

Characterization of the Ethanol-Water Blend by Acoustic Signature Analysis in Ultrasonic Signals

ANDRÉ BERNARDI¹, LUIZ E. BORGES DA SILVA, (Senior Member, IEEE),
GISCARD F. C. VELOSO¹, AND JOSÉ A. FERREIRA FILHO¹

Institute of Systems Engineering and Information Technology, Federal University of Itajubá, Itajubá, Minas Gerais 37500-903, Brazil

Corresponding author: André Bernardi (andrebernardi@unifei.edu.br)

ABSTRACT The use of ethanol as fuel in Brazil stimulated the competition between distribution companies and resellers, which aggravated the practice of adulteration of fuels, aiming for illicit gains and tax evasion. The most common practice of adulteration in fuel alcohol is the addition of water. The classic techniques for measuring the water content in ethanol offer good precision and good detailing as to the presence of water. However, they present disadvantages such as the need for sample collection, long analysis time, in addition to the need for specialized laboratory and labor. This work aims to propose digital signal processing techniques to analyze and quantify the presence of water in ethanol fuel using a combination of Principal Component Analysis (PCA), and Singular Spectrum Analysis (SSA) applied on ultrasonic signals. This method resulted in the proposal of a new score, which relates to the ethanol/water ratio information present in the mixture. The results were promising when relating the proposed score to the presence of moisture in ethanol to a greater or lesser degree. Experiments performed prove the technique's feasibility and pave the way for a new method for real-time monitoring.

INDEX TERMS Acoustic signature, ethanol, principal component analysis (PCA), singular spectrum analysis (SSA), singular value decomposition (SVD), ultrasound.

I. INTRODUCTION

Ethanol is an alcoholic substance widely used as a fuel or additive to gasoline in the fuel industry. Due to its high solubility in water, the fast and effective determination of the water content in ethanol is essential to guarantee the quality of the product. The alcohol content below the allowable makes combustion more difficult due to the greater amount of water present in the fuel. Levels above the established increase the volatility of the fuel and can cause problems in the functioning of engines. As ethanol is added to water, the interactions between the molecules lead to changes in the molecular structure of the compounds [1]. As a result, there is a variation in the final sample volume. Due to volume dependence, some physical-chemical parameters, including the speed of sound in the sample and the density, show a nonlinear behavior depending on the ethanol concentration in the mixture [2]. The classic technique, like Karl Fischer's titration, is used for measuring the water content in ethanol [3], offering good precision and good detailing as to the presence of water. However, they present disadvantages, such as the need for

sample collection, long analysis time, in addition to the need for specialized laboratory and labor.

Ultrasound technology has been used to evaluate the composition of materials, such as insulating oil [4], [5], fuel oil [6], [7], foods, among others. There is a relationship between the ultrasonic velocity and the composition of the mixtures at a given temperature [8]. Since the change in ultrasonic velocity with temperature differs from substance to substance, measuring the ultrasonic velocity in a mixture at two temperatures would supply information about the volume fraction of the ingredients.

This work intends to find ways to discover the adulteration in ethanol fuel through the acoustic signature in the single-source ultrasound signal, with the measurement at only one temperature, and in a non-invasive way to allow online applications. The main objective of this work is to find a mechanism to characterize and classify the percentage of water present in ethanol fuel using digital signal processing techniques. The parameters used as the acoustic signature of the mixture are obtained by combining digital signal processing techniques applied to the signal captured from an ultrasonic pulse that passed through the sample to be classified. The method used in the development of this

The associate editor coordinating the review of this manuscript and approving it for publication was Manuel Rosa-Zurera.

focuses on applying a set of mathematical techniques in ultrasonic signals to obtain these parameters. These parameters will be used to classify the proportions of moisture in ethanol-water blends.

This work proposes a new methodology to reveal parameters in the acoustic signature of a mixture, based on the digital manipulation of data from ultrasonic transducers associated with a combination of mathematical techniques, such as Singular Value Decomposition (SVD), Singular Spectrum Analysis (SSA) and Principal Component Analysis (PCA), successfully used to estimate the water ratio in the ethanol-water mixture.

The paper is organized as follows: the related works used as a basis for its development are presented in Section II. In Section III, the mathematical concepts of SVD, PCA, and SSA are briefly presented and the proposed methodologies. In section IV, the proposed methodologies are applied in conducting experiments to generate data to prove the hypothesis of this work. A brief discussion of the results obtained in the experiments is also presented. Finally, the relevant conclusions are made in Section V.

II. RELATED WORKS

Veloso [9] proposed the PCA analysis through scores on acoustic signals from partial discharges in power transformers to find patterns of contamination in insulating oil through parameters obtained with the Discrete Wavelet Transform. A system using the same techniques in ultrasonic signals that can evaluate the contamination by moisture in insulating oil of power transformers was developed by Noronha [4]. Ultrasonic techniques were also applied in the fuel industry by Teixeira *et al.* [6] for the online monitoring of the degree of water contamination in heavy fuel oil, using the FFT of ultrasonic pulses captured by the sensor, after passing through the liquid having the mixture, as input data for the PCA. A principle like this is used to find a pattern in the acoustic signature of ultrasonic signals collected in water-ethanol fuel mixtures at different degrees of contamination. Thus, is used a combination of Fourier Transform and SSA parameters applied in the PCA.

Studies related to the analysis of ethanol-water mixtures have also been reported using ultrasound techniques. Vatandas and Koc [8] performed the identification of the type of alcohol and the volume concentration in binary water-ethanol/methanol mixtures using ultrasonic velocity measurements. D'Arrigo and Paparelli [10] analyzed ultrasonic velocity measurements in aqueous ethanol solutions from -40°C to $+30^{\circ}\text{C}$ over the entire range of composition of the mixture (0%-100% v/v), and in the range of frequency of 10-70 MHz. As a result, they proposed a model to explain the volumetric properties of ethanol mixtures at low temperatures and low solute concentration ranges. Possetti *et al.* [11] proposed a heterogeneous measurement system to determine the concentration of ethanol-based solutions on optical fibers and ultrasonic sensors to solve the ambiguous problem in the measures presented by the water-ethanol mixture. The

application of a transducer to check the quality of biofuels, combining optical fibers and ultrasound, was proposed by Kawano [1]. The works cited show that the ethanol-water mixture presents an ambiguous behavior about its physicochemical properties and proposes hybrid mechanisms to solve this nonlinearity. This work seeks to find a mechanism to resolve this ambiguity and nonlinearity in measurements using the acoustic signature revealed by digital signal processing techniques.

Maddirala and Shaik [12] introduced the SSA technique combined with Independent Component Analysis (ICA) to separate sources from single-channel EEG signals, proposing a new way of grouping for reconstructed SSA signals. Kuang *et al.* [13] proposed an efficient and adaptive denoising method based on the multistage SSA. Multistage SSA applies basic SSA with small window length recursively to the noisy signal. As an adaptive stop criterion, they used a correlation measure. Moreover, the method exploits the use of SVD on small matrices, making it more efficient. These studies show that new forms of grouping in SSA and even elementary grouping can be exploited to separate sources into signals in an optimized way.

Fu *et al.* [14], [15] successfully used a combination of techniques of PCA and SSA for feature extraction and data classification in hyperspectral imaging, proving to be a method with better performance when a small set of training is available. Compared to other state-of-the-art techniques, they still obtained a higher classification and accuracy in their proposed method, proving that the combination of these techniques has great potential to be explored.

III. PROPOSED METHODOLOGY

The proposed method is based on the combination of SSA-PCA techniques to derive an expression that relates the acoustic signature to the ethanol-water ratio.

A. MATHEMATICAL PRELIMINARIES

The SVD is the mathematical tool used in the PCA and SSA techniques, which are part of the signal analysis method exploited to reveal the acoustic signature. Properties related to the presence of water in the ultrasonic signal applied to the ethanol-water mixture are discussed below.

