OBTAINING OF CRITERIA FOR CONTROLLABLE PARAMETERS AND THEIR TRENDS' ESTIMATION VIA OPERATIONAL DATA STATISTICAL TREATMENT

I. Davidenko
Boris N.Yeltsin Ural Federal University, mail box 261, Ekaterinburg, 620057, Russia
*Email: <jera@r66.ru>

Abstract: One of the key condition for accurate estimate of technical state of electric equipment is the objectivity of criteria for its estimation. Not all controllable equipment parameters have values regulated by management directives or manufacturing company, while those that have them need revision for a number of reasons. At present databases of energy companies have a large reserve of values of equipment controllable parameters saved throughout the period of its operation. Therefore there is a need to obtain objective criteria of diagnostics on one hand, while on the other there is an opportunity to obtain such criteria using the methods of Knowledge Discovery in Databases. The article cites the method of long-term results' analysis of controllable parameters with the aim to obtain their permissible and maximum permissible values.

1 INTRODUCTION

Successful operation of electrical equipment depends on objectivity of controllable parameters estimation, i.e. on values accuracy, regulating parameter values and dynamics of their changes. The following methods are available to obtain permissible and maximum permissible values (PV, MPV) of parameters and their trends, applicable as a kind diagnostics criteria:

- obtained from experts, based on conceptions of physics of processes occurring in equipment and taking into account experts' practical experience;
- based on the results of mathematical and natural modeling;
- via statistical technology.

Realization of all these methods is linked with certain difficulties. For a such complicated object as electrotechnical power equipment the valid mathematical models, that take into account the full scope of processes' variety and peculiarity occurring in it, are not developed yet. Acceptable solutions are available at the moment refer to this equipment separate units and elements only.

At present databases of energy companies contain large data array of equipment controllable parameters, allowing obtaining of parameters' estimation criteria by means of statistical technology. According to IEC 60599 [1] PV of oil dissolved gas concentrations are defined by the integral distribution function F_x on the 0.90 level of its value for transformers and 0.95 for bushings. As relative cell frequency of parameter on the interval is often distributed through partition intervals unequally, this approach to F_x calculation leads to distorted PV definition. Distortion character depends on character of inequality. Fig.1 shows possibility for both desired value overestimation and underestimation.

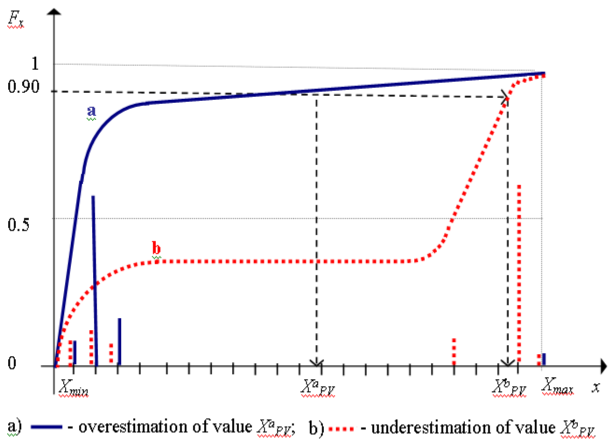

Figure 1: Origin of the distortion of PV definition.

Below you will find methods to calculate PV and MPV of controllable parameters, suggested by the author.

2 ISSUE OF DETERMINATION ACCURACY OF CONTROLLABLE PARAMETER PV

For the equipment that is subjected to increased frequency inspection, there is a need for value averaging per time slice equal to regular measurements' frequency. This is to obtain identical frequency of results occurrence in data array of defect-free equipment with normal periodicity of control and equipment with increased frequency of expected defect development.

It is advisable to increase sampling capacity via the use of all equipment base measurements within the period of its operation (measurements taken the last year only limit sampling representativeness).

It is required to inspect observable values range X_{min} – X_{max} for natural restrictions with the aim of range reduction. For example, all observable zero values of gas concentrations and those less than sensitivity bound for this gas detection on chromotograph can be regarded as equal to sensitivity bound for the chromotograph.

