywords: process equipment, detection of abnormalities, predictive analytics, least potentials method, regions of normal and abnormal operation, state space, functioning indicators

DOI: 10.1134/S0040601521090032

In operating process equipment, events (accidents) that may adversely affect it or cause its failure inevitably occur. A model able to predict a future emergency would make it possible to timely take measures for eliminating it, thus helping to achieve more efficient use of process equipment. Development and investigation of such models is the subject of predictive analytics [1, 2].

The main idea of predictive analytics in the power industry, according to the authors, is that the occurrence of an accident may be predicted with some probability based on a continuous analysis of a set of data that characterize the functioning of the diagnosed equipment and are measured by standard monitoring instruments.

To determine the state of process equipment, various classification methods are used [3] that are based on multiclass and one-class classification.

The main idea of methods based on multiclass classification is to build classifiers of normal and abnormal data. Some examples of these methods include:

1. Neural network classifiers [4];

2. Statistical classifiers, in particular the Bayesian approach, analysis of distributions (as applied to the equipment of a power unit, it is studied in [5, 6]);

3. Machine learning (decision trees, SVM, etc.) [7].

Methods in which learning is based on precedents, able to diagnose specific types of accidents, but their main drawback is the complexity, and often the fact that data on defects and accidents that have occurred at the power unit, cannot be obtained. Therefore, from the standpoint of initial detection of possible deviations in the state of process equipment, the application of methods whose training does not require data on precedents seems advantageous.

The main idea of methods based on the one-class classification is to build boundaries of the region of normal data. All the data beyond this region are considered abnormal. Examples of such methods include:

1. Neural network classifiers (autoencoders, etc.) [8];

2. Machine learning (one-class SVM, one-class Fisher Discriminant, MSET [9], etc.).

However, the above methods are rather difficult to implement while they feature a high ambiguity of the result for power equipment.

The predictive diagnostics, which is focused on a single-class classification, can lead to technically incorrect results of the analysis of a specific case since increasing the distance of these data on this case from the boundaries of a known (previously determined) region of normal operation will not necessarily be an entry into the real region of abnormal data and, accordingly, a warning about an emergency. This may be an entry in a normal region that has not yet been recorded. Therefore, such a system can develop false messages about a defect, and, in order to reduce their probability, it is highly desirable to "retrain" the system against the verified data measured in the diagnosed equipment during a selected certain period of operation with known defects (hereinafter referred to as the training on a period).

The method described in this paper is intended for diagnosing process equipment in power units' equipment that have individual measurement instruments, which, on the one hand, can directly diagnose the occurrence of defects/malfunctions, but, on the other hand, cannot detect defects in the early stages of their development and also when these defects do not manifest themselves all in individual measurements until an accident occurs. The method gives a chance for early detection of defects by analyzing both the measured parameters themselves in their totality and the statistical derivatives of these parameters. In a particular case, the method does not require the mandatory availability of verified DPE periods with known faults or defects, i.e., learning may be carried out only against the normal periods of DPE operation. Checking the model for adequacy with this approach consists in the fact that the abnormality criterion resulting from the model testing against another (nonlearning) period of normal operation should be within the specified settings, i.e., should not give false positive predictions of abnormality.

FORMULATION OF THE PROBLEM

Initial data are archives of firmware complexes of automatic control systems of power units that store data arrays for a long period of operation of DPE: $\mathbf{P}(t) = [P_1(t),\ldots,P_m(t)]^T$, where $\mathbf{P}(t)$ is the vector of indicators at the time moment t; $P_1(t),\ldots,P_m(t)$ are the indicators of DPE operation at the time moment t; and m is the number of these indicators. In what follows, by the indicators of DPE functioning, we mean not only individual measurements but also their statistical derivatives (average values over a period, variances, correlation functions, etc.). Vectors $\mathbf{P}(t_r)$, taken at time moments t_r ($r = 1, 2, \ldots$), form points in the m-dimensional space. In addition, a finite set of precedents (accidents, failures, defects) is given, each of which can be classified, i.e., attributed to a certain type of DPE defect. Each precedent is assigned a time stamp of its detection, $t_{i,j}$, and $t_{i,j}$ is the time moments of detection of the ith precedent of class (defect type) j ($j = 1, \ldots, l$; l in the number of classes; $i = 1, \ldots, k$; k is the number of precedents in class j).

