Applications of Reference Cycle Building and K-shape Clustering for Anomaly Detection in the Semiconductor Manufacturing Process

by

Han He

Master of Science in Mechanical Engineering

University of Illinois at Urbana-Champaign, 2017

Submitted to the Department of Mechanical Engineering in fulfillment of the requirements for the degree of Master of Engineering in Advanced Manufacturing and Design

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2018

Author Signature redacted

Department of Mechanical Engineering
August 14, 2018

Certified by Signature redacted

Duane S. Boning
Clarence J. LeBel Professor, Electrical Engineering and Computer Science
Thesis Supervisor

Certified by Signature redacted

David E. Hardt
Ralph E. and Eloise F. Cross Professor, Mechanical Engineering
Thesis Reader

Accepted by Signature redacted

Rohan Abeyaratne
Quentin Berg Professor, Mechanical Engineering
Chairman, Committee for Graduate Students

This page is intentionally left blank.

Applications of Reference Cycle Building and K-shape Clustering for Anomaly Detection in the Semiconductor Manufacturing Process

by
Han He
Master of Science in Mechanical Engineering
University of Illinois at Urbana-Champaign, 2017
Submitted to the Department of Mechanical Engineering
on August 14, 2018 in partial fulfillment of the requirements for the degree of
Masters of Engineering in
Advanced Manufacturing and Design

Abstract

Early and accurate anomaly detection plays a key role in reducing costs and improving benefits, especially for complicated and time-consuming manufacturing such as semiconductor production. A case study of detecting anomalies from several monitored parameters during one plasma etching process is presented in this thesis. The thesis focuses on optimized ways to build reference cycles, or centroids of univariate parameters, a critical component to determine clustering accuracy and to facilitate process engineers' offline anomaly detections and diagnoses.

Three time series centroid building methods are discussed and evaluated in the thesis, arithmetic, the Dynamic Time Warping Barycenter Averaging (DBA), and the soft-DTW-based centroid. As a result, DBA is chosen considering its comprehensive performance of accuracy and calculation time. Optimizations on DBA is further discussed to reduce calculation time. The window constraint, as well as the recalculation method of combining the previous centroid and new datasets, substantially reduce calculation time with slight accuracy loss.

Based upon one centroid building method, shape extraction, a novel clustering method, k-shape, is implemented and applied to the plasma etching process. It is found that it achieves great accuracy with substantially shorter calculation time than one mainstream clustering method, k-means.

Thesis Supervisor: Duane S. Boning

Title: Clarence J. LeBel Professor, Electrical Engineering and Computer Science

Thesis Reader: David E. Hardt

Title: Ralph E. and Eloise F. Cross Professor, Mechanical Engineering

This page is intentionally left blank.

Acknowledgements

I would like to take this opportunity to deliver my sincere appreciation to the people helping me in this project.

First I would like to thank MIT and ADI for providing such a challenging and meaningful joint project to practice my skills and expand my expertise. The experience I have learnt from this project will be a solid background for my future career.

Thanks to Mr. Jack Dillon for his tremendous and prompt coordination and assistance in this project. I feel so supported in ADI with his help. Thanks to Mr. Ken Flanders for sharing his insights and leading me to the theoretical world of the semiconductor industry. Thanks also to Mr. Leslie Green and Mr. Charles Mathy for their expertise and inspirations.

Thanks to Dr. Duane S. Boning for his patient and thorough instructions and guidance during the project. Without his encouragement, I cannot easily jump out of the traditional anomaly detection methods, apply and optimize a series of innovative methodologies in this project.

Thanks to Mr. Jose J. Pacheco and Dr. David E. Hardt for the great operation and organization of the MEng program during the past year. It is a fabulous experience enriching my life and memory.

Thanks to my teammates, Tiankai Chen and Oumaima Makhlouk, for your seamless cooperation and strong support during the project.

Finally, I would like to thank my parents for their consistent understanding. Thanks to all the people who are loving and kind to me. You make my life brighter than I think.

This page is intentionally left blank.

Contents

List of Tables

List of Figures

1. Introduction

This introduction describes the value and the objective of the Machine Health Project from the level of ADI and the research team, then introduces individual tasks, and finally outlines the organization of the thesis.

1.1 Project Overview

Analog Devices Inc. (ADI) is an American international semiconductor company, specializing in the design, manufacture, and marketing of high performance analog, mixed-signal, and digital signal processing integrated circuits. The company's products play a fundamental role in converting, conditioning, and processing real-world phenomena such as temperature, pressure, sound, light, speed, and motion into electrical signals to be used in a wide array of electronic devices (Inc., 2015)[1]. Its products have been widely used in instrumentation, automation, communications, healthcare, automotive and numerous other industries [2].

In view of the fabrication cost, ADI faces the cost pressure to optimize manufacturing systems and processes, and is therefore exploring new approaches including machine learning to improve manufacturing. Before chips are finally diced, packaged and shipped, each wafer has to go through multiple process cycles such as plasma etching and implantation. Overall the process is long and suffers from low yield. Commonly when one certain cycle goes wrong, it is hardly possible to detect the anomaly in time and hence the entire process continues until the entire process is finished, at which point wafers may be found defective at the end. Such process failures during the production cause losses in both production costs and time, which are significant for the capital-intensive semiconductor industry. If anomalies could be detected in earlier and accurately, the plant could either scrap the work-in-progress part or adjust the following processes accordingly to compensate for the loss. Therefore it is meaningful to explore expanded monitoring of the process parameters, and accurately alarm and notify operators in time so that they can respond to anomalies at an early stage. ADI is thus interested in such in-time anomaly detections and gets involved in a series of innovative projects.

One such project is the Machine Health Project. The project intends to improve the way ADI collects and analyzes the data from the process. Its objective is to monitor process parameters, provide timely alerts on anomalies and thus enable efficient and effective response in time. It seeks improvements on machine reliability, productivity, quality and cost. In addition to the traditional methods such as traditional Statistical Process Control (SPC), ADI is seeking the feasibility of applying advanced methods such as machine learning to their production systems for further improvements.

1.2 Project Scope

Our project is to optimize the health monitoring model for machines and processes in ADI's Wilmington Fabrication Plant. Our team consists of three MIT students engaged in thesis research as part of the Master of Engineering in Advanced Manufacturing and Design Program. The objective is to evaluate the appropriate analytic techniques for timely anomaly detection.

The project requires us to implement analytics using R within the SQL server, which needs careful and efficient data preprocessing. ADI expects some "real-time" analytic tools and best practices, which could be either integrated into the online analytic platform or be provided to process engineers for their daily and off-line work.

The objective of the project includes:

- Evaluate key parameters and methods for anomaly classification
- Evaluate methodologies and algorithms for timely anomaly detections
- Find solutions to anomaly detection methodologies and algorithms
- Determine best practices for implementing "real-time" analysis
- Provide efficient and useful assistance tools for process engineer's daily analysis

According to the objective, the framework of the Machine Health Project is summarized in Figure 1. The data received from the machine's sensors is first preprocessed in the pre-processing section. The preprocessed data can be presented in plots by the visualization section. Then the pre-processed data goes through the data analytics section. The data be captured as reference cycles, or can go through the anomaly detection part. The built reference cycles in turn facilitate anomaly detection. Results of both reference cycle and anomaly detection can be shown in plots by the visualization section. After the data analytics section, the result goes through the interface section. The anomaly is alerted to the process engineers and all results and records are sent back to the database. Through the improvement of anomaly detection, the plant gains multidimensionally in terms of cost, yield and machine life. The algorithms themselves are not process, recipe and machine dependent, so that they could be extensively applied conveniently.

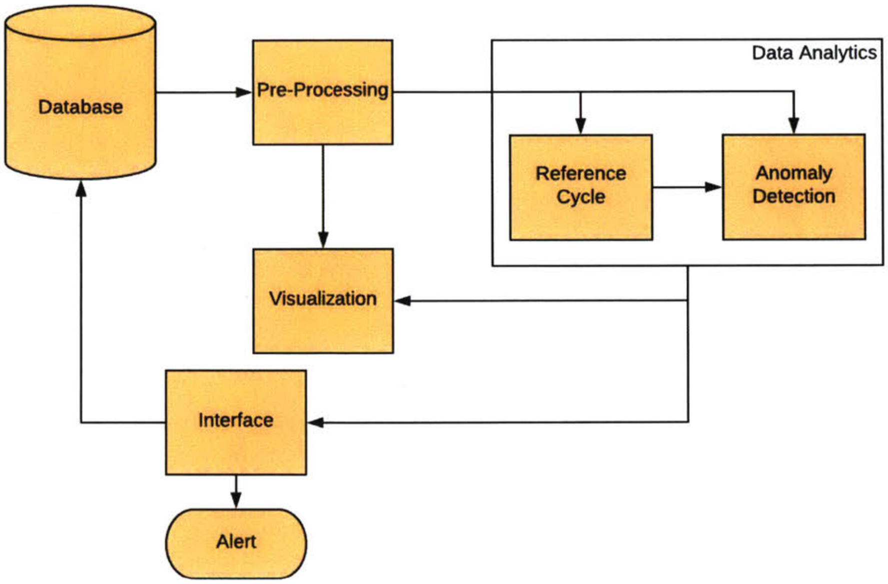


Figure 1. Machine Health Project Framework

1.3 Problem Statement

The problem statement explains the part of the project this thesis focuses on. The major case the team has worked on is first presented, and then the specific individual task to be solved in this thesis is described.

1.3.1 Case Overview

The major case studied in the project is related to unconfined plasma excursions happening during the plasma etching process. Plasma etching is one major process of semiconductor production. It involves plasma of an appropriate gas mixture generated and exposed to a sample. The plasma consists of etch species, which are either charged or neutral. The charged species are composed of ions and the neutral consist of atoms and radicals. During the process, the reactive species produced by the plasma react with the materials to be etched and produce volatile etch products at room temperature. Finally, the charged species are accelerated vertically to the wafer substrate by the applied electric field and embed themselves at or just below the target surface. In this way the target's physical properties are modified [5]. The mechanism is shown in Figure 2.

