A New Approach for Event Detection in Smart Distribution Oscillograph Recorders

André E. Lazzaretti, Vitor H. Ferreira, Hugo Vieira Neto, Luiz F. R. B. Toledo and Cleverson L. S. Pinto

***Abstract*—This paper presents some new results for a fundamental step in automatic oscillography analysis: transient detection. We performed experiments with usual detection methods, such as the Kalman filter (KF) and autoregressive (AR) models, and we are proposing a new method based on the Discrete Wavelet Transform (DWT) and Support Vector Data Description (SVDD). Data simulated in the Alternative Transient Program (ATP) was generated for comparison and validation of detection performance. The results presented here demonstrate that the proposed detection method based on DWT and SVDD yields better overall performance for the transient detection process when compared to currently used methods such as KF and AR models. These results show the potential for possible embedded applications in automatic oscillographic recorders in smart distribution networks, in which identification, characterization, and mitigation of events is critical for network operation and maintenance.**

***Index Terms*—Oscillography Segmentation, Transient Detection, Wavelet Transform, Support Vector Data Description, Automatic Waveform Analysis.**

I. Introduction

The analysis of voltage and current oscillographies is one of the most important tools for event discrimination in electrical power system networks. This analysis provides a basis for mitigation processes, maintenance, and fault characterization, resulting in a support system for decision making by power utilities [1].

The voltage and current variations that occur due to the transition between two events are critical for event segmentation and characterization in a recorded waveform. Basically, the detection process involves two distinct stages: the identification of instants at which transitions occur (triggering), and segmentation of the waveform in disjoint time blocks. The segmentation stage splits the signal in quasi-stationary states, allowing independent waveform analysis in each one of them [2].

In figure 1 there are examples of transients in an oscillography, which are marked by rectangles that characterize the triggering process. Between transients are the segments of an oscillographic record, or quasi-stationary states. The joint analysis of triggering and segmentation provides a complete interpretation of each event, although it is noticeable in figure 1 that most of the events have their main characteristics in quasi-stationary states, once they are related to short-circuits or switching events.

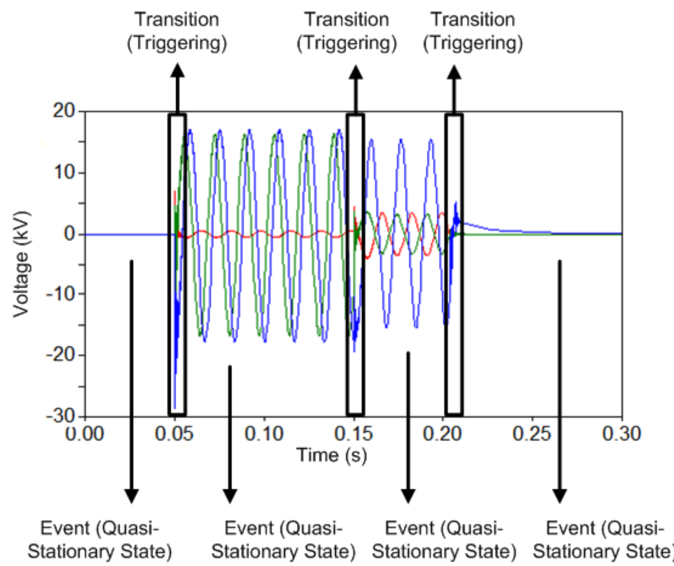

Fig. 1. Triggering and segmentation of an oscillographic record.

There are several methods for triggering and segmentation in oscillographies, which are usually dealt with simultaneously. Generally, these methods are grouped according to the nature of preprocessing applied to the signals under analysis. The main groups are [2]:

- Time domain analysis of variations;
- Frequency domain filtering;
- Time-frequency domain analysis;
- Spectral estimation analysis of most prominent residuals;
- Machine learning methods.

In all methods, the simplest and most commonly used involve time domain signal analysis. These methods determine the beginning and the end of an event through the extraction of features such as the RMS value of the signal. In [3], it is shown that the performance of these methods is suitable only for events with large variations in the RMS value in quasi-stationary states. However, for small variations in the RMS value, the performance of time domain methods can affect the temporal localization accuracy of events, especially for short duration transients.

A first alternative to overcome the difficulties that appear in time domain methods is the use of filters for better characterization of transitions between segments of a waveform [4]. In

A. E. Lazzaretti and Luiz F. R. B. Toledo are with the Institute of Technology for Development (LACTEC), Avenida Comendador Franco 1341, Curitiba-PR, Brazil.