We may assume that the domain of space with the points of normal functioning of DPE differs from the domains of space that are formed by the points of functioning of the same object in the periods of time that coincide with the time moments of precedent occurrence, $t_{i,j}$, just as domains of the precedent space belonging to different classes differ from each other. Moreover, for the process equipment of power units, it is quite reasonable to consider these sets to be linearly separable. The period of equipment operation of equipment with a duration from $(t_{i,j} - \Delta t_{b,j})$ to $(t_{i,j} + \Delta t_{a,j})$ can be called a period of abnormal operation in class j. Here, $\Delta t_{b,j}$ is the period of time before the precedent, when the presence of a developing defect begins to affect the array of indicators of the DPE functioning; $\Delta t_{a,j}$ is the period of time from the occurrence of a precedent to the elimination of its consequences.

The set of all available precedent descriptions is called a training sample. It should be used to find general dependencies enabling us to determine the abnormal functioning of DPE with an indication of the class of abnormality. In doing so, a generalized criterion of abnormality, which offers the possibility of a precedent to be determined with some certainty based on its change, should be constructed.

CONSTRUCTION OF A MODEL FOR THE GENERAL CASE

Let several sets from the n-dimensional Euclidean space of indicators of the DPE state be given. Each point of this space is formed by the values of the indicators obtained at certain time intervals. The first set of points was composed in the period when the equipment was healthy. Other sets are formed during periods of abnormal operation of the equipment operation, and each of them is associated with a certain type of abnormality characterizing a specific failure. The sets of normal and abnormal functioning are subsequently called base sets. Figure 1 shows the distance in the two-dimensional space (P_1 and P_2 are the first and second coordinates of the vector) from the point of the process at the current time moment t to the points of the subsets of normal (subscript "norm") and abnormal (subscripts "a1" and "a2") functioning of DPE; N_{norm}, M_{a1}, M_{a2} are the number of points in the subsets of normal and abnormal functioning of DPE, and the

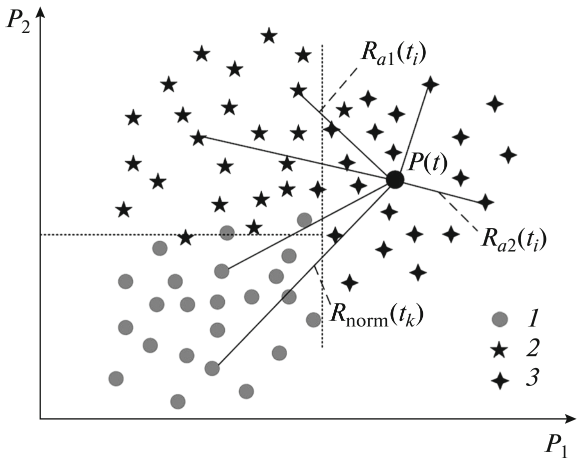

Fig. 1. Space of equipment functioning parameters. Training set—points of the subset of DPE functioning: *1*—normal; *2*—abnormal with type 1 defect; *3*—abnormal with type 2 defect. Dashed curves illustrated the linear separation of the normal and abnormal regions.

values of N_{norm}, $M_{\text{a}1}$, $M_{\text{a}2}$ correspond to the duration of the normal and abnormal operation periods.

