Mathematical model and regression analysis of acoustic emission signals generated by partial discharges

Daria Wotzka

Faculty of Electrical Engineering, Automatic Control and Informatics, Opole University of Technology, Opole, Poland

Email address:
d.wotzka@po.opole.pl

Abstract: An improved mathematical model describing acoustic emission (AE) signals generated by different types of partial discharges (PD) that occur in electric power transformer insulation system is presented in the paper. AE signals are analyzed within the AE method as applied for power transformer failure detection due to occurrence of PD. There are several types of basic defects, which are characterized by different types of PD. The mathematical model presented here is crucial for numerical analyses and simulations, where it acts as the function describing the acoustic source in an acoustic model of power transformer insulation system. The regression procedure was performed based on empirical AE signals, registered in a laboratory experiment. The AE signals are described by a mathematical model being a multi-parameter function, which involve both the time domain and the frequency domain. Goodness of the model was evaluated based on analysis of 480 data samples in the time, frequency and time-frequency domains. Also coherence between the registered and modeled signals was calculated. It was stated that the improved model fits very well to the real data, although, due to high level of noise embodied in signals registered in experiments, the coherence values remain low. Moreover, analyses of the estimated data were performed and some example results are presented in this paper. Based on the achieved outcomes a collection of parameter values was prepared for each of the eight considered PD basic types. One can simple use it now in a numerical model for simulation of AE signal source generated by specified type of PD, what corresponds to a particular power transformer insulation system failure. Furthermore, the regression procedure presented in this paper can be easily transferred to any other types of AE sources including processes of compression, tension and cracking.

Keywords: Modeling, Regression Analysis, Acoustic Emission, Partial Discharges

1. Introduction

Technical condition monitoring and diagnosis of electric power transformers constitutes a sophisticated issue, because it is connected with many physical phenomena, occurring during operation of the device [1]. Electrical partial discharges (PD) are examples of such phenomenon, and are reason for many power transformer failures [2]. Generation of PD induces chemical reactions, acoustic and electric signals generation, emission of light and other phenomena [3,4]. Scientific research performed to date, which concerns recognition of defects caused by various forms of PD, is conducted with the use of spark gaps, which allow their experimental modeling; however, this requires application of high voltage assay systems [5,6]. From the viewpoint of the practical usefulness such experiments are relatively dangerous to persons performing the measurements, time-consuming, costly, and

results obtained are susceptible to a number of external factors that may influence their repeatability and reproducibility. Therefore, there is a need for design and practical implementation of an arrangement to allow modeling of the physical phenomena accompanying the generation of PD, without the need to conduct research with the use of high voltage. In this way it will be possible to carry out a detailed assessment of the impact of various parameters connected with PD generation on the measurement results. This in turn may contribute to improvement of the diagnosis method applied and increase its accuracy while reducing the costs associated with it. In our study we develop a multiphysical model of the insulation system, in which different phenomena are integrated into a single simulation model. For the simulation purposes the AE wave source is needed to be described by a mathematical function. Up till now, other scientists have used Heaviside function, linear function or Gauss type function. There exist also models, which consider combination of linear function

and multiplication of exponentially decreasing function by a sine or cosine function [7,8]. The advantage of such functions lies in their simplicity, but for our purposes they are not applicable, since either they give only a week mapping to the real world signals or the functions are not continuous, what causes numerical calculation troubles (the function must be differentiable). Therefore we have developed our own model for AE source, which constitutes a combinations of a sigmoid type function and superposition of cosine type functions, each having own frequency and amplitude. During research works, the model was further improved in terms of giving a better correlation between computer calculated and registered (real) signals and was then applied in simulations of acoustic wave propagation in insulation systems. The simulations consider AE signal path estimation while the model parameters are changed. In previous studies, performed in our research team, AE signals generated by different kinds of PD occurring in electric power transformer insulation systems have been analyzed by use of sophisticated statistical and numerical methods, wavelets (CFT, DWT), FFT, STFT, and many others [9,10,11]. Based on those analyses, the team has developed a number of indicators, which were then applied for classification purposes performed among others by use of neural networks. The study resulted with a classification of the PD AE signals into eight various classes, where each class corresponds to different damage occurring in the insulation system [11]. The considered forms of PD are described in Table 1. The PD types causing failures in insulation systems and the measurement procedure are beyond the scope of this paper. If interested, one might find further details in e.g. [12,13,14].