V. H. Ferreira is with the Federal University Fluminense (UFF), Rua Passo da Pátria 156, Niterói-RJ, Brazil.

H. Vieira Neto is with the Federal University of Technology – Paraná (UTFPR), Avenida Sete de Setembro 3165, Curitiba-PR, Brazil.

Cleverson L. S. Pinto is with Companhia de Energia do Paraná (COPEL Distribuição), Rua José Izidoro Biazetto 158, Curitiba-PR, Brazil.

[5], a procedure for filtering the signal in the frequency domain was proposed, whose main feature is to make the transitions between states more prominent, even in noisy conditions. Moreover, the proposed filter has a low computational cost, which makes it more attractive than conventional time domain methods.

Still in this context, one of the major innovations in event detection was the use of time-frequency transforms, in particular the Discrete Wavelet Transform [2]. In this type of transform, the temporal information available in the signals is preserved in different frequency bands. This is an important feature, once the observed transients in electrical power systems events have a fairly wide range of frequency content.

For wavelet-based detection methods, the approaches proposed by Ukil and colleagues [4], [6]–[9] are noteworthy. They consist of computing a wavelet decomposition of the signals being analyzed and then automatically segmentating disturbances. For this, the universal threshold presented in [10] is computed and followed by a smoothing filter. This procedure guarantees that the whole detection process is done without the need of empirical thresholds, avoiding a supervised processing stage normally present in waveform analysis [11].

Regarding spectral estimation methods, we highlight the use of the Kalman filter [12]–[14] and autoregressive models [3], whose performance is comparable to the performance of time-frequency methods. Accordingly, spectral estimation and time-frequency methods can be considered the most frequently used methods in recent studies involving event detection [2].

Another group of techniques used for event detection, not yet completely explored in electrical power systems, is based on machine learning [15]. However, for signals from other domains, these methods are well exploited, mainly by novelty detection methods. In [16]–[18], one-class Support Vector Machines (SVM) were applied for signals in different applications and the results are comparable to spectral estimation methods, both in terms of performance and computational complexity. An interesting advantage of this approach, although, is the possibility of a combined representation, by using a DWT as preprocessing stage and an SVM-based novelty filter as processing stage, exploiting all the benefits provided by time-frequency and novelty detection representations.

Therefore, this paper compares currently used spectral estimation methods – Kalman filter and autoregressive models – and a machine learning technique based on Support Vector Data Description and a Discrete Wavelet Transform, following the idea of the one-class SVM proposed in [18] and the segmentation methods proposed in [14]. This new approach for event detection in electrical power systems using a combined DWT and SVDD representation yields great performance for automatic segmentation when compared to currently employed techniques.

It is also important to note that the oscillograph event detection may be of fundamental importance in a smart distribution network – an automatic process of detection may allow faster identification and analysis of faults in the network, supporting the process of decision making by operation engineers.

II. Theoretical Aspects

A. Discrete Wavelet Transform

The main motivation for DWT is the time-scale signal decomposition in frequency sub-bands, using orthonormal bases obtained from digital filter banks [19]. The signal is processed by a series of low- and high-pass filters that separate low- and high-frequency components in different subspaces. The filters construction is based on the wavelet function properties, as defined in [20] and [19]. Thus, orthonormal bases of discrete wavelet functions are not only associated with the mother wavelet function, but also with the scale function. The mother wavelet function is associated with details or high-pass filters, whereas the scale function is related to approximations or low-pass filters, forming an orthonormal basis.

Decomposition of the input signal into approximation and detail coefficients is the foundation of multi-resolution analysis [19], and can be done using a pair of finite impulse response filters – a high- and a low-pass filter for the decomposition process, as well as their conjugates for the reconstruction process. In this way, the resolution analysis can be associated with filtering operations, and the scale analysis can be associated with downsampling and upsampling operations during decomposition and reconstruction, respectively. Once the signal is decomposed, the most prominent frequency components result in high amplitudes in the DWT sub-bands that include these particular frequencies, retaining the temporal localization of the frequency components, different from that which occurs in the Fourier Transform.

The wavelet decomposition procedure presents good time-domain resolution for high-frequency components and good frequency-domain resolution for low-frequency components. These properties make an alternative spectral representation to the one given by the Fourier Transform possible, using nonlinearly spaced frequency sub-bands, allowing specific components localization of the signal under analysis, which consists of a very important characteristic for power distribution waveform analysis. More details of the procedure of DWT can be found in [20] and [19].