Each point from the base sets can be associated with a function similar in form to the electric potential, i.e., having a maximum at this point and decreasing in all directions from it (thus, the point will appear as if it were a source of potential). For example, such a function may be,

$$\varphi(R) = \frac{W}{1 + R^2},$$

where W is the weight of a point; R is the distance between the source point and the point at which the potential is calculated, which is calculated by the formula

$$R(t) = \sqrt{\sum_{i=1}^{m} \left[P(t)_i - P_i\right]^2};$$

here $P(t)_i$ is the ith coordinate (DPE attribute) of the tth point of the m-dimensional set (points at a time moment t) with reference to which potential $\varphi(t)$ of a point with coordinates P_i is determined.

The value of function $\varphi(t)$ at each point of the space can be considered a measure of the proximity of this point to a source point.

Let the sources be the points of the space of normal functioning. Then the average potential created at a given point of the space by all points, i.e., the total potential divided by the number of points (potential of normal functioning) will characterize the proximity of this point to the entire overall base set:

$$\varphi_{\text{norm}}^{\text{av}} = \frac{1}{N_{\text{norm}}} \sum_{t=1}^{N_{\text{norm}}} \varphi\left[R(t)\right].$$

Similarly, the potentials of the jth base set can be determined at the same point of the space, where j = 1, ..., l is the type of a failure (l is the number of failure types (classes of defects) in the training sample):

$$\varphi_j^{\text{av}} = \frac{1}{M_{\text{a}j}} \sum_{t=1}^{M_j} \varphi\left[R(t)\right].$$

Here, $M_{\text{a}j}$ is the number of points forming the base set of abnormal operation with a Class j defect.

Since the sets of normal and abnormal functioning are assumed to be linearly separable in the calculation, it will be natural to place the point in the base set whose potential at this point is the greatest.

However, as already noted, it is extremely difficult to obtain data on abnormal classes, and it cannot be done at all in most cases. Therefore, a modification to the stated general approach will be considered below, namely, the construction of a model based on the data only for the normal operation period.

CONSTRUCTION OF A MODEL BASED ON THE DATA ONLY FOR THE NORMAL OPERATION PERIOD

In this particular case, there is only one basic set, i.e., the set of normal functioning $\mathbf{N}_{\text{norm}}$ with the number of points N_{norm}. The learning algorithm is then modified as follows. For each kth point of the base set, where $k \in (1, N_{\text{norm}})$, the average potential of this point with reference to the rest $N_{\text{norm}} - 1$ points should be calculated by the formula

$$\varphi_k^{\text{av}} = \frac{1}{N_{\text{norm}} - 1} \sum_{t=1}^{N_{\text{norm}}} \varphi\left[R(t)_k\right],$$

where $t \in (1, N_{\text{norm}})$; $t \neq k$; $R(t)_k$ is the distance from the tth to kth point.

Then, the minimal potential of the base set $\mathbf{N}_{\text{norm}}$ is calculated

$$\varphi^{\min} = \min\left[\varphi_k^{\text{av}}, k \in (1, N_{\text{norm}})\right].$$

In the run-time mode, the potential at each new time moment will then be determined by the following formula:

$$\varphi^{\text{av}}(t) = \frac{1}{N_{\text{norm}}} \sum_{t=1}^{N_{\text{norm}}} \varphi\left[R(t)_k\right],$$

where t is the current time moment in time and $R(t)_k$ is the distance from the point formed by the values of indicators at the current time moment t to the kth point in the training period.

At this point, the construction of the model can be considered finished. If it now turns out that the potential $\varphi^{\text{av}}(t)$for the points formed by the current values of DPE parameters with reference to the points of the training set is less than the value of $\varphi^{\min}$, we can assume that the DPE has entered the abnormal functioning zone. By monitoring the tendency for a further decrease in $\varphi^{\text{av}}(t)$, we can predict the risk of an accident. The calculated potential varies within (1, 0) and, in fact, is a criterion for abnormality.