Table 1. *Description of the considered basic PD types [11].*

PD type	Description of basic forms of PD
1	PD occurring in insulation system with particles of undefined potential moving in oil, which can model discharges occurring in oil containing particles of cellulose fibers created due to aging and gradual degradation of the paper-oil insulation
2	PD occurring in a point-plane oil system, which can be related to discharges generated by insulation damage of two neighbouring turns of transformer windings
3	PD occurring in a surface system of two flat electrodes with paper - oil insulation between them, which models the most common discharge type occurring between electrode surfaces and solid or liquid dielectric
4	PD occurring in a surface system with one flat and one multipoint electrode with a paper - oil insulation between them, which models different electric field intensity distribution as compared to the surface system of two flat electrodes
5	PD occurring in a multipoint-plane system in oil, which can be related to discharges occurring between a multipoint damage of transformer winding insulation and its flat earthed elements
6	PD occurring in a multipoint-plane oil system with gas bubbles moving in oil, which can reflect discharges in gassy oil, due to multipoint winding insulation damage
7	PD occurring in a point-point oil system, which can model discharges occurring between earthed flat parts and damaged winding insulation of the transformer
8	PD occurring in a point-point oil system with gas bubbles moving in the oil, which can be related to discharges in gassy oil, due to insulation damage of two turns of transformer windings

2. Mathematical Model Definition

The mathematical model presented in this paper has been created on the basis of AE signals registered in experiments performed under laboratory conditions, while changing the setup for particular insulation defects induced by PD. One needs to mention that due to the very unsystematic and stochastic nature of PD occurrence in the negative half period of the supply voltage, the studies considered only the positive voltage half period, thus the first 10 ms of the registered signals. 480 selected AE timeruns, each of 25 000 samples, which are divided into mentioned PD classes, were applied during parameter estimation and optimization processes. The model is composed of two parts: Part1 is combination of a sigmoid and exponential type functions and describes envelop of the considered AE signal; Part2 constitutes a sum (superposition) of 60 cosine functions, each with own magnitude and frequency and relates to the frequency components involved in the real AE signal. The frequencies are set a priori: f = {10, 20, 30,., 600} (kHz). These values were selected on the basis of earlier gathered indicators and correspond to components contained in signals registered in real measurements. The proposed mathematical model is presented in eq. (1).

$$y(t) = \underbrace{\frac{A}{1+e^{-B(t-\mu_1)}} e^{-C(t-\mu_2)}}_{Part\ 1} \underbrace{\sum_{i=1}^{60} A_{\cos\_i} \cos\left(2\pi f_i t\right)}_{Part\ 2} \qquad (1)$$
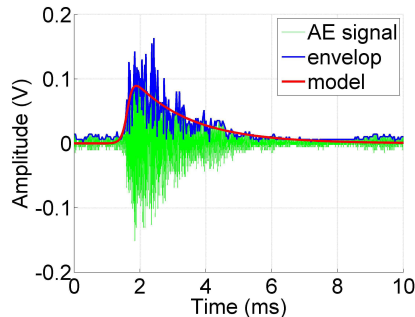
where: $y(t)$ – AE signal generated by a PD (V), A – amplitude parameter of the sigmoid function (V), B – scale parameter of sigmoid function (-), μ_1 - localization parameter of the sigmoid function (s), C – scale parameter of exponential function (-), μ_2 - localization parameter of the exponential function (s), $A_{\cos\_i}$ – amplitude of the i-th cosine function, (-), f – frequency of the i-th cosine function, f $\in$ {10, 20, 30, ., 600} (kHz), t – independent variable, time (s).

