Electric Power Systems Research 147 (2017) 245–253

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

On the use of hypothesis tests as statistical indexes for frequency response analysis of electric machinery

Wilson Cesar Sant'Ana [a,b,*], Camila Paes Salomon [a,b], Germano Lambert-Torres [b], Luiz Eduardo Borges da Silva [a], Erik Leandro Bonaldi [b], Levy Ely de Lacerda de Oliveira [b], Jonas Guedes Borges da Silva [b]

[a] Universidade Federal de Itajuba, Av. BPS, 1303, Itajuba, MG 37500-903, Brazil
[b] Instituto Gnarus, rua Cel. Francisco Braz, 185, sl. 302, Itajuba, MG 37500-052, Brazil

ARTICLE INFO

Article history:
Received 14 October 2016
Received in revised form 13 January 2017
Accepted 4 March 2017
Available online 19 March 2017

Keywords:
AC machines
Condition monitoring
Frequency response analysis
Rotating machine insulation testing

ABSTRACT

This work proposes a new statistical index, based on hypothesis tests, to be used on frequency response analysis (FRA) of electric machines. This index is, particularly, interesting on the detection of early failures (which appear as small deviations from the baseline spectrum), even when the measurements have high variability (such as the case of online measurements in rotating machines). The majority of the indexes proposed so far in the FRA literature only compares two spectra at a time and mix data from the measurements at different frequencies. The index proposed in this paper calculates one hypothesis test for each frequency (in order to assure randomness, independence and normality) in the spectra of several measurements (in order to take into account the variability on the measurements) and performs the normalized sum of the individual tests. Experimental results are presented on a synchronous generator (both offline and online) to validate the proposed index.

1. Introduction

Frequency response analysis (FRA) has been widely used as an offline condition based monitoring tool for transformers [1]. It consists in the comparison of a spectrum against its historical data. Based on the pattern of deviation from the historical data, a specialist is able to identify several types of failures: axial displacement [2], buckling deformation [3], bushing faults and oil degradation [4] and winding short circuit [5]. The literature also presents comparisons on the use of FRA against other popular methods, such as VM (Vibroacoustic-Method) [6] and FDS (Frequency Domain Dielectric Spectroscopy) [7]. Concerning rotating machines, the application of FRA is still under study [8], mainly due to the increased complexity of the high frequency circuit in these machines [9,10].

It is of great interest to use FRA as an online tool, without the need to stop the machine's operation to perform the tests. Ref. [11] presented a review on the efforts towards online application of FRA in transformers and discusses its issues, i.e., the coupling of the FRA equipment to the energized transformer. In the case of transformers, literature shows coupling through the capacitive layers of the bushings (forming a high pass filter): either for bushings with a voltage divider tap (as seen in Refs. [12–14]) or without a voltage divider tap (in this case, a capacitive sensor is adapted on the surface of the bushing, as proposed in Ref. [15]). Concerning rotating machines, the high pass filter must be installed externally to the machine. Ref. [16] has proposed a C-L-C filter as the coupling circuit to an energized induction generator and discussed in depth the trade-off between the attenuation of the power signal (to avoid damages to the FRA electronics) and the useful area obtained in the spectra.

Some studies [17,18] have shown that temperature and moisture may have some influence on the repeatability of FRA results in transformers. Specifically for rotating machines, some studies [19,20] have shown that the angular position of the rotor may also have influence on FRA results. This is of particular concern in case of online analysis of motors/generators, as the rotor may assume several positions within a same FRA sweep. Thus, it is of great importance to achieve a meaningful diagnosis even when the measurements have high uncertainty.

In order to reduce subjectivity of a visual comparison between spectra (usually being performed by a specialist), statistical indexes have been introduced in the literature. A survey on these indexes is provided in Ref. [21]. All of these indexes aimed to compare the

* Corresponding author at: Universidade Federal de Itajuba, Av. BPS, 1303, Itajuba, MG 37500-903, Brazil.
E-mail address: wilson_santana@ieee.org (W.C. Sant'Ana).

http://dx.doi.org/10.1016/j.epsr.2017.03.005
0378-7796/

magnitude spectra. Recently, Ref. [22] proposed a new index based on both magnitude and phase spectra. Either way, the result of any of these indexes is a number that gives a more objective indication of the machine condition. However, in case of external factors (temperature, humidity, rotor position, etc.) influencing the measurements, the calculated indexes would also carry the uncertainty, and this could result in a wrong diagnosis. To mitigate such errors, Refs. [20,21] proposed the analysis of a trend curve of the calculated indexes. However, when the uncertainty is large and the failures are very subtle (as the case analyzed in Section 5.2), more sophisticated techniques are required.

Statistical hypothesis tests are well suited to compare two sets of data and decide if their means are different or not, even when these sets have relatively high standard deviations. Some varieties of hypothesis tests have been used as FRA indexes: 2-sample *t*-test [23] and *f*-test anova1 [24]. However, the assumptions for these tests are normality, randomness and independence [25, p. 353], [26] and the tests performed in Refs. [23,24] were performed in an entire range of frequencies, thus intrinsically containing a trend. Ref. [26] proposed the use of non-parametric tests, which do not require the above assumptions. However, non-parametric tests are considered to be less efficient than the parametric ones (*t*-test, *f*-test, etc.), as they do not utilize all the information provided by the sample [25, p. 337].