B. Support Vector Data Description

The main feature of the SVDD model is a representation of the input data in a high dimensional space without the need of large additional computational effort [21]. This representation allows more flexible descriptors of the input data, following the same idea of Support Vector Machines [22].

It is assumed that for a given input pattern $\mathbf{x}$, there is a closed surface – more precisely, a hypersphere – that surrounds it and is characterized by its center $\mathbf{a}$ and radius R. For the particular case where all the input patterns (training set) are located within a single hypersphere, the empirical risk is null. Thus, one can define the structural risk as:

$$\varepsilon(R, \mathbf{a}) = R^2, \tag{1}$$

which must be minimized with the following restrictions:

$$\|\mathbf{x}_i - \mathbf{a}\|^2 \leq R^2, \forall i. \tag{2}$$

When one takes into account the possibility that there are outliers in the training set, it is necessary to include an additional term in the initial formulation. This additional term penalizes patterns farther from the edge of the hypersphere (outliers or novelties), keeping the trade-off between empirical and structural risks, since it results in a training error that is different than zero. With the new formulation, the minimization problem becomes:

$$\varepsilon(R, \mathbf{a}, \xi) = R^2 + C \sum_i \xi_i, \tag{3}$$

grouping most patterns within the hypersphere:

$$\|\mathbf{x}_i - \mathbf{a}\|^2 \leq R^2 + \xi_i,\ \xi_i \geq 0,\ \forall i. \tag{4}$$

where C is the model structural control, which keeps the trade-off between the volume of the hypersphere and the empirical risk.

The parameters R, $\mathbf{a}$ and ξ can be obtained by optimizing the functional (5), whose restrictions were included using Lagrange multipliers (α_i and γ_i).

$$L(R, \mathbf{a}, \xi, \alpha, \gamma) = R^2 + C \sum_i \xi_i - \tag{5}$$

$$\sum_i \alpha_i [R^2 + \xi_i - (\mathbf{x}_i.\mathbf{x}_i - 2\mathbf{a}.\mathbf{x}_i + \mathbf{a}.\mathbf{a})] - \sum_i \gamma_i \xi_i.$$

In order to adapt the problem to quadratic programming [23], partial derivatives are calculated from functional (5) with respect to R, $\mathbf{a}$, ξ and equated to zero, which results in the following formulation:

$$L = \sum_i \alpha_i (\mathbf{x}_i.\mathbf{x}_i) - \sum_{i,j} \alpha_i \alpha_j (\mathbf{x}_i.\mathbf{x}_j), \tag{6}$$

subject to the following restrictions:

$$0 \leq \alpha_i \leq C, \quad \forall i. \tag{7}$$

Through Lagrange multiplier constraint analysis, it is possible to establish the location of a given pattern with respect to the edges of the hypersphere. That is, a pattern may be located within, on the edge, or outside the edge (outlier) of the hypersphere. The patterns located on the edge of the hypersphere with nonzero α_i are called support vectors, since they are responsible for edge characterization of the hypersphere. Other patterns can be discarded, as their α_i are equal to zero.

In order to develop the hypersphere in feature space, it is necessary to perform a mapping Φ of $\mathbf{x}$ to the new space, resulting in mapped patterns $\mathbf{x}^*$. Therefore, the equation can be rewritten:

$$L = \sum_i \alpha_i \Phi(\mathbf{x}_i).\Phi(\mathbf{x}_i) - \sum_{i,j} \alpha_i \alpha_j \Phi(\mathbf{x}_i).\Phi(\mathbf{x}_j). \tag{8}$$

Because the mapping is applied only for the vectors on which the inner product is evaluated, it is possible to use the Mercer kernel representation [22] without the need of explicitly mapping calculations, as $K(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i).\Phi(\mathbf{x}_j)$. This results in the following modification:

$$L = \sum_i \alpha_i K(\mathbf{x}_i, \mathbf{x}_i) - \sum_{i,j} \alpha_i \alpha_j K(\mathbf{x}_i, \mathbf{x}_j). \tag{9}$$