The most important parameters in Part1 of the model are the amplitude and scale of the sigmoid as well as the scale parameter of the exponential function, since they represent the shape of the AE signal envelop. The localization parameters are important only during the estimation process for to enable a well matching between the estimated and given data samples. Often the signals registered are shifted in time and this is due to the very stochastic nature of PD occurrence. However, in the numerical simulations one can assume that the signal is not shifted, thus the localization parameters might be ignored. In Part2 of the model the frequencies are set a priori and only the value of magnitude (amplitude) of each frequency component is estimated. Part2 represents the frequency domain, while Part1 the time domain of AE signals generated by PD.
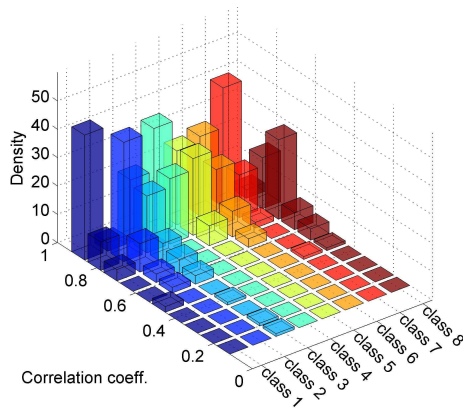
During the studies also other functions were considered for the model envelop (Part1) [15], e.g. T-Student type function, Lognormal type function, Generalized Extreme Value type function and Tangent-Hyperbolic type function. Nevertheless the sigmoid type function achieved the highest values of goodness indicators and thus was applied for the improved

model, presented in this paper.

3. Parameter Estimation and Optimization Procedure

Figure 1. *Example timerun of registered AE signal (green) with theirs calculated (blue) and computer-modeled (red) envelops.*
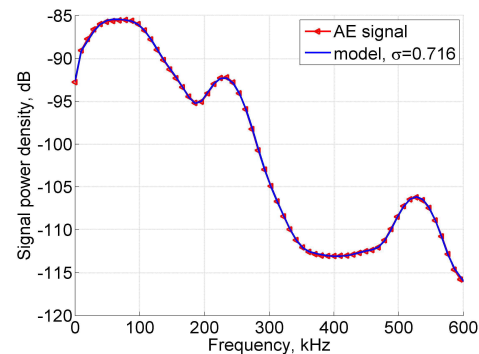

Figure 2. *Histograms of correlation coefficients calculated between the registered and modeled signal envelops (only Part1 of (1)) for the eight considered PD classes.*

The regression procedure consists of several steps. First, for each registered AE signal its envelope was determined. The gathered envelops were then used as reference dataset for the model Part1 calculation. Then, parameters of model Part1 were estimated by means of Least Squares method (LSM) for rough results. In the next step, the estimates were optimized by use of the Nelder-Mead-Simplex (NMS) method in order to meet possible lowest value of the residual norm, which served as goodness indicator and also as a stop criterion in the optimization routine. In this way envelops of AE signals generated by eight PD classes were calculated. In Fig. 1 an example of timerun of AE signals registered in experiments is shown in green. Its envelop is marked with blue and the envelop calculated with model Part1 is shown in red. During the first step of regression procedure we used the Pearson correlation coefficient as a second goodness indicator, next to the residual norm. The correlation coefficient values, calculated between the two runs (envelops of the registered and modeled AE signals) indicate mostly very high correlation, above 0.8. In Fig. 2, histograms determined for all regression steps (for all 480 data-sample pairs) are depicted. Similarly good outcome we gathered for

the residual norm, which for the most analyzed data-samples did not exceed the value of two, and in some rare cases the norm was higher but always less than 3.5.

The second step of the regression procedure regarded to parameter estimation of the model Part2 and was related to the frequency domain. For the given 60 frequency values the appropriate magnitudes were determined by means of Least Squares analysis. In contrast to the first step, not the timeruns but the PSD (Power Spectra Densities) were used when comparing the regression result. For determination of the signal power at each specified frequency, the Welch's method was used. It is based on the concept of using periodogram spectrum estimates, which are the result of converting a signal from the time domain into the frequency domain. Application of the Welch's method enable for reduction of noise contained in the considered signals in exchange for reducing the frequency resolution. Due to the noise level present in the registered AE signals, the noise reduction from approach given by the Welch's method was desired in our regression analysis.

