Statistical Assessment of Abrupt Change Detectors for Non-Intrusive Load Monitoring

Sarra Houidi
Université de Tunis El Manar, ENIT LSE
IREENA, Université de Nantes
Tunis, Tunisia
sarra.houidi@etu.univ-nantes.fr

François Auger
IREENA, Université de Nantes
Saint-Nazaire, France
francois.auger@univ-nantes.fr

Houda Ben Attia Sethom
Université de Tunis El Manar
ENIT, LSE LR11ES15
Tunis, Tunisia
houda.benattia@enicarthage.rnu.tn

Laurence Miègeville
IREENA, Université de Nantes
Saint-Nazaire, France
laurence.miegeville@univ-nantes.fr

Dominique Fourer
UMR STMS (IRCAM - CNRS - UPMC)
Paris, France
dominique.fourer@ircam.fr

Xiao Jiang
South China University of Technology
Guangzhou, China
xiaojiangscut@qq.com

***Abstract*—Non-Intrusive Load Monitoring (NILM) is the process of monitoring Home Electrical Appliances (HEAs) from voltage and current measurements at the Point of Common Coupling (PCC). The resulting information can be useful to raise greater awareness among customers about their HEAs' usage for a more efficient energy behaviour. NILM methods can be roughly categorized into event-based and non event-based approaches, depending on whether or not they rely on the detection of HEAs' significant state transitions (e.g. On/Off or state change). In this paper, we focus on event-based methods and describe a test bench to statistically assess three abrupt change detectors. Then, we define several evaluation metrics for judging on performance and quality of the detection algorithms in order to carry out a comparison. Finally, the three abrupt change detectors are applied to real data coming from a scenario of HEAs switched on and off by hand.**

***Index Terms*—Abrupt change detection, Statistical assessment, Non Intrusive Load Monitoring (NILM), Home Electrical Appliances (HEAs).**

I. Introduction

Monitoring electricity consumption in the residential sector is a major way to make customers informed of their energy consumption, for helping them to change their behaviour and to reduce their electricity bills. Furthermore, for the grid operator, having access to information on consumer usage represents more possibilities of actions such as off loading, while keeping the stability and reliability of the grid, especially with the increasing penetration of renewable energy sources. Non-Intrusive Load Monitoring (NILM) [1] is a promising approach to estimate the electrical consumption of individual Home Electrical Appliances (HEAs) from current and voltage measurements at the Point of Common Coupling (PCC) by restricting the metering equipment to one measuring device including current and voltage sensors associated with an acquisition system connected to the house main power line. There are different approaches for NILM that can be classified into supervised or unsupervised methods. Supervised approaches rely on pattern recognition based machine learning techniques and require labeled data to train a classifier [2]. Unsupervised methods, conversely, are free from this constraint and rely on clustering procedures or blind source separation techniques. Reviews of NILM methods can be found in the literature [3]–[5]. Another dichotomy of the field is that NILM approaches can be categorized into event-based and non event-based approaches [6]. The former use an event detection algorithm to locate and identify significant HEA events (e.g. On/Off or state transitions) in the power consumption curve, thanks to the HEAs' signatures available in the HEA features database [7]. The latter disaggregate every sample of the load curve to provide the contribution of each identified HEA in the total consumption [3]. In the present paper, an interest is paid to event-based approaches, in particular to the detection process. Several event detection methods have been investigated for event-based NILM approaches. In [8], Jin *et al.* propose a robust change-point detection approach based on a Goodness-of-Fit (GoF) test. Paper [9] proposes an algortihm which tracks the standard deviation variation of the signal envelope using a moving window. A comprehensive review of some of these methods can be found in [6], [10]. As a part of their contribution, the authors propose herein to statistically assess three abrupt change detection algorithms classically used in several applications areas other than NILM, namely:

1) The Effective Residual algorithm applied to sensor fault detection in [11] and to NILM in [12].
2) The CUmulative SUM (CUSUM) algorithm used for biomedical engineering [13] as well as for NILM applications [14].
3) The Bayesian Information Criterion (BIC) algorithm used for acoustic change detection [15], [16].

The remainder of this paper is structured as follows:
Section II proposes a formal mathematical problem statement to the detection of abrupt changes in the NILM framework. Section III shortly presents the three studied detection algorithms. Several criteria to compare the algorithms' performances through a statistical assessment are introduced in

Section IV. Section V deals with the application of these algorithms to real HEAs' signals. Finally, Section VI leads to the conclusion and the prospectives of the current work.

