PHOTONIC SENSORS / Vol. 6, No. 2, 2016: 158–168

Shift Endpoint Trace Selection Algorithm and Wavelet Analysis to Detect the Endpoint Using Optical Emission Spectroscopy

Sihem BEN ZAKOUR[1*] and Hassen TALEB[2]

[1]*Higher Institute of Management Tunis, University of Tunis, Tunisia*

[2]*Higher institute of Business and Accounting Bizerte, University of Carthage, Tunisia*

[*]Corresponding author: Sihem BEN ZAKOUR E-mail: Sihembenzakour@yahoo.com

Abstract: Endpoint detection (EPD) is very important undertaking on the side of getting a good understanding and figuring out if a plasma etching process is done on the right way. It is truly a crucial part of supplying repeatable effects in every single wafer. When the film to be etched has been completely erased, the endpoint is reached. In order to ensure the desired device performance on the produced integrated circuit, many sensors are used to detect the endpoint, such as the optical, electrical, acoustical/vibrational, thermal, and frictional. But, except the optical sensor, the other ones show their weaknesses due to the environmental conditions which affect the exactness of reaching endpoint. Unfortunately, some exposed area to the film to be etched is very low (<0.5%), reflecting low signal and showing the incapacity of the traditional endpoint detection method to determine the wind-up of the etch process. This work has provided a means to improve the endpoint detection sensitivity by collecting a huge numbers of full spectral data containing 1201 spectra for each run, then a new unsophisticated algorithm is proposed to select the important endpoint traces named shift endpoint trace selection (SETS). Then, a sensitivity analysis of linear methods named principal component analysis (PCA) and factor analysis (FA), and the nonlinear method called wavelet analysis (WA) for both approximation and details will be studied to compare performances of the methods mentioned above. The signal to noise ratio (SNR) is not only computed based on the main etch (ME) period but also the over etch (OE) period. Moreover, a new unused statistic for EPD, coefficient of variation (CV), is proposed to reach the endpoint in plasma etches process.

Keywords: Dimension reduction; OES; plasma etching process; wavelet analysis; CV; SNR

Citation: Sihem BEN ZAKOUR and Hassen TALEB, “Shift Endpoint Trace Selection Algorithm and Wavelet Analysis to Detect the Endpoint Using Optical Emission Spectroscopy,” *Photonic Sensors*, 2016, 6(2): 158–168.

1. Introduction

Plasma is partially ionized gas [1]. Therefore, it contains electron energy which excites the atoms and molecules then de-energizes in emitting photons. Under those circumstances, the plasma thus emits light. On the temperature scale, plasma has the three following classical states, solid, liquid, and gas [2]. Plasma is used for the surface treatment through transforming the electrical energy into a chemical energy by separating molecules [3]. Thus, it contains not only radicals and reactive atoms but also ions which can be accelerated by an electric field applied to bombard surfaces. The plasma process is used in many industrial fields such as biomedical, food, textile, automotive, and micro-electronics. In the biomedical sector, plasma is used to sterilize instruments or modify surface properties to make

Received: 10 September 2015 / Revised: 20 January 2016
© The Author(s) 2016.This article is published with open access at Springerlink.com
DOI: 10.1007/s13320-016-0280-5
Article type: Regular

them bio-compatible, thus limiting the risk rejection by the human body. Plasma is also used for the deposition of protective layers on the biomedical tool surfaces. During the etch process, when the desired layer material is clear, the gas of plasma should be stopped to avoid the over etch of the underlying layer. At this moment, a signal will appear indicating that the required clearing is done [4]. The most popular method for detecting the endpoint is to monitor the trace of the reactive species emission or volatile products emission through optical emission spectrometer (OES) [5 – 8]. At the start of the endpoint phenomenon, the augmented intensity in a particular channel signal corresponds to a growth in the concentration of reactant in the plasma etch process, considering that the reactant species is less used in the surface reaction of the integrated circuit. In contrast, any decrease in the intensity of wavelength channel is assigned by a slack in product concentration, because the under product species is contrived in the integrate circuit (IC) surface reaction [9]. As the etched surface becomes more and more small, the collection of huge number of spectra is unavoidable in the aim of improving the detection of endpoint. The implementation of EPD system allows having multi-OES and then a precise stop procedure in a specific layer, which increases throughput and yield [10, 11]. In this paper, a new algorithm is proposed to select the important fifty endpoint traces named shift endpoint trace selection (SETS) from the full spectra in the first section. Then the linear and nonlinear dimension reduction techniques are applied named principal component analysis (PCA), factor analysis (FA), and wavelet analysis (WA), in Section 3, respectively. The results and the sensitivity analysis is done based on mean and coefficient of variation (CV) statistics through the use of signal to noise ratio (SNR) in Section 4. Finally, the concluding remarks are given in Section 5. Table 1 shows the list of abbreviations used in this work.

Table 1 Abbreviation lists.