Figure 3. *Example power spectra density of AE signal registered in the laboratory (red) and calculated on the basis of the model (blue). The curves mostly overlap, what is confirmed by the low value of σ-Euclidean distance.*

After achieving rough estimates with the LSM, the parameter values were further optimized using the NMS method, similar as it was in step one. The residual norm was used as stop criterion in the optimization routine. For all considered data-samples we gathered almost overlapping curves, thus the correlation coefficient calculated over the PSD was mostly equal to 1. In Fig. 3 an example of PSD calculated for the registered (red) and modeled (blue) AE signal is shown. The curves mostly overlap, what indicates very well mapping of the signal modeled to the signal registered. This was confirmed by low value of σ, which is the Euclidean distance calculated over residuals.

4. Regression Analysis Results

For all considered data-samples the appropriate signals were calculated based on the mathematical model developed. In Fig. 4 an example of timerun of the AE signal calculated by the model is depicted in green, while its envelop (just Part1) is marked with red.
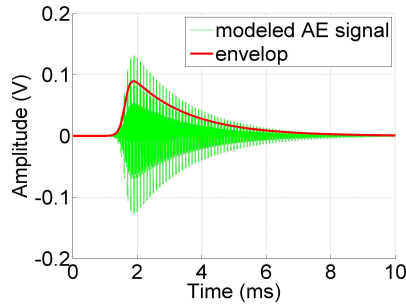
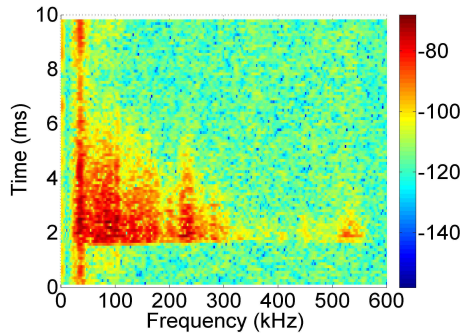**Figure 4.** *Example timerun of computer-modeled AE signal (green) and its envelop (red).*

Figure 5. *Examples of time-frequency spectrograms of AE signals registered in the laboratory.*

We have then further analyzed the results in the time-frequency domain by use of STFT (Short Time Fourier Transform) method. This analysis was based on pure observation of the time-frequency structures. An example of spectrogram calculated over the registered AE signal is presented in Fig. 5. In Fig. 6. A spectrogram of the modeled AE signal is shown. Based on achieved results we recognized that the time-frequency structures contained in the modeled signals corresponded well to the particular data gathered from laboratory experiments for all types of PD.
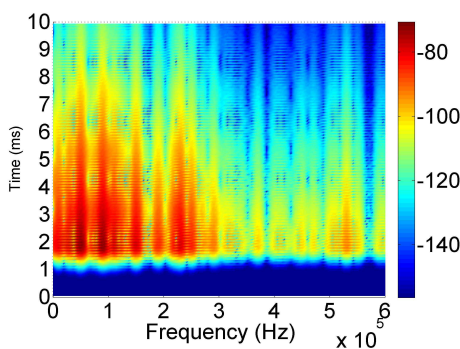

Figure 6. *Examples of time-frequency spectrograms of AE signals computer-modeled.*

In the next step we evaluated also the coherence between each of data-sample pairs. This analysis did not give so satisfactory outcomes. For the most of PD types we recognized that the signals do not stay in very good relation, as the coherence values in average, when regarded all data, did not exceed 0.1, as presented in Table 2. Only for AE signals generated by PD type 3 and type 4 the coherence values are for

given frequency bands above 0.6, indicating coherent structures between the measured and modeled signals. The explanation for such behavior may be in that these signals are less noisy despite to other measured signals. An example magnitude squared coherence graph is presented in Fig. 7.