It is advisable to check spikes of bounds at the ends of observable value range (X_{min} – X_{max}) and apply spike values range smoothing, in case they are observed. In case under investigation "lazy" smoothing is sufficient [2], i.e. the bound value is equalled to the previous one. If samling is quite extensive, the smoothing may be repeated for a number of times so that number of values involved in smoothing would not exceed 1% of the sampling capacity.

The recommended number of partition intervals equals 50 – 70 pieces.

Unequal intervals and data array smooting technique have to be used for experimental data destribution levelling. We recommend median interval calculations as the most efficient and suitable for levelling of density of data array distribution of investigated parameter as per observable intervals. Median calculation implies data array ranging in increments and equal number of values being distributed into every partition cell. Therefore, the resulting intervals are of different size. Sliding medians can be applied later, where smoothing occurs in triplets of ranged values [2].

Median distribution in cells and smoothing reduce dispersion and distortion of both PV and expectation value of investigated parameter.

3 ISSUE OF LEVEL F_X=0.9 VALIDITY AT PV ESTIMATION

Level 0.9 is used to determine oil gas concentrations PV for power transformers by means of integral distribution function, level 0.95 – for instrumental transformers and bushings [1].

The following approaches are known to be used for MPV and PV levels selection.

- According to item [1] "choice of standard level is often the matter of intuitive guess, based on equipment users experience".

- In engineering calculations with similar purposes the "three sigmas" rule is often used, where the PV would be determined at F_x =0.9972 level of integral function. For the considered samples of the results of dissolved gases analysis (DGA) and physical-chemical analysis (PCA) this rule was not applicable as it is applied to normal distribution law only, whereas the DGA and PCA parameters distribution differ from the normal one. Moreover, through detection of PV for level F_x=0.9972, as it will be stated later the equipment with developing defects will be disregarded.

- In cases when the distribution law is not known Klein and Macclintock [3] recommended to use the "chances 20:1" criterion which the level of integral function F_x =0.95 corresponded to.

In the above listed approaches the choice of value level is not justified. Therefore they do not guarantee accuracy in "operable" and "defective" equipment differentiation and may result in cases of "oversight" and "reinsurance".

The viewing sampling contains long-term measured results of controllable parameters for the whole base of operational equipment, including equipment with developing defects. Every year it is registered a definite number of equipment "failures" and "imperfections". The "failed" equipment is deactivated because of breakdowns or disconnections by transformer protections.

We need to find out the number of "clear" failures R_o, when analyzed equipment work becomes inoperative. In case of imperfections (defectiveness) the personnel deactivates equipment on the basis of deteriorating controllable parameters for the subsequent repair.

Defects can be subdivided into quickly, medium and slowly developing ones. Furthermore, defects may appear in different stages of their development: non-dangerous and critical, which is identified as well by defect location as its character. In the non-dangerous stage characteristics degradation has not transformed into defect yet, i.e. the equipment is operable. On the critical stage equipment failure is highly possible because of degradation of parameters which are responsible for equipment operability. In case of quickly developing defects (hours) the non-dangerous stage is practically absent. For slowly developing defects, caused by aging and degradation of equipment material, the non-dangerous stage may last for decades. Typically equipment with slowly developing faults on the non-dangerous stage is not determined as imperfection. The number of rejected equipment R_d. consists of the equipment with dangerous defects R_{do}, that need quick problem-solving and adequate steps and equipment with developing defects R_{dd}, that have not entered the dangerous stage of development.

The first equipment group (R_{do}) includes the equipment with quickly developing defects, and defects of medium and slow rate of development in critical stage. The second (R_{dd}) group is the group of equipment with non-dangerous defects as well as with medium and slow rate of defect development that belong to the non-dangerous stage of development (e.g. aging of insulation materials).

If the flow value of damageability of the given equipment type λ is known, then it is suggested to consider the value of *1–λ/100* as the level of integral function of distribution F_x that will separate equipment samples into operable and those with defect under development.

$$\lambda = \frac{R_o + R_d}{N * t} \qquad (1)$$

where: N = number of equipment;
R_o =failure number;
R_d = number of equipment imperfections;
t = observation period, years.