Abbreviation	Definition
IC	Integrated circuit
SETS	Shift endpoint trace selection
PCA	Principal component analysis
FA	Factor analysis
WA	Wavelet analysis
CV	Coefficient of variation
SNR	Signal to noise ratio
EPD	Endpoint detection
ME	Main etch
OE	Over etch
M	Mean
SD	Standard deviation

2. Shift endpoint trace selection (SETS) algorithm

2.1 Endpoint states and traces

Endpoint detection is employed to identify when the etched film has been cleared to the underlying film. At this moment, the process can be stopped or modified to a more selective etch. To detect the endpoint, when the film will be removed, without falling on over etch state in other words without damaging or removing the underlying film, and being sure about avoiding also the under etch state, that is the film being etched has not been completely removed, as shown in Fig. 1.

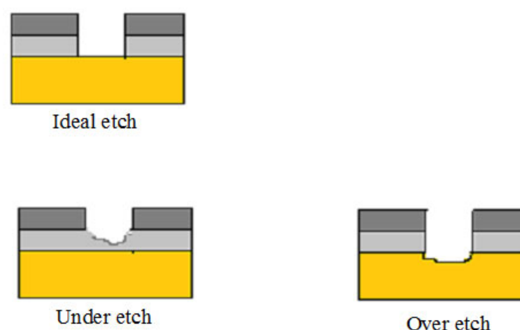

Fig. 1 Devices have been ideally etched, of which one has been over etched and the other under etched.

The ideal endpoint trace for an etch process has plotted intensity like a step change [12] as depicted in Fig. 2. This ideal case in the plasma etch process has no noise, no drift, and with uniform clearing of features across the wafer. In reality, the etch process is affected by some variations, and those variations in the etch rate will produce non-uniform clearing. Hence, the endpoint trace will contain error and drift

as shown in the aforementioned figure. In general, the endpoint detection does not occur at a specific time but refers to the range of times over which the film is cleared. The starting of endpoint is named the start of clear, and the finishing of endpoint is named the end of clear. During any chemical process, there is typically a transient state which starts at the beginning of any plasma process, which refers to the initial transient. Then, the signal generally obtains a steady state before detecting the endpoint, named the main etch [4].

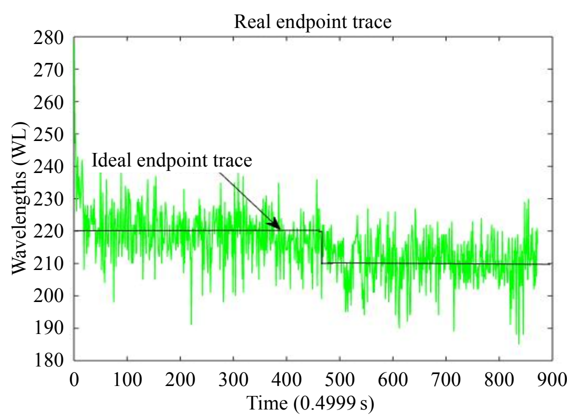

Fig. 2 Real endpoint trace can have various sources of noise and drift, with a change that occurs not in a single step, but gradually over some time span, due to non-uniform clearing of features and the ideal endpoint trace has step-wise change at endpoint and no noise in the signal.

2.2. Proposed algorithm named shift endpoint trace selection (SETS) algorithm

The growth on the collected data leads to a very large databases, high complexity, and long time execution [4, 6, 12]. The size reduction is one of the main tasks on the multivariate analysis [12]. It abates a large observed set of dimensions into a smaller features set. The major and the significant purposes of dimensionality reduction techniques are to visualize, compress, de-noise, and reduce the size of the data. As the importance of plasma etch process on the production of integrated circuit (IC) and on the side to understand and detect endpoint in the plasma etching process, collecting a huge number data (about of 1201×872×5=4695910 intensities) is unavoidable. All spectra intensities are presented in time resolution and spectral resolution. Despite the benefits of having a lot of information about all process details and progresses, this collection could handle the exactness of monitoring the endpoint. For this reason, the selection of the most important OES light is a decisive and essential task. A new proposed algorithm, named shift endpoint trace selection (SETS) to select the nearly meaningful time traces, is given as follows:

For all run Plasma etch step
  For time endpoint trace
  Compute |difference| between endpoint range
  Rank Difference with an increasing order
  Selecting the first fifty differences
  ENDFOR
ENDFOR.