When using the model in practice, it is possible to set the potential threshold setting and such a time period T_{a} that, if $\varphi(t) < \left(\varphi^{\min} - \Delta a\right)$ for a period of time longer than T_{a}, an alarm will be triggered (Δa is the maximum allowable decrease in the conditional potential). After that, a process engineer must assess the adequacy the model behavior and, if the prediction and the actual state of DPE coincide, take necessary measures to eliminate the failure. Should the model prediction be false, then the model should be retrained by expanding the space of normal operation by adding new points to it.

ALGORITHM FOR THE CALCULATION OF THE CRITERION FOR ABNORMALITY OF DPE OPERATION

Since the proposed method features a computational load increasing exponentially with an increase in the number of points in the base set, we have to impose limits on their number without losing the points that control the configuration of this set.

Formation of the Base Set of Points in the Space of Indicators

1. Assume that N_{norm} is the total number of points (time moments) derived from the archived data in the period of normal operation of DPE (set $\mathbf{N}_{\text{norm}}$).

2. N_F is the number of points in the specified period, which have no unreliable indicators (see below the section Accounting for unreliability).

3. N_L is the limiting number of points at which all calculations required by the method are completed within an acceptable time.

4. m is the number of indicators (coordinates of the base set space), $m << N_L$.

Then:

if $N_F \le N_L$, the base set is identical to the set of archived data on normal functioning of DPE;

Otherwise, the base set should be formed as follows to include in it

1. All points of the set $\mathbf{N}_{\text{norm}}$ having $\min_{t=1,\dots,N_{\text{norm}}} P(t)_i$, i.e., the minimal value of the coordinate (indicator) P_i of all N_{norm} points in the set $\mathbf{N}_{\text{norm}}$, $i = 1, \dots, m$.

2. All points of the set $\mathbf{N}_{\text{norm}}$ having $\max_{t=1,\dots,N_{\text{norm}}} P(t)_i$, i.e., the maximum value of the coordinate (indicator) P_i of all N_{norm} points in the set $\mathbf{N}_{\text{norm}}$, $i = 1, \dots, m$;

3. Of the remaining $N_{\text{norm}} - 2m$ points, we should take additional points at equal intervals of $\Delta t = \dfrac{N_{\text{norm}} - 2m}{N_L - 2m}$ to increase their number to N_L.

The final number of points in the base set can be denoted by N_B. Then it follows from the above-mentioned that

$$N_B = N_F \quad \text{for} \quad N_F \le N_L;$$
$$N_B = N_L \quad \text{for} \quad N_F > N_L.$$

Determining Weights W(t)

The weight is an attribute of a point, i.e., an attribute of each of the time moments of the base set. It seems reasonable to give greater weight to those time moments at which normal operation can be assumed with a higher degree of confidence than at other points in the historical data record. These can be, for example, the periods of DPE operation immediately after the repair.

Developing Indicators

The indicators that characterize the functioning of DPE are, first of all, measured parameters of the DPE.

To enhance the ability to capture the effects of abnormal functioning of DPE, it is possible to supplement the list of indicators with derivatives of the measured parameters, for example, their variances $D(t)$. It is assumed that such an extension may be useful when a developing defect initially affects not so much the parameters themselves or their combination but their variability:

$$D(t) = \frac{1}{T_w} \sum_{t-T_w}^{t} \left[P(t)_{\text{av}} - P(t)\right]^2,$$

where $P(t)_{\text{av}} = \dfrac{1}{T_w} \sum_{t-T_w}^{t} P(t)$. Here, T_w is the width of the window for determining the statistical characteristics of the time series; t is- the discrete time moments at which the values of DPE indicators are recorded. The width of the window can be chosen as equal to the longest period in the spectrum of fluctuations of the DPE parameters on a daily segment of the stationary mode in the normal period of DPE operation, for example, $T_w = 8$ h.

The values of certain parameters are kept at their setpoints by the action of automatic control systems. In this case, with the development of a defect that theoretically affects such a parameter, in a fairly wide range, the parameter proper will not change significantly. However, the position of the control element, which keeps this parameter at the setpoint, will change. This means that it is advisable to include the readings of the position indicators of the regulators in the list of DPE indicators.