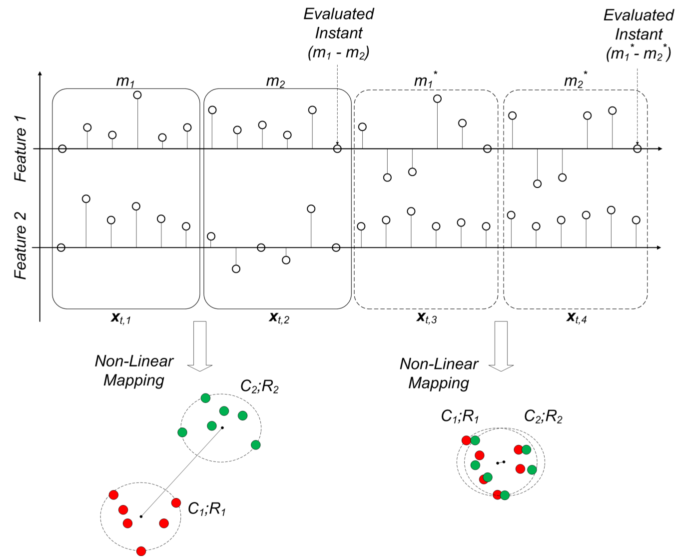

Fig. 2. Distance computation using SVDD hyperspheres.

By using the Gaussian kernel, the Lagrangian is given by:

$$L = -\sum_{i,j} \alpha_i \alpha_j K(\mathbf{x}_i, \mathbf{x}_j). \tag{10}$$

In this representation, a pattern $\mathbf{z}$ is considered a novelty, if

$$\sum_i \alpha_i \exp\left(\frac{-\|\mathbf{z}-\mathbf{a}\|^2}{\sigma}\right) > \tag{11}$$

$$\frac{1}{2}\left[\sum_{i,j} \alpha_i \alpha_j \exp\left(\frac{-\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma}\right) - R^2\right],$$

otherwise the pattern is within the hypersphere in feature space.

Finally, the center of the hypersphere may be computed using the expression:

$$\mathbf{a} = \sum_i \alpha_i \mathbf{x}_i, \tag{12}$$

and its radius given by the distance between $\mathbf{a}$ and one of the support vectors.

C. SVDD for Event Detection

Assuming a pattern recognition problem in which input data is composed by two attributes, for a given time instant t, two windows are extracted for posterior analysis, generating the sequences $\mathbf{X}_1 = (\mathbf{x}_{t-m_1}, \ldots, \mathbf{x}_{t-1})$ and $\mathbf{X}_2 = (\mathbf{x}_t, \ldots, \mathbf{x}_{t+m_1-1})$, whose dimensions are $(2, m_1)$ and $(2, m_2)$, respectively [18]. Assuming also that on each of these sequences an SVDD model is applied, the input data for each sequence ($\mathbf{X}_1$ and $\mathbf{X}_2$) is represented by the hyperspheres with radius R_1 and R_2, and centers $\mathbf{c}_1$, $\mathbf{c}_2$, as shown in figure 2.

Figure 2 shows two different instants, in which the proposed analysis is applied. The first instant is calculated for windows m_1 and m_2 – the SVDD model is applied for each window and the result is schematically presented in the figure. The result reflects the behavior of input patterns into a new domain, whose representation is associated to a hypersphere of center and radius determined by the SVDD algorithm.

The same procedure can be adopted to a later instant in time, as shown by windows m_1^* and m_2^* in figure 2. In this case, it is noted that the mapping generates very similar representations for both windows, since the resulting hyperspheres are virtually superimposed. Observing the features, it is clear that there is a very similar behavior between windows m_1^* and m_2^*, which is automatically reflected in the high dimensional space in which the hyperspheres are constructed. This scenario is quite different for windows m_1 and m_2, since features of the input data are significantly different and result in distant hyperspheres in feature space.

Whereas the attributes of features 1 and 2 are a result of preprocessing (e.g, using a DWT), one can use the method described to determine abrupt changes in a given signal. This process was initially applied to musical signals using one-class Support Vector Machines in [24] and its main foundation lies on the distance between the hyperspheres, calculated as:

$$D(t) = \frac{\sum\limits_{features} \|\mathbf{c}_1 - \mathbf{c}_2\|^2}{R_1 + R_2}. \tag{13}$$

The hyperspheres are updated for each time step t, taking into account the analysis for each window and resulting in a value for $D(t)$ that is similar to the Euclidean distance between the centers of the hyperspheres in the high-dimensional feature space $(\mathbf{c}_1, \mathbf{c}_2)$. The greater the difference between subsequent windows, the greater the distance in feature space between the hyperspheres and, consequently, the greater the probability of existing an abrupt variation on the signal within the two windows considered in the analysis.

After computing the distance, a threshold is applied in order to separate transition and quasi-stationary states of the signal. That is, instants whose distance value exceeds the threshold value are considered as transitions between states. The instants of time whose distance value is below the threshold are characterized as quasi-stationary states, or even the steady-state signal. In this work we have used different thresholds for model comparison.