3. Dimension reduction techniques

The use of multivariate methods for endpoint detection is unavoidable to monitor multi-wavelength channels. In this section, the multivariate tools are investigated. The matrices notations are given as a basic fact to master the multivariate analysis. And an introduction to matrix (linear) algebra is essential in order to better understand the next coming multivariate algorithms. The endpoint optical data are arranged in two-dimensional array (matrix) and given by the matrix below:

$$\mathbf{X}=\begin{bmatrix} x_{11} & x_{12} & \cdots\cdots & x_{1n} \\ x_{21} & x_{22} & \cdots\cdots & x_{2n} \\ . & & & \\ . & . & \cdots\cdots & . \\ . & & & \\ x_{m1} & x_{m2} & \cdots\cdots & x_{mn} \end{bmatrix} \tag{1}$$

where $\mathbf{X}$ is the intensity matrix having m time samples and n wavelength channels. Each sample x represents the spectra intensity for the ith time sample and the jth wavelength channel. It is often commodious to divide the matrix into row and column vectors. The column of the matrix $\mathbf{X}$ refers

to a particular wavelength trace, noted as x_j. Hence, the endpoint traces can expressed with **X** by

$$\mathbf{X}=[x_1 \quad x_2 \quad \cdots \quad x_j \quad x_{n-1}\, x_n]. \qquad (2)$$

The row vector of the matrix, $x_{i.}$, refers to spectrum at a specific time sample i. The matrix **X** could be expressed by using row vectors as follows:

$$\mathbf{X}=\begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ x_m \end{bmatrix}. \qquad (3)$$

As it was mentioned previously, the endpoint occurs seldom instantaneously, and in the most cases it occurs during a small time interval not on a specific point. On all occasions, the endpoint represents a mean shift from the main etch mean to the over-etch mean [4]. If this shift is much larger than some boundary which is computed from the etch data state, the endpoint is detectable. The matrix formulation of endpoint problem is given by the matrix **X** as a matrix containing two partitions, the main etch data and the endpoint data.

$$\mathbf{X}=\begin{bmatrix} X_{(\mathrm{ME})} \\ X_{(\mathrm{EP})} \end{bmatrix} \qquad (4)$$

where $X_{(\mathrm{ME})}$ contains the main etch data and $X_{(\mathrm{EP})}$ contains the endpoint data. The starting idea of principal component analysis (PCA) is to fractionate correlated data into a new set of uncorrelated measurements. The principal component analysis (PCA) is the most used method to reduce data [13−15]. References [16, 17] employed PCA to analyze in-situ spectroscopy data, and PCA is also used as a feature selection by [18, 19] in order to have information about processes and detect faults when there is no sufficient historical data. While the major aims of factor analysis (FA) is to identify the most significant data set to explain correlations among factors. There are several references that treat the factor analyses [20]. Reference [21] employed FA to evaluate of semiconductor ray spectra. Hence, the factor analysis serves to identify the correlation between the process variables and the common factors (latent variables). The main difference between PCA and FA is that the first relates variables into a small number of PCs and studies all variance while the second produces the factors and analyzes only the shared variance. The employment of PCA and FA which transform data on linear combinations of variables to analyze OES data represents a constraint themselves of linearity. A common form of multivariate non-linear analysis is the wavelet analysis. A wavelet is a waveform, with limited duration and having an average value of zero, and with irregular and asymmetric properties. As a result, there are different types of wavelets such as the Haar, Daubechies, Coiflets, Symlet sand, and biorthogonal wavelets [22]. For each aforementioned wavelet, they have their wavelet filters (low pass and high pass) while the Haar is the most simplest and its filter has only two coefficients in both low pass and high pass. The others such as Daubechies and Coiflet, have more vanishing moments not symmetric and also more coefficients both in low pass and high pass side. The Haar wavelet is a perfect choice in studying the time domain (compactly supported, small support, only 2 taps) but not in the frequency domain. In addition, the Haar wavelet has an efficient memory exactly reversible (easy reconstruction) and it is computationally the cheapest one. Wavelet theory, discovered by [23], has been employed in different scientific fields, such as physics, engineering and mathematic, data compression, and speech analysis. The wavelet analysis decomposes a function into frequency components that represent different degrees of function smoothness, with high frequency components capturing the least smooth function behavior while low frequency components capture the most smooth function behaviors, which makes it easy to extract the information exclusively in the time-frequency domain, as shown in Fig. 3.

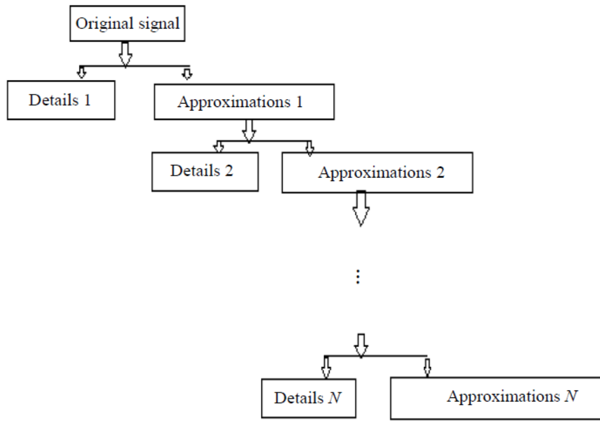

Fig. 3 Multi-resolution analysis showing the decomposed signal into approximation and details at Level *N*.