Filtering of Interference by Averaging over an Interval

The time series of indicators of DPE functioning are known in the form of a discrete series of values taken at regular intervals.

In statistics, there are three components of a time series: trend, seasonal fluctuations, and random component. To solve this problem, it is the trend that is of interest, and the random component is harmful noise, which should be filtered by averaging over the interval, while seasonal fluctuations for most features can hardly be found.

Averaging the indicator over a time interval:

$$\overline{P(t)_l} = P(t)_i \quad \text{for} \quad t = 1,\ldots,T_m;$$

$$\overline{P(t)_l} = \frac{1}{T_m}\int_{t-T_m}^{t} P(t)_i\,\mathrm{d}t \quad \text{for} \quad t > T_m.$$

Here, T_m is the averaging interval; $i = 1, \ldots, m$ is the number of indicators.

Since data in the archives of the firmware complex of the automatic process control system are stored with discreteness in time, the integral is, in fact, the sum of the values of each indicator.

The value of T_m is determined by the period of interference in the frequency spectrum of time series of process parameters. Based on the available archive data, it seems reasonable to assume $T_m = 10$ min.

Normalizing Data Samples

The need to normalize data samples stems from the different nature of the applicable indicators, which can vary in a wide range and, being diverse in physical meaning, differ from one another by several orders of magnitude. For example, the temperature is expressed in three digits, while the pressure difference can be expressed in one digit.

The operation of models with such indicators will be incorrect since the imbalance among the values of indicators can lead to completely inadequate results.

When calculating the distances between points or vectors in practice, Z-scaling of the coordinates of these vectors is most often used:

$$\overline{P(t)_l} = \frac{P(t)_i - P(t)_{i\text{av}}}{\sqrt{D_i}},$$

where

$$P_{i\text{av}} = \frac{1}{N_B}\sum\nolimits_{i=1}^{N_B} P(t)_i;$$

$$D_i = \frac{1}{N_B}\sum\nolimits_{i=1}^{N_B}\left[P_{i\text{av}} - P(t)_i\right]^2;$$

N_B is the total number of points of the base set; $t = 1,\ldots,N_B$; $i = 1,\ldots,m$.

The normalization reduces all the numerical values of the input indicators to the same range of their variation, i.e., to a certain narrow range. This enables them to be brought together in a single model and ensures proper operation of computational algorithms.

Considering the Unreliability of Measurements

Construction of an adequate model requires that unreliable measurements be filtered out using one or another method. Since all modern firmware systems used as data sources for any predictive model provide information not only about the values of measured parameters but also about their reliability (so-called validity bit), then unreliable values can be processed using this information thereby avoiding learning errors if these measurements belong to the training period (base set), or false predictions, when these measurements and the DPE performance indicators based on them are taken when the model is running in run-time mode.

When forming the base set, all the points whose coordinates includes at least one unreliable indicator, are rejected.

When using the model in the run-time mode, the best solution is to replace the value of an unreliable parameter with its last valid value.

RESULTS OF EXPERIMENTAL STUDIES

To improve the model sensitivity and find the process assembly with a developing defect, the entire set of taken measurements was divided into groups in accordance with the technology.

The experimental studies were carried out on historical data on the operation of a gas turbine unit (GTU) recorded over 3 years. At each time moment, with an interval of $\Delta t = 10$ min, $m = 45$ indicators of the GTU functioning (including active power, bearing vibration, temperature of bearing white metal, etc.) were recorded with an interval of $\Delta t = 10$ min. Figure 2 shows a fragment of the initial data.