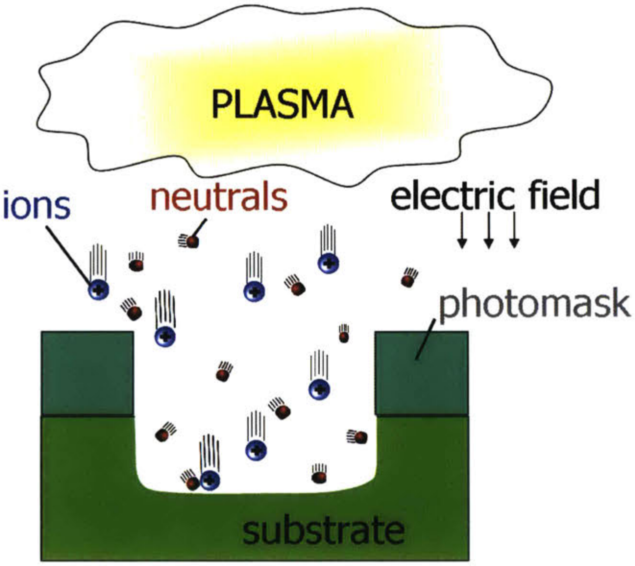


Figure 2. Schematic of Plasma Etching Process [6]

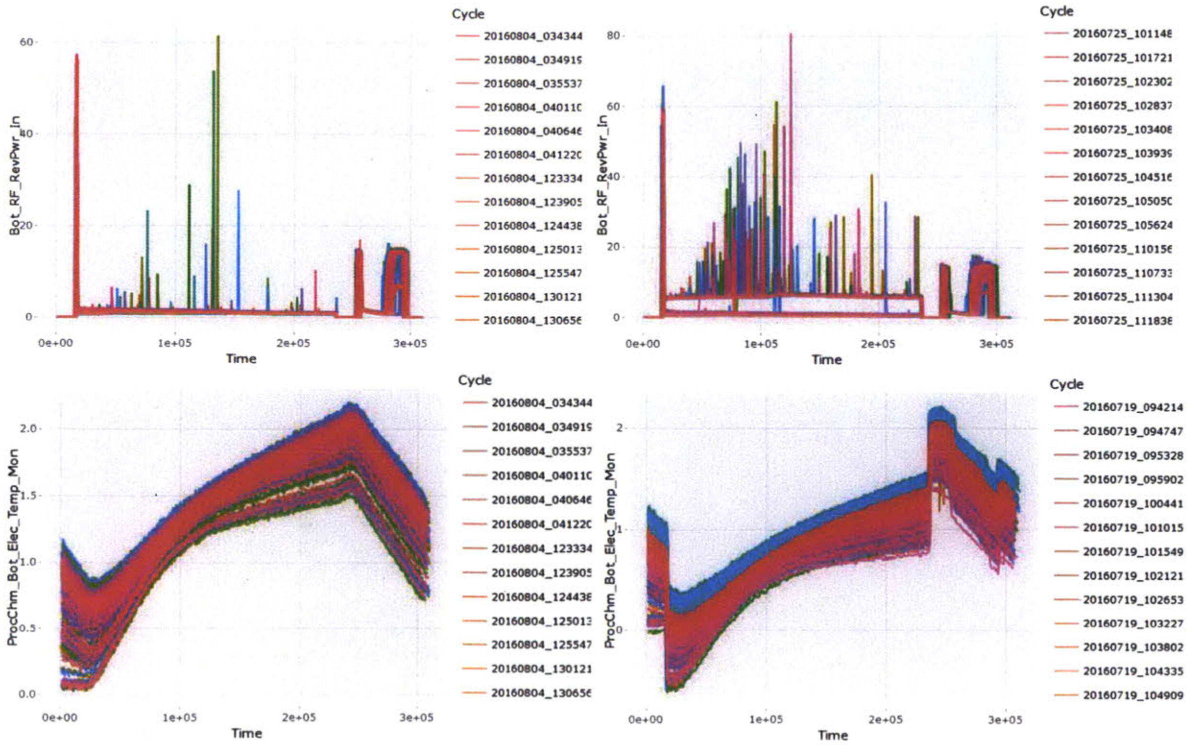
The unconfined plasma excursion causes unplanned etching on the wafer. Currently for the machine 'OXLR7_LAMAL1' three major related parameters are monitored and analyzed: 'BOT_RF_RevPwr_In', 'ProcChm_Bot_Elec_Temp_Mon' and 'ProcChm_EndPt_ChanC_In'. These parameters come with respectively definitions and meanings:

BOT_RF_RevPwr_In: The amount of the reflected Radio Frequency (RF) power reflected back to the supply. Radio Frequency (RF) power is used in plasma etching to ionize the gas generating a plasma. The RF power is a programmable parameter that is part of a recipe for the particular etch. Forward power is the power delivered to the load and reflected power is the power that is reflected back to the supply. Ideally reflected power is 0 W. However, due to inefficiencies of the RF match network or the physics of the chamber, there are losses. Reflected power is an indicator of how efficiently the power is being delivered to the chamber or the load.

ProcChm_Bot_Elec_Temp_Mon: An indicator of the temperature of the wafer during processing. The work-in-progress wafer sits on a chunk during processing with temperature monitored and recorded.

ProcChm_EndPt_ChanC_In: A parameter used to determine whether the etch is complete. It provides a signal, expressed in counts, of the plasma intensity at a specific wavelength. As the target material being etched is completed, the next layer is exposed, which results in a change in the spectrum along with a change in amplitude of the endpoint signal. A parameter value of zero indicates the completion of the etching.

Plots for each parameter are shown in Figure 3 for Recipe 920, with good-behaved cycles and mix-behaved cycles listed separately. A total of 341 cycles' data is shown in the figures.



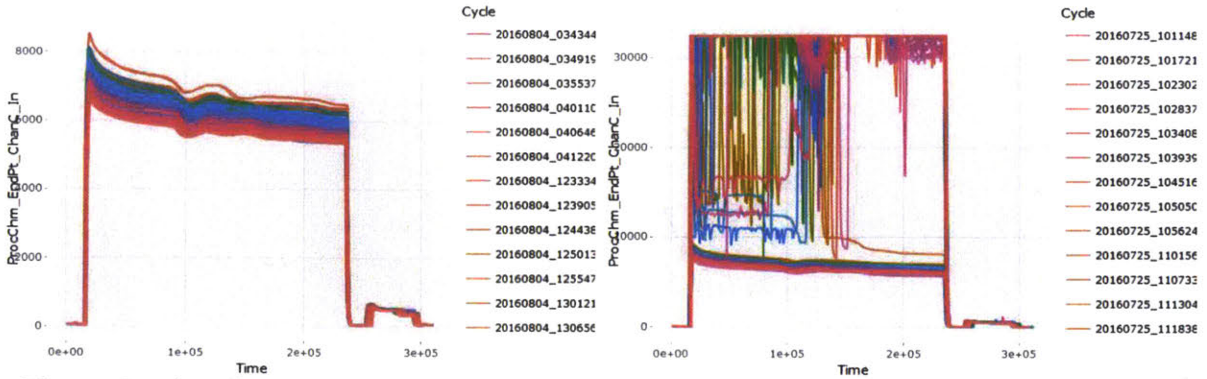


Figure 3. Plots for Parameter Behaviors for the Recipe 920, with Left the Good Cycles and Right the Bad Cycles

As can be seen, parameters behave normally in a cycle within ranges of amplitude and phase. For the parameter 'BOT_RF_RevPwr_In', it can be seen that the major anomaly is a step increase in the power in the first 80% section of the cycle. For the parameter 'ProcChm_Bot_Elec_Temp_Mon', major anomalies are sudden and drastic increases and decreases in the temperature at turning points rather than smooth changes. For the parameter 'ProcChm_EndPt_ChanC_In', the major anomaly is the much higher saturation point. Normally the value should be below 8000. In anomaly cases, the value reaches over 30,000 and stays at such high level until the end.

1.3.2 Individual Task

Individually, the initial objective in this thesis is to provide an accurate and efficient tool to assist engineers in building reference cycles, or centroids for their daily data comparison and analysis. Then, related and extensive applications of the centroids for anomaly detection and classification are explored, mainly around the k-shape clustering method.

When meeting with process engineers in ADI, we found that they preferred to manually track back anomalies using Excel. Normally, they plotted each process parameter and reasoned which parameter went wrong according to their experience, or rules of thumb. This method mainly caused two problems. Initially, without a standard centroid for each parameter quantitatively and geometrically in mind, engineers could easily incorrect false judgements and they tend not to make consistent judgements about the same cycle. This issue could be worse for engineers unfamiliar with the process.

In addition, without a mutually agreed-upon standard cycle, different engineers apply their individual rules of thumb and form diverse centroids in their minds. As a result, they sometimes cannot reach agreement on the anomaly detection for the same cycle. In short, currently process engineers cannot make consistent judgements on anomaly detection personally and interpersonally. One of the examples is the analysis of the parameter 'ProcChm_EndPt_ChanC_In', with the data shown in Figure 4. Among the 11 cycles, the three low-amplitude cycles are correct. However, it is hard for engineers to describe the amplitude and shape of such a good cycle accurately and thus they tend to make mistakes in later anomaly detections and analysis. It is necessary to help engineers build accurate centroids from normal cycles. These cycles improve their anomaly detection accuracy and analysis quality when they work offline. With a mutually agreed-on centroid, it is also smoother and more efficient for engineers to reach agreement on anomaly detection results.

Furthermore, an accurate centroid also improves the quality of semi-/automatic anomaly classification and detection methods, such as clustering. In addition to providing practical and accurate centroid building tools for process engineers, extensive applications of the centroid in anomaly detection and classification are also explored in the thesis.

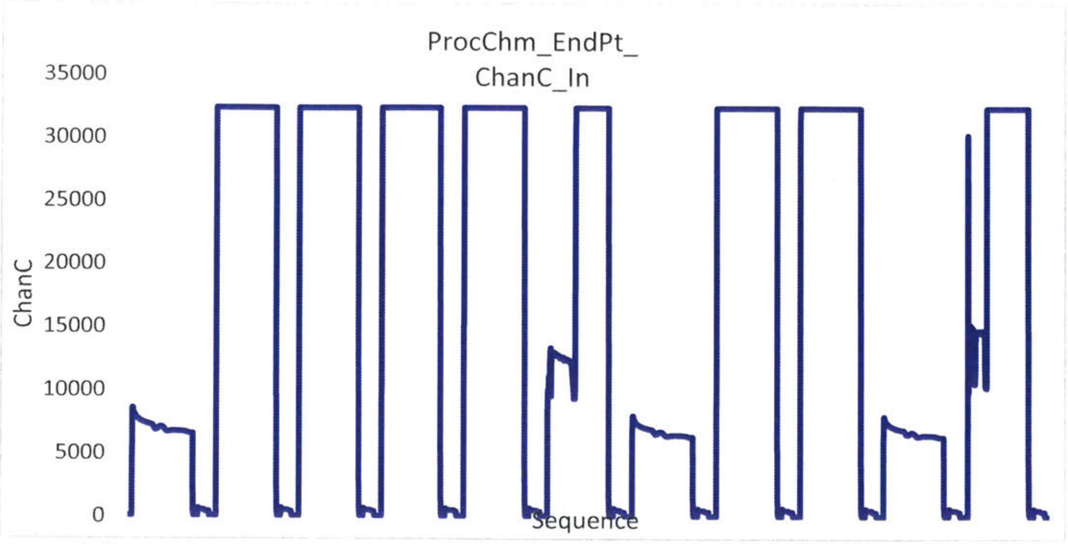


Figure 4. Mixture of Normal and Bad Time Series Data of Parameter 'ProcChm_EndPt_ChanC_In'

1.4 Thesis Outline

After the project's objectives and the problem statement are presented in Chapter 1, introductions to theory and algorithms are summarized in Chapter 2. These include the methods to measure the similarity between time series, centroid building methods, and one novel clustering method, the k-shape. Additional and necessary theories are introduced, such as the resampling for data pre-process, and the confusion matrix for clustering performance evaluation.