1) SVD

The Singular Value Decomposition (SVD) of a matrix is a core matrix decomposition method in linear algebra. It is referred to as the “fundamental theorem of linear algebra” [16] because it can be applied to all matrices, not just square matrices, and always exists. The SVD of a matrix \mathbf{A} , with rank r , is a linear transformation $\Phi: \mathbf{V} \rightarrow \mathbf{W}$, where $\mathbf{V} \in \mathbb{R}^n$ and $\mathbf{W} \in \mathbb{R}^m$, and can be interpreted as to its decomposition into three operations. In general, SVD performs a change of basis via \mathbf{V}^T followed by a scale and an increase (or reduction) in dimensionality through the matrix of singular values $\mathbf{\Sigma}$. Finally, it performs a second base change via \mathbf{U} .

In addition, the bases are orthonormal, the best type of base possible.

The SVD Theorem [16] can be written in the form:

$$A = U \Sigma V^T \tag{1a}$$

$$A = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_m \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_r & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{1a}$$

$$\times \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix} \tag{1b}$$

$$A = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_r \end{bmatrix} \begin{bmatrix} \sigma_1 \mathbf{v}_1^T \\ \sigma_2 \mathbf{v}_2^T \\ \vdots \\ \sigma_r \mathbf{v}_r^T \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{1c}$$

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T \tag{1d}$$

where: σ_i is a scalar, \mathbf{u}_i is a m -dimensional column vector, and \mathbf{v}_i is a n -dimensional column vector, each $\sigma_i \mathbf{u}_i \mathbf{v}_i^T$ is an $m \times n$ matrix, and the SVD equation decomposes matrix A into r matrices with the same shape ($m \times n$).

2) SSA

Singular Spectrum Analysis (SSA) is a method for decomposing time series and can be used efficiently to decompose signals to categorize their oscillation signatures over time [17]. It is a time series analysis technique that incorporates the elements of classical time series analysis, multivariate statistics, multivariate geometry, dynamic systems, and signal processing. This method helps analyze data with complex seasonal patterns and non-stationary trends, mainly where single-channel measurement is available. The SSA technique is a non-parametric method. One of its advantages is that it can be used without any assumptions, such as stationarity and normality of data [18].

This technique has been used successfully in different areas of time series analysis, such as biological signals such as EEG and ECG ([12], [19]–[23]), climatic and hydrological data [24], acoustic signals [25]–[28], economic and financial data and application on radars [29]. It consists of two complementary stages: *decomposition* and *reconstruction*. Each stage can be divided into two steps, as illustrated in Fig. 1, and these stages are discussed below.

a: DECOMPOSITION

The decomposition stage includes an embedding operation followed by an SVD.

Consider a time series with real value $Y = Y_N = (y_1, \dots, y_N)$ of length N . Suppose that $N > 2$ and Y is a non-zero series; that is, there is at least one i such that $y_i \neq 0$. Let L ($1 < L < N$) be an integer called window length and $K = N - L + 1$.

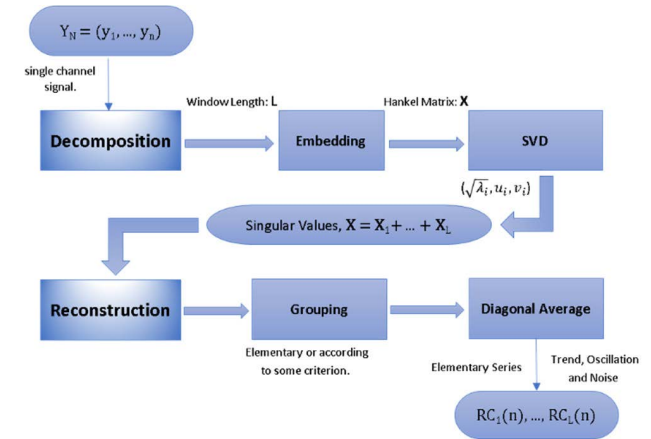


FIGURE 1. Representation of the singular spectrum analysis algorithm (SSA).

The one-dimensional time series Y is mapped into a multidimensional series by sliding a window of length L over the observed data Y_N of size N . A matrix is produced by stacking the signal segments that may have overlapping $L-1$. This procedure is called embedding and results in the trajectory matrix

$$X = [\mathbf{x}_1, \dots, \mathbf{x}_k] = (x_{ij})_{i,j=1}^{L,K} = \begin{bmatrix} y_1 & y_2 & \cdots & \cdots & y_K \\ y_2 & y_3 & \cdots & \cdots & y_{K+1} \\ \vdots & \vdots & \dots & \dots & \vdots \\ y_L & y_{L+1} & \cdots & \cdots & y_N \end{bmatrix} \tag{2}$$

with dimensions of L by K .

The embedding can be considered a mapping that transfers a one-dimensional time series $Y_N = (y_1, \dots, y_N)$ to the multidimensional series $\mathbf{x}_1, \dots, \mathbf{x}_K$ with vectors $\mathbf{x}_i = (y_i, \dots, y_{i+L-1})^T \in \mathbb{R}^L$. The \mathbf{x}_i vectors are called L -lagged vectors (or, simply, lagged vectors). The result of this step is the trajectory matrix X (2), which is a Hankel matrix, which means that all the elements along the diagonal, $i + j = \text{constant}$, they are equal.

The second step, the singular value decomposition step (SVD), decomposes the trajectory matrix X into its orthonormal bases and represents it as a sum of elementary bi orthogonal matrices of rank = 1. Denote $\lambda_1, \dots, \lambda_L$ as the eigenvalues of $C_X = X X^T$ in decreasing the order of magnitude ($\lambda_1 \geq \dots \lambda_L \geq 0$) and $\mathbf{u}_1, \dots, \mathbf{u}_L$ the corresponding eigenvectors forming an orthonormal system i.e., $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$ for $i \neq j$ (the orthogonality property) and $\|\mathbf{u}_i\| = 1$ (the

unit norm property). $\langle \mathbf{u}_i, \mathbf{u}_j \rangle$ is the internal product of the vectors \mathbf{u}_i and \mathbf{u}_j and $\|\mathbf{u}_i\|$ is the norm of the vector \mathbf{u}_i . Since $d = \text{rank}(\mathbf{X}) = \max(i, \text{where } \lambda_i > 0)$, and denoting $v_i = X^T u_i / \sqrt{\lambda_i}$, then the SVD of the trajectory matrix \mathbf{X} can be written as:

$$\mathbf{X} = \mathbf{X}_1 + \dots + \mathbf{X}_d \tag{3a}$$

$$\sum_{i=1}^d \mathbf{X}_i = \sum_{i=1}^d \sqrt{\lambda_i} u_i v_i^T = \sum_{i=1}^d u_i u_i^T \mathbf{X} \tag{3b}$$

where: $\mathbf{X}_i = \sqrt{\lambda_i} u_i v_i^T = u_i u_i^T \mathbf{X} (i = 1, \dots, d)$. Matrices \mathbf{X}_i have rank 1, so they are elementary matrices, \mathbf{u}_i and \mathbf{v}_i represent the left eigenvectors and right eigenvectors of the trajectory matrix \mathbf{X} . In SSA literature, these eigenvectors are commonly called “empirical orthogonal functions” (EOF) and “principal components” (PC), respectively. The term $\mathbf{u}_j \mathbf{u}_j^T$ forms a subspace for the j th component present in \mathbf{X} . Mapping \mathbf{X} in that subspace results in the trajectory matrix of the j th component in \mathbf{X} . In other words, \mathbf{u}_j is the base vector to extract the j th component present in the time series Y [16].

The collection $(\sqrt{\lambda_i}, u_i, v_i)$ is called the eigentriple of the matrix \mathbf{X} , $\sqrt{\lambda_i} (i = 1, \dots, d)$ are the singular values of the matrix \mathbf{X} , and the set $\{\sqrt{\lambda_i}\}_{i=1}^d$ is called the spectrum of matrix \mathbf{X} . If all eigenvalues have multiplicity one, then the expansion (3a) is uniquely defined.