The wavelet analysis employs a linear combination of basis functions (wavelets), including time and frequency resolutions. For any function in L^2, the wavelet could be presented as follows [23]:

$$f(t)=\sum_{k\in Z}C_k\Phi_k(t)+\sum_{j=L}^{\infty}\sum_{k\in Z}d_{j,k}\Psi_{j,k}(2^j t-k) \tag{5}$$

where j and k are dilation and translation indices, respectively, and C_k and $d_{j,k}$ refer to the approximation and detail coefficients, respectively. $\Phi_{L,k}(t)$ is the father wavelets representing the low frequency and smooth parts of a signal, however, $\Psi_{j,k}(t)$ refers to the mother wavelet having high frequency and detail part of a signal. Their expressions of father and mother wavelet functions are given, respectively, as follows:

$$\Phi_{L,k}(t)=2^{L/2}\Phi(2^L t-k) \tag{6}$$

$$\Psi_{j,k}(t)=2^{j/2}\Psi(2^L t-k) \tag{7}$$

where $j,k,L\in Z$ and $2^{L/2}$ (and $2^{j/2}$) are needed to normalize the function $f(t)$, where, L (and j) corresponds to the level of time resolution (i.e, the width of the time interval) and k corresponds to the shift in the time location. The wavelet coefficients $c_{L,k}$ and $d_{j,k}$ are defined as inner products of $f(t)$ and the corresponding wavelet functions (mother and father) are called the discrete wavelet transformation of the signal $f(t)$, respectively. They are expressed as follows:

$$C_{L,k}=\langle f(t),\Phi_{L,k}\rangle=\int f(t)\Phi_{L,k}dt \tag{8}$$

$$d_{j,k}=\langle f(t),\Psi_{j,k}\rangle=\int f(t)\Psi_{j,k}dt. \tag{9}$$

The highest level of decomposition corresponds to the level after which there is a significant drop in the energy content, and the expression of energy content is given as follows:

$$EN_j=\int f_j^2(t)=\sum_{k=1}^{n}\left[d_{j,k}\Psi(2^j t-k)\right]^2. \tag{10}$$

The reconstructed signal is accurate only if the criterion of threshold selection is optimized. The threshold value using the Visushrink method (or [24−27] universal threshold rule) is given as follows:

$$t_j=\sigma_j\sqrt{2\lg(n)} \tag{11}$$

where n is the signal length and σ_j is the standard deviation of the noise at scale j. Only the significant wavelet coefficient situated outside of the threshold limits are extracted by applying soft or hard thresholding. In hard thresholding, the wavelet coefficient (at each level) above threshold will be unchanged (keep the same value for the coefficients that exceed the threshold), and the values which are lower than the threshold are made zero, which can cause large variance in the reconstructed signal and sometimes artifacts with an roughness appearance of the signal after reconstruction. However, it can better represent peaks and discontinuities. While the soft thresholding is an extension of hard thresholding, of which the thresholded coefficients are set to zero when the absolute values of wavelet coefficients are lower than the threshold (t_j) and adjusted by the following expression $\mathrm{sign}(d_{j,k})(|d_{j,k}|-t_j)$ if coefficients are upper than t_j. This method of thresholding gives better visual filtering quality. Indeed, it affects the detail threshold coefficients in a smooth way without making a radical change in its value. And the final step in the wavelet analysis is the reconstruction. Through inverse wavelet transforms, the signal $f(t)$ is reconstructed from the threshold wavelet coefficients. After determination of the threshold details and approximation at Level j, they will be used as inputs, to calculate the coefficients at Level (j–1) until getting the signal with the noise eliminated.

The summary of the main three steps in wavelet analysis:

Decompose: Choose a wavelet. Choose the Level J. Calculate the wavelet decomposition of the signals at the Level J.

Threshold: For each level from 1 to J, select a threshold and apply soft thresholding to the detail coefficients.

Reconstruct: Through the approximation coefficients of Level J and the thresholded detail coefficients the wavelet reconstruction is done.

4. Experimental results and discussion

4.1 Results

In this paper, the optical emission spectrometer (OES) is employed. And physically, the root of the optical emission is the light emitted through a chemical element, when the high energy state decreases to the lower one. In the plasma etch process, many chemical species have several emission spectra. The observed optical emission spectra display the chemical species and their variations. An optical emission spectroscopy should be able to resolve three components of plasma gas: (1) spectral resolution, (2) temporal resolution, and (3) spatial resolution. Hence, the study of the full spectral range OES is a challenging task. In this work, the sensor collects an array of measurements having 1201 channels of data, with over 827 units of time, since there are about approximately million data points in a single processing step. In other words, an optical emission spectroscopy is implemented in order to scan 1201 wavelengths (200 nm – 800 nm) from 0.4999 s to 435999 s. Given the extra data size, it is logical to ameliorate the sensitivity of the endpoint detection. And it is recommended to compress the data into a smaller subset that contains the most valuable information about the process, and at the same time minimizing the space on the hard drives by using dimension reduction techniques. The collected channels are gathered and analyzed in order to reach the real EP.