Following the discussion of theoretical background, methodologies for developing experiments evaluating diverse centroid building methods are discussed in Chapter 3. This chapter details the structure of experiments and important factors considered in designing experiments.

Results and discussions are presented in Chapter 4. The discussion is not only on the type of the centroid building method preferred, but also on optimization methods for further improvements. In addition, a brief evaluation of the k-shape clustering is presented. Finally, Chapter 5 makes conclusions on experiment results, value to ADI and recommendations on further tests and improvements.

2. Theoretical Backgrounds

Chapter 2 focuses on theoretical backgrounds the thesis is based on. Section 2.1 introduces metrics of judging time series similarities and types of centroid building methods. Section 2.2 introduces clustering theories and then the k-shape clustering method. Section 2.3 describes resampling, one necessary signal pre-process step. Finally, Section 2.4 introduces an important clustering evaluation method, the confusion matrix.

2.1 Time Series Similarity Measurement and Centroid Building Methods

This section first introduces a fast but inaccurate way to calculate centroids, the arithmetic method. In view of the arithmetic method's inaccuracy, metrics to evaluate time series similarities and build centroids are then introduced. One quantitative metric measuring the distance, Dynamic Time Warping (DTW) distance, is introduced in Section 2.1.2. The centroid building method based upon DTW distance, DTW Barycenter Averaging (DBA), along with a variant of DTW, soft-DTW, and the centroid building method based upon soft-DTW, are each presented in Sections 2.1.3 through 2.1.5. Another metric comparing shape similarity, Shape-based Distance (SBD), is then described in Section 2.1.6. In the end, the centroid building method derived from SBD, shape extraction, is presented in Section 2.1.7.

2.1.1 Arithmetic

The arithmetic method calculates the centroid based upon the mean/median of the time series data. Similar to the calculation of mean and median for arrays, the mean takes the average/median of each time-point *i* across all variables of the considered time-series [7]. Then for a cluster *C* of size *N*, the time-series mean μ is calculated by Equation 1, where $x_{c,i}^{v}$ is the *i*-th element of the *v*-th variable from the *c*-th time series which belongs to cluster *C*.

$$\mu_i^v = \frac{1}{N}\sum_c x_{c,i}^v \quad \forall \mathrm{c} \in \mathrm{C} \tag{1}$$

The median takes the median value rather than the mean value across series in the *C*. It is more robust to outliers across time series. Alternatively, winsorization could be used to obtain more robust series means, as shown in Figure 5. Unlike simply removing outliers in the trimmed mean method, the winsorization limits effects of outliers by replacing the smallest *k* values with the (*k*+*1*)-th smallest and the largest *k* values with the (*k*+*1*)-th largest [8]. The tightness of the winsorization can be adjusted by changing the upper and lower percentile of the boundary. Although it still brings bias to the result, the bias is better than simply removing all outliers and calculating the trimmed mean.

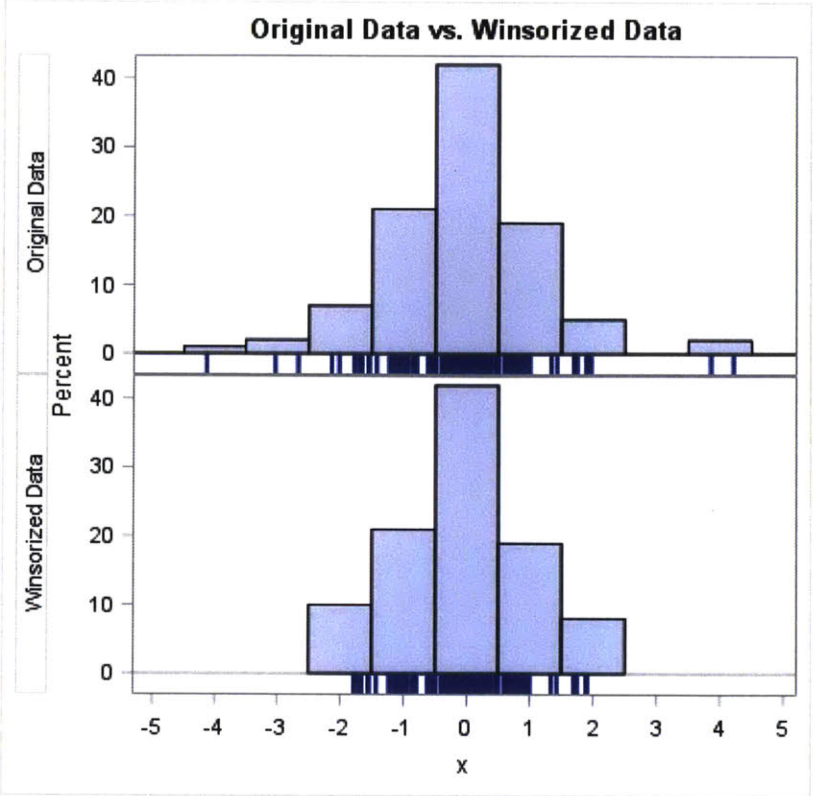


Figure 5. Original vs. Winsorized Data [8]

Overall, the arithmetic method is the simplest and the fastest. However, it is quite sensitive to phase-shift values and outliers. It is also restricted to applications on time series data with the same length and number of variables. From the perspective of these two concerns, dynamic time warping and shape-based distance methods are introduced to create more accurate and descriptive centroids.

2.1.2 Dynamic Time Warping (DTW)

Dynamic Time Warping (DTW) is a times series alignment algorithm calculating and comparing the dissimilarity between two time series based upon a distance measure. The shorter the DTW distance is, the more similar the two series are. It aims at warping two time series iteratively until optimally minimizing the DTW distance between the two time series and mapping one (query) onto the other (reference). For two time series, $A = (a_1, a_2, \ldots, a_n)$ and $B = (b_1, b_2, \ldots, b_m)$, with lengths of n and m respectively, it initially creates an n-by-m distance matrix. The time series A and B could be either univariate or multivariate time series, but the two should have the same number of parameters. Each element in the matrix is a cumulative distance of a minimum of the three surrounding neighbors. The (i,j) element $Y_{i,j}$ in the matrix is defined as:

$$Y_{i,j} = |a_i - b_j|^p + min\{ Y_{i-1,j-1}, Y_{i-1,j}, Y_{i,j-1}\} \quad (1 \le i \le n,\ 1 \le j \le m,\ Y_{0,0}=0,\ Y_{i,0}=Y_{0,j}=\infty) \qquad (2)$$

Here $Y_{i,j}$ is the summation of the distances between the i-th point in the A series and the j-th point in the B series, $|a_i - b_j|^p$, and the minimum of the three minimum distances around the (i, j) element. Variable p is the dimension of the $|a_i - b_j|$-norms. Normally p is chosen to be 2 so that the Euclidean distance is used to measure the distance between two points. The cumulative distance between the two series are finally determined by $Y_{i,j}$. An example of the mapping is shown in Figure 6, where the query series, $A = \{2, 3, 8, 2, 3, 1,3\}$ is aligned to the reference series, $B = \{3, 1, 2, 3, 8, 3,2\}$.

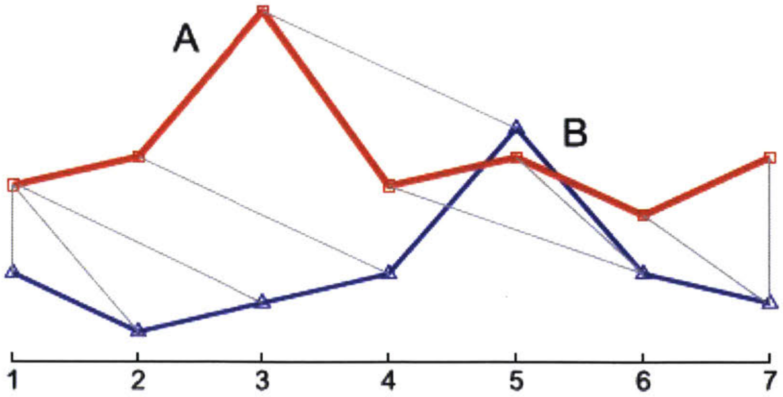


Figure 6. Mapping between the Two Time Series [9]

DTW can find an optimal global alignment between series and thus is probably the most popular measure to quantify the dissimilarity between sequences [10–14]. It has been shown to be one of the most effective distance measurement methods for time series [15]. Besides, unlike the arithmetic method, the two time series do not need to be of equal lengths. Therefore it is introduced in this thesis to generate centroids from a cluster of time series, which will be discussed in detail in the next section. However, since an n-by-m distance matrix needs to be created for the DTW and the computational complexity is O(mn), calculation becomes expensive and time-consuming. As a result, possible optimized methods will be discussed in Chapter 4. A variant of the DTW, soft-DTW, uses a differentiable measurement algorithm to calculate the distance between the two series and is more robust to shifts or dilatations across the time dimension [15], which will also be discussed along with the centroid building methods derived from the soft-DTW.

2.1.3 DTW Barycenter Averaging (DBA)

A warping path between the query and the reference time series is generated during the warping. The original query time series is warped with each point corresponding to a specific point in the reference time series. Multiple query points can refer to the same reference point. The DTW Barycenter Averaging (DBA) is introduced to generate a centroid from a cluster of time series based upon the DTW. This is an iterative and global method. The latter word means that the order the series get input into the function is not related to the result. During DBA, a centroid is initially selected for the cluster. Normally this begins by randomly selecting a time series from the cluster. On each iteration, the DTW alignment between each time series in the cluster and the centroid is recalculated and updated. All points in the cluster corresponding to the same point in the centroid are grouped and then are averaged to get the new value of that centroid point. Iterations continue until either the upper limit of the iteration time is reached or the centroid is converged.

2.1.4 Soft-DTW

As introduced in Section 2.1.2, soft-DTW uses a differentiable distance measurement algorithm, where both the value and gradient can be computed with quadratic time/space complexity. In contrast, the traditional DTW has quadratic time but only linear space complexity. As a result, soft-DTW builds smoother and more detailed centroids.