The SVD of \mathbf{X} is based on the spectral decomposition of the covariance matrix $\mathbf{C}_X \in \mathbb{R}^{L \times L}$. It should be noted that the \mathbf{C}_X matrix is symmetric, positive definite, or positive semidefinite. Consequently, it has a complete set of eigenvectors and can be diagonalized in the form $\mathbf{U}\mathbf{\Sigma}\mathbf{U}^T$, where $\mathbf{\Sigma}$ is the diagonal matrix $L \times L$ of the eigenvalues and $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_L)$ is an orthogonal matrix of eigenvectors of the \mathbf{C}_X matrix. The name “singular spectrum analysis” comes from this property of the SVD technique and is a vital part. The whole process focuses on obtaining and analyzing this spectrum of singular values to find and differentiate between signal and noise in each time series.

b: RECONSTRUCTION

The reconstruction stage includes grouping and diagonal average operations to recreate the one-dimensional time series. Once the expansion (3b) is obtained, the grouping procedure partitions the set of indices $\{1, \dots, d\}$ into m disjoint subsets I_1, \dots, I_m . Let $I = \{i_1, \dots, i_p\}$. Then, the resulting matrix \mathbf{X}_I corresponding to group I is defined as $\mathbf{X}_I = \mathbf{X}_{i_1} + \dots + \mathbf{X}_{i_p}$. The resulting matrices are calculated for groups $I = I_1, \dots, I_m$ and the expansion (3b) leads to decomposition

$$\mathbf{X} = \mathbf{X}_{I1} + \dots + \mathbf{X}_{Im} \tag{4}$$

The grouping step corresponds to dividing the elementary matrices into several groups and the sum of the matrices within each group. For a given group I , the contribution of component \mathbf{X}_I is measured by the participation of the

corresponding eigenvalues:

$$\sum_{i \in I} \lambda_i / \sum_{i=1}^d \lambda_i \tag{5}$$

Several grouping criteria for the SSA technique are found in the literature, primarily based on the size of the eigenvalues of the covariance matrix \mathbf{C}_X [18], [22]. Maddirala and Shaik [12] proposed an alternative form of grouping based on the frequency components of the eigenvectors of the covariance matrix \mathbf{C}_X . This paper explores a grouping criterion that, in addition to considering the size of the eigenvalues, also considers the frequency information contained in the eigenvectors, combining the two methods of grouping the literature. Elementary grouping is also used.

After proceeding with the desired groupings, the next step for the signal reconstruction is calculating the diagonal average because the matrices obtained may not be in the form of a Hankel matrix. The diagonal average transforms a generic matrix into a Hankel matrix that can later be converted into a time series.

If z_{ij} is an element of a matrix \mathbf{Z} of dimensions $L \times K$, then the k th term of the resulting series is obtained by the mean of z_{ij} over all i, j such that $i + j = k + 1$. This corresponds to the average of the matrix elements on the antidiagonal $i + j = k + 1$: for $k = 1$ results $z_1 = z_{1,1} / 1$, for $k = 2$, $z_2 = (z_{1,2} + z_{2,1}) / 2$, for $k = 3$, $z_3 = (z_{1,3} + z_{2,2} + z_{3,1}) / 3$ and so on. Let $L^* = \min(L, K)$, $K^* = \max(L, K)$ and $N = L + K - 1$. Let $z_{ij}^* = z_{ij}$ also be $L < K$ and $z_{ij}^* = z_{ji}$ otherwise. Making the diagonal average, the matrix \mathbf{Z} is transferred to the series $z(1), \dots, z(N)$ using the formula:

$$z(n) = \begin{cases} \frac{1}{k} \sum_{m=1}^k z_{m,k-m+1}^* & \text{for } 1 \leq k < L^* \\ \frac{1}{L^*} \sum_{m=1}^{L^*} z_{m,k-m+1}^* & \text{for } L^* \leq k \leq K^* \\ \frac{1}{N-k+1} \sum_{m=k-K^*+1}^{N-L+1} z_{m,k-m+1}^* & \text{for } K^* < k \leq N \end{cases} \tag{6}$$

Applying the diagonal average to the matrices obtained through the SVD supplies reconstructed signals corresponding to the realized groupings. Therefore, the initial series x_1, \dots, x_N is decomposed into a sum of m reconstructed series using the SSA technique. The reconstructed series produced by the elementary cluster is called the reconstructed elementary series.

3) PCA

Principal Component Analysis (PCA) is one of the most used tools in modern data analysis in several areas [9]. It supplies a method of finding patterns in data and expressing them to highlight their similarities and differences [7].

Given a set of data \mathbf{X} , an $m \times n$ matrix with a zero-mean and standardized, where m is the number of different parameters and n corresponds to the number of measurements for each

parameter. The principal components of \mathbf{X} are the eigenvectors of its covariance matrix \mathbf{C}_X . When calculating the SVD of \mathbf{X} , the columns of matrix \mathbf{V} hold the eigenvectors of \mathbf{C}_X . Therefore, the columns of \mathbf{V} are the principal components of \mathbf{X} .

PCA seeks a linear combination of the original variables that produces orthogonal axes forming a new meaningful basis that can re-express the associated data set, filter noise, and reveal hidden information. It assumes that the data to be analyzed have a high signal-to-noise ratio (SNR). Principal components with more significant associated variances are exciting information, while those with lower variances represent noise or undesired information [6], [34], [35].

A detailed description of how the Principal Component Analysis technique can be deduced and how it can be applied is found in Shlens [35]. It is possible to deduce the concept of PCA intuitively and then, through mathematical rigor, derive its algebraic solution by EVD and SVD.

B. PROPOSED SSA-PCA TECHNIQUES

The earlier section shows that the PCA uses the singular vectors on the right of SVD, \mathbf{v}_i , and the SSA uses the singular vectors on the left \mathbf{u}_i . As both form the orthonormal basis of \mathbf{A} , this work explores techniques based on the projection of the X signal in both bases and compares them.

A combination of parameters from the PCA of the Fourier Transform and the SSA is used in this paper to search for patterns that can be used to classify the level of water contamination in ethanol.

1) INTERPRETATION OF THE PRINCIPAL COMPONENTS

As the eigenvectors of a covariance matrix have the characteristic of being orthogonal to each other, it implies data independence. It is seen that the principal components are orthogonal vectors in which the original data can be projected, and such projection reveals independent information present in the data. The interpretation of the principal components can be eased by defining *scores* as the projection of the original data in the eigenvectors of the covariance matrix, i.e., in the principal components. Each data set will generate components that must be analyzed to get to any interpretation. The search for patterns is done by associating each of the principal components to each of the scores, and it done using a bar graphs.

2) APPLICATION OF THE PCA TECHNIQUE TO THE RECONSTRUCTED ELEMENTARY SIGNALS OF THE SSA

The SSA technique decomposes a time series into elementary signals called reconstructed signals (\mathbf{RC}_i). To find patterns in the ultrasonic data collected, the SSA technique was used to decompose the pulse into its elementary spectral components associated with the singular vectors on the left (Columns of the \mathbf{U} matrix, \mathbf{u}_i). The Fourier Transform was applied to each reconstructed signal to get the parameters for PCA.

The block diagram of this approach is shown below in Fig. 2.

This method decomposes the digital signal corresponding to the ultrasonic pulses using SSA and the Fourier Transform. All frequency bins are considered the input parameters of the PCA. In this way, it is found which principal component has a linear behavior and can separate the acoustic signatures obtained from mixtures with different concentrations of ethanol-water.

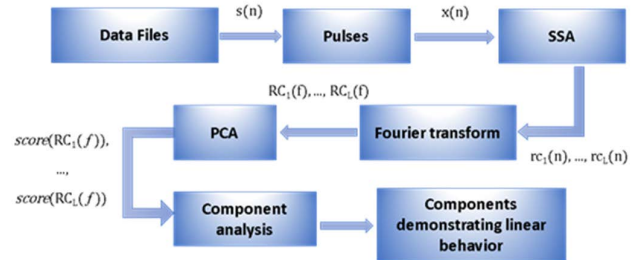


FIGURE 2. Methodology of application of the PCA analysis on the parameters of the FFT of the reconstructed signals.