II. Abrupt change in the NILM framework and mathematical problem statement

Abrupt changes represent fast transitions that occur between stationary states in a signal [13]. The abrupt change detection requires tools that decide dichotomously whether such a change occurred or not in the signal [17]. NILM systems use this definition of abrupt changes that is mainly applicable for On/Off appliances and Finite State Machines (FSM), which are HEAs with a series of changes in power draw, such as washing machines. Let $X_n = \{x_m \in \mathbb{R}, m = n-k+1, \dots, n\}$ be a vector of the last k available samples of a signal x_m at the current time n. For NILM purposes, this signal can be the active power, or any other electrical feature. Each signal sample follows a Probability Density Function (PDF) $p_\theta(x_m)$ depending on a deterministic parameter vector θ (e.g. the mean μ and the standard deviation σ of a signal following a Gaussian distribution). An abrupt change, modeled by an instantaneous modification of θ, may occur at a change time n_c. Two statistical hypotheses H_0 and H_1 are then considered to represent the "without change" and "with change" assumptions, respectively:

$$\text{under } \mathrm{H}_0,\ \theta = \theta_0 \quad \text{for} \quad n-k+1 \le m \le n \tag{1}$$

$$\text{under } \mathrm{H}_1,\ \theta = \begin{cases} \theta_{1a} & \text{for} \quad n-k+1 \le m \le n_c - 1 \\ \theta_{1b} & \text{for} \quad n_c \le m \le n \end{cases} \tag{2}$$

Within an on-line framework, a dichotomous decision between H_0 and H_1 (i.e. a decision "to reject H_0 in favor of H_1") has to be made at each time n [13], [18], by comparing a decision function g_n to a threshold value h

$$\text{decide } \mathrm{H}_1 \quad \text{if} \quad g_n > h \tag{3}$$

$$\text{decide } \mathrm{H}_0 \quad \text{if} \quad g_n \le h \tag{4}$$

where h is generally adjusted according to some desired decision probabilities.

III. Abrupt change detection algorithms

This section is devoted to short presentations of the three investigated abrupt change detection algorithms.

A. Effective residual processing method

Paper [12] proposes a HEA event detection algorithm, based on an approach used for sensor fault detection and isolation [11]. The absolute variation between two consecutive signal samples is defined by:

$$\delta_m = |x_m - x_{m-1}| \quad \text{for} \quad n-k+2 \le m \le n \tag{5}$$

The residual r_m representing the difference between two consecutive variations is calculated as:

$$r_m = |\delta_m - \delta_{m-1}| \quad \text{for} \quad n-k+3 \le m \le n \tag{6}$$

For a lesser sensitivity to measurement noise, the sum of the last three residuals is computed and constitutes the Effective Residual decision function:

$$g_n \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} h \quad \text{with} \quad g_n = r_n + r_{n-1} + r_{n-2} \tag{7}$$

where h is a positive threshold of the same physical dimensionality as x_m. This detector aims at detecting a mean change at $n_c = n$ from the last $k = 5$ samples of the signal. As we aim at assessing the change detection ability of all the algorithms in strictly the same conditions, a sliding window of $k = 5$ samples will also be considered for the following CUSUM and BIC algorithms.

B. CUmulative SUM (CUSUM) algorithm

Since the signal samples x_m are supposed to be statistically independent, the PDFs of X_n under hypotheses H_0 and H_1 are [13]:

$$p(X_n|\mathrm{H}_0) = \prod_{m=n-k+1}^{n} p_{\theta_0}(x_m) \tag{8}$$

$$p(X_n|\mathrm{H}_1) = \prod_{m=n-k+1}^{n_c-1} p_{\theta_{1a}}(x_m) \prod_{m=n_c}^{n} p_{\theta_{1b}}(x_m) \tag{9}$$

If θ_0 and θ_{1a} are supposed equal, then the log-likelihood ratio $L(X_n, n_c)$ is a cumulative sum of values s_m, which corresponds to the instantaneous log-likelihood ratio:

$$L(X_n, n_c) = \ln\left(\frac{p(X_n|\mathrm{H}_1)}{p(X_n|\mathrm{H}_0)}\right) = \sum_{m=n_c}^{n} s_m \tag{10}$$

$$\text{with} \quad s_m = \ln\left(\frac{p_{\theta_{1b}}(x_m)}{p_{\theta_{1a}}(x_m)}\right) \tag{11}$$