The difference between the DTW and the soft-DTW will be explained in detail. Equation 3 shows the algorithm for the DTW. It only involves (min, +) operations and thus holds linear complexity only. Given the distance matrix $\Delta(x,y) = [\delta(x_i,y_j)]_{ij} \in R^{n\times m}$ and the inner product

$<A, \Delta(x,y)>$, where A is the alignment matrix in $A_{n,m}$, the distance formulas below are used to generalize the total distance for two time series via the DTW and the soft-DTW methods, respectively. The distance between the two time series given the alignment is $<A, \Delta(x,y)>$. Equation 4 refers to the original DTW discrepancy [16] and Equation 5 refers to the Global Alignment kernel (GAK) [17]. The GAK is for the soft-DTW method.

$$DTW(x,y) = \min_{A \in A_{n,m}} < A, \Delta(x,y) > \tag{3}$$

$$k_{GA}^{\gamma}(x,y) = \sum_{A \in A_{n,m}} e^{-<A,\Delta(x,y)>/\gamma} \tag{4}$$

Compared with the traditional DTW algorithm, the GAK replaces all inner products with their neg-exponentials and uses (+, ×) operations. The GAK integrates over all alignments. Consider a list of n aligned distances $<A, \Delta(x,y)>$, $\{a_1, a_2, \ldots, a_n\}$, a unified minimum operator can be generalized as:

$$\min{}^{\gamma}\{a_1, a_2, \ldots, a_n\} = \begin{cases} \min_{i \le n} a_i & (\gamma = 0) \\ -\gamma \log \sum_{i=1}^{n} e^{-\frac{a_i}{\gamma}} & (\gamma > 0) \end{cases} \tag{5}$$

We define a unified distance algorithm:

$$\mathrm{dtw}_{\gamma}(\mathrm{x},\mathrm{y}) = \min{}^{\gamma} \{<A, \Delta(x,y)>, \mathrm{A} \in \mathrm{A}_{\mathrm{n,m}}\} \tag{6}$$

The result is controlled by the smoothness factor, γ. It can be seen that when γ approaches infinity, the dtw_{γ} converges to the sum of all aligned distances. When the distances $<A, \Delta(x,y)>$ are concave, dtw_{γ} (x,y) also turns concave gradually as the γ grows. Therefore the soft-DTW algorithm with γ >0 smooths out local minima and provides a better optimization landscape.

2.1.5 Averaging with the soft-DTW geometry

The centroid time series based on the soft-DTW geometry directly applies Fréchet means [18] to the dtw_{γ} algorithm. Given a cluster of N time series, $\{y_1, y_2, \ldots, y_N\}$ with fixed p parameters and lengths $\{m_1, \ldots, m_N\}$, the goal is to find a barycenter time series $x \in R^{p\times n}$ with length of n for all p parameters. With normalized weights for each time series, $\{\lambda_1, \lambda_2, \ldots, \lambda_N\}$ and sum of the weights of 1, the centroid x is built in such a way:

$$\min_{x \in R^{p\times n}} \sum_{i=1}^{N} \frac{\lambda_i}{m_i} dtw_{\gamma}\ (x, y_i) \tag{7}$$

2.1.6 Shape-based Distance (SBD)

The shape-based distance (SBD) is a faster alternative to the DTW algorithm. Compared to the DTW, it compares the shape similarity rather than the quantitative distance. It is based upon the cross-correlation with coefficient normalizations (NCCc) between the two time series. In short, a global shift is made onto the query series x to maximize the cross-correlation between the query series (x) and the reference (y). Equation 8 considers the cross-correlation between the query series

(x) and the reference (y), both with lengths of m, and maximizes cross-correlation value when the query shifts by k. For simplicity, we consider the alignment for two series with the same length though the alignment also works for series with different lengths.

$$R_k(x,y) = \begin{cases} \sum_{l=1}^{m-k} x_{l+k}y_l \ (0 \le k \le m-1) \\ R_{-k}(y,x) \ (-m \le k < 0) \end{cases} \quad k \in Z \tag{8}$$

Then the shifted x series $\overrightarrow{x(k)}$ is:

$$\overrightarrow{x(k)} = x_{1*m} = \begin{cases} 0, \dots, 0, x_1, x_2, \dots, x_{m-k} \ (s \ge 0) \\ x_{1-k}, \dots, x_{m-1}, x_m, 0, \dots, 0 \ (s < 0) \end{cases} \tag{9}$$

Cross-correlation is sensitive to the scales. Normally the series are z-normalized and the NCCc is defined as:

$$NCC(x, y) = \frac{R_k(x,y)}{\sqrt{R_0(x,x)}\sqrt{R_0(y,y)}} \tag{10}$$

The SBD is then defined as:

$$SBD(x, y) = 1 - max(NCC(x, y)) \tag{11}$$

The SBD value is between 0 and 2. A value of 0 means that the two series are perfectly identical in shape. In comparison with the DTW algorithm, the SBD algorithm is much faster. With the application of fast Fourier-transformation in the SBD, the complexity is O($mlog$(m)) instead of O-(mn) and hence the speed is substantially improved, especially for long time series. However, since the time series in comparison have to be normalized, the SBD can describe shape similarity/dissimilarity but cannot indicate differences in amplitudes.

2.1.7 Time-series Shape Extraction

The DBA method is used to capture representative and shared characteristics of a cluster of time series and to build up centroids based upon the DTW distance measurement. Similarly, the shape extraction method is used to build a centroid for a cluster of normalized time series based upon the SBD measurement. Unlike the way the DBA builds centroids based upon average of the points grouped to the same centroid point according to the DTW alignment, the shape extraction uses numerical optimization [7]. Given a series of *N* normalized and shifted time series vectors after the SBD alignment with a *1*-by-*m* centroid *C*, $\{x'_1, x'_2, \ldots, x'_N\} \in R^{1 \times m}$, the process of finding the centroid is given in Equation 12 [19]:

$$\begin{cases} S = X'^{T}X' \\ Q = I - \frac{1}{N}O \\ M = Q^{T}SQ \\ C' = Eig(M,1) \end{cases} \tag{12}$$

In detail, the *X'* is an *N*-by-*m* matrix with $\{x'_1, x'_2, \ldots, x'_N\}$ spanning each row. The symbol ' means transpose. The *I* is the *m*-by-*m* identity matrix and the *O* is the *m*-by-*m* matrix with all ones. The output *C'* is the first eigenvector of the matrix *M* and the new centroid for the cluster.

Commonly the shape extraction operation begins with randomly selecting one time series from the cluster as the initial centroid. Then all series in the cluster are shifted and aligned to it before shape extraction. As can be seen from Equation 13, the centroid is not calculated iteratively and therefore the derived centroid may not get as good results as the DBA, since the latter includes iterations in the algorithm. Unlike the arithmetic method for centroid building, the SBD operation is not constrained to equal-length time series. However, currently the SBD is applicable to univariate time series analysis only. In addition, since each time series is normalized locally, shape-extracted centroids cannot be denormalized and restored with actual amplitudes.

2.2 Clustering

In extension to the centroid building, clustering theory is first summarized and then the k-shape clustering method is introduced in this section.

2.2.1 Introduction to Clustering

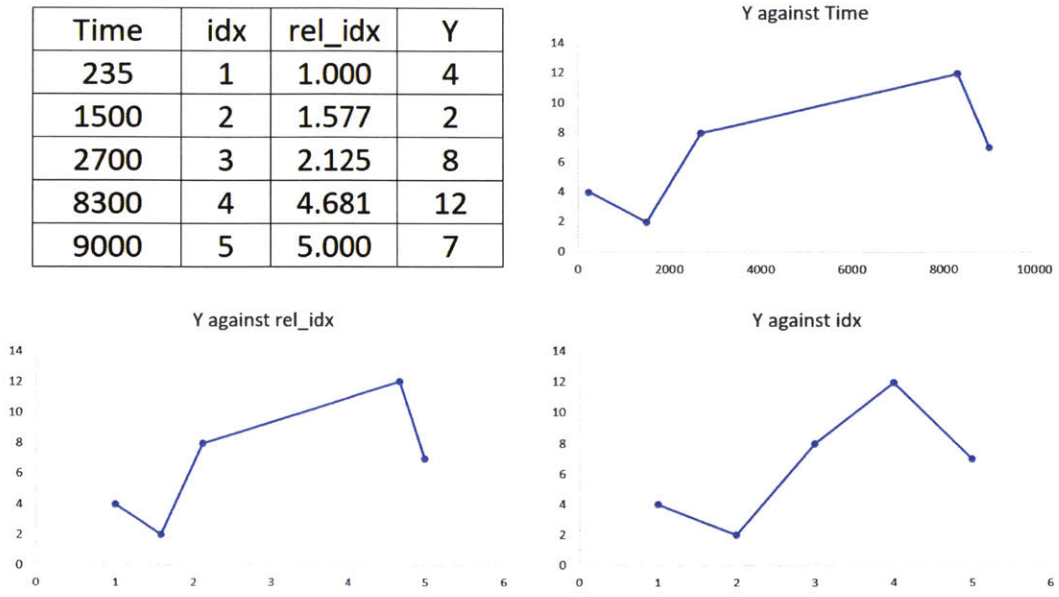
Clustering is the task of dividing a set of objects into several clusters, in such a way that each cluster is characterized with homogeneity and separation. The former refers to the similarity of observations within the same cluster, and the latter refers to the dissimilarity across different clusters [20]. Two mainstream types of clustering are hierarchical and partitional methods. Both methods rely on distance/dissimilarity measurement algorithms to optimize the similarity and dissimilarity iteratively, and form homogeneous and well-separated clusters eventually. The difference lies in whether clusters are nested or not. As for the hierarchical method, clusters are nested, while for the partitional method, time series are divided into non-overlapping clusters. Each method has respective advantages and disadvantages. The hierarchical method calculates iteratively and eventually forms an optimized number of clusters without requirement of presetting the quantity of clusters, which is good for taxonomy. However, since the distance/dissimilarity has to be calculated pairwise for every two time series, the calculation is particularly complex and expensive for a large set of data. The partitional method requires a preset quantity of clusters but has lower calculation complexity and cost.

2.2.2 K-shape Clustering Method

The k-shape clustering method uses the SBD to compare time series' similarities and calculate their distances, and then update the assignment of time series to clusters as well as the cluster centroids [19]. It is processed through two cycling steps: (1) assign each time series to the closest centroid with the shortest SBD; (2) use the shape extraction to update cluster centroids. Once the reassignment of time series is stopped or the iteration limit is reached, the k-shape clustering is completed. Through this iterative procedure, the k-shape minimizes the sum of the squared distances between time series and their centroids. The k-shape clustering's advantage is that it scales linearly with the number of time series [19]. However, since the SBD only works on univariate time series, the k-shape clustering does not support multivariate clustering.

2.3 Resample

Resampling is used to convert time series to uniform lengths for convenience of comparison. The time series is marked with either time or the consecutive indices. An index multiplier should be defined to decide the new quantity of indices for the data. Interpolation is used to update index spacings, as well as a list of relative index compared to the old one. Then the data value is interpolated according to the relative index value. One example of resampling data to the new relative indices is shown in Figure 7. Time lengths remain the same for resampled series but sampling frequencies vary.