3) PROJECTION OF THE ORIGINAL SIGNAL ON THE BASE FORMED BY THE RECONSTRUCTED SIGNALS USING THE SSA

The SSA technique can decompose a time series X representing a single channel signal into a sum of L reconstructed elementary signals. Each reconstructed elementary signal (\mathbf{RC}_i) is a component of the original signal X revealed by the SVD incorporated into the SSA [16], [18], [30]. As each component is obtained by projecting the original signal onto the left orthogonal vectors (\mathbf{u}_i) of the SVD, they are linearly independent [30]. This paper proposes a new scoring model based on input data projection on the base formed by reconstructed elementary signals by the SSA technique. In analogy to the score based on the principal components, a score based on the reconstructed elementary signals (\mathbf{RC}_i) obtained in the reconstruction of the time series due to the application of the SSA technique is submitted. Each obtained component is put in a column of the RC matrix. After being orthogonalized by the Gram Schmidt process, each project the original signal and obtains the *scoreRC* parameter. In matrix form, this proposed new score, called *scoreRC*, can be written as:

$$scoreRC = RC' * X \tag{7}$$

where the RC matrix has its columns composed by the reconstructed elementary signals of the SSA, after being orthogonalized by the Gram Schmidt process, and X is the time series holding the ultrasonic pulse captured by the acquisition system after crossing the mixture and aggregating information about the medium.

The proposed *scoreRC* solves the problem of nonlinearity, observed in the PCA score for certain training sets, noticed during the performed experiments.

The block diagram of this approach is shown below in Fig.3.

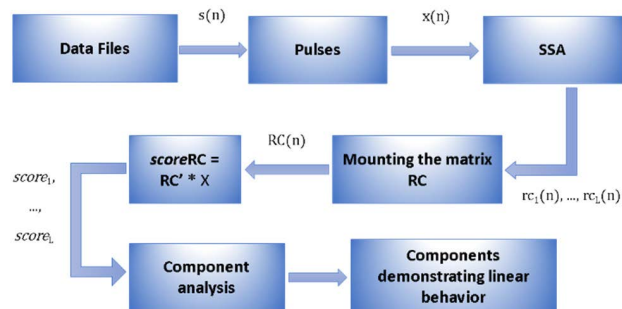


FIGURE 3. Methodology for obtaining standards in the scoreRC, through the analysis of its behavior.

IV. EXPERIMENTS AND RESULTS

Based on the methodologies described in the previous sections, several experiments were performed with ultrasonic signals obtained from the ethanol-water mixture in various concentrations.

A. EXPERIMENTAL SETUP

This section presents the system’s characteristics used to obtain the ultrasonic signals used in this work and essential information about the collected signals. The test bench description can be found in Noronha [4]. A schematic drawing is shown in Fig. 4.

Two ultrasonic transducers of the single crystal type S9208-AF84, from Physical Acoustic Corporation, are used. They work in pulsed mode, instrumenting a cuvette, one for transmitting and the other for receiving the signal. Both are positioned facing the liquid of the mixture. Each pulse applied to the transmitter is collected by the receiver after going through the liquid phase mixture. The resulting analog signal is converted to digital using a National Instruments system, with eight independent and simultaneous channels, variable sampling up to 2.5MSamples/s, and 14-bit resolution at -10V to +10V. Two channels were used, one for the acoustic signal and another for the trigger timing.

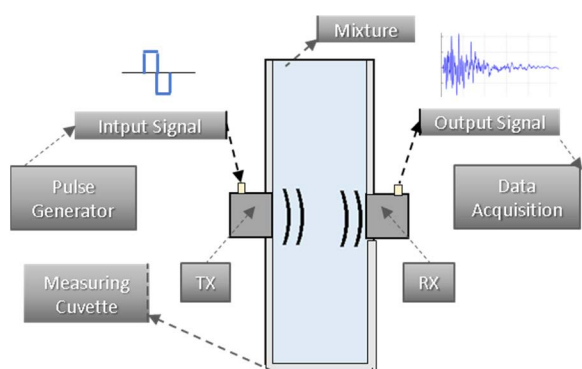


FIGURE 4. Simplified diagram of the experimental setup.

A series of mixtures with proportions, as shown in Table 1, were prepared from an initial volume of ethanol.

TABLE 1. Proportions of the mixtures.

| Mixture | % ethanol | % water |
|---------|-----------|---------|
| A | 100 | 0 |
| B | 99 | 1 |
| C | 98 | 2 |
| D | 97 | 3 |
| E | 96 | 4 |
| F | 95 | 5 |
| G | 90 | 10 |
| H | 85 | 15 |
| I | 80 | 20 |
| G | 75 | 25 |
| K | 70 | 30 |
| L | 65 | 35 |
| M | 50 | 50 |

A volume is placed in the stirrer. It is then drained to the measuring cuvette until it is filled. The first measurement is made. In each measurement, ten signals are collected. The cuvette is then emptied, and a new flow is made from the initial volume until filling the cuvette again. Another measurement is made. This is repeated until the fifth measurement. After that, a new volume with the new proportion to be measured is done by adding a certain amount of water to give the desired proportion, and the whole process is repeated.

The collected signals are composed of 100,000 points sampled at 2.5 MHz, having a certain number of acoustic pulses (Fig. 5 (a)). Thus, the first step in treating these signals is to extract each pulse. An algorithm based on the detection of the signal envelope is applied, and it results in signals like the one in Fig. 5 (b), where 4096 points are needed to record the part of the interest of each ultrasonic pulse. Under these conditions, the frequency resolution of the Fourier Transform is 610.35 Hz. Fig. 5(c) illustrates part of the frequency spectrum of a normalized ultrasonic pulse. For the energy contained in the pulse to be unitary, they are normalized by multiplying each sample of the signal by the inverse of the square root of its energy value.

Applying the SSA technique to the pulse shown in Fig. 5 (b), the reconstructed elementary signals (RC_i) are obtained, and the first six are shown in Fig. 6.

Each pulse is considered an experiment performed to be applied in the two techniques proposed in this work. The first technique decomposes the digital signal corresponding to the ultrasonic pulses using SSA and the Fourier Transform. All frequency bins are considered the input parameters of the PCA. In a second approach, the SSA will generate the reconstructed elementary signals corresponding to the ultrasonic pulse, which, represented in matrix form, are used to obtain a score for the % of water in the mixture.

B. SAMPLE CLASSIFICATION

The classification of the samples is shown by applying the PCA technique on samples resulting signals from

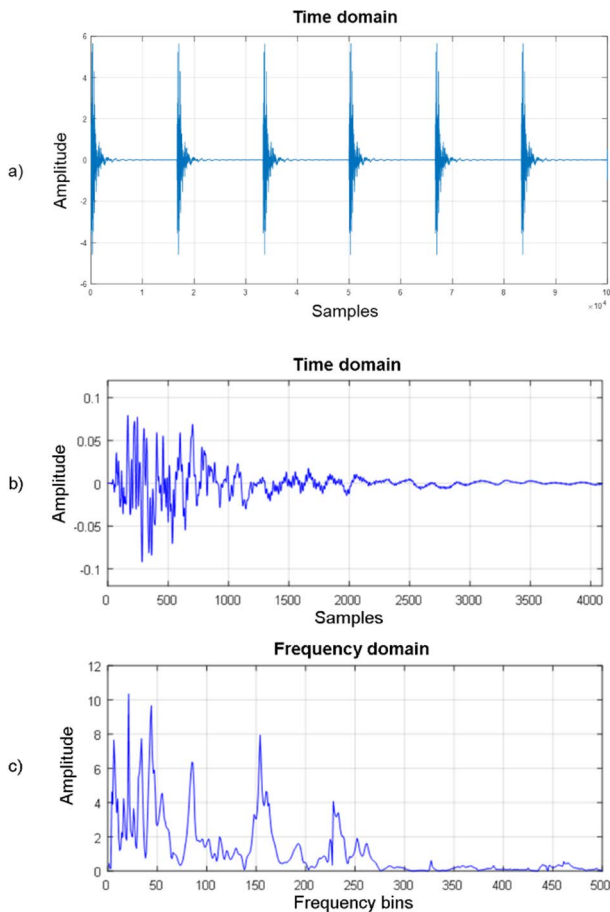


FIGURE 5. (a) Acoustic signal in the time domain, (b) Pulse extracted from the acoustic signal in the time domain, and (c) In the frequency domain by the FFT, for an ethanol-water mixture with a proportion of 50% each.