The first fifty rows (from 0.499 s to 24.999 s) referring to the initial state of plasma etch (Fig. 4) will be suppressed in order to avoid bias results (Fig. 5). Based on the new proposed algorithm named shift endpoint trace selection (SETS), only the first fifty endpoint traces having the highest intensity difference are selected. As the experimental OES data are coming from 5 etch steps, the total retained endpoint traces are equal to one hundred (50×5). Then reduction dimension techniques noted before will be applied to the retained traces to improve the picked-out endpoint traces. Moreover, the spectra are pre-processed to remove noise and reduce dimensionality.

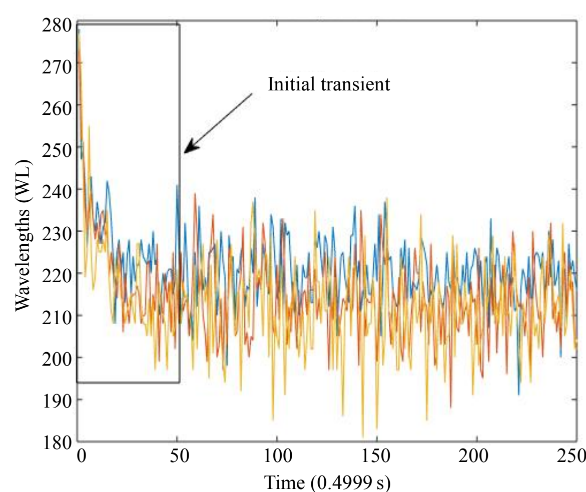

Fig. 4 Plot of an endpoint traces showing the initial transition state.

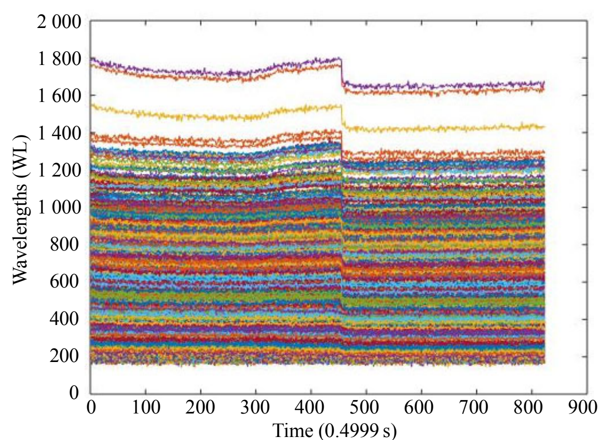

Fig. 5 Plasma etch endpoint traces from Step 3 after suppressing intensities from 0.499 s to 24.999 s.

The PCA is commonly used in the data analysis. The first fifty principal components are studied which catch most of the original data variation, even for large numbers of wavelengths (>1000). After applying PCA, the five retained endpoint traces from the fourth etch run notice that the endpoint is detected in 250.999 s to 252.499 s. The same procedure done on PCA is done on FA, hence the new proposed algorithm is preceded then FA is applied (Fig. 6).

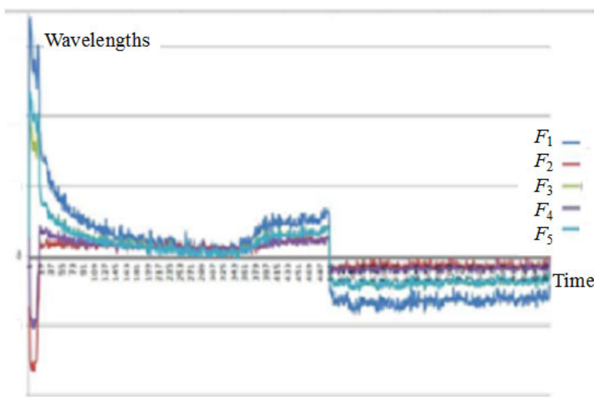

Fig. 6 Plotting the five significant retained factors.

The SETS algorithm is applied to the optical endpoint traces then the denoising procedure is applied by using the wavelet analysis. The chosen wavelet here is Haar wavelet. As mentioned previously, it is the most appropriate to describe the step change. Here, the obtained endpoint traces from the shift endpoint trace selection algorithm are then denoised and decomposed by using the wavelet analysis. The mean and CV of each endpoint trace of ME and OE are computed separately. From the obtained mean column presenting mean of all kept spectra, the mean will be decomposed at Level 3. This level is chosen based on the energy function drop. It should be noted that based on the gathered data, if the level of decomposition increases signal at a higher level, the signal will be smoother and may lose a lot of information about the right moment of endpoint detection and the species (gas) of the plasma etch process. Also, to plot endpoint traces, the reconstructed approximation coefficients will be used for those reasons noted below. (1) It is the denoised reconstruct original signal. (2) The endpoint detection is done based on the mean shift. (3) The detail coefficients represent high variance