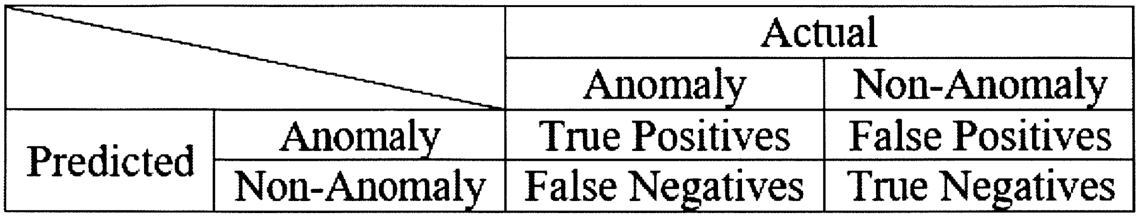
Time	idx	rel_idx	Y
235	1	1.000	4
1500	2	1.577	2
2700	3	2.125	8
8300	4	4.681	12
9000	5	5.000	7

Figure 7. Resampling the Time Series According to Relative Indices

Since time sampling rates and time lengths vary within and among time series, sampling frequencies can vary from time series to time series. However, sampling rates mainly result from round-up errors during the measurement and time lengths of series for the same process and recipe mostly vary within 5-percent range in the cases we studied. Therefore we assume that the resampling has insignificant loss on the time series data.

2.4 Confusion Matrix

The parameter to judge clustering quality is the accuracy, considering the percentage of making type-I error (false negative) and type-II error (false negative). The confusion matrix is used here to measure the clustering accuracy, as shown in Figure 8.



		Actual	
		Anomaly	Non-Anomaly
Predicted	Anomaly	True Positives	False Positives
	Non-Anomaly	False Negatives	True Negatives

Figure 8. Confusion Matrix

The result is divided into four parts: true positives (TP), false positives (FP), false negatives (FN) and true negatives (TN), where the FN counts Type-I errors and the FP counts Type-II errors. Additional parameters can be calculated based upon the confusion matrix, and measure the clustering accuracy regardless of scale:

$$Precision = \frac{TP}{TP+FP} \tag{13}$$

$$Recall = \frac{TP}{TP+FN} \tag{14}$$

$$F_1 = \frac{2}{\frac{1}{Precision}+\frac{1}{Recall}} \tag{15}$$

The precision considers the possibility of Type-II error and omittance of anomalies. The recall considers the accuracy of detecting normal cycles. The *F_1* value is the harmonic weighted mean of the two parameters. For all three parameters, the higher they are, the better performance of the clustering is. The parameter precision is particularly important since the omittance of anomalies causes more issues and extra costs rather than the false alarm.

3. Centroid Building Test Methodology

A series of experiments and comparisons are set up to determine appropriate applications of centroid building methods and parameter settings for the parameter analysis of the plasma etching the thesis works on, which could also be extensively applied to a wide variety of other recipes and processes. This chapter describes factors, parameters and metrics that the experiment design takes into account. The experiment scheme is described in the end.

3.1 Parameters

Multiple factors should be taken into account when constructing experiment processes and structures. Initially key parameters should be chosen for the centroid building. Among tens of process parameters in one set of time series data, the most representative parameter evaluating the process behavior should be primarily considered. In addition, parameters which are hard to describe and standardize quantitatively and qualitatively should be taken into consideration. Even when behaving normally, time series of such parameters vary greatly in terms of amplitude and phase.

3.2 Resampling

As discussed in Section 2.3, the influence of resampling is negligible on the datasets we are concerned with, and hence resampling is used to extend every set of time series to the same length, in order to cancel out the influence of data length differences on the distance measurement.

3.3 Sample Size

Centroid building methods should be measured under different sample sizes. A well-behaved method should deliver desirable results over variable sample sizes.

3.4 Randomness

Samples should be chosen randomly for each test so that the behavior of the centroid building methods is not influenced by the order and the subset of the selected series chosen for the test.

3.5 Performance Rating

The primary metric evaluating a centroid building method is that whether it can extract a representative time series optimizing the similarity from the samples it is based upon. The distance between the centroid and the time series is a good way to judge, since distance is a linear variable describing the similarities directly.

Minkowski and DTW distances are considered for measuring the similarity from the perspective of distance. The former measures the distance between two equal-length and equal-dimension time series $\{x_1, x_2, \ldots, x_N\}$ and $\{y_1, y_2, \ldots, y_N\}$ with Equation 16 [21]:

$$d(x,y) = \left(\sum_{i=1}^{N} \lceil x_i - y_i \rceil^p\right)^{1/p} \tag{16}$$

The case where $p = 1$ is equivalent to the Manhattan distance and the case where $p = 2$ is equivalent to thc Euclidean distance. Normally the Euclidean distance is more widely accepted. The Minkowski distance is a fast way to calculate the distance, or the dissimilarity, between the two time series. However, it is not robust to off-phase time series and tends to result in large distances even when the two time series have minor offsets in phase. The DTW distance, as described in Equation 3, is used instead, since it warps the query time series locally and is less sensitive to off-phase time series. Since the SBD measures the shape similarity between the two time series, it is used as the secondary metric evaluating the similarity/dissimilarity between the centroid and the time series. It is used to evaluate from the perspective of shape rather than the quantitative distance. In addition to the similarity/distance measurement, calculation time is also an essential metric. It is critically important when operators need to extract centroids of dozens of parameters from thousands of sample datasets.

3.6 Centroid Building Method

Arithmetic, DBA and soft-DTW-based centroid are the three methods to be discussed and compared. The shape extraction method is not considered since each time series is z-normalized locally and thus the centroid cannot be denormalized. As discussed earlier, DBA and soft-DTW-based centroid methods are more advantageous since they mitigate the influence of off-phase time series. The soft-DTW-based centroid is an advanced method based upon the soft-DTW. It is logical to compare and choose one of the three methods first and then discuss the possibility of further improving the chosen method.

3.7 Centroid Building Scheme

Two parameters are chosen from the case reviewed in Section 1.3, the 5th parameter 'Bot_RF_RevPwr' and the 19th parameter 'ProChm_EndPt_ChanC'. Good cycles of 'Bot_RF_RevPwr' are shown in Figure 9. This parameter is chosen considering its complicated variances in amplitude and phase and thus the necessity of building a uniform centroid. Good cycles of 'ProChm_EndPt_ChanC' are shown in Figure 10. This parameter is chosen since it is one of the key parameters typically used by engineers to determine whether the process behaves normally.

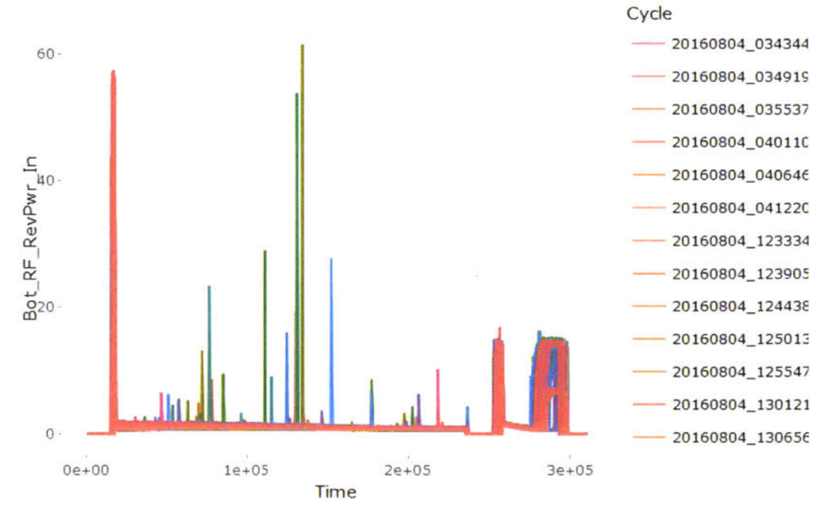


Figure 9. Good Cycles of 'Bot_RF_RevPwr'

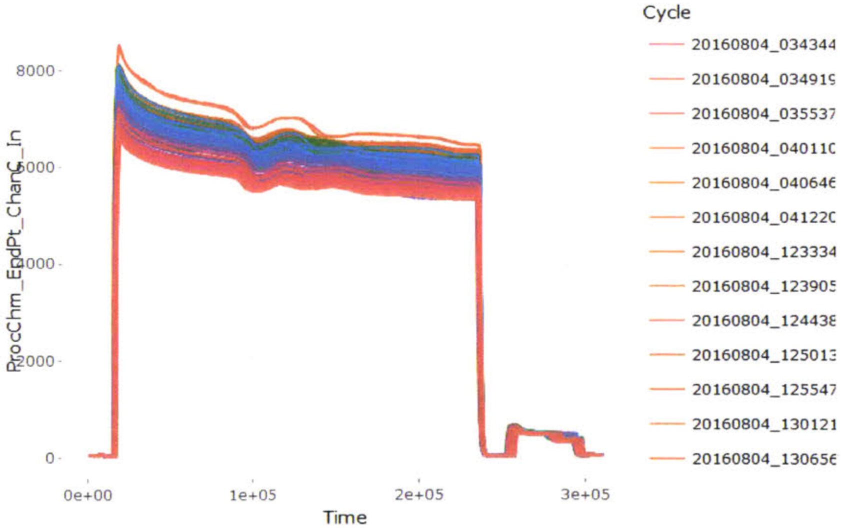


Figure 10. Good Cycles of 'ProChm_EndPt_ChanC'

Sample sizes of 20 and 50 are chosen for each of the two parameters. For each sample size for each parameter, 5 packages of random-picked data are analyzed. In terms of the centroid's accuracy, the DTW distance and the SBD are applied. The former prefers to quantitatively compare the distance between the centroid and the time series, and the latter prefers to indicate how similar the centroid's shape is to the time series'. Calculation time of each method is also listed and compared. Initially the methods in comparison will be arithmetic (winsorized mean with winsorization level of 0.05 on each side, untrimmed mean, and median), DBA (without any constraints), and soft-DTW-based centroid (smoothing parameter set at 0.001, 0.01, 0.1 and 1). The smoothing parameter has a large influence on the soft-DTW-based centroid method so that the range of the smoothing parameter is expanded widely to comprehensively show the method's performance. After comparing the DTW distance and calculation time from the level of each centroid method, the well-behaved method will be chosen and ways to further improve its performance will be discussed later.

4. Results and Discussion

This chapter summarizes the result of the centroid comparison experiment discussed in Chapter 3 first, and chooses the most appropriate method for the plasma etching process discussed in this thesis. Extensive optimization approaches to further improve this method are then explored. Finally, application of the k-shape clustering method is introduced, along with its result compared with the mainstream k-means method.

4.1 Method Selection

Representative centroids built from the three models discussed above are shown in Figures 11 through 19 for the case of 20 time series of the parameter 'Bot_RF_RevPwr', with black lines indicating the original datasets and the red dashed lines the centroid.