Thus, the PV of tested parameter, defined by F_x at *1-λ/100*, designates the “risk area”. The equipment, where values of controllable parameters exceed PV may remain in operation but must be subjected to more frequent inspection to prevent miss of stage defect development.

To find “danger space ” you require knowledge of both failure number R_o and dangerous imperfection number R_{do}. The flow value of both failure and dangerous imperfection is calculated by the following formula:

$$\lambda_o = \frac{R_o + R_{d\,o}}{N * t} \qquad (2)$$

where: N = number of equipment;
R_o =failure number;
R_{do} = number of dangerous imperfections;
t = observation period, years.

So, to detect the dangerous defect equipment, the $1\text{-}\lambda_o/100$ value provides us with the MPV level estimated by the integral function F_x. In case the equipment with the values of controllable parameters exceeds their MPV, it requires immediate solutions and/or actions, such as deactivating of the equipment from operation to prevent failures.

It should be noted that to estimate F_x levels by damageability flows λ and λ_o, only defects which affect the tested parameters directly or indirectly, should be taken into account. Besides, the estimation of flow failure must be synchronized with the measurement interval of tested parameter.

According to the damageability analysis of the 35-220 kV transformers based on the author data researchers of four large energy companies it is suggested to use F_x levels equal to 0,96 and 0,986 respectively for estimation of PV and MPV.

It is worth mentioning that transformer defects were disregarded for calculation of damageability flow, which are not detected by DGA: sudden failures caused by natural disasters, acts of vandalism, false operation of protective circuits, etc. Due to a variety of reasons collected data cannot be considered complete. For example, in these cases representatives of the party that operates the equipment and those of the manufacturing company aim to protect their own interests.

Also it should be noted that estimating PV and MPV of the dissolved-gas analysis and physical-chemical analysis which are used for diagnostics of bushings, current and voltage transformers, different levels of cumulative distribution function F_x must be used in accordance with the damageability flows of the equipment.

4 ESSENTIAL USE OF DIFFERENT F_X LEVELS FOR CALCULATIONS OF PERMISSIBLE VALUES IN ACCORDANCE WITH EQUIPMENT OPERATION LIFE

Damageability flow of the equipment with different operation periods is not identical. In general case, it is supposed that damageability curve has three specific segments. These periods are as follows: run-in period, standard operation period and operation of the old equipment with high risk of fault development.

However, according to the data acquired by the author there are much more specific segments. Figure 2 shows the diagram of the transformer damageability depending on the period of operation.

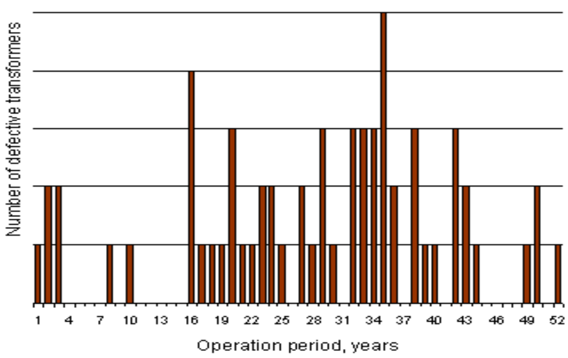

Figure 2: Power transformers damageability determined by the period of operation

According to the diagram, the run-in period is over by the third year. Most of the detected defects are caused by manufacturing, delivery and assembly faults. Then, after the run-in period and within the next sixteen years the equipment operation is sufficiently reliable. Increase of damageability within the period from 16 to 34 years is caused first of all by the damageability of the component units (such as tap changers, bushings), dynamic instability of windings and the results of low-quality repair and operation. The period when the damageability is caused by the transformer aging itself (e.g. aging of turn insulation and oil decomposition) starts in the 46th – 48th year. At this time gradually developing defects which are associated with deterioration of the transformer insulation materials (R_{dd}) turn into the dangerous stage of development.

Thereby, considering the evident damageability dependency of the transformer operation period, it is suggested to calculate PV and MPV for this equipment with different values of level F_x for four groups of the equipment with the following operation periods: to 3 years, from 3 to 16 years, from 16 to 39 years, over 39 years.