The CUSUM decision rule is based on a maximization of the log-likelihood ratio over n_c, such as:

$$g_n \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} h, \quad \text{with} \quad g_n = \max_{n-k+1 \le n_c \le n} \sum_{m=n_c}^{n} s_m \tag{12}$$

where h is an adimensional positive threshold. The most common form of the CUSUM algorithm is a statistical test for the detection of a mean change in a Gaussian process $\mathcal{N}(\mu, \sigma)$ [18]. The only changing parameter is the mean value μ such that:

$$\text{under } \mathrm{H}_0,\ \mu = \mu_0 \quad \text{for} \quad n-k+1 \le m \le n \tag{13}$$

$$\text{under } \mathrm{H}_1,\ \mu = \begin{cases} \mu_{1a} & \text{for} \quad n-k+1 \le m \le n_c - 1 \\ \mu_{1b} & \text{for} \quad n_c \le m \le n \end{cases} \tag{14}$$

with $\mu_{1a} = \mu_0$. The instantaneous log-likelihood ratio s_m then becomes:

$$s_m = -\frac{(x_m - \mu_{1b})^2}{2\sigma^2} + \frac{(x_m - \mu_{1a})^2}{2\sigma^2} \tag{15}$$

$$= \frac{\Delta\mu}{\sigma^2}\left(x_m - \frac{\mu_{1b} + \mu_{1a}}{2}\right), \tag{16}$$

where $\Delta\mu = \mu_{1b} - \mu_{1a}$ is the jump height on the mean. It can be observed that if $x_m = \mu_{1a}$, $s_m = -\Delta\mu^2 / \left(2\sigma^2\right)$ and if $x_m = \mu_{1b}$, $s_m = +\Delta\mu^2 / \left(2\sigma^2\right)$. This means that the decision function g_n expressed by (12) is most likely negative before the mean change, and becomes most likely positive after the mean change. If considering an abrupt change occurring inside a sliding window of $k = 5$ samples at $n_c = n$, the Maximum Likelihood Estimations (MLEs) of the mean values $\hat{\mu}_{1b} = x_n$ and $\hat{\mu}_{1a} = \frac{1}{4} \sum_{m=n-4}^{n-1} x_m$ are used. Then according to (16), the CUSUM decision function becomes:

$$g_n = s_n = \frac{(x_n - \hat{\mu}_{1a})^2}{2\,\hat{\sigma}^2}, \text{ with } \hat{\sigma}^2 = \frac{1}{4} \sum_{m=n-4}^{n-1} (x_m - \hat{\mu}_{1a})^2 \quad (17)$$

This decision function is used to find which of the two hypotheses H_0 and H_1 is the most likely, by comparing it to a threshold h, as described in (12).

C. The Bayesian Information Criterion

The Bayesian Information Criterion (BIC) is at the heart of numerous works in audio segmentation [15], [16], [19]. It consists in dividing the sequence of observed random samples into homogeneous segments by performing a hypothesis test. For each potential change point, two hypotheses $\mathrm{H_i}$, $i \in \{0, 1\}$ can be drawn. H_0 supposes that, on both sides of this point, the signal follows the same probabilistic model, while H_1 supposes that a change of model occurs. The BIC of X_n under hypothesis H_i is defined as a likelihood criterion penalized by the model complexity [19], [20], proportional to the number M of parameters in the probabilistic model:

$$\mathrm{BIC}(H_i) = \ln(p(X_n|H_i)) - \frac{\lambda}{2} M \ln(k) \quad (18)$$

where $p(X_n|H_i)$ is the maximized data likelihood for the given model (8) or (9) and λ a penalty factor, ideally equal to 1 [19]. In our application, the probabilistic model is defined as follows:

$$\mathrm{H}_0 \; : \; x_{n-k+1}, \ldots, x_n \sim \mathcal{N}(\mu_0, \sigma_0) \quad (19)$$

$$\begin{aligned} \mathrm{H}_1 \; : \; & x_{n-k+1}, \ldots, x_{n_c-1} \sim \mathcal{N}(\mu_{1a}, \sigma_{1a}); \\ & x_{n_c}, \ldots, x_n \sim \mathcal{N}(\mu_{1b}, \sigma_{1b}) \end{aligned} \quad (20)$$