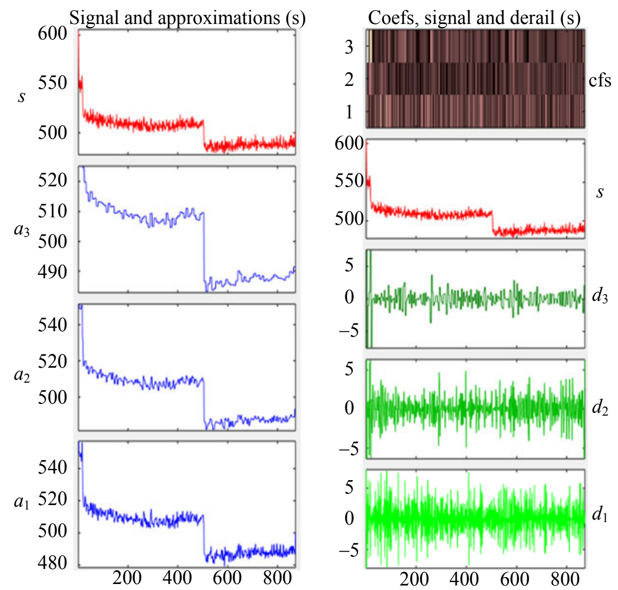

Fig. 7 Mean approximation signal at Level 3.

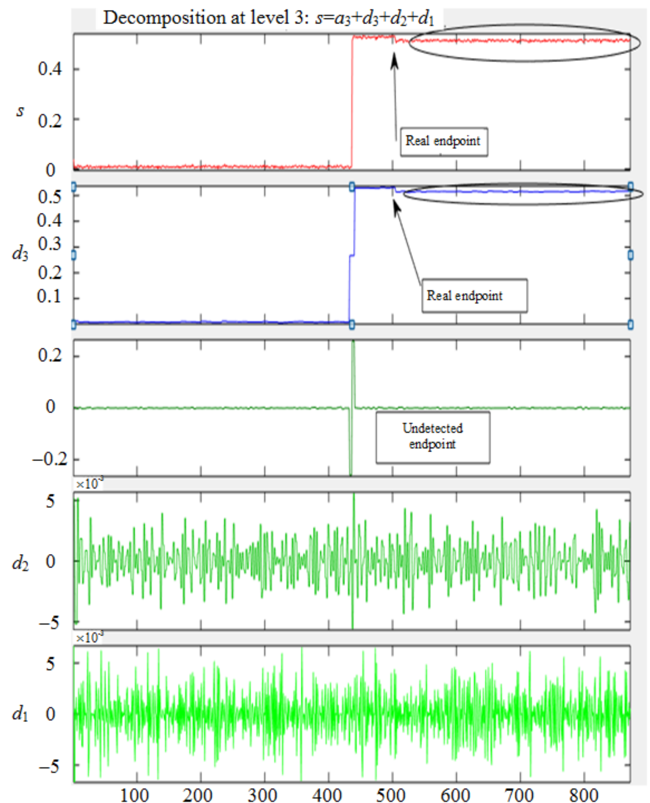

Fig. 8 CV approximation and details signal at Level 3.

and noise. The mean and the CV of each endpoint trace of ME and OE are computed separately, both of which will be decomposed at Level 3. Figure 7 shows the approximated mean wavelength at Level 3 that the endpoint is reached at the interval 250.999 s to 252.499 s. Based on Fig. 8, the WA-CV-approx at Level 3 records a meaningful shift before the real endpoint (under etched device). Therefore, the endpoint should be monitored based on the OE interval. The latter is more stable, and the first significant shift is detected at the real endpoint (250.999 s to 252.499 s). The WA-CV-details do not allow the detection of endpoint while the WA-CV-approx can detect endpoint if it is computed based on the coefficients of approximation in the OE interval.

4.2 Comparing result

As mentioned before, after initial transient and during the main etch step, a stable signal exists during the ME for each of the whole channels, but the intensity of the signal changes (decreases or increases) after the onset of endpoint. Any increase in the intensity of the signal refers to an increase in reactant in the plasma chamber, while the decrease in intensity of spectral channel refers to product. The SNR for the main etch period is the amount of signal compared with the noise on the main etch, which is used to compare the performance of the aforementioned preprocessing methods and is expressed as follows:

$$\left|SNR_{ME}\right| = \left|\frac{M_{OE} - M_{ME}}{\sigma_{ME}}\right|. \tag{12}$$

When the SNR for the over etch period is the amount of signal compared with the noise on the over etched device, it is expressed as follows:

$$\left|SNR_{OE}\right| = \left|\frac{M_{OE} - M_{ME}}{\sigma_{OE}}\right|. \tag{13}$$

A comparative result is summarized in Tables 2 and 3. Table 2 computes M, SD, and CV based on the main and over Etch intervals. Table 3 compares PCA, FA, WA-mean, WA-CV-approx, and WA-CV-details based on SNR. The SNR is computed during the main etch period and the over etch. CV coefficients are computed based on the approximated signal (approximation) and details.

Table 2 Mean, SD, and CV results based on Main and over etch intervals.