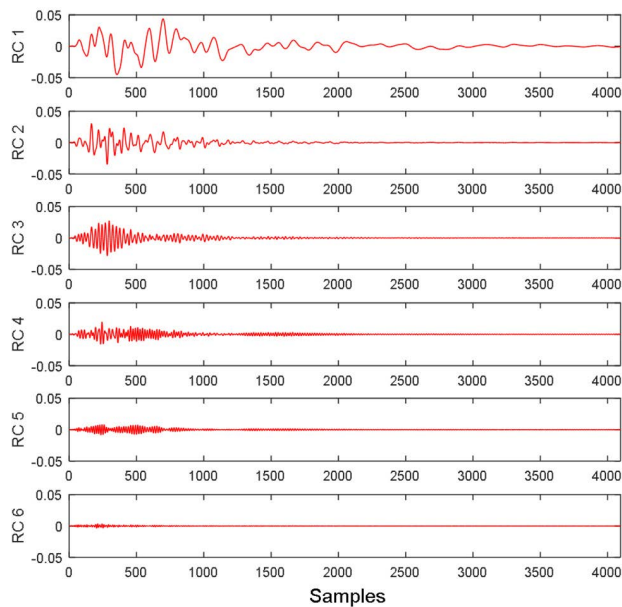


FIGURE 6. Reconstructed elementary signals (RC1 to RC6), applying the SSA technique.

ethanol-water blends with quite different levels of water. The signals that are obtained from mixtures A and M in Table 1 are used in this experiment. For each of the two contamination levels, 100 pulses are analyzed. The parameters obtained by applying the Fourier Transform on the pulses are considered column vectors and placed in the matrix for the PCA analysis. The reconstructed elementary signals (RC_i) are placed in columns of the matrix to generate the proposed *scoreRC*.

Applying the methodology described in III.B.2, a clear separation between the first 100 pulses from mixture A and the subsequent 100 pulses from mixture M is obtained, as shown in Fig. 7.

This same classification using the *scoreRC* defined in (7) is obtained as shown in Fig. 8, where different levels are evident between the 100 pulses corresponding to each level of contamination, in this case, about the first elementary signal reconstructed with the SSA technique.

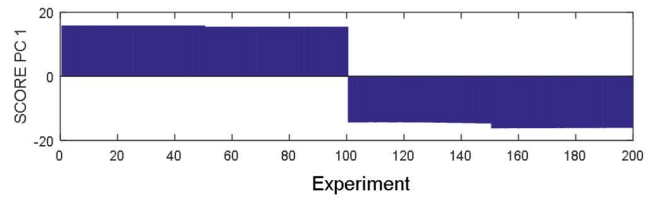


FIGURE 7. Scores of the first principal component (FFT).

From these results, the same techniques are applied to ultrasonic signals obtained from different proportions of the ethanol-water mixture described in Table 1, and their results are discussed below.

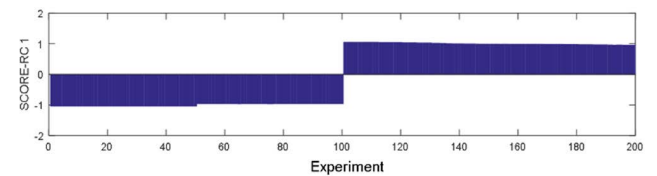


FIGURE 8. ScoreRC of the first reconstructed elementary signal.

C. ESTIMATION OF THE PROPORTION OF WATER IN THE ETHANOL-WATER MIXTURE

Experiments are conducted with different proportions in the mixture to analyze different ranges of moisture contamination in ethanol. The first experiment for a narrow range of moisture contamination ranging from 0 to 5% of 1 in 1% is represented by mixtures A to F in Table 1. The second experiment in mixtures with water proportions ranging from 0 to 35% every 5% and a third combining the signals used in the two initial experiments. Each of them is detailed in the next sections.

1) ADDITION OF WATER FROM 0 TO 5% IN 1% STEPS

As input data for the proposed methods, 50 ultrasonic pulses of each level of water contamination in ethanol is used, corresponding to six different ethanol-water proportions with the addition of 0%, 1%, 2%, 3%, 4%, and 5% water for the volume of ethanol, referring to mixtures A to F in Table 1.

The complete score analysis on PC#1 arising from the FFT parameters of ultrasonic pulses has shown a direct relationship between the percentage of water in the ethanol-water mixture, as shown in Fig. 9.

Pulses 1 to 50 come from the ethanol volume sample without added water, corresponding to Mixture A. The level obtained with the pulses from 51 to 100 comes from the score of the pulses captured in the ethanol-water mixture with the addition of 1% of water in the volume of ethanol (Mixture B). Pulses 101 to 150 come from Mixture C samples with 2% water added to the volume of ethanol. In experiments from 151 to 200, signals from Mixture D are used, which stands for an addition of 3% water in the volume of ethanol. From 200 to 250 (Mixture E) and from 251 to 300 (Mixture F), the levels refer to the additions of 4 and 5% of water in the volume of ethanol, respectively.

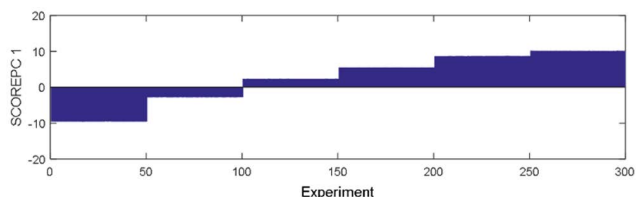


FIGURE 9. Score of the 1st principal component for each of the contamination levels (FFT).

The results for the *scoreRC*, represented in Fig. 10, show that the separations between the levels can be seen unequivocally for the third and fourth elementary signals. The fourth elementary signal is the one that would present the best results, as it shows the most significant difference of value between the levels obtained for each group in Table 1, used in this experiment, allowing a better approximation in the linear regression.

The *scoreRC* related to the first and second elementary signals have ambiguity in the threshold values, as shown in Fig. 10.

2) ADDITION OF WATER FROM 0 TO 35% IN 5% STEPS

In this experiment, 50 ultrasonic pulses of each level of water contamination in ethanol, corresponding to eight different ethanol-water proportions with the addition of 0%, 5%, 10%, 15%, 20%, 25%, 30%, and 35 % water in the volume of ethanol. The pulses are obtained by applying the ultrasonic signal at each level, represented by Mixtures A and F-L in Table 1.

The complete analysis of the score on PC#1 of the parameters resulting from the FFT of the ultrasonic pulses proved

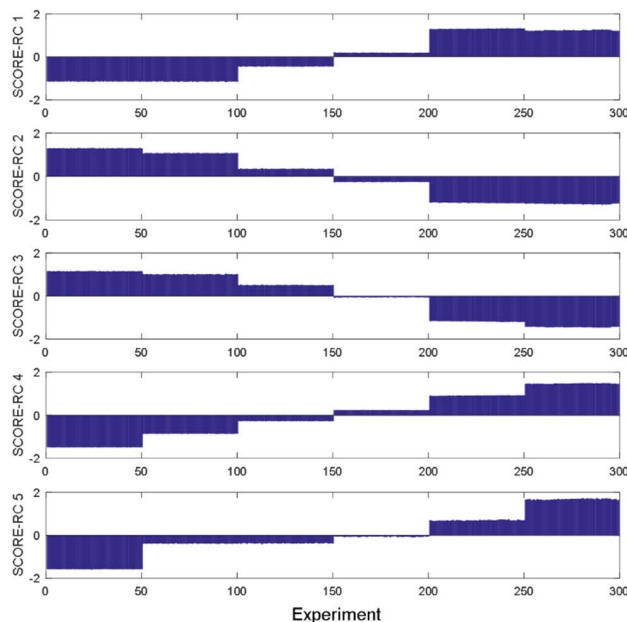


FIGURE 10. ScoreRC of the first five elementary signals reconstructed from the SSA.

a direct relationship between the percentage of water in the ethanol-water mixture, as can be seen in Fig. 11.

Pulses 1 to 50 come from the ethanol sample without adding water (Mixture A). The level obtained with pulses from 51 to 100 comes from the score of the pulses captured in Mixture F with the addition of 5% water to the volume of ethanol. The visible levels of 101 to 150, 151 to 200, 201 to 250, 251 to 300, 301 to 350, and 351 to 400 are, respectively, derived from the addition of 10%, 15%, 20%, 25%, 30%, and 35% water in the volume of ethanol, represented by mixtures of G to L.