Under H_0, the model parameters are μ_0 and σ_0, so $M = 2$. Under H_1, the model parameters are $\mu_{1a}, \sigma_{1a}, \mu_{1b}, \sigma_{1b}$, so $M = 4$. It can be shown that $p(X_n|H_i)$ and $\mathrm{BIC}(H_i)$ are maximized when μ and σ^2 are replaced by their MLEs $\hat{\mu}$ and $\hat{\sigma}^2$ [19], [20]. The maximized BIC values under H_0 and H_1 are respectively:

$$\mathrm{BIC}(H_0) = -\frac{k}{2}\ln(2\pi) - \frac{k}{2}\ln(\hat{\sigma_0}^2) - \frac{k}{2} - \lambda \ln(k) \quad (21)$$

$$\begin{aligned} \mathrm{BIC}(H_1) = & -\frac{k}{2}\ln(2\pi) - \frac{(n_c - n + k - 1)}{2}\ln(\hat{\sigma}_{1a}^2) \\ & - \frac{(n - n_c + 1)}{2}\ln(\hat{\sigma}_{1b}^2) - \frac{k}{2} - 2\lambda\ln(k) \end{aligned} \quad (22)$$

The BIC variation for a given value of n_c is given by:

$$\Delta\mathrm{BIC}(n_c) = \mathrm{BIC}(H_1) - \mathrm{BIC}(H_0) \quad (23)$$

$$\begin{aligned} &= \frac{k}{2}\ln(\hat{\sigma}_0^2) - \frac{(n_c - n + k - 1)}{2}\ln(\hat{\sigma}_{1a}^2) \\ &\quad - \frac{(n - n_c + 1)}{2}\ln(\hat{\sigma}_{1b}^2) - \lambda\ln(k) \end{aligned} \quad (24)$$

If $\Delta\mathrm{BIC}(n_c) > 0$, the model of two Gaussians is favored (i.e. the signal can be segmented into two parts at n_c). Consequently, the decision function can be expressed as:

$$g_n \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} h \quad \text{with} \quad g_n = \max_{n-k+1 \le n_c \le n} \Delta\mathrm{BIC}(n_c), \quad (25)$$

where h is an adimensional threshold set around zero. If considering an abrupt change occurring inside a sliding window of $k = 5$ samples at $n_c = n - 1$ (since we need at least two samples to estimate σ_{1b}), the BIC decision rule becomes:

$$g'_n \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} h' \text{ with } g'_n = \frac{1}{2}\ln\left(\frac{\hat{\sigma}_0^{10}}{\hat{\sigma}_{1a}^6 \hat{\sigma}_{1b}^4}\right), \; h' = h + \lambda\ln(5), \quad (26)$$

$$\hat{\sigma}_0^2 = \frac{1}{5}\sum_{m=n-4}^{n}(x_m - \hat{\mu}_0)^2, \quad \hat{\mu}_0 = \frac{1}{5}\sum_{m=n-4}^{n} x_m, \quad (27)$$

$$\hat{\sigma}_{1a}^2 = \frac{1}{3}\sum_{m=n-4}^{n-2}(x_m - \hat{\mu}_{1a})^2, \quad \hat{\mu}_{1a} = \frac{1}{3}\sum_{m=n-4}^{n-2} x_m, \quad (28)$$

$$\hat{\sigma}_{1b}^2 = \frac{1}{2}\sum_{m=n-1}^{n}(x_m - \hat{\mu}_{1b})^2 \text{ and } \hat{\mu}_{1b} = \frac{1}{2}\sum_{m=n-1}^{n} x_m \quad (29)$$

where h' is a real valued threshold, which gathers all the settings of h and λ.

IV. Statistical Assessment of Abrupt Change Detection Algorithms

This section is devoted to the statistical assessment of the three considered detection algorithms using performance metrics.

A. Test bench description

To reach this purpose, the following experiment is repeated 100 000 times: X_n is filled with 5 samples drawn from an independent zero mean Gaussian process with a standard deviation $\sigma = 1$. Under H_1, a mean change is simulated by adding $\Delta\mu$ to the last sample for the Effective Residual and CUSUM algorithms, and to the last two samples for the BIC. A Signal to Noise Ratio (SNR) is then defined as $\mathrm{SNR} = \Delta\mu/\sigma$. This makes it possible to compare the step height with respect to the noise level. Both hypotheses H_0 and H_1 are tested using the decision functions expressed in (7), (17) and (26). Decisions are taken using 400 logarithmically spaced values of the threshold. The assessment is made for fixed values of the SNR, and for a uniformly distributed SNR ranging from 0.5 to 10 to simulate realistic HEAs of different powers.