	M_{ME}	M_{OE}	SD_{ME}	SD_{OE}	CV_{ME}	CV_{OE}
PCA	40.48	41.017	53.451	6.7949	1.320	0.165
FA	5042.627	6917.3	6566.9	455.80	1.302	0.0659
WA-mean	511.1609	487.17	9.651	3.6795	18.88^{-3}	7.552^{-3}
WA-CV -approx	1.945	1.953	59^{-3}	14.1^{-3}	30.74^{-3}	7.2^{-3}
WA-CV -details	0.5309	0.5143	272.24	252.965	512.722	491.873

Table 3 SNR results based on Main etch interval and the Over etch interval.

	SNR_{ME}	SNR_{OE}
PCA	0.01034	0.079
FA	0.285	4.113
WA-mean	2.4857	6.52
WA-CV-approx	0.1099	0.496
WA-CV-details	6.095^{-5}	6.05^{-5}

4.3. Discussion of results

Based on the obtained results and with the most compelling evidence, the wavelet analysis method for the mean outperforms all other methods. And this is due to the characteristics of the WA-mean method, which is based on the flexible signal with no restriction about linearity, stationarity and symmetry of studied traces. In contrast, the wavelet coefficient of variation does not give us a better understanding about the endpoint detection, especially if it is computed based on the detail coefficients. As the CV is computed in terms of SD which tends to be non-stationary statistic having an increasing then decreasing trend, then the endpoint detection is not possible with statistic CV. The high variance hinders the endpoint detection. FA preserves the shape of most data, therefore, all the

retained five factors detect endpoint clearly. While the data are mean centered in PCA, the FA detects mean shift better than PCA based on SNR. The obtained result could be explained according to the differences between PCA and FA given below: (1) PCA works in the variable space while FA surpasses the variable space; (2) PCA resolves non-segmented variance while FA resolves common variance only; (3) PCA is an empirical summarizing technique keeping *m* components while FA is a theoretical modeling method suiting fixed number *m* factors to the data; (4) PCA is a dimension reduction technique only when FA is also a clustering technique which tries to find coherent variables; (5) FA is a more statistical technique used to translate an observed dataset into new axes, similarly to PCA. While PCA refines to combine variables into tiny PCs, FA examines the structure underlying the original variables. The SNR is a relative measure of the magnitude of a data set to the standard deviation. If the SNR is larger, the magnitude of the signal is relatively larger than the amount of noise which is quantified by the standard deviation. Then in this case, the studied signal is deemed to be significant signal. There is a negative relation between CV and SNR, such as the WA-CV-details for ME is 512.722 and its SNR is 6.095 e^{-6}. Hence, an inverse correlation is detected between them. The small peaks with SNR give a large CV while the largest SNR gives small details and approximations. Using details coefficients there is a high variance compared with the mean which is very small, therefore the SNR is very low. Also, there is a significant improvement of SNR for all used methods if this ratio is computed based on the variance of the OE period.

5. Conclusions and future perspective

Based on Fig. 9, the worst result is given by WA-CV-details because the details are generally used to monitor variance, and the monitored variance is very small compared with mean shift. Based on [28], for CV<0.5, the influence function response will have negative values. There is a negative correlation between CV and SNR. The WA-CV-approx surpasses PCA because the former is computed based on the mean and the variance of the approximate signal which are proportionally significant. In addition to that, the WA-CV-approx has no constraints such as linearity and mean centering data, which are the main postulates of PCA. Furthermore, WA-CV-approx has less performance than FA, because both methods do not need to mean central data. While for the linearity assumption, both are also appropriate but FA is the most appropriate because it is already designed for linear transformation. The ratio CV has a small amount compared with mean, therefore FA gives us better SNR results than WA-CV-approx. To detect EP, it is advantageous to use directly the approximation coefficients which identify quickly the mean shift (EP). Those results remain the same in both intervals (ME and OE) but it should be noted that there is an improvement of SNR for OE range because the variance during the aforementioned period is more stable and smaller compared with the ME period. In relation to our current results, one can investigate more OE periods to detect the EP and hence consider the plasma etch process CV for moving from unstable to stable one.

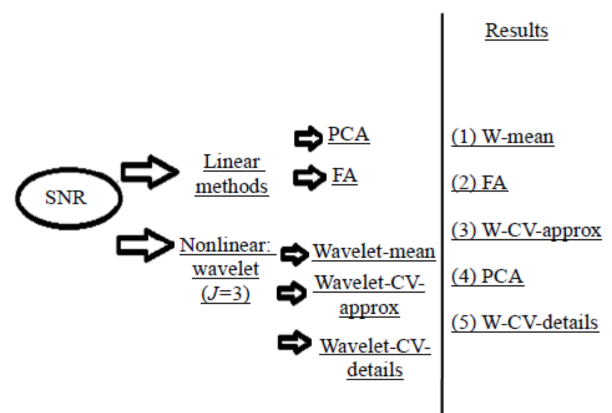

Fig. 9 Schematic presenting the obtained results.