Once the threshold calculation is done, it is possible to perform the waveform segmentation. Spectral estimation methods require an additional stage before the complete segmentation, as shown in [7]. In these methods, a previous smoothing of the signal is necessary for a complete segmentation, as there are cases in which residual values are below the calculated threshold even when they reflect transitional state changes. For the model proposed here, the smoothing stage is unnecessary, since the obtained mapping is itself smoothed.

In this type of representation, it is possible to exploit the advantages that other transformations provide for event detection. By preprocessing the signal with a DWT, a more robust event detection can be performed because the distances between hyperspheres are dependent on the DWT at different levels (input features).

III. Experimental Setup

In order to describe the experimental setup used in this work, we chose to present the characteristics of the simulated database used in the comparison of detection methods. Then, we also present details of the DWT and SVDD detection method that was used, so that comparisons to spectral methods presented in [4], [6]–[9] can be made.

A. Simulated Events

In order to test and validate the segmentation models, we decided to perform simulations on the Alternative Transient Program (ATP) environment [25]. In the basic ATP model used for the generation of simulated events, we have considered the basic elements of a distribution substation from COPEL, as well as the structures needed for the simulation of the events [26]. These elements are: substation transformer, capacitor bank, grounding transformer, four bar feeders, and the equivalent of the electric circuit up to the substation transformer. Only one feeder was modeled, being segmented in several parts, as shown in figure 3, with different cable types (CA336, Cu 35mm^2, and Cu 120mm^2). For other feeders, only the equivalent load was used. All parameters used in the model were based on real data obtained from COPEL.

The following events were considered: single-line-to-ground faults, two-phase-line-to-ground faults, three-phase-line-to-ground faults, two-phase faults, three-phase faults, feeder circuit breaker switch-off, automatic feeder reclosing, capacitor bank switching, and start of heavy-duty engines. The instant and the measurement location of the events generated were varied, and when concerning short-circuits, the fault resistance was also varied. Moreover, we considered the number of events per oscillography and the interval between occurrences when consecutive events were considered [2].

We selected two instants of occurrence for each event. The first being at the zero crossing of the voltage signal and the second at the maximum positive signal. For fault resistance, we also selected two values: 5Ω and 300Ω. For the measurement location, we chose to record the signals at the substation bar and at the middle of the feeder. For the number of events per oscillography, the following sequences were used:

- Single-line-to-ground faults can be: isolated, followed by feeder circuit breaker switch-off, followed by feeder circuit breaker switch-off and automatic feeder reclosing, followed by two-phase-line-to-ground fault, or followed by three-phase-line-to-ground fault;
- Two-phase-line-to-ground faults can be: isolated, followed by feeder circuit breaker switch-off, followed by feeder circuit breaker switch-off and automatic feeder reclosing, or followed by three-phase-line-to-ground fault;
- Two-phase faults can be: isolated, or followed by feeder circuit breaker switch-off;
- Capacitor bank switching can be: isolated, followed by single-line-to-ground fault, followed by two-phase-line-to-ground fault, or followed by three-phase-line-to-ground fault.

The intervals between events were selected as 1/4 of a cycle and 1.5 cycle, in order to assess the segmentation for very close consecutive events or even widely separated ones. Using these features, we generated 170 waveforms for analysis. For the ground-truth, we chose to segment transients and quasi-stationary states of each disturbance visually.