B. Performance evaluation tools

1) Basic performance metrics: Different evaluation metrics are used to compare the performance of abrupt change detection rules [21]:

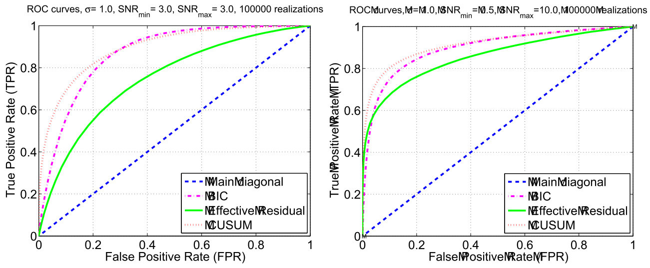
- True Positive (TP) counts the times a test detects a change in the sequence when there is really one,
- True Negative (TN) counts the times a test does not detect a change in the sequence when there is not,
- False Positive (FP) counts the times a test detects a change in the sequence when there is not,
- False Negative (FN) counts the times a test does not detect a change in the sequence when there is really one.

Different rates or probablities are computed from these performance metrics:

- The True Positive Rate (TPR), which measures the ability of a test to detect a truly present change, is defined as the fraction of the correctly detected events over all the H_1 cases:
$$\mathrm{TPR} = \mathrm{TP}/(\mathrm{TP} + \mathrm{FN}) \tag{30}$$
- The False Positive Rate (FPR), which denotes the fraction of wrongly detected “non events” over all the H_0 cases:
$$\mathrm{FPR} = \mathrm{FP}/(\mathrm{TN} + \mathrm{FP}) \tag{31}$$
- The Precision $\mathrm{P_R}$ which expresses the fraction of the correctly detected changes over all the times the decision rule favors H_1:
$$\mathrm{P_R} = \mathrm{TP}/(\mathrm{TP} + \mathrm{FP}) \tag{32}$$

2) Receiver Operating Characteristics: The plot of the TPR versus the FPR obtained for varying values of the threshold h is called the Receiver Operating Characteristic (ROC) and is acknowledged in the literature as a performance measure of detectors [21]. As the test detection ability increases, the curve shifts toward the upper left corner, where the TPR is always greater than the FPR [19]. On the opposite, a test with a weak detection ability will produce a curve on or below the main diagonal. Fig. 1 represents the three ROC curves obtained for the CUSUM, BIC and Effective Residual algorithms for a fixed and a variable SNR. It appears that the CUSUM ROC curve tends the most toward the “optimal” point (TPR=1; FPR=0). Then follows the BIC the Effective Residual curves.

3) Area Under the Curve (AUC): Another performance measure is the Area Under the ROC Curve (AUC) [22], obtained by a most often numerical integration of the ROC curve. AUC lies between 0 and 1. Values below 0.5 represent worthless detection rules, while values close to 1 imply nearly perfect detection. In Table I the AUC values obtained for the three considered detection algorithms are reported. It can be observed that the CUSUM algorithm shows the best detection ability with an AUC value of 89% for $\mathrm{SNR} = 3$ and reaching up to 91 % (very close to the ideal value) for a random SNR uniformly distributed between 0.5 and 10. Fig 2 depicts the AUC values versus the SNR. As expected, a higher SNR leads to a better AUC (close to 1) for the three investigated detectors. However, for very low values of SNR which are representative of small abrupt changes buried into noise, the CUSUM is relatively robust against noise. For example, for a SNR=0.1, its AUC value is equal to 57% in comparison to the Effective Residual and the BIC algorithms which are much more sensitive to noise for low values of SNR.



(a) (b)

Fig. 1: ROC curves of the CUSUM, BIC and Effective Residual algorithms for a constant SNR equal to 3 (a) and for SNR values uniformly distributed between 0.5 and 10 (b).

TABLE I: AUC values computed for the three considered detection algorithms

	Effective Residual	**CUSUM**	**BIC**
$\mathrm{SNR} = 0.5$	0.51	0.59	0.53
$\mathrm{SNR} = 3$	0.75	0.89	0.87
$\mathrm{SNR} = 6$	0.96	0.98	0.97
variable SNR	0.85	0.91	0.89

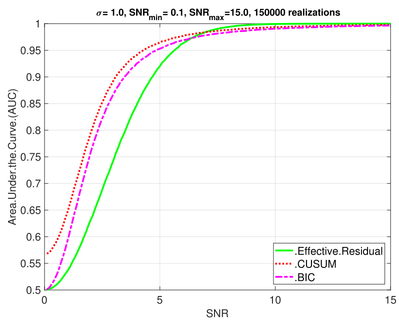


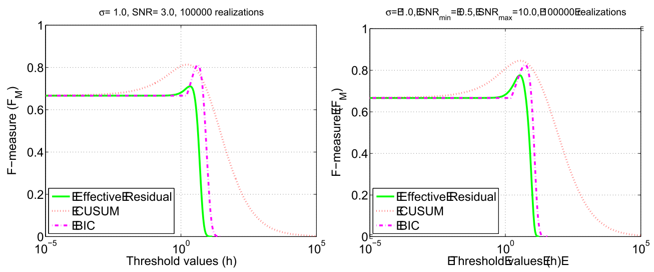
Fig. 2: SNR-AUC profile of the Effective Residual, CUSUM and BIC algorithms.