Table 2. *Magnitude squared coherence values calculated for PD types.*

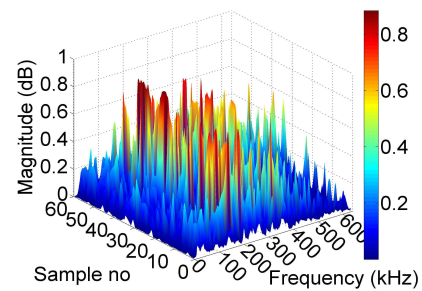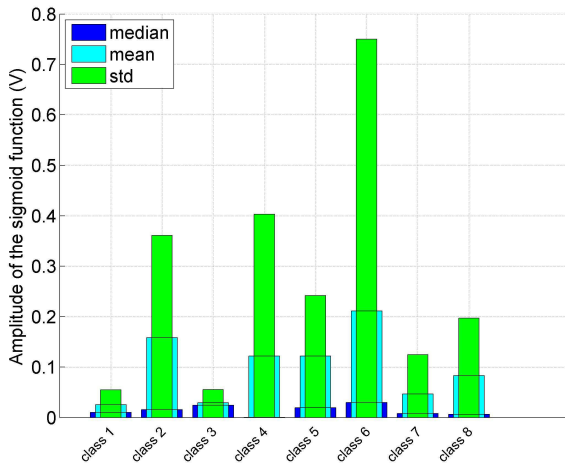| PD type	Maximal value	Arithmetic average
1	0.2621	0.0235
2	0.4296	0.0335
3	0.6191	0.0237
4	0.8841	0.1094
5	0.3657	0.0297
6	0.4690	0.0237
7	0.3351	0.0281
8	0.4754	0.0232

Figure 7. *Example Magnitude squared coherence graph calculated for the analyzed registered and computer-modeled AE signal.*

5. Analysis of the Estimated Parameter Values for Simulation Purposes

As mentioned earlier, the main aim of the model development, presented in this paper, is to enable research works regarded to numerical simulations of AE signal propagation in a power transformer insulation system model. In order to implement an AE signal source that corresponds to a specified defect (PD class) one needs to consider the estimated parameter values. For this purpose, for each parameter estimates set and for each PD class we have calculated the arithmetic mean, variance and standard deviation values. We have also calculated data distribution histograms. This statistics give us certain information about the various differences in model parameters corresponding to particular PD types. Next, we have determined the data distribution functions over all parameter estimates for each PD class. For this task, one can take into account e.g. a Gauss function, if data is well and symmetrically distributed, or in other cases a Nakagami, Levy, GEV, Birnbaum-Sanders, Weibull, Gamma, Lognormal or other distribution function. These types of distribution functions are defined by just a few parameters, which, when estimated, can serve as indicators about the overall value of the considered model parameter. Other possibility is to take the mean or median value. But in situations when there does not exist a well diversity between parameter values of the particular PD classes, application of distribution function estimates provides a better solution.
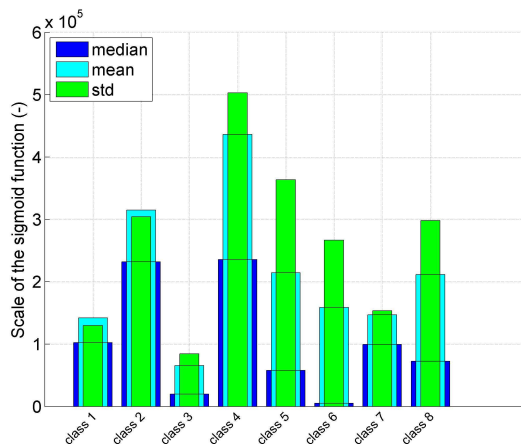
In Fig. 8 the mean, variance and standard deviation values calculated for the estimates of parameter A (the amplitude of the sigmoid function) for all of the eight considered PD classes are shown. The mean and standard deviation values depict more or less diversity between the columns. The medians remain rather low.
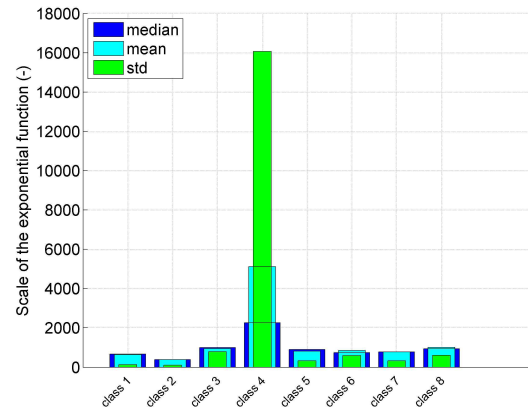

Figure 8. *The A parameter values estimated in the regression process for all eight classes of PD. Histogram depict mean, median and standard deviation values calculated over the gathered data.*