Acknowledgment

The authors would like to thank reviewers and the editors for their valuable remarks and their time and every person helped us to do this work.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

[1] M. Mitchner and C. H. Kruger, *Partially ionized gases*. New York: Wiley, 1973.

[2] ChemWiki: The dynamic chemistry hypertext. http://chemwiki.ucdavis.edu/Analytical_Chemistry/Qualitative_Analysis/Classification_of_Matter.

[3] W. Taylor, "Technical synopsis of plasma surface treatments," Dissertation for the Degree of University of Florida, Gainesville, FL, December, 2009.

[4] B. Goodlin, "Multivariate endpoint detection of plasma etching processes," Ph.D. dissertation, Dept. Massachusetts Institute of Technology, U. S. A., 2002.

[5] R. Yang and R. Chen, "Real-time plasma process condition sensing and abnormal process detection," *Sensors*, 2010, 10(6): 5703−5723.

[6] J. Yang, C. McArdle, and S. Daniels, "Dimension reduction of multivariable optical emission spectrometer datasets for industrial plasma processes," *Sensors*, 2014, 14(1): 52−67.

[7] J. Yang, "Multivariable OES data analysis for plasma semiconductor etching process," Ph.D. dissertation, Dept. Dublin City University, Republic of Ireland, 2014.

[8] D. Mercier, M. Bouttemy, J. Vigneron, P. Chapon, and A. Etcheberry, "GD-OES and XPS coupling: a new way for the chemical profiling of photovoltaic absorbers," Applied Surface Science, 2015, 347 (2015): 799–807.

[9] H. K. Chiu, "Method of controlling plasma etch process," US patent 6703250 B2, 2004 March 9.

[10] P. M. Cederstav. "Increase vacuum processing throughput and yield using advanced downstream pressure control methods," in *The 44th Annual Technical Conference Proceedings*, Philadelphia, pp. 501−502, 2001.

[11] Z. Fekete, "Development and characterisation of silicon microfluidic components and systems." Ph.D. dissertation, Dept. Institute for Technical Physics and Materials Science Research Centre for Natural Sciences, Hungarian Academy of Sciences, Budapest, 2012.

[12] C. J. Pugh, "End point detection in reactive ion etching," Ph.D. dissertation, Dept. University College London, London, 2013.

[13] B. M. Wise, N. B. Gallagher, S. W. Butler, D. D. White, and G. G. Barna, "A comparison of principal component analysis, multiway principal component anal ysis, trilinear decomposition and parallel factor analysis for fault detection in a semiconductor etch process," *Journal of Chemometrics*, 1999, 13(13): 379−396.

[14] B. M. Wise and N. B. Gallagher, *PLS Toolbox Version 2.0*, Manson, WA: Eigenvector Research, Inc., 1998.

[15] I. T. Joliffe, *Principal component analysis*, New York: Springer-Verlag, 2002.

[16] B. Kim and M. Kwon, "Optimization of PCA-applied in-situ spectroscopy data using neural network and genetic algorithm," *Applied Spectroscopy*, 2008, 62(1): 73−77.

[17] H. S. Kim, Y. J. Sung, D. W. Kim, T. Kim, M. D. Dawson, and G. Y. Yeom, "Etch end-point detection of GaN-based device using optical emission spectroscopy," *Materials Science and Engineering: B*, 2001, 82(1−3): 159−162.

[18] J. Chen and J. Liu, "Derivation of function space analysis based PCA control charts for batch process monitoring," *Chemical Engineering Science*, 2001, 56(10): 328−3304.

[19] S. Yoon and J. F. MacGregor, "Statistical and causal model-based approaches to fault detection and isolation," *Aiche Journal*, 2000, 46(9): 1813−1824.

[20] R. L. Gorsuch, *Factor analysis (2nd ed.)*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1983.

[21] S. Krnac and P. P. Povinec, "Factor analysis of semiconductor γray spectra," *Applied Radiation and Isotope*, 1996, 47(9−10): 905−910.

[22] E. Brannock, M. Weeks, and R. Harrison, "The effect of wavelet families on watermarking," *Journal of Computers*, 2009, 4(6): 554−566.

[23] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Communications on Pure and Applied Mathematics*, 1988, 41(7): 909–996.

[24] D. L. Donoho, I. M. Johnstone, G. Kerkyacharian, and D. Picard, "Wavelet shrinkage: asymptopia?" *Journal of the Royal Statistical Society-Series B*, 1995, 57(2): 301−337.

[25] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrica*, 1994, 81(3): 425−455.

[26] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *Journal of the American Statistical Association*, 1995, 90(432): 1200−1224.

[27] D. L. Donoho and I. M. Johnstone, "Minimax estimation via wavelet shrinkage," *The Annals of Statistics*, 1998, 26(3): 879−921.

[28] R. A. Groeneveld, "Influence functions for the coefficient of variation, its inverse, and CV comparisons," *Communications in Statistics-Theory and Methods*, 2011, 40(23): 4139−4150.