4) F-Measure: The F-Measure is a statistical quality score widely used in the detection framework [23]. It is defined as the harmonic mean of the Precision $\mathrm{P_R}$ and the True Positive Rate TPR, as given in (33) and is computed at various threshold settings.

$$\mathrm{F_M} = \left(\frac{P_R^{-1} + \mathrm{TPR}^{-1}}{2}\right)^{-1} = 2 \times \frac{P_R \times \mathrm{TPR}}{P_R + \mathrm{TPR}} \tag{33}$$

Fig. 3 depicts the plot of $\mathrm{F_M}$ versus the threshold value h. For small values of h, $P_R = 1/2$ and $\mathrm{TPR} = 1$, so

$F_M = 2/3$. For large values of h, $\mathrm{TPR} = 0$, so $F_M = 0$. For the three detectors, the maximum F_M score is reached for a specific threshold value which is the optimal one. For threshold values higher than the optimal one, the F_M score significantly decreases until reaching the zero value, meaning consequently that the detectors are no more effective. For a fixed SNR, the CUSUM detector shows a relative robustness to a wide range of threshold values as its F_M score remains high and constant until reaching the maximum F_M score. In comparison to the BIC and the Effective Residual algorithms which obtain almost the same curve shape but reach maximum F_M scores for lower threshold values. For a variable SNR, the same observation is made. Fig. 4 plots the optimal threshold values h that maximize F_M for different values of SNR. It can be noted that the optimal threshold value strongly depends on the SNR.



(a) Fixed SNR (b) Variable SNR

Fig. 3: F-measure (F_M) depending on threshold values h of each detector: for a fixed SNR (a) and for a variable SNR (b).

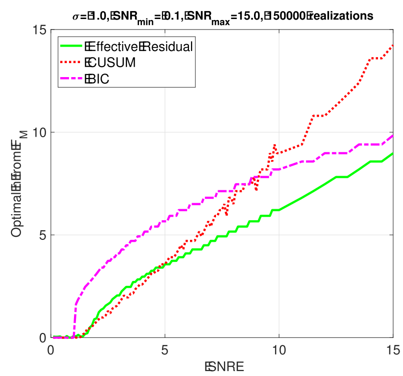


Fig. 4: Optimal value of h from F_M depending on SNR for the Effective Residual, CUSUM and BIC algorithms.

V. Experimental results

In this section, an abrupt change detection example is proposed using the three handled algorithms with real current and voltage measured in a power network at $F = 50$ Hz by our own acquisition device. This measurement system consists of an analog circuit for voltage and current sensing, associated with an Arduino Nano microcontroller with a 10-bit analog to digital converter. This allows the creation of real consumption scenarios with the advantage of providing ground truth time instants for the On/Off events occurrences. The measured current $i(t)$ and voltage $v(t)$ are sampled at $F_s = 1.2$ kHz. The abrupt change detectors are applied on the active power signal $P[n] = \frac{1}{M}\sum_{k=n-M+1}^{n} v[k]i[k]$, with $M = Fs/F$ the number of samples per period. The sliding window of 5 successive values of the active power computed every M samples has then a width of 80 ms. In the considered scenario, a hair dryer is first turned on then turned off. Then successively, a low-energy fluorescent lamp, an incandescent lamp and an iron are turned on. Finally, the iron followed by the incandescent lamp and the low-energy light bulb are turned off. Fig. 5 plots the active power time profile and the three handled algorithms decision function outcomes.