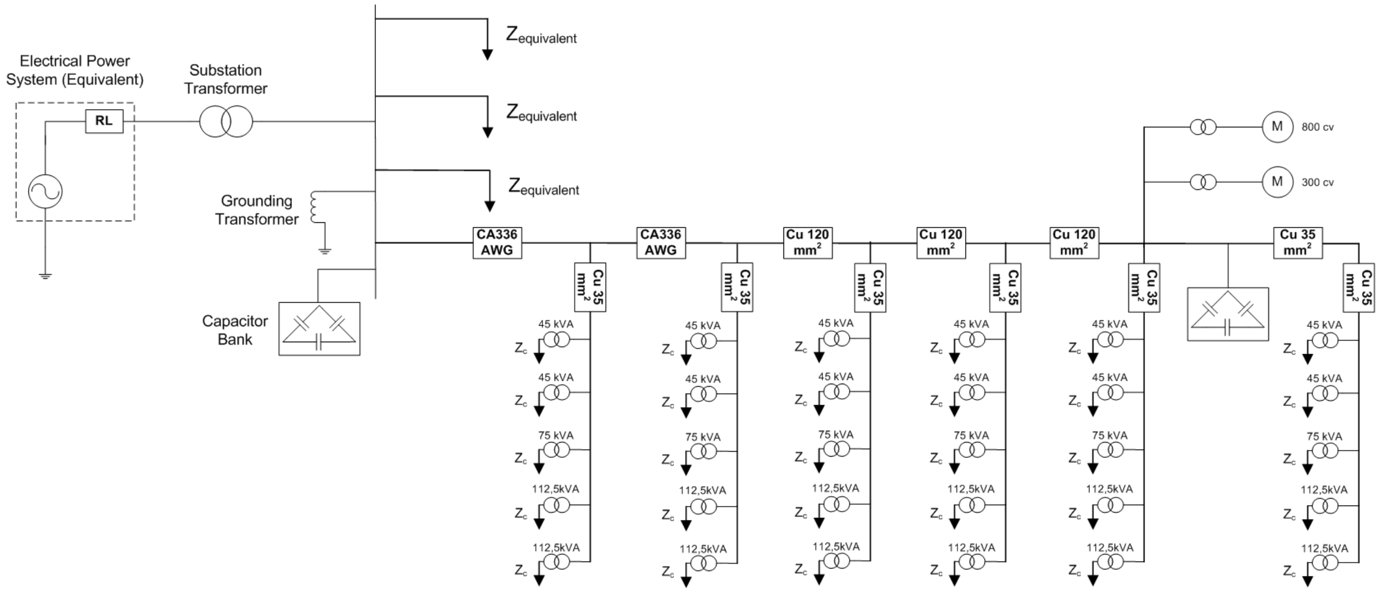

Fig. 3. Schematic Diagram of the Simulation Model.

B. Waveform Segmentation

Before we present the method applied for waveform segmentation (or event detection) using DWT and SVDD, we will present all modeling parameters. These parameters include the sampling frequecy, the analysis window widths (e.g. m_1 and m_2 in figure 2), the shift between the windows, the SVDD parameters (σ and C), and the number of decomposition levels and wavelet function used for the DWT.

The wavelet function used was the Daubechies-20. As shown in [7], this wavelet family presents the best results for detection problems. The order of the wavelet function was selected to be 20 because it allows a better localization of transients on different sub-bands, given the high Q factor of the corresponding filter. The number of levels was chosen to be four, with signals sampled at 7,680 Hz. These levels have shown to contain the most prominent frequency component variations at the instants of transients observed in simulated data.

Regarding the analysis windows, we used 1/8 of a cycle as width for each observation window. Thus, we obtained a resolution of 1/4 of a cycle. For parameters σ and C, we chose the values 20 and 0.2, respectively. This choice was based on the method for automatic selection of these parameters presented in [27]. Since the computational cost for this automatic selection method is too high for our application, it was used only to define parameters – the parameters selected during an optimization for a reduced set of oscillographic records containing one disturbance for each class were later used for the segmentation of other records.

After defining the parameters, distances were computed and thresholded to perform the segmentation. The entire process is shown in figure 4.

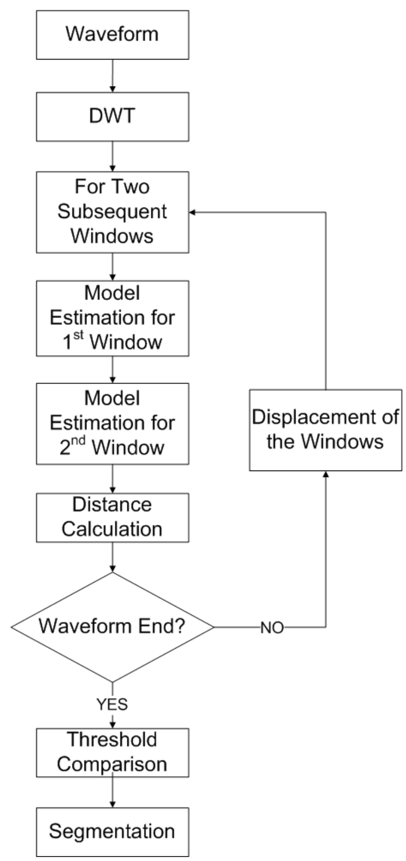

Fig. 4. Segmentation method using DWT and SVDD.

Initially, the wavelet decomposition was performed for the entire waveform. Then, the SVDD model was estimated for each window, determining the center and radius of hyperspheres in the wavelet domain. After that, distance calculation between the hyperspheres corresponding to each window was performed. Upon reaching the end of the waveform, threshold comparison and signal segmentation was performed, highliting the samples that corresponded to transitions and quasi-stationary states.