The analysis in the *scoreRC* shows that the separations between the levels can be seen unambiguously for the second elementary signal, where a single value represents each proportion of water addition in the ethanol-water mixture, as seen in Fig. 12.

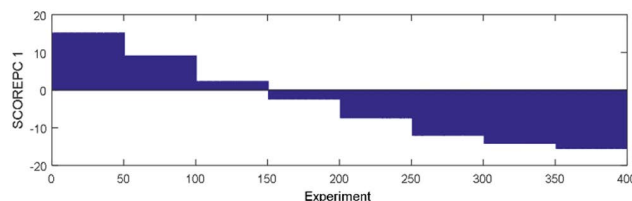


FIGURE 11. Score of the 1st principal component for each of the contamination levels (FFT).

3) ADDITION OF WATER FROM 0 TO 5% IN STEPS OF 1% AND FROM 5 TO 35% IN STEPS OF 5%

In this experiment, 50 pulses from each of Mixtures A to L in Table 1 were used, being the 12 proportions of water that

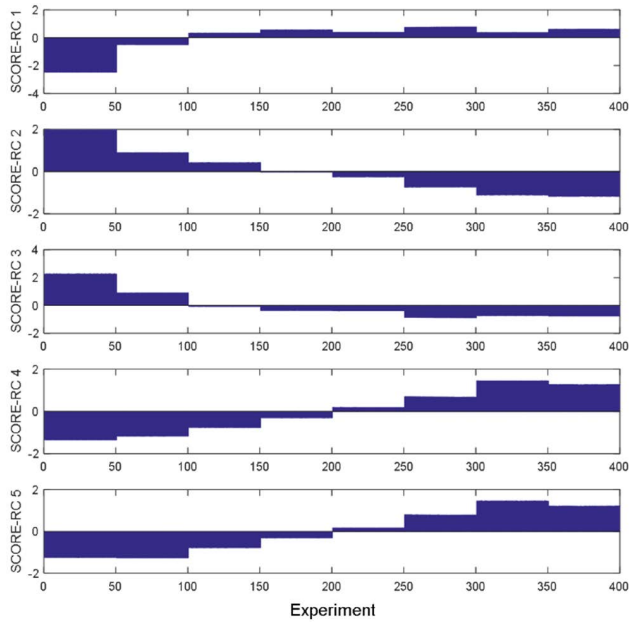


FIGURE 12. ScoreRC of the first five elementary signals reconstructed from the SSA.

were used in the mixture with ethanol, which corresponds to the addition of 0%, 1%, 2%, 3%, 4%, 5%, 10%, 15%, 20%, 25%, 30%, and 35% of water in the volume of ethanol. Fig. 13 illustrates the values of the score of the parameters coming from the FFT of the ultrasonic pulses in PC#1. It is possible to observe different proportions of water in the ethanol-water mixture related to levels of the same value, presenting an ambiguity for mixtures A and C, making it impossible to classify with this technique.

In Fig. 14, pulses 1 to 50 come from the ethanol sample without adding water (Mixture A). The level obtained with pulses from 51 to 100 comes from the *scoreRC* for pulses captured in Mixture B, which has a proportion of 1% of water addition in the volume of ethanol. The subsequent levels were obtained using the pulses obtained from mixtures C to L, which represent, respectively, the proportions of 2%, 3%, 4%, 5%, 10%, 15%, 20%, 25%, 30%, and 35% of addition of water in the volume of ethanol.

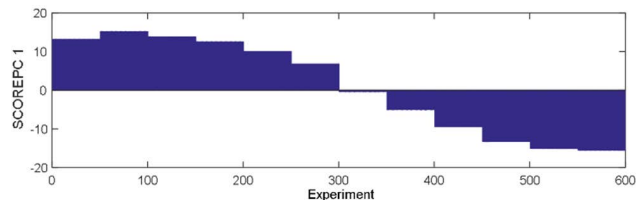


FIGURE 13. Score of the 1st principal component for each contamination level (FFT).

Unlike what occurred in the FFT PCA analysis seen in Fig. 13, using the analysis of *scoreRC*, the threshold for the proportions of water in the ethanol-water mixture of

the second elementary signal present uniqueness in values, enabling the classification of pulses in this hybrid set of proportions of water in the ethanol-water mixture.

D. SAMPLE CLASSIFICATION

Observing Fig. 7 to 14 presented in the earlier section, it is possible to prove the feasibility of applying the proposed techniques to classify the proportions of water in ethanol fuel, both in a narrow range of contamination (0% to 5%, steps of 1%) and in a broader range (0% to 35%, steps of 5%). For both ranges, the average values of the score obtained with the PCA of each of the groups can be used. As each group is related to a ratio of water to ethanol, using linear regression, one can obtain an equation where the x-axis corresponds to the value of the score. The y-axis is the ethanol-water ratio.

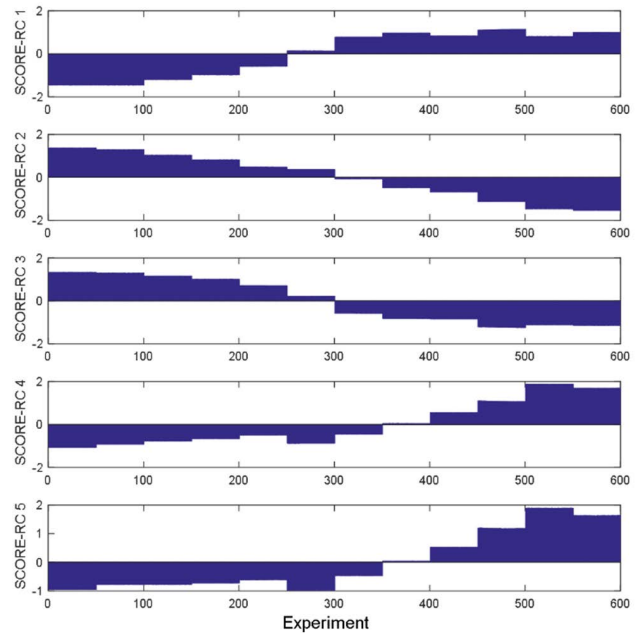


FIGURE 14. ScoreRC of the first five elementary signals reconstructed from the SSA.

It is possible to use the results presented in Fig. 14 to obtain a general expression combining the narrow and broad ranges. The average values of the *scoreRCs* of the second reconstructed elementary signal using Singular Spectrum Analysis (SSA) can be related to each contamination group. Likewise, each group is related to a proportion of water in ethanol, so one can also obtain an equation that relates the value of the *scoreRC* found with the ethanol-water ratio, allowing the classification of a particular sample.

The average *scoreRC* value for each group can thus be associated with the corresponding ethanol/water ratio value, as presented in Table 2. Thus, it is possible to obtain a relation between these values.

A third-order linear regression, shown in Fig. 15, results in the formula

$$y = -3.77398x^3 + 4.77176x^2 - 5.94885x + 7.45136 \quad (8)$$

TABLE 2. Water addition (%) versus average ScoreRC value.

| Mixture | % water | Average ScoreRC |
|---------|---------|-----------------|
| A | 0 | 1,19431 |
| B | 1 | 1,17267 |
| C | 2 | 1,10948 |
| D | 3 | 1,00439 |
| E | 4 | 0,76900 |
| F | 5 | 0,46091 |
| G | 10 | -0,37506 |
| H | 15 | -0,71887 |
| I | 20 | -0,81263 |
| G | 25 | -1,19046 |
| K | 30 | -1,26151 |
| L | 35 | -1,35221 |

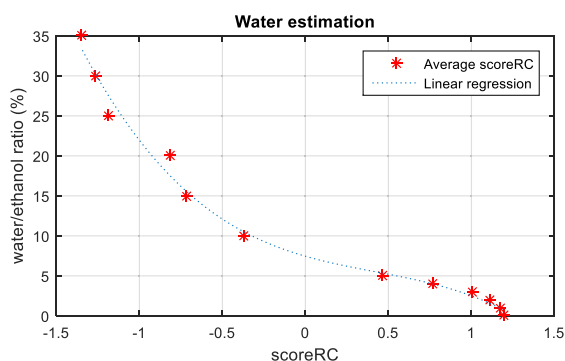


FIGURE 15. Water estimation according to scoreRC values for the second elementary signal reconstructed from the SSA, obtained after laboratory experiments.

where y corresponds to the water addition, given in %, and x corresponds to the $scoreRC$ for the second reconstructed elementary signal.