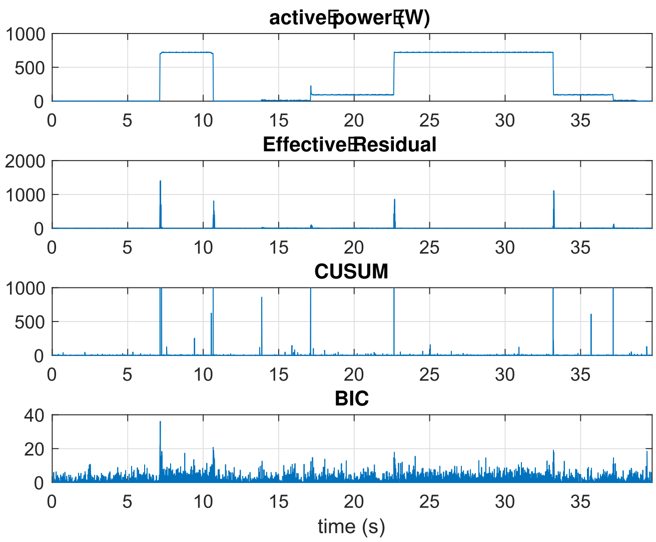


Fig. 5: Active Power time profile and corresponding decision function outcomes of the Effective Residual, CUSUM and BIC algorithms.

For each detector, a threshold value h is set, which allows the detection of the lowest active power variation represented here by the low-energy lamp switch-on. Table II reports the true positive TP, the false positive FP and the precision rate P_R values of the three studied detection algorithms applied to the active power signal with fixed threshold values h. It should be noticed that the CUSUM algorithm appears to be the most effective with a high precision rate and a small number of false positve cases. On the other hand, the number of detected events for both the Effective Residual and the BIC algorithms, is relatively large. This is due to the noisy steady-states leading to false positives and consequently to inaccurate detections. Fig. 6 highlights the iron switch on event with detected events in red lines for the three studied algorithms. At the top of Fig. 6, a zoom on the iron steady-state active power signal after being turned On is also displayed. At the top of Fig. 6, a zoom on the iron steady-state active power signal after being turned on is also displayed. Although it corresponds to the iron steady state operation, the noise makes the Effective Residual method detecting abrupt changes. The same observation is made for the BIC algorithm.

TABLE II: Threshold values h and number of TP, FP and P_R of the three considered detection algorithms applied to the active power signal.

	Effective Residual	**CUSUM**	**BIC**
h	31.0	850.0	12.0
Number of detected events	32	8	33
TP	7	7	7
FP	25	1	26
$\mathbf{P}_R$	21.3%	87.5%	21.2%

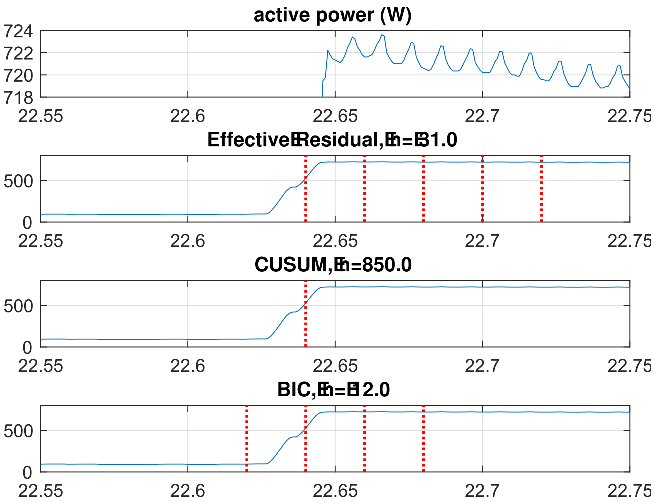


Fig. 6: Zoom on the Iron switch On from the active power signal and detected events of each handled algorithm.

VI. Conclusion

In this paper, three abrupt change detectors were considered, namely the Effective Residual, the CUSUM and BIC within an event-based NILM approach. A test bench was also proposed to statistically assess these three abrupt change detectors under the same conditions. To reach this purpose, several metrics were defined, in order to judge the detectors performances. These metrics have outlined the noise level effects on abrupt change detectors performances. Finally, using our own acquisition system, the detectors were tested on real data related to HEAs power consumption and representing a controlled scenario of HEAs switched on and off. The CUSUM algorithm has shown itself more effective for detecting abrupt changes compared to the Effective Residual and BIC algorithms. For future work, we are intend to study the probability density distribution of each decision function outcomes for computing the probability of detection and probability of false alarms as a function of the SNR. The decision function will also be extended to multidimensional signals to make a decision from several features. Lastly, other detection methods will be considered [9].