This current paper proposes one calculation of a parametric test for each frequency in the spectra (thus respecting the assumptions above – as different data are not mixed together) and, then, the normalized sum of all the scores performed at the single frequencies. This results in an index with great potential to detect incipient failures in electric machinery, even in cases with high measurement uncertainty (such as the case of online measurements in rotating machines).

Section 2 presents an overview of the FRA technique and its interpretation through statistical indexes. Section 3 briefly explains the concepts of hypothesis testing. Section 4 presents the proposed index, based on the calculation of a hypothesis test for each of the frequencies in the spectra and their further normalized sum. Section 5 presents experimental results collected on a synchronous generator, both offline and online (at $220V_{RMS}/50Hz$), with inserted failures representing less than 1% of its stator winding.

2. Frequency response analysis and its interpretation through statistical indexes

Ref. [20] presented a detailed description of the FRA technique. Basically, it consists in the comparison of two spectra, taken at a different instant of time of the machine's lifespan. These spectra may be impedance spectra, admittance spectra or the spectra of any relation between two measurable quantities of the machine. The comparisons involve a reference set of measurements (called *baseline*) and the subsequent measurements. Differences between the spectra indicate that something has changed withing the machine, and may indicate the onset of a failure (either mechanical or electrical). However, an expert is usually required to analyze the differences in spectra and to provide a diagnosis about the machine condition.

Some statistical indexes have been proposed in the literature of FRA, in order to reduce the subjectivity of the analysis. The result of theses indexes is a quantitative score that indicates the machine condition. A survey on these indexes is provided in Ref. [21]. Among them, the ASLE (Absolute Sum of Logarithmic Error) seems to be the one preferred in the literature. The ASLE index can be calculated as Eq. (1). Also, recently, Ref. [27] investigated the susceptibility to measurement uncertainty of several statistical indexes and it was identified that the SSE (Sum of Squared Error)[1] was the one less vulnerable to uncertainties. The SSE index can be calculated as Eq. (2).

$$ASLE = \frac{\sum_{i=1}^{n} |20\log_{10} y_i - 20\log_{10} x_i|}{n}, \tag{1}$$

where $X = \{x_1, x_2, \ldots, x_n\}$ and $Y = \{y_1, y_2, \ldots, y_n\}$ are two FRA spectra, with n number of frequencies in each of them.

$$SSE = \frac{\sum_{i=1}^{n} (y_i - x_i)^2}{n}. \tag{2}$$

The use of the ASLE or the SSE indexes (as well as the majority of the indexes reviewed in Ref. [21]), however, only gives a numerical score on how much a spectrum Y differs from the spectrum X (the baseline) – it does not take into consideration the variability from one measurement to another (due to factors unrelated with the failures, such as temperature, humidity, etc.). Thus, in cases with high variability, such as the one analyzed in Section 5.2, the use of single comparisons between only two spectra at a time may result in wrong diagnosis.

An index based on hypothesis testing would be better suited to deal with measurement variability, as several samples are required to reach a conclusion within a certain degree of confidence. It is important to note that Refs. [23,24] have already proposed indexes using hypothesis tests, however, they still compared only two spectra at a time – not taking into consideration the measurement variability. Section 3 discusses the concepts of hypothesis testing, while Section 4 presents the proposed index.

3. Hypothesis testing

A hypothesis test is a statistical way to reject (or to accept) a research hypothesis (e.g., a population Y has different mean than a population X) within a certain degree of confidence. Chapters 9 and 10 of Ref. [25] discuss in depth the several types of hypothesis tests. In case of FRA, the most appropriate test would be the *2-sample t-test* – where the goal is to determine if the average measurements at a given condition is significantly different from the average measurements at the baseline condition. Considering two sets of data X and Y, with unknown population variances and considered to be not equal, a T-type statistic and the degrees of freedom must be calculated as Eqs. (3) and (4), respectively. In these equations, n_X and n_Y are the number of samples on each of the data sets X and Y and $\bar{X}, \bar{Y}, S_X^2$ and S_Y^2 are the sample means and sample variances of sets X and Y, respectively.[2] The sample means and sample variances are calculated as Eqs. (5) and (6), respectively.

$$t_0 = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}}. \tag{3}$$

$$\nu = \frac{\left(\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}\right)^2}{\frac{\left(\frac{S_X^2}{n_X}\right)^2}{n_X - 1} + \frac{\left(\frac{S_Y^2}{n_Y}\right)^2}{n_Y - 1}}. \tag{4}$$

$$\bar{X} = \frac{\sum_{i=1}^{n_X} x_i}{n_X}, \qquad \bar{Y} = \frac{\sum_{i=1}^{n_Y} y_i}{n_Y}. \tag{5}$$

[1] This same index is called MSE, Mean Squared Error, in Ref. [21].

[2] The sample means ($\bar{X}$ and $\bar{Y}$) and the sample variances (S_X^2 and S_Y^2) are estimates of the population means (μ_X and μ_Y) and population variances (σ_X^2 and σ_Y^2), which are unknown.