The performance assesment was made using ROC (Receiver

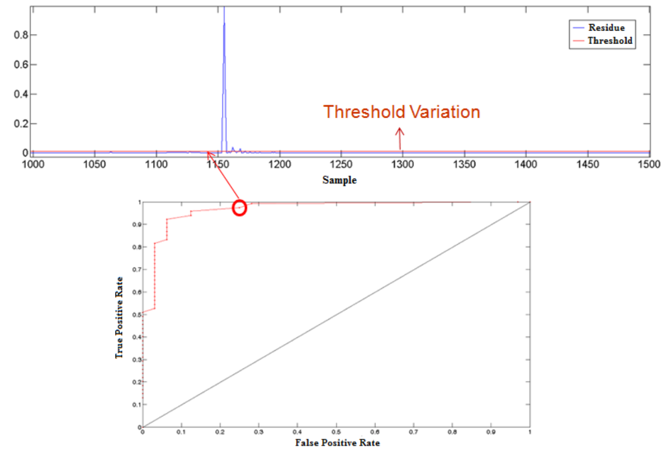

Fig. 5. Threshold variation and ROC analysis.

Operating Characteristic) analysis [28], based on the ground-truth established for each event. To this end, the residuals or distances (resulting from the mapping model) were analyzed and it was possible to establish the estimation behavior according to the selected threshold value. We observed the residual or distance behaviors for different threshold values, measuring the quality of the model output – different threshold values generated different points in the ROC curve, as shown in figure 5. We chose 50 different thresholds linearly separated in the range $[0, 1]$, considering normalized residuals and distances.

We have used the area under the ROC curve (AUC) [28] as the performance measurement for each method. As a single value, the use of the AUC statistic simplifies performance comparisons.

IV. Results

Table I summarizes the performance for the three detection models evaluated, where the superior performance of the DWT and SVDD method for different thresholds can be noticed. In some cases, the detection performance was considerably lower than the average performance for the 170 waveforms. These cases correspond to events with more than one transient state in the waveform, since residues or calculated distances – and hence the threshold – may be significantly influenced by the first large transient in the signal, hindering the characterization of other (less impacting) transients.

TABLE I
Performance Segmentation for Simulated Data.

Model	AUC for Different Thresholds (Mean ± Std. Dev.)
KF	0.92 ± 0.14
AR	0.93 ± 0.11
DWT and SVDD	$\mathbf{0.97 \pm 0.07}$

Moreover, in cases related to the start of a heavy-duty engine, it is usual to characterize the transient period as the complete duration of the voltage sag. Nevertheless, the methods evaluated here aim to detect only the higher frequency transients observed between different quasi-stationary states – this compromises the detection of any slow changes in the waveform.

From the viewpoint of a subsequent classification, only the initial instant is necessary to implement classification algorithms [26]. However, what is proposed here is the extraction of the whole transient between quasi-stationary states. This extraction maximizes the probability of event start detection.

From a global perspective, the method based on DWT and SVDD yielded better results for different thresholds. Analyzing only the efficiency of the mapping generated by residuals or distances, we observe a far superior performance for this method, with AUC over 0.95. This demonstrates that the generated mapping is quite efficient, even for different events in the same waveform.

V. Conclusions

This paper presented the results of a new approach for a fundamental step in automatic waveform analysis in distribution networks: detection of transient and quasi-stationary states. The proposed approach was compared with commonly used detection methods for simulated data in ATP.

Regarding event segmentation, performances were verified for different thresholds, in order to verify the response of the mapping performed by detection methods and the performance of a fully automatic segmentation method. It was observed that the proposed method, based on DWT and SVDD, yields better overall performance for simulated data, considering different threshold levels. This indicates that the proposed method can be a very efficient way to extract transitions between quasi-stationary states.

Further work includes the investigation of automatic threshold calculation, in which referral to a specialist for choosing the threshold for segmentation will be unnecessary – sometimes in the literature this process is called “fully automatic”. We also intend to extend this analysis to real data.

VI. Acknowledgment

This work was partially supported by Companhia Paranaense de Energia (Copel), within the Research and Development Program of the Agência Nacional de Energia Elétrica (Brazilian Electrical Energy Agency).

References

[1] R. C. Dugan, M. F. McGranaghan, S. Santoso, and H. W. Beaty, *Electrical Power System Quality*, 2nd ed. New York: McGraw-Hill, 2002.

[2] M. Bollen and I. Gu, *Signal Processing of Power Quality Disturbances:*. John Wiley & Sons Inc., 2006.