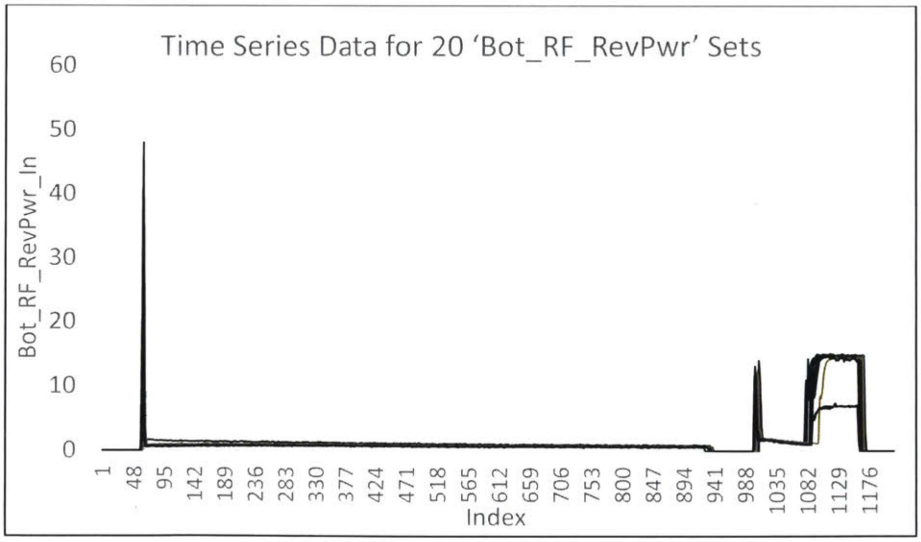


Figure 11. Time Series for Parameter 'Bot_RF_RevPwr', with Sample Size of 20

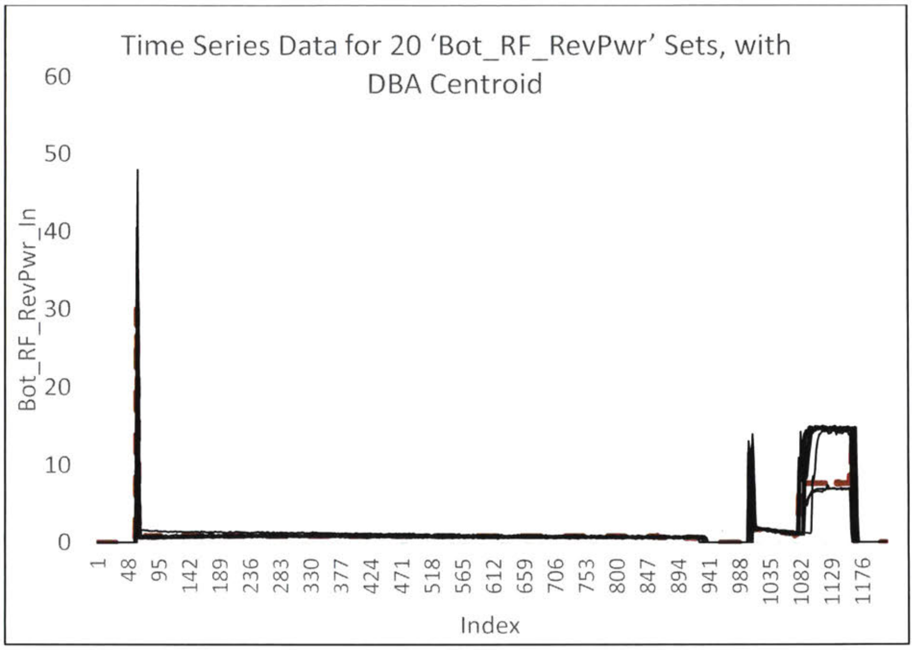


Figure 12. Time Series and Centroid for Parameter 'Bot_RF_RevPwr', with Sample Size of 20 and DBA Centroid Building Method

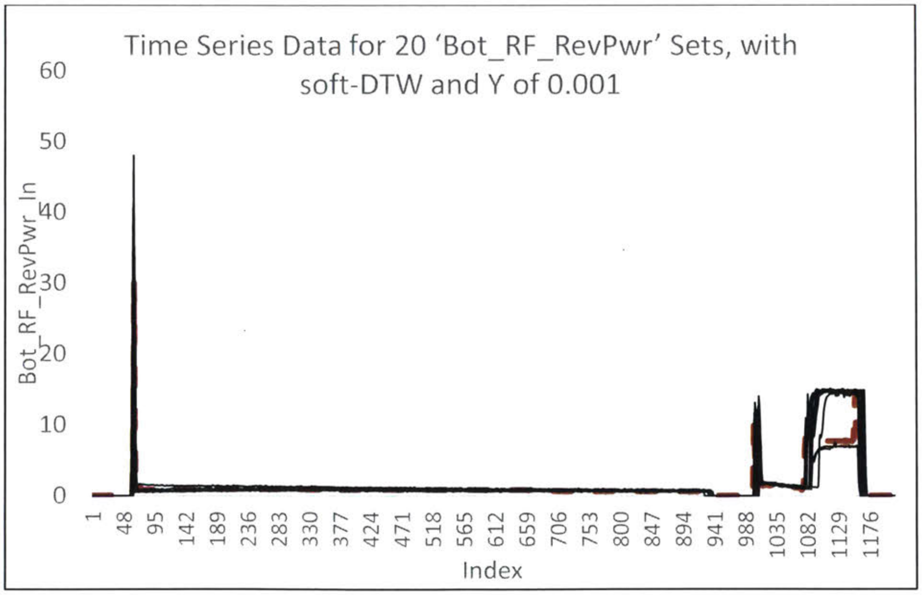


Figure 13. Time Series and Centroid for Parameter 'Bot_RF_RevPwr', with Sample Size of 20 and soft-DTW Centroid Building Method with γ of 0.001

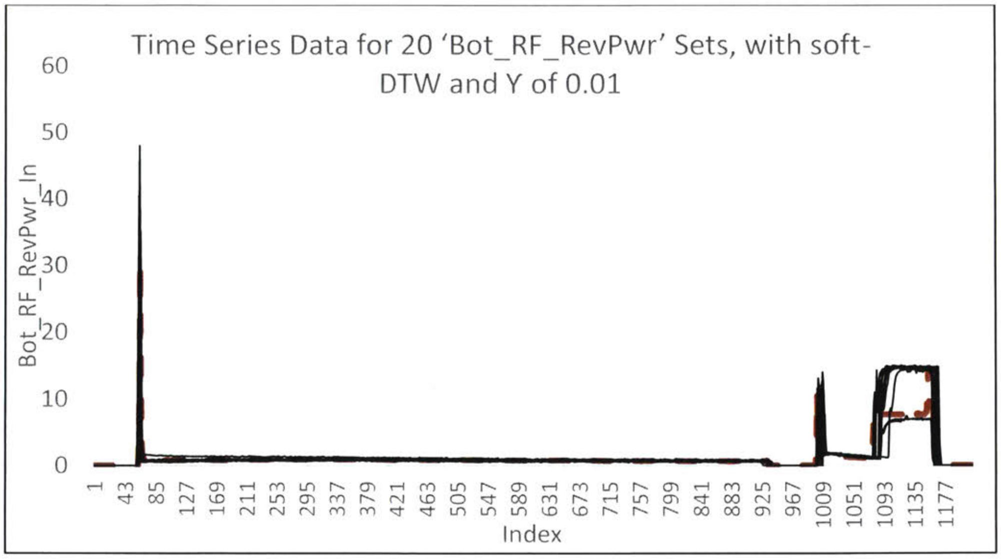


Figure 14. Time Series and Centroid for Parameter 'Bot_RF_RevPwr', with Sample Size of 20 and soft-DTW Centroid Building Method with γ of 0.01

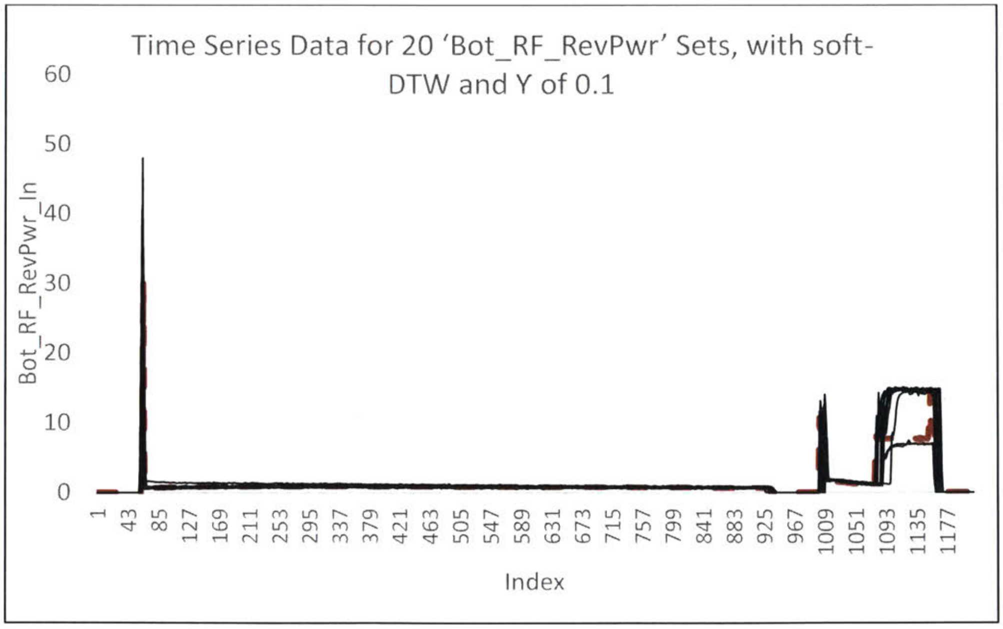


Figure 15. Time Series and Centroid for Parameter 'Bot_RF_RevPwr', with Sample Size of 20 and soft-DTW Centroid Building Method with γ of 0.1