The % of water present in the ethanol-water mixture can be estimated using (8).

With these results, it was possible to obtain a non-invasive method using digital signal processing techniques. The method is based on the analysis of ultrasonic signals that crossed the ethanol-water mixture in different proportions and undergone changes in its characteristics due to the presence of water in it. So, this method provides a basis for classification by the acoustic signature.

V. CONCLUSION

This work aims to contribute to monitoring the quality of ethanol fuel related to moisture contamination through non-invasive techniques that allow online monitoring. The initial objective of verifying the possibility of detection through the analysis of ultrasonic signals that came into contact with the mixture in the liquid phase and had its characteristics altered by the presence of water was achieved.

To identify and quantify the water levels in the ethanol-water mixture, a combination of the Principal Component Analysis (PCA) and Singular Spectrum Analysis (SSA) techniques were used. This method resulted in the proposal of

a new score linked to the ethanol/water ratio information present in the mixture, which was used to classify the pulses of mixtures with different degrees of concentration.

The proposed methodologies aimed to determine which parameters were more sensitive to water in the mixture, making it possible to relate these acoustic signatures and obtain expressions that supply the ethanol/water ratio estimation.

Experimental results show that the proposed methodologies could indicate the moisture level in ethanol samples. The results obtained with the SSA algorithm led to an expression that links the $scoreRC$ and the ethanol/water ratio, given in %, allowing the classification of a particular sample.

Mathematical transformations were also applied to the signals to obtain parameters for the analysis. The tests performed with the Fourier Transform proved to generate acoustic signatures for the classification of contaminated samples in steps of 1% and 5% in isolation. The results obtained with the PCA technique related the scores and the ethanol-water ratio. These same results were obtained by applying the SSA technique and, using the $scoreRC$ proposed in this paper, i.e., the projection of the ultrasonic signal in the reconstructed elementary signals from the SSA.

A significant result was that when combining data from samples contaminated with steps of 1% with data from samples contaminated with steps of 5%, it was still possible to find parameters that could associate the proposed $scoreRC$ with the ethanol/water ratio, in this case, the analysis of the PCA score of the Fourier Transform acoustic signature presents ambiguities in the results, not allowing its use to generate an expression that directly relates the score to the ethanol-water ratio.

An application of this method, in practice, would consist of equipment similar to the one used in the experiments, but more compact, which would be coupled to the fuel supply line of a machine/engine and take measurements during fueling. The sensors would be previously calibrated in the laboratory using the same method described in this article. The measurement results would allow evaluating whether the fuel is acceptable and even inferring its yield/efficiency. Another use could be portable equipment for on-site verification of fuel quality by collecting samples in tanks, presenting the result instantly.

Finally, the proposed techniques also provide alternative methodologies for analyzing other contaminants and other mixtures.

REFERENCES

- [1] M. S. Kawano, "Aplicação de transdutores óticos para monitoração da qualidade de biocombustíveis," Ph.D. dissertation, Dept. Elect. Eng., Federal Univ. Technol. Paraná, Curitiba, PR, Brazil, 2015.
- [2] G. R. Possetti, G. G. Terada, R. J. Daciuk, C. Y. Ofuchi, F. Neves, J. L. Fabris, M. Müller, and L. V. R. Arruda, "Optical-ultrasonic heterogeneous sensor based on soft-computing models," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 9, pp. 2338–2346, Sep. 2015, doi: 10.1109/TIM.2015.2415071.
- [3] *Ethanol—Determination of Water Content—Volumetric Karl Fischer Method*, Standard ABNT NBR 15531:2009 Ing, Associação Brasileira de Normas Técnicas, Rio de Janeiro, SP, Brazil, 2009.