References

[1] G.W. Hart, “Nonintrusive appliance load monitoring”, Proc. of the IEEE, vol. 80, 1870–1891, December 1992.
[2] C. Bishop, “Pattern Recognition and Machine Learning”, Information Science and Statistics, Springer-Verlag New York, 2004.
[3] I. Abubakar, S. N. Khalid, M. W. Mustafa, H. Shareef, M. Mustaph, “Recent approaches and applications of non-intrusive load monitoring,” ARPN Journal of Engineering and Applied Sciences, vol. 11, pp. 4609–4618, April 2016.
[4] M. Zeifman, K. Roth, “Non intrusive appliance load monitoring : Review and outlook,” IEEE Trans. on consumer electronics, vol. 57, pp. 76–84, February 2011.
[5] S. M. Tabatabaei, S. Dick, W. Xu, “Toward non-intrusive load monitoring via multi-label classification,” IEEE Trans. on smart grid, vol. 8, pp. 26–40, June 2016.
[6] K. D. Anderson, M. E. Berges, A. Ocneanu, D. Benitez, J. M. F. Moura, “Event Detection for Non-Intrusive Load Monitoring,” 38th Annual Conf. on IEEE Industrial Electronics Soc. (IECON), October 2012.
[7] S. Houidi, F. Auger, H. Ben Attia Sethom, D. Fourer, L. Miègeville, “Relevant feature selection for home appliances recognition”, Proc. of Electrimacs, July 2017.
[8] Y. Jin, E. Tebekaemi, M. Berges, L. Soibelman, “A time-frequency approach for event detection in nonintrusive load monitoring,” in Proc. of the Signal Processing, Sensor Fusion, and Target Recognition, May 2011.
[9] M. N. Meziane, P. Ravier, G. Lamarque, J-C. Le Bunetel, Y. Raingeaud, “High accuracy event detection for non-intrusive load monitoring”, Proc. of the IEEE Int. on Acoustics, Speech and Signal Processing (ICASSP), March 2017.
[10] J. Froehlich, E. Larson, S. Gupta, G. Cohn, M.S. Reynolds, S.N. Patel, “Disaggregated End-Use Energy Sensing for the Smart Grid”, IEEE Pervasive Computing, vol. 10, no. 1, pp. 28-39, 2011.
[11] H. Berriri, M.W. Naouar, and I.S. Belkhodja, “Easy and fast sensor fault detection and isolation algorithm for electrical drives,” IEEE Trans. Power electronics, vol. 27, pp. 490-499, February 2012.
[12] F. Mouelhi, H. Ben Attia Sethom, I. Slama-Belkhodja, L. Miègeville, P. Guérin, “A fast event detection algorithm for residential loads within normal and disturbed operating conditions,” European Journal of Electrical Engineering, vol. 18, pp. 95–116, February 2016.
[13] M. Basseville, I. V. Nikiforov, “Detection of abrupt changes - Theory and application”, published by Prentice Hall, 1993.
[14] K. N. Trung, E. Dekneuvel, B. Nicolle, O. Zammit, “Event Detection and Disaggregation Algorithms for NIALM System,” 2nd Int. Non-Intrusive Load Monitoring (NILM) Workshop, June 2014.
[15] A. Tritschler, R. Gopinath, “Improved speaker segmentation and segments clustering using the bayesian information criterion,” in Proc. 1999 European Speech Processing, vol. 2, pp. 679–682, 1999.
[16] J. Ajmera, I. McCowan, H. Bourlard,“Robust speaker change detection,” IEEE Signal Processing Letters, vol. 11, pp. 649-651, July 2004.
[17] B. Wild, K. Said Barsim, B. Yang, “A new unsupervised event detector for non intrusive load monitoring,” IEEE Global Conf. on Signal Processing Applications in Smart Buildings GlobalSIP, December 2015.
[18] M. Basseville, “Detecting changes in signals and systems - a survey,” Automatica, vol. 24, pp. 309-326, May 1988.
[19] P. Delacourt, C. Wellekens, “DISTBIC: a speaker-based segmentation for audio data indexing,” In Speech Communication, vol. 32, pp. 111-126, September 2000.
[20] B. Zhou, J. H. L. Hansen,“Efficient audio stream segmentation via the combined T^2 Statistic and Bayesian Information Criterion,” IEEE Trans. on speech and audio processing, vol. 13, July 2005.
[21] H. Stanislaw, N. Todorov, “Calculation of signal detection theory measures,” Behavior research methods, instruments and computers, vol. 31, pp. 137-149, March 1999.
[22] S. Atapattu, C. Tellembura, H. Jiang,“Analysis of Area under the ROC Curve of Energy Detection”, IEEE Trans. on wireless comunications, vol. 9, March 2010.
[23] C. Ma , J. Yu, B. Huang, “A Rapid and Robust Method for Shot Boundary Detection and Classification in Uncompressed MPEG Video Sequences”, Int. Journal of Computer Science Issues (IJCSI), Vol. 9, Issue 5, No 2, September 2012.