The dependence of the calculated criterion φ for abnormality of the gas turbine's operation (curve *1*) is presented in Fig. 3. The horizontal line indicates the minimum value of potential obtained in learning the model. Based on the expert opinion and the assessment of each individual measurement, one more period of normal equipment operation was selected as the test period. It follows from Fig. 3 that the abnormality indicator diagnoses the normal state of the studied object during the test period of operation. Thus, in the test period, the developed method

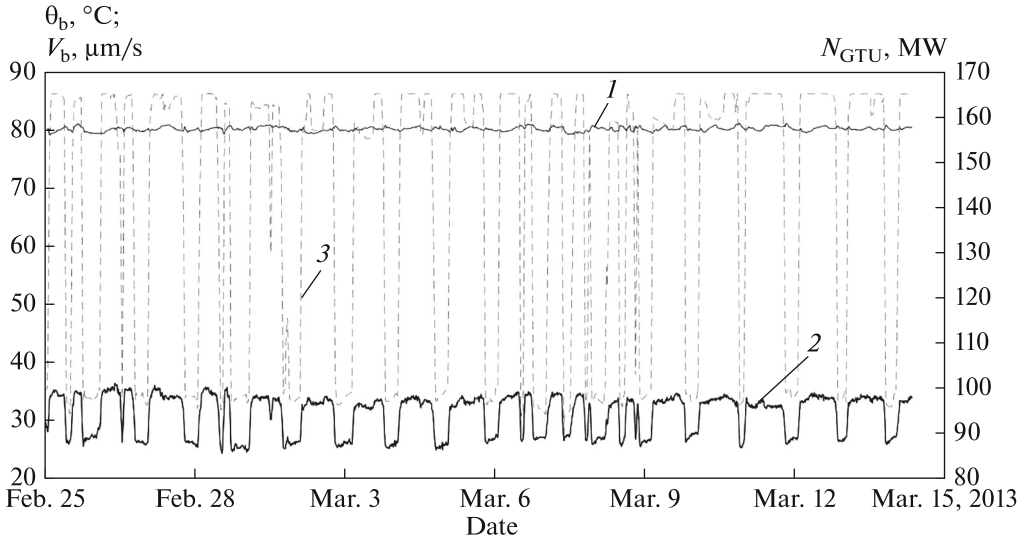

Fig. 2. Gas turbine performance in the period between February 25, 2013, and March 15, 2013, characterized by process engineers as a period of normal functioning. *1*—Bearing temperature θ_b; *2*—bearing relative vibration V_b; *3*—GTU active power N_{GTU}.

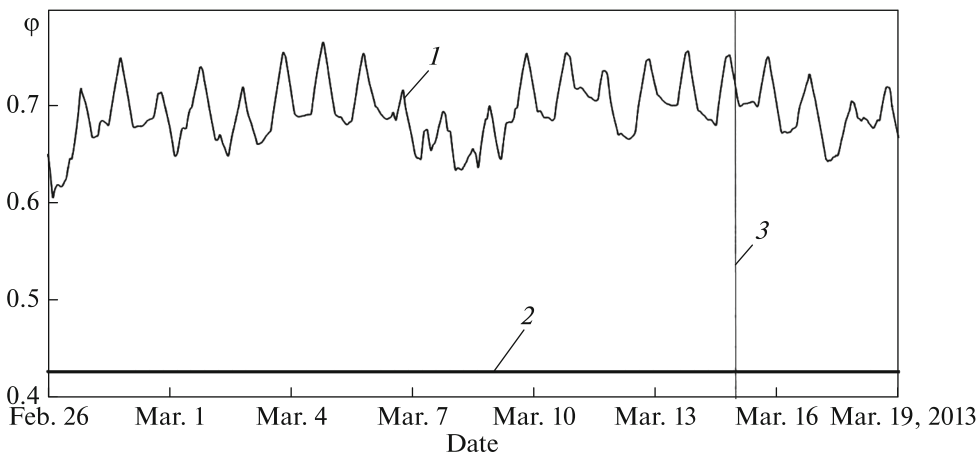

Fig. 3. Criterion (potential) of abnormality of the gas turbine operation, φ, in the period between February 26, 2013, and March 19, 2013, including normal (until March 15, 2013) and test (after March 15, 2013) operation periods. *1*—Calculated criterion (potential) of abnormality; *2*—earlier determined minimum potential; *3*—March 15, 2013.

demonstrates consistent results of determining the normal or abnormal functioning of DPE.