- [4] I. Noronha, "Sistema de avaliação da contaminação, por umidade, de óleo isolante de transformador usando técnicas de ultrassom," Ph.D. dissertation, Dept. Elect. Eng., Federal Univ. Technol. Itajubá, Itajubá, MG, Brazil, 2012.
- [5] V. C. M. N. Leite, G. F. C. Veloso, L. E. B. da Silva, G. Lambert-Torres, J. G. B. da Silva, and J. O. P. Pinto, "Moisture estimation in power transformer oil using acoustic signals and spectral kurtosis," *Meas. Sci. Technol.*, vol. 27, no. 3, Mar. 2016, Art. no. 035301, doi: [10.1088/0957-0233/27/3/035301](https://doi.org/10.1088/0957-0233/27/3/035301).
- [6] C. E. Teixeira, L. E. B. da Silva, G. F. C. Veloso, G. Lambert-Torres, M. M. Campos, I. Noronha, E. L. Bonaldi, and L. E. L. de Oliveira, "An ultrasound-based water-cut meter for heavy fuel oil," *Measurement*, vol. 148, Dec. 2019, Art. no. 106907, doi: [10.1016/j.measurement.2019.106907](https://doi.org/10.1016/j.measurement.2019.106907).
- [7] C. E. Teixeira, "Sistema de monitoramento da contaminação por umidade em óleos combustíveis através da análise de sinais ultrassônicos," Ph.D. dissertation, Dept. Elect. Eng., Federal Univ. Itajubá, Itajubá, MG, Brazil, 2019.
- [8] M. Vatasdas, A. B. Koc, and C. Koc, "Ultrasonic velocity measurements in ethanol-water and methanol-water mixtures," *Eur. Food Res. Technol.*, vol. 225, nos. 3-4, pp. 525-532, Jun. 2007, doi: [10.1007/s00217-006-0448-0](https://doi.org/10.1007/s00217-006-0448-0).
- [9] G. F. C. Veloso, "Identificação de padrões de contaminação em óleo de transformador através da monitoração acústica das descargas parciais," Ph.D. dissertation, Dept. Elect. Eng., Federal Univ. Itajubá, Itajubá, MG, Brazil, 2011.
- [10] G. D'Arrigo and A. Paparelli, "Sound propagation in water-ethanol mixtures at low temperatures. II. Dynamical properties," *J. Chem. Phys.*, vol. 88, no. 12, pp. 7687-7697, 1988.
- [11] G. R. C. Possetti, G. G. Terada, R. J. Daciuk, C. Y. Ofuchi, F. N. Junior, L. V. R. de Arruda, M. Müller, and J. L. Fabris, "Heterogeneous measurement system based on optical fiber and ultrasonic sensors to determine ethanol concentration," in *Proc. IEEE Sensors Conf.*, Limerick, Ireland, Oct. 2011, pp. 1169-1172, doi: [10.1109/ICSENS.2011.6127310](https://doi.org/10.1109/ICSENS.2011.6127310).
- [12] A. K. Maddirala and R. A. Shaik, "Separation of sources from single-channel EEG signals using independent component analysis," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 2, pp. 382-393, Feb. 2018, doi: [10.1109/TIM.2017.2775358](https://doi.org/10.1109/TIM.2017.2775358).
- [13] W. Kuang, S. Wang, Y. Lai, and W.-K. Ling, "Efficient and adaptive signal denoising based on multistage singular spectrum analysis," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1-20, 2021, doi: [10.1109/TIM.2020.3010426](https://doi.org/10.1109/TIM.2020.3010426).
- [14] H. Fu, G. Sun, J. Ren, A. Zhang, and X. Jia, "Fusion of PCA and segmented-PCA domain multiscale 2-D-SSA for effective Spectral-Spatial feature extraction and data classification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Nov. 2020, Art. no. 5500214, doi: [10.1109/TGRS.2020.3034656](https://doi.org/10.1109/TGRS.2020.3034656).
- [15] H. Fu, G. Sun, J. Zabalza, A. Zhang, J. Ren, and X. Jia, "A novel Spectral-Spatial singular spectrum analysis technique for near real-time *in situ* feature extraction in hyperspectral imaging," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 2214-2225, 2020, doi: [10.1109/JSTARS.2020.2992230](https://doi.org/10.1109/JSTARS.2020.2992230).
- [16] G. Strang, "The fundamental theorem of linear algebra," *Amer. Math. Monthly*, vol. 100, no. 9, pp. 848-855, Nov. 1993, doi: [10.1080/00029890.1993.11990500](https://doi.org/10.1080/00029890.1993.11990500).
- [17] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed. New York, NY, USA: Academic, 1990.
- [18] N. Golyandina, V. Nekrutkin, and A. Zhigljavsky, *Analysis of Time Series Structure: SSA and Related Techniques*, vol. 90. London, U.K.: Chapman & Hall, 2001.
- [19] S. M. Mohammadi, S. Kouchaki, M. Ghavami, and S. Sanei, "Improving time-frequency domain sleep EEG classification via singular spectrum analysis," *J. Neurosci. Methods*, vol. 273, pp. 96-106, Nov. 2016, doi: [10.1016/j.jneumeth.2016.08.008](https://doi.org/10.1016/j.jneumeth.2016.08.008).
- [20] F. Ghaderi, H. R. Mohseni, and S. Sanei, "Localizing heart sounds in respiratory signals using singular spectrum analysis," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 12, pp. 3360-3367, Dec. 2011, doi: [10.1109/TBME.2011.2162728](https://doi.org/10.1109/TBME.2011.2162728).
- [21] A. K. Maddirala and R. A. Shaik, "Removal of EOG artifacts from single channel EEG signals using combined singular spectrum analysis and adaptive noise canceler," *IEEE Sensors J.*, vol. 16, no. 23, pp. 8279-8287, Dec. 2016, doi: [10.1109/JSEN.2016.2560219](https://doi.org/10.1109/JSEN.2016.2560219).
- [22] S. Sanei and A. Hosseini-Yazdi, "Extraction of ECG from single channel EMG signal using constrained singular spectrum analysis," in *Proc. 17th Int. Conf. Digit. Signal Process. (DSP)*, Corfu, Greece, Jul. 2011, pp. 1-4, doi: [10.1109/ICDSP.2011.6004876](https://doi.org/10.1109/ICDSP.2011.6004876).
- [23] J. Cheng, L. Li, C. Li, Y. Liu, A. Liu, R. Qian, and X. Chen, "Remove diverse artifacts simultaneously from a single-channel EEG based on SSA and ICA: A semi-simulated study," *IEEE Access*, vol. 7, pp. 60276-60289, 2019, doi: [10.1109/ACCESS.2019.2915564](https://doi.org/10.1109/ACCESS.2019.2915564).
- [24] R. Vautard and M. Ghil, "Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series," *Phys. D, Nonlinear Phenom.*, vol. 35, no. 3, pp. 395-424, 1989, doi: [10.1016/0167-2789\(89\)90077-8](https://doi.org/10.1016/0167-2789(89)90077-8). [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0167278989900778>
- [25] D. Y. Mohammed, "Overlapped speech and music segmentation using singular spectrum analysis and random forests," Ph.D. dissertation, School Comput., Sci. Eng., Salford Univ., Manchester, U.K., 2017.
- [26] F. G. Encinas, L. A. Silva, A. S. Mendes, G. V. Gonzalez, V. R. Q. Leithardt, and J. F. De Paz Santana, "Singular spectrum analysis for source separation in drone-based audio recording," *IEEE Access*, vol. 9, pp. 43444-43457, 2021, doi: [10.1109/ACCESS.2021.3065775](https://doi.org/10.1109/ACCESS.2021.3065775).
- [27] M. Kutsenko, V. Ovcharuk, and D. B. Solovov, "Application of singular value decomposition method for acoustic emission data analysis," in *Proc. Int. Sci. Technol. Conf. 'EastConf'*, Vladivostok, Russia, Mar. 2019, pp. 1-4, doi: [10.1109/EastConf.2019.8725314](https://doi.org/10.1109/EastConf.2019.8725314).
- [28] Y. Lu and J. Saniie, "Singular spectrum analysis for trend extraction in ultrasonic backscattered echoes," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Taipei, Taiwan, Oct. 2015, pp. 1-4, doi: [10.1109/ULTSYM.2015.0440](https://doi.org/10.1109/ULTSYM.2015.0440).
- [29] H.-G. Ma, Q.-B. Jiang, R. Lei, Z.-Q. Liu, and X.-Y. Kong, "Determine a proper window length for singular spectrum analysis," in *Proc. IET Int. Conf. Radar Syst. (Radar)*, Glasgow, U.K., 2012, pp. 1-6, doi: [10.1049/cp.2012.1722](https://doi.org/10.1049/cp.2012.1722).
- [30] G. Strang, *Matrix Methods in Data Analysis, Signal Processing, and Machine Learning*. Cambridge, MA, USA: MIT, Spring 2018. [Online]. Available: <https://ocw.mit.edu>
- [31] B. Pang, G. Tang, and T. Tian, "Complex singular spectrum decomposition and its application to rotating machinery fault diagnosis," *IEEE Access*, vol. 7, pp. 143921-143934, 2019, doi: [10.1109/ACCESS.2019.2945369](https://doi.org/10.1109/ACCESS.2019.2945369).
- [32] J. Harmouche, D. Fourer, F. Auger, P. Borgnat, and P. Flandrin, "The sliding singular spectrum analysis: A data-driven nonstationary signal decomposition tool," *IEEE Trans. Signal Process.*, vol. 66, no. 1, pp. 251-263, Jan. 2018, doi: [10.1109/TSP.2017.2752720](https://doi.org/10.1109/TSP.2017.2752720).
- [33] J. B. Elsner and A. A. Tsonis, *Singular Spectrum Analysis: A New Tool in Time Series Analysis*. New York, NY, USA: Springer, 2013, p. 184.
- [34] I. T. Jolliffe, *Principal Component Analysis* (Springer Series in Statistics), 2nd ed. New York, NY, USA: Springer, 2002.
- [35] J. Shlens, *A Tutorial on Principal Component Analysis*. Mountain View, CA, USA: Cornell Univ., 2014.
- [36] J. Fu, F. Cai, Y. Guo, H. Liu, and W. Niu, "An improved VMD-based denoising method for time domain load signal combining wavelet with singular spectrum analysis," *Math. Problems Eng.*, vol. 2020, Apr. 2020, Art. no. 1485937, doi: [10.1155/2020/1485937](https://doi.org/10.1155/2020/1485937).



ANDRÉ BERNARDI was born in Bragança Paulista, Brazil. He received the B.Sc. and M.Sc. degrees in electrical engineering from the Federal University of Itajubá (UNIFEI), Itajubá, Minas Gerais, Brazil, in 1994 and 1999, respectively, where he is currently pursuing the Ph.D. degree.

He is currently a Professor with the Federal University of Itajubá. His research interests include time-frequency analysis, signal processing, digital signal processing applied to biomedical problems and instrumentation, and machine learning.



LUIZ E. BORGES DA SILVA (Senior Member, IEEE) was born in Passa Quatro, Brazil. He received the B.Sc. and M.Sc. degrees in electrical engineering from the Federal University of Itajubá (UNIFEI), Itajubá, Brazil, in 1977 and 1982, respectively, and the Ph.D. degree from the École Polytechnique de Montréal, Montréal, QC, Canada, in 1988.

He was a Visiting Professor with the University of Tennessee, Knoxville, in 1998. He is currently a Professor with the Electronic Engineering Department, UNIFEI, where he is also the Head of Power Electronics. He has directed many projects in the field of industrial electronics and coauthored over 200 technical papers. He has supervised more than 30 master and doctoral theses and dissertations. His research interests include power electronics, electronic power systems, power converters, and adaptive and intelligent control applications in industrial problems.



JOSÉ A. FERREIRA FILHO received the B.Sc. and M.Sc. degrees in electrical engineering from the Federal University of Itajubá (UNIFEI), Itajubá, Minas Gerais, Brazil, in 1986 and 1992, respectively.

He is currently a Professor with the Department of Electronic Engineering, UNIFEI. His research interest includes biomedical engineering.

...



GISCARD F. C. VELOSO received the B.Sc., M.Sc., and Ph.D. degrees from the Federal University of Itajubá (UNIFEI), Itajubá, Brazil, in 2003, 2006, and 2011, respectively, all in electrical engineering.

He is currently a Professor with the Department of Electronic Engineering, UNIFEI. His research interests include predictive maintenance and fault monitoring on partial discharges in power transformers and moisture contamination in power transformer oil.