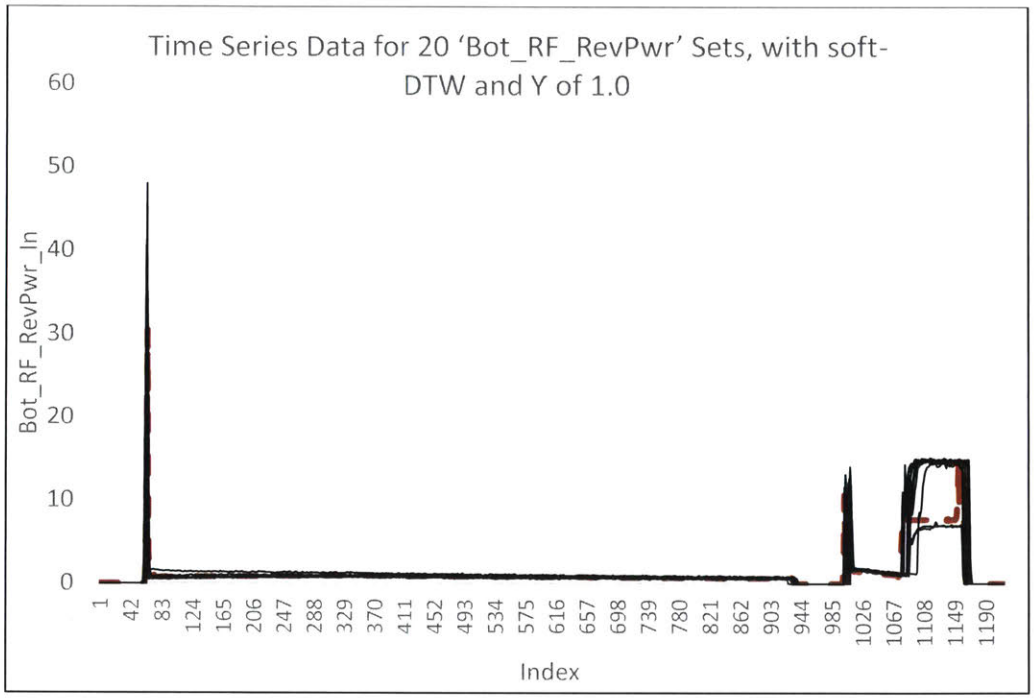


Figure 16. Time Series and Centroid for Parameter 'Bot_RF_RevPwr', with Sample Size of 20 and soft-DTW Centroid Building Method with γ of 1.0

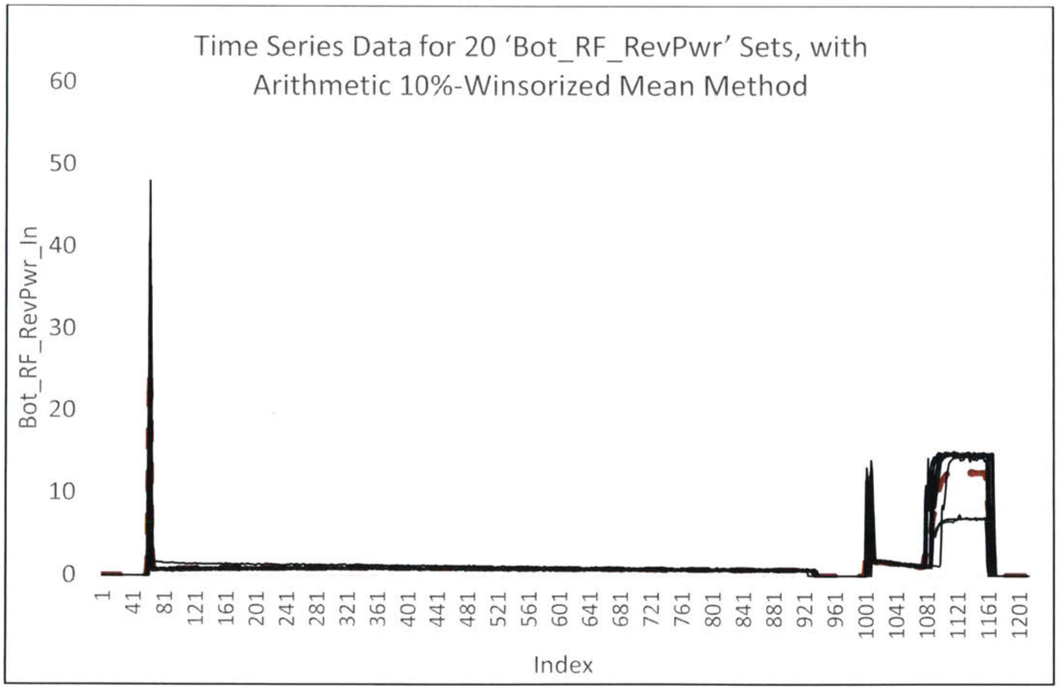


Figure 17. Time Series and Centroid for Parameter 'Bot_RF_RevPwr', with Sample Size of 20 and Arithmetic Winsorization Method

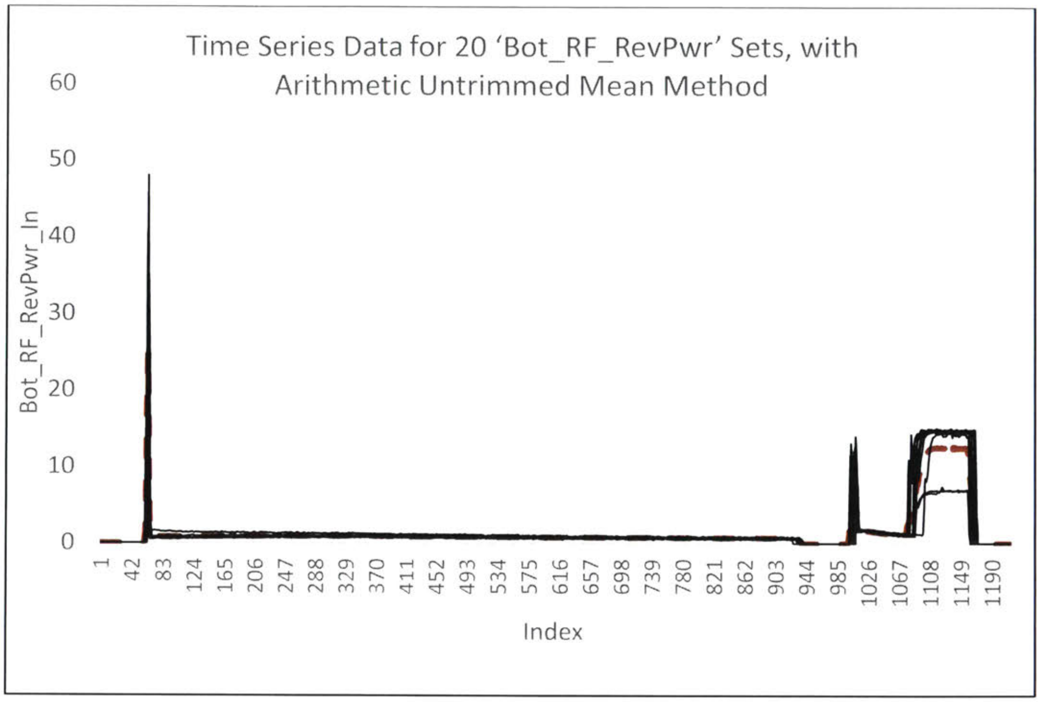


Figure 18. Time Series and Centroid for Parameter 'Bot_RF_RevPwr', with Sample Size of 20 and Arithmetic Untrimmed Mean Method

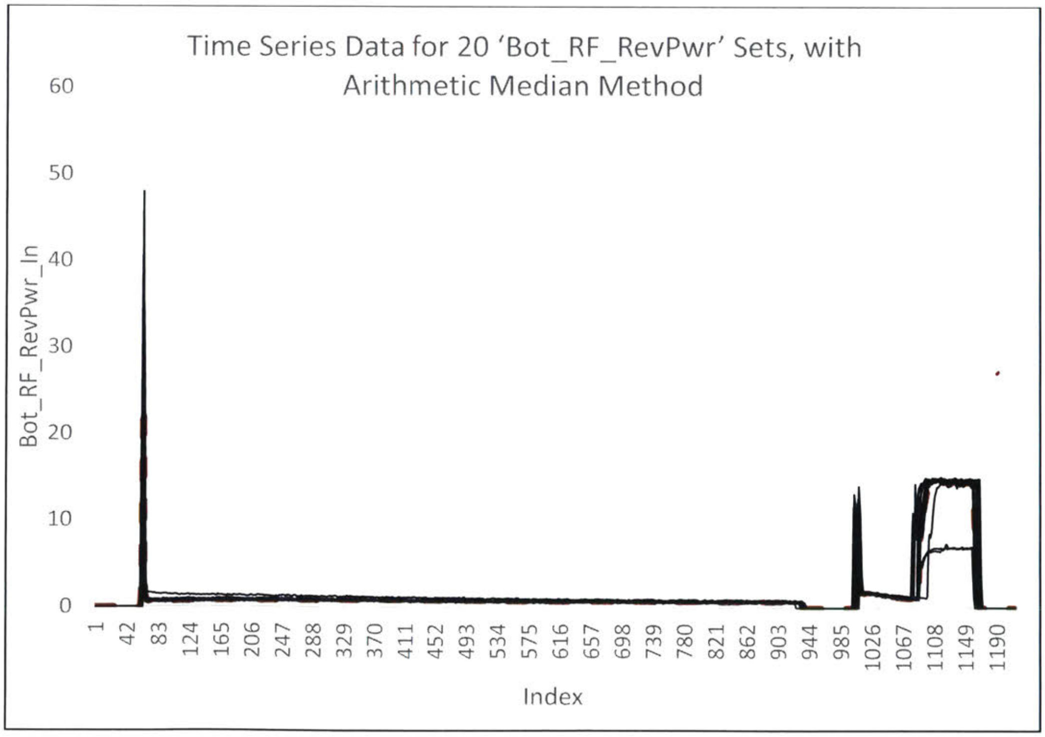
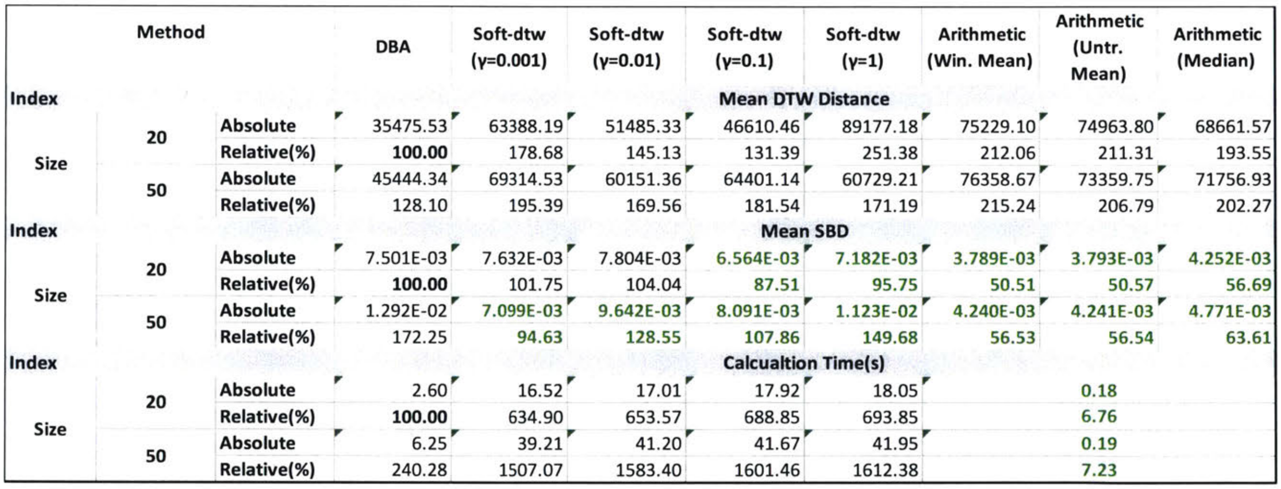