Within in operation several cases have been registered, when while “long-operated” transformers had tested parameters equal to those of the new transformers they demonstrated rapid development of failures. This observation proves the necessity to use different levels of F_x for groups of equipment with different operation periods. The decrease in level F_x when calculating PV and MPV for “long-operated” transformers will enable to consider the potential risk of such failures.

It should be noted that specific periods and damageability flows for other types of oil-filled equipment differ from the periods and damageability flows obtained for the transformers.

5 THE NECESSITY FOR PV DIFFERENTIATION BASED ON THE EQUIPMENT DESIGN PHILOSOPHY AND OPERATION PERIOD

To make decision on the necessity of tested parameter PV differentiation with respect to any factor, it is required to estimate the factor influence on the parameter. It is suggested to use dispersion analysis [4] for estimating of the factor influence.

The essence of dispersion analysis is in comparison of residual and factor dispersions by the Fisher–Snedecor criterion for a certain level of significance (generally for 0,05 and 0,01).

On the first stage, the residual dispersion is determined on every level of the factor:

$$D_R = \frac{\sum_{i=1}^{L} \sum_{k=1}^{Ki} (x_{ik} - \overline{M}_i)^2}{\sum_{i=1}^{L} K_i - L} \qquad (3)$$

where: $\overline{M}_i$ = mathematical expectation of the *i-th* factor level;
L = number of factor levels;
X_{ik}= parameter value at the *i-th* level;
K_i = capacity of the *i-th* factor level.

Further, factor (external) dispersion is calculated by the formula:

$$D_S = \frac{\sum_{i=1}^{L} K_i * (\overline{M}_i - \overline{M})^2}{L-1} \qquad (4)$$

where: $\overline{M}$ = expectation of the whole data array;
L = number of factor levels;
K_i = capacity of the *i-th* factor level.

Then the ratio of estimated dispersions is determined by the formula:

$$W = \frac{D_S}{D_R} \qquad (5)$$

The number of degrees of freedom for factor *N-L* (where *N* is the capacity of the data array) and residual *L-1* dispersions, specified in tables for 0,05 and 0,01 significance levels, defines the boundary of the right critical region: $W_{0,05}$ and $W_{0,01}$. According to the Fisher-Snedecor criterion, if the following inequality is fulfilled:

$$\mathrm{W} > \mathrm{W}_{0,05} (\mathrm{W}_{0,01}), \qquad (6)$$

so, the viewed factor makes a appreciable effect on the tested parameter.

6 THE ISSUE OF OBTAINING PV AS A WHOLE FOR CORPORATION (BRANCH)

Calculating PV (MPV) of the controllable parameters for national, regional and corporate standards there appeared a question about possibility of data arrays union of these energy companies because in most cases arrays are heterogeneous.

Data arrays heterogeneity of various enterprises may be caused by:

- difference of base according to operation life, composition and design philosophy;
- influence of climatic zone and specific work mode;

- different load of equipment;
- difference in applied technologies, methods and instrumentation;
- different level of maintenance management and personnel qualification.

It is recommended to study factors influencing the PV parameters and their effect separately for each data array; then comparing the obtained results between each other and check accuracy of obtained inferences with the data arrays of other enterprises.

When the factors which significantly influence the tested parameters are defined, the data arrays of each enterprise shall be divided into samples according to found factors of influence and damageability periods. With such division in samples the heterogeneity subsides, caused by factors of influence. Before consolidation the samples taken from the data arrays of different enterprises but with equal levels of influence factors are checked and compared for homogeneity.

The following criterion [5] is suggested to check two independent representative randomly distributed samples:

if $Z_{observe} < Z_{crit}$, so the compared data samples may be regarded as homogeneous.

Observed value of the criterion is calculated by the following formula:

$$Z_{\text{observe}} = \frac{|M_1 - M_2|}{\sqrt{D_2(x_1)/N_1 + D_2(x_2)/N_2}} \qquad (7)$$

where: M_1 and M_2 = middle meaning of samples;
$D_1(x_1)$ and $D_2(x_2)$ = sample dispersions;
N_1 and N_2 = sample sizes;

Value of critical point (Z_{crit}) is calculated by the Laplace function and the equation:

$$\Phi(Z_{кр}) = (1-\alpha)/2\,, \qquad (8)$$

where: α = significance level.