CONCLUSIONS

(1) The proposed method featuring simplicity and clarity has demonstrated the adequacy and consistency of the results of determining the normal or abnormal functioning of DPE based on the available archived data.

(2) Should appropriate initial (archive) data on recorded failures in DPE be available, the model can be trained to search for them. However, even in the absence of information sufficient to train the model to search for specific failures (which is almost always the case for all power units), the model can be used successfully having been trained only against the data obtained during normal operation of DPE.

(3) The disadvantages of the method include exponentially increasing computational complexity as the training data array increases. However, this problem is present in almost all available predictive analytics models, and this paper proposes a method for solving it.

(4) The presented method, as all other predictive analytics methods used for diagnostics of process equipment, is general with respect to the type of this equipment. The only condition is that the equipment must have a number of measured parameters that

characterize the equipment operation and are sufficient for diagnostics.

(5) The proposed method requires additional testing against other archive samples containing (that is desirable) the period of abnormal operation verified by an expert in both offline and online modes on the current data, optimization of the training dataset, improvement of predicting the time to attain the preemergency and emergency technical state, and determination of probabilistic estimates of the fact of the occurrence and development of a defect in time.

REFERENCES

1. E. Siegel, *Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die* (Wiley, Hoboken, N. J., 2016). https://doi.org/10.1002/9781119172536
2. M. A. Waller and S. E. Fawcett, "Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management," J. Bus. Logist. **34**, 77–84 (2013). https://doi.org/10.1111/jbl.12010
3. S. S. Khan and M. G. Madden, "One-class classification: Taxonomy of study and review of techniques," Knowl. Eng. Rev. **29**, 345–374 (2014). https://doi.org/10.1017/S026988891300043X
4. A. G. Trofimov, K. E. Kuznetsova, and A. A. Korshikova, "Abnormal operation detection in heat power plant using ensemble of binary classifiers," in *Neuroinformatics 2018: Advances in Neural Computation, Machine Learning, and Cognitive Research II*, Ed. by B. Kryzhanovsky, W. Dunin-Barkowski, V. Redko, and Y. Tiumentsev (Springer-Verlag, Cham, 2019), in Ser.: *Studies in Computational Intelligence*, Vol. 799, pp. 227–233. https://doi.org/10.1007/978-3-030-01328-8_27
5. A. A. Korshikova and A. G. Trofimov, "Model for early detection of emergency conditions in power plant equipment based on machine learning methods," Therm. Eng. **66**, 189–195 (2019). https://doi.org/10.1134/S0040601519030042
6. A. A. Korshikova and A. G. Trofimov, "Predictive model for calculating abnormal functioning power equipment," in *Cyber-Physical Systems: Industry 4.0 Challenges*, Ed. by A. Kravets, A. Bolshakov, and M. Shcherbakov (Springer-Verlag, Cham, 2020), in Ser.: *Studies in Systems, Decision and Control*, Vol. 260, pp. 249–259. https://doi.org/10.1007/978-3-030-32648-7_20
7. H. Lingjun, R. A. Levine, J. Fan, J. Beemer, and J. Stronach, "Random forest as a predictive analytics alternative to regression in institutional research," Pract. Assess., Res. Eval. **23**, 1 (2018). https://doi.org/10.7275/1wpr-m024
8. T. G. Dietterich, "Ensemble learning," in *The Handbook of Brain Theory and Neural Networks*, 2nd ed., Ed. by M. A. Arbib (Massachusetts Inst. of Technology Press, Cambridge, Mass., 2002), Vol. 3, pp. 110–125.
9. M. Lipatov, "New reality dictates new rules," Prana, Nov., May 25 (2020). https://prana-system.com/novosti/novosti/novaya-realnost-diktuet-novye-pravila/. Accessed November 23, 2020.

Translated by T. Krasnoshchekova