Figure 19. Time Series and Centroid for Parameter 'Bot_RF_RevPwr', with Sample Size of 20 and Arithmetic Median Method

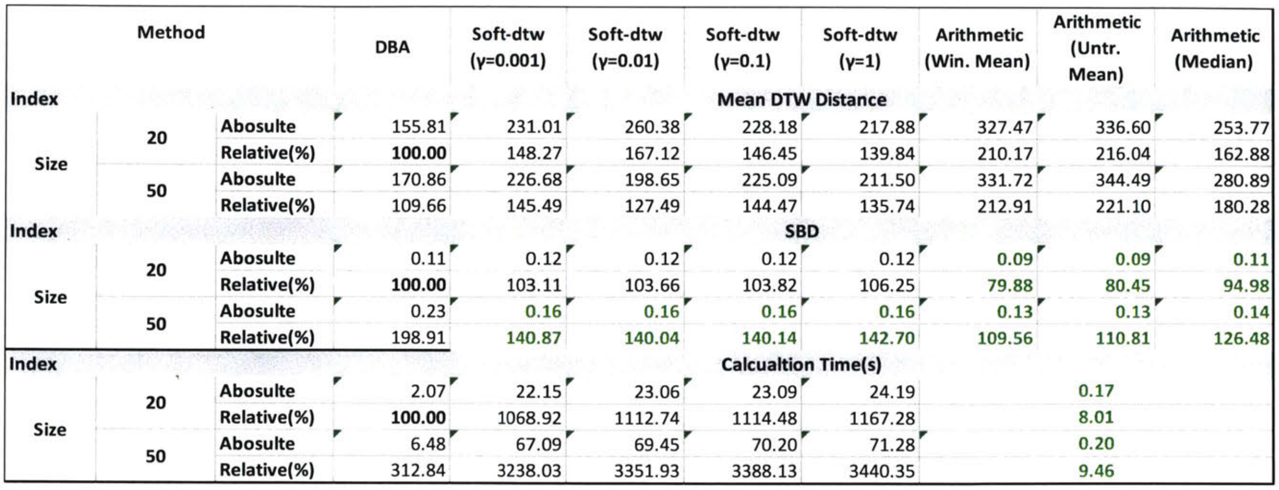
The performance for each method is summarized in Tables 1 and 2 for each parameter. The mean DTW distance is the average DTW distance between the centroid and the time series it is built upon, indicating the quantitative similarity in terms of the distance. The mean SBD is the average SBD between the centroid and the time series it is built upon, indicating the shape similarity. Calculation time is the average time each method spends on each sample size. Both absolute and relative values are listed, with relative percentage value compared to the value for the case where sample size is 20 and the centroid building method is the DBA. Detailed results are listed in Table A.1 and A.2 in the Appendix.

Table 1. Average Performance of Centroid Methods for the Parameter 'ProChm_EndPt_ChanC'



	Method		DBA	Soft-dtw (γ=0.001)	Soft-dtw (γ=0.01)	Soft-dtw (γ=0.1)	Soft-dtw (γ=1)	Arithmetic (Win. Mean)	Arithmetic (Untr. Mean)	Arithmetic (Median)
Index						Mean DTW Distance				
Size	20	Absolute	35475.53	63388.19	51485.33	46610.46	89177.18	75229.10	74963.80	68661.57
		Relative(%)	**100.00**	178.68	145.13	131.39	251.38	212.06	211.31	193.55
	50	Absolute	45444.34	69314.53	60151.36	64401.14	60729.21	76358.67	73359.75	71756.93
		Relative(%)	128.10	195.39	169.56	181.54	171.19	215.24	206.79	202.27
Index						Mean SBD				
Size	20	Absolute	7.501E-03	7.632E-03	7.804E-03	6.564E-03	7.182E-03	3.789E-03	3.793E-03	4.252E-03
		Relative(%)	**100.00**	101.75	104.04	87.51	95.75	50.51	50.57	56.69
	50	Absolute	1.292E-02	7.099E-03	9.642E-03	8.091E-03	1.123E-02	4.240E-03	4.241E-03	4.771E-03
		Relative(%)	172.25	94.63	128.55	107.86	149.68	56.53	56.54	63.61
Index						Calcualtion Time(s)				
Size	20	Absolute	2.60	16.52	17.01	17.92	18.05		0.18	
		Relative(%)	**100.00**	634.90	653.57	688.85	693.85		6.76	
	50	Absolute	6.25	39.21	41.20	41.67	41.95		0.19	
		Relative(%)	240.28	1507.07	1583.40	1601.46	1612.38		7.23	

Table 2. Average Performance of Centroid Methods for the Parameter 'Bot_RF_RevPwr'



Method			DBA	Soft-dtw (γ=0.001)	Soft-dtw (γ=0.01)	Soft-dtw (γ=0.1)	Soft-dtw (γ=1)	Arithmetic (Win. Mean)	Arithmetic (Untr. Mean)	Arithmetic (Median)
Index						Mean DTW Distance				
Size	20	Abosulte	155.81	231.01	260.38	228.18	217.88	327.47	336.60	253.77
		Relative(%)	**100.00**	148.27	167.12	146.45	139.84	210.17	216.04	162.88
	50	Abosulte	170.86	226.68	198.65	225.09	211.50	331.72	344.49	280.89
		Relative(%)	109.66	145.49	127.49	144.47	135.74	212.91	221.10	180.28
Index						SBD				
Size	20	Abosulte	0.11	0.12	0.12	0.12	0.12	0.09	0.09	0.11
		Relative(%)	**100.00**	103.11	103.66	103.82	106.25	79.88	80.45	94.98
	50	Abosulte	0.23	0.16	0.16	0.16	0.16	0.13	0.13	0.14
		Relative(%)	198.91	140.87	140.04	140.14	142.70	109.56	110.81	126.48
Index						Calcualtion Time(s)				
Size	20	Abosulte	2.07	22.15	23.06	23.09	24.19		0.17	
		Relative(%)	**100.00**	1068.92	1112.74	1114.48	1167.28		8.01	
	50	Abosulte	6.48	67.09	69.45	70.20	71.28		0.20	
		Relative(%)	312.84	3238.03	3351.93	3388.13	3440.35		9.46	

It can be seen clearly that the DBA method outperforms both the soft-DTW centroid and arithmetic methods in terms of the mean DTW distance. Although literature suggests that the soft-DTW could deliver smoother centroids [15], especially when the smooth parameter γ is large enough, the result from the experiment points out that the mean DTW distance is not improved by the soft-DTW method and the mean DTW distance does not seem to be related to the value of γ. The soft-DTW seldomly outperforms the DBA on the centroid accuracy in view of the mean DTW distance. It still outperforms the arithmetic centroid methods in terms of the mean DTW distance.

When judging the accuracy in terms of shape similarity by the mean SBD, all the three types of methods perform well and extract most features of the time series data. The soft-DTW slightly outperforms the DBA especially when the sample size has a higher value of 50, but the smoothing norm γ is not related to the mean SBD in all cases. It is interesting that the arithmetic methods have the best SBD performance among the three even though they are poor in terms of the mean DTW distance.

On top of that, it is interesting to see that among the three arithmetic centroid methods (winsorized mean, untrimmed mean and median), the median method always delivers the most accurate centroid, considering its lower mean DTW distance. Although it may not extract the most representative shape feature considering its higher SBD relative to the other two arithmetic methods, the median method overall is still the most accurate among the three arithmetic methods since it has substantial advantage in the mean DTW distance, which matters strongly for clustering and other machine anomaly detection methods. The arithmetic methods overall provides the best SBD, which is not quite meaningful since all centroids have extracted the shape accurately with SBD below 0.16, regardless of the building method.

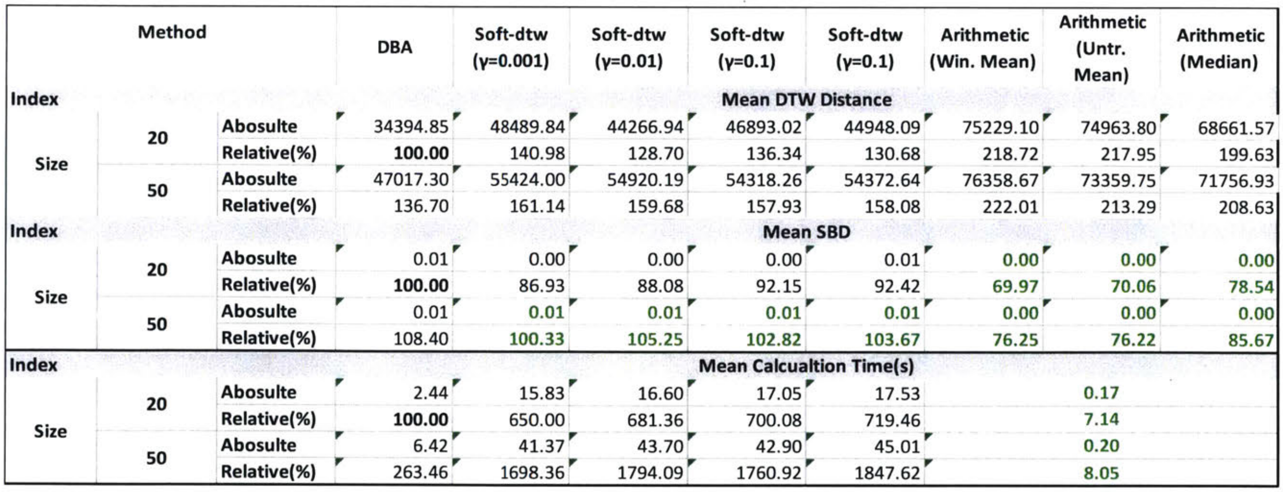
In terms of the calculation time, even when summing all three arithmetic centroid methods' time together, the combined time is still much lower than those of the DBA and the soft-DTW-based centroid methods in all cases. In each case, the time that all the three arithmetic methods take up is about 15% of the time that the DBA method spends. The soft-DTW-based centroid method, in contrast, always requires much longer calculation time. In the worst case, on average it takes 12x the time that the DBA method spends, in the case dealing with 20 samples of the parameter 'Bot_RF_RevPwer'. Similarly, the larger the group size is, the longer time each method spends and the larger the time differences among methods are.

The soft-DTW-based centroid behaves inaccurately in the experiment, in contrast to the results of Cuturi et al. [15]. Several factors contribute to the observed results. The method randomly selects one of the time series it is based upon as the initial centroid and updates the centroid only once later without iteration. Therefore the method cannot obtain centroids as optimized and converged as those derived from the DBA. Since the result has shown that