[3] M. H. J. Bollen, I. Y. H. Gu, P. G. V. Axelberg, and E. Styvaktakis, “Classification of Underlying Causes of Power Quality Disturbances: Deterministic versus Statistical Methods,” *EURASIP Journal on Advances in Signal Processing*, vol. 2007, pp. 1–18, 2007.

[4] A. Ukil and R. Zivanovic, “Abrupt Change Detection in Power System Fault Analysis using Wavelet Transform,” in *International Conference on Power Systems Transients*, 2005.

[5] D. Wiot, “A New Adaptive Transient Monitoring Scheme for Detection of Power System Events,” *IEEE Transactions on Power Delivery*, vol. 19, pp. 42–48, 2004.

[6] A. Ukil and R. Zivanovic, "Detection of abrupt changes in power system fault analysis: A comparative study," in *Southern African Universities Power Engineering Conference*, 2005.

[7] ——, "Abrupt change detection in power system fault analysis using adaptive whitening filter and wavelet transform," *Electric Power Systems Research*, vol. 76, pp. 815–823, 2006.

[8] A. Ukil and R. Ivanovi, "Application of Abrupt Change Detection in Power Systems Disturbance Analysis and Relay Performance Monitoring," *IEEE Transactions on Power Delivery*, vol. 22, pp. 59–66, 2007.

[9] A. Ukil and R. Zivanovic, "Adjusted Haar wavelet for application in the power systems disturbance analysis," *Digital Signal Processing*, vol. 18, pp. 103–115, 2008.

[10] D. L. Donoho and I. M. Johnstone, "Ideal Spatial Adaptation by Wavelet Shrinkage," *Biometrika*, vol. 81, p. 425, 1994.

[11] F. B. Costa, B. A. Souza, and N. S. D. Brito, "Detecção de faltas evolutivas e múltiplos distúrbios em registros oscilográficos baseada na transformada wavelet discreta," *Revista Controle e Automação*, vol. 21, pp. 173–184, 2010.

[12] M. Moreto and J. G. Rolim, "Segmentação automática de registros oscilográficos: um método voltado para geradores de energia," in *XVIII Congresso Brasileiro de Automática*, 2010, pp. 1272–1279.

[13] C. D. Le, I. Y. H. Gu, and M. H. J. Bollen, "Joint Causal and Anti-Causal Segmentation and Location of Transitions in Power Disturbances," in *Power and Energy Society General Meeting*, 2010.

[14] A. Ukil and R. Zivanovic, "The detection of abrupt changes using recursive identification for power system fault analysis," *Electric Power Systems Research*, vol. 77, pp. 259–265, 2007.

[15] T. Abdel-Galil, "Power quality event detection using Adaline," *Electric Power Systems Research*, vol. 64, pp. 137–144, 2003.

[16] M. Davy, F. Desobry, a. Gretton, and C. Doncarli, "An online support vector machine for abnormal events detection," *IEEE Transactions on Signal Processing*, vol. 86, pp. 2009–2025, 2006.

[17] J. Ma and S. Perkins, "Time-series novelty detection using one-class support vector machines," in *Proceedings of the International Joint Conference on Neural Networks*. Ieee, 2003, pp. 1741–1745.

[18] F. Desobry, M. Davy, and C. Doncarli, "An online kernel change detection algorithm," *IEEE Transactions on Signal Processing*, vol. 53, pp. 2961–2974, 2005.

[19] S. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed. Academic Press, 1999.

[20] I. Daubechies, *Ten Lectures on Wavelets*. SIAM, Philadelphia, 1992.

[21] D. M. J. Tax, "One-class classification," Ph.D. dissertation, Technische Universiteit Delft, 2001.

[22] V. N. Vapnik, *Statistical Learning Theory*. New York: Wiley-Interscience, 1998.

[23] B. Schölkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond*. MIT Press, 2001.

[24] F. Desobry and M. Davy, "Support vector-based online detection of abrupt changes," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2003.

[25] EMTP, *Alternative Transients Program (ATP) Rule Book*, Can Am EMTP User's Group, 1995.

[26] A. E. Lazzaretti, V. H. Ferreira, H. Vieira Neto, R. J. Riella, and J. Omori, "Classification of events in distribution networks using autonomous neural models," in *15th International Conference on Intelligent System Applications to Power Systems*, 2009.

[27] R. P. W. Duin, "Uniform Object Generation for Optimizing One-class Classifiers," *Journal of Machine Learning Research*, vol. 2, pp. 155–173, 2001.

[28] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, pp. 861–874, 2006.