In Fig. 9 the statistics calculated for the estimates of parameter B (the scale of the sigmoid function) for eight PD classes are shown. All values depict a well diversity between the columns.
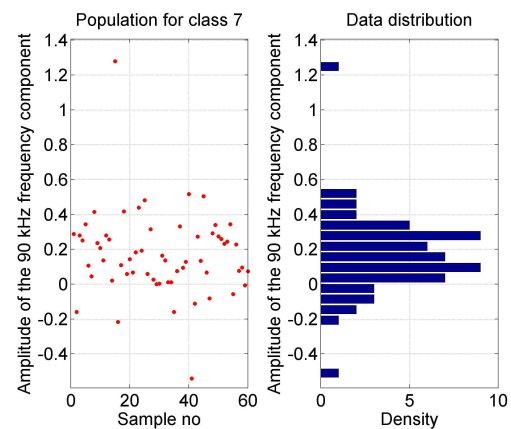

Figure 9. *The B parameter values estimated in the regression process for all eight classes of PD. Histogram depict mean, median and standard deviation values calculated over the gathered data.*

In Fig. 10 the statistics calculated for the estimates of parameter C (the scale of the exponential function) for eight PD classes are shown. All values depict smaller diversity between the columns except the characteristics for PD class 4, which are much higher in this case. This is clear evidence showing significant difference between signals generated by the particular PD types, which can be applied in the numerical simulations.

Figure 10. *The C parameter values estimated in the regression process for all eight classes of PD. Histogram depict mean, median and standard deviation values calculated over the gathered data.*
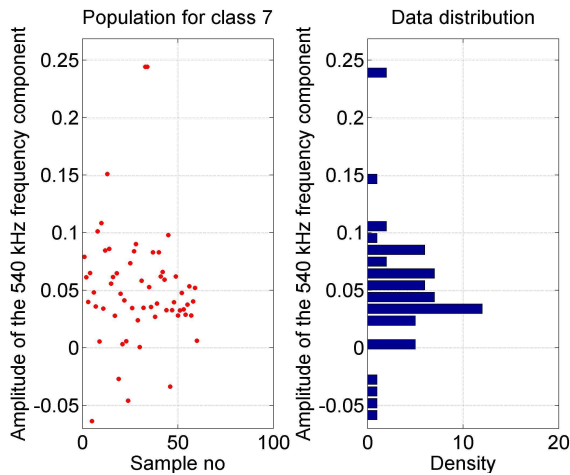

Figure 11. *Example result showing population and distribution of the amplitude values estimated in the regression process related to PD class 7, for the frequency equal to 90 kHz.*

In Fig. 11 and 12 examples of the $A_{\_cos\_i}$ parameter estimates of (1), which correspond to signal magnitudes at frequency values equal to 90 kHz and 540 kHz are shown.

In the last step we have created a collection of indicators containing distribution function estimates, mean, median and standard deviation values, calculated for each parameter and each PD class. These values are currently applied in numerical simulations regarding multiphysical model of power transformer insulation system, in which different types of damages are simulated.

6. Conclusion

In this paper an improved mathematical model describing AE signals generated by eight different classes of PD occurring in electric power transformer insulation systems was presented. Goodness of the model was evaluated on the basis of timeruns, power density spectra, time-frequency spectrograms and coherence values. Based on the Welch and Fourier analyzes in the frequency domain and on the correlation analyzes of signal envelops in the time domain, we stated that the proposed model fits very well to the measured data.

Figure 12. Example result showing population and distribution of the amplitude values estimated in the regression process related to PD class 7, for the frequency equal to 540 kHz.

We have also performed an analysis of parameter estimates and have selected appropriate indicators, which enable one to distinguish between the particular PD types. In such a manner the collected values can be involved in numerical simulations regarding AE signal propagation studies, where different types of defects are modeled.

Another application possibility of the model created is to use it in experiment with a physical power transformer insulation system. Here the AE signal can be generated by a wideband piezoelectric transducer, controlled by a generator, which in turn transmits signals in accordance to our model. These types of experiments enable for AE propagation path analyses, without the need of initiation the natural phenomena generating the AE signal. For instance in [16] such type of experiment is presented.

The procedure of signal modeling and parameter estimation shown in this paper might be easily applied for modeling of AE signals generated by any other kind of source. It is e.g. possible to create a mathematically described source function of AE signals generated due to compression, tension and cracking occurring in steel, ceramic or other materials but is not limited to this.