As a rule significance level values 0,05 and 0,01 are used.

It is recommended to check the homogeneity of data samples in pairs, in case where more than two data samples are combined. After this test, the samples which demonstrate homogeneity are combined. Then there is an estimation of PV of parameters for each combination of homogeneous samples.

Next it is suggested to calculate PV as a whole for corporation (industry) by PV obtained for combined data samples accounting their weight:

$$X_{PV} = \frac{\sum_{i=1}^{L} X_{PVi} * N_i}{\sum_{i=1}^{L} N_i} \qquad (9)$$

where: X_{PVi} = PV of the observed parameter for the *i-th* combination of homogeneous samples;
N_i = capacity of the *i-th* combined sample;
L = number of combined samples.

7 FINDINGS

The introduced procedure may be applied for preparation of enterprise standards in the field of technical diagnostics. Local standards cover wider set of controllable parameters; also, local standards are subjected to more frequent revisions and thus may timely respond to new types of equipment, diagnostic facilities and changes in understanding processes happened in the equipment.

Procedure automation allow minimizing of timetable of this laborious process and increasing of the immediacy of obtaining results in accordance with the existent technical and economic changes.

The procedure has been implemented in one of the subsystems of “Albatros”, expert-diagnostic and information system (EDIS) for technical state evaluation of oil-filled equipment. During EDIS operation in dozens of energy companies, considerable set of observations resulting from DGA and PCA was accumulated in its database. This allowed the author to carry out a number of researches [6-9].

With the help of method offered by the author and implemented in EDIS there were calculated evaluation criteria of DGA and PCA for power and instrument transformers and also for bushing. There were determined PV (MPV) of gas concentrations and rates of their change, also there was a PV (MPV) of moisture content, oil dielectric loss tangent, acid number and oil breakdown voltage.

Besides the described procedure, EDIS also includes other elements of *KDD* technology, such methods, which allow defining criteria of defect identification.

Criteria of technical state evaluation of oil-filled equipment, which were obtained by *KDD* technologies, are applied in EDIS "Albatros". Successful operation of the system used within a number of years at 250 work sites in Russia, Ukraine, Latvia and Moldova has proven objectivity of the criteria.

In conclusion, it should be noted that operative discovery of knowledge in enterprise database and its application provide enterprise with competitive advantage to optimize the use of other traditional resources and facilitate the management of equipment maintenance.

8 REFERENCES

[1] IEC 60599. Mineral oil-impregnated electrical equipment in service – Guide to the interpretation of dissolved and free gases analysis, 1999-03.
[2] G.Tyuki: "Analysis of Observation Results", M.:Mir, 693 p, 1981.
[3] Shenk H.: "Theory of Engineer Experiment", M.:Mir, M.: Мир, 376 p, 1972.
[4] Melentiev E.K.: "Elementa of Dispersion Analysis", M.: Nauka, 72 p, 1967.
[5] Gmourman V.E.: "Theory of Probability and Mathimatical Statistics", M.: Vysshaya Shkola, 432 p, 1997.
[6] I. V. Davidenko: "The Development of Diagnostics Criteria of Power Transformers and Instrumant Transformers Using the Results of the Chromatograph Analysis of Gases Dissolved in Oil", CMD 2006 - International Conference on Condition Monitoring and Diagnosis, Changwon, Korea, April 2006
[7] I. V. Davidenko:" Specification of Insulation State Estimating Criteria for Oil-filled Bushing and Power Transformers Basing on Dissolved Gas Analysis", ISH 15-th- International Simposium on Higt Voltage Engineering, Slovenia, Ljubljana, Paper T8-55, 2007
[8]. I. V. Davidenko: "Graphical model of imperfection identification in instrumental transformers by DGA", ISH 16-th - International Simposium on Higt Voltage Engineering, Cape Town, South Africa, 2009.
[9]. I. V. Davidenko: "Criteria of Fault Type Identification in Bushings Based on DGA", CMD 2010 - International Conference on Condition Monitoring and Diagnosis, Tokyo, Japan, Vol.1, pp.137-140, 2010.