References

[1] S. Borucki, "Diagnosis of technical condition of power transformers based on the analysis of vibroacoustic signals measured in transient operating conditions", IEEE Trans. Pow. Del. Vol. 27, pp. 2012, 670–676. DOI: 10.11000009/TPWRD.2012.2185955

[2] L. E. Lundgaard, "Partial discharge - part XIV: Acoustic PD detection – practical application". IEEE El. Insul. Mag., Vol. 8, No. 5, 1992, pp. 34–43.DOI: 10.1109/57.156943

[3] P. Frącz, "Influence estimation of the voltage value on the measurement results for the optical radiation generated by partial discharges on bushing isolator", Acta Phys. Pol. A, Vol. 120, 2011, pp. 604-607,

[4] A. Cichoń, P. Frącz, D. Zmarzły, "Characteristic of acoustic signals generated by operation of on load tap changers", Acta Phys. Pol. A, Vol. 120, 2011, pp. 585-589.

[5] J. Gubio-Serrano, M.V. Rojas-Moreno, J. Posada, J.M. Martinez-Tarifa, G. Robles, J.A. Garcia-Souto, „Electro-acoustic detection, identification and location of PD sources in oil-paper insulation systems", IEEE trans. Dielectr. Electr. Insul., Vol. 19, No. 5, 2012, pp. 1569-1578, DOI: 10.1109/TDEI.2012.6311502

[6] M. Pompili, R. Bartnikas, "On PD measurement in dielectric liquids", IEEE Trans. Dielectr. Electr. Insul, Vol. 19, No. 5, 2012, pp. 1476-1481. DOI: 10.1109/TDEI.2012.6311489

[7] M. Shibata, "A theoretical evaluation of AE signals - the rise time effect of dynamic forces". Mat. Eval., Vol. 42, No. 1, 1984, pp. 107–115.

[8] L. Bolin, "A model for estimating the signal from an acoustic emission source". Ultrason., Vol. 17, No. 2, 1979, pp. 67–70.

[9] T. Boczar, D. Zmarzły, "The application of correlation analysis to acoustic emission pulses generated by partial discharges", Mat. Eval., Vol. 62, No. 9, 2004, pp. 935-942.

[10] T. Boczar, D. Zmarzły, "Multiresolution analysis of the AE pulses generated by PD", INSIGHT, Vol. 45, No. 7, 2003, pp. 488-492.

[11] T. Boczar, A. Cichoń, S. Borucki "Diagnostic expert system of transformer insulation systems using the acoustic emission method", IEEE Trans. Dielectr. Electr. Insul., Vol. 21, No. 2, 2014, pp. 854–865. DOI 10.1109/TDEI.2013.004126

[12] S. Rudd, S.D.J. Mcarthur, M.D. Judd, "A generic knowledge-based approach to the analysis of PD data", IEEE Trans. Dielectr. Electr. Insul., Vol. 17, No. 1, 2010, pp. 149 – 156. DOI: 10.1109/TDEI.2010.5412013

[13] K.X. Lai, B.T. Phung, T.R. Blackburn, "Application of data mining on PD part I: predictive modelling classification", IEEE Trans. Dielectr. Electr. Insul., Vol. 17, No. 3, 2010, pp. 846–854. DOI: 10.1109/TDEI.2010.5492258

[14] T. Pinpart, M.D. Judd, "Differentiating between PD sources using envelope comparison of ultra-high-frequency signals", IET Science, Measurement & Technology, Vol. 4 , No. 5, 2010, pp. 256–267. DOI: 10.1049/IET-SMT.2009.0064

[15] A. Cichoń, S. Borucki, D. Wotzka, "Modeling of acoustic emission signals generated in on load tap changer" Acta Phys. Pol. A, Vol. 125, No. 6, 2014, pp.1396-1399. DOI: 10.12693/APhyspolA.125.1396

[16] D. Wotzka, A. Cichoń, T. Boczar, "Modeling and experimental verification of ultrasound transmission in electro insulation oil", Arch. Acoust., Vol. 37, No. 1, 2012, pp. 19-22, DOI: 10.2478/v10168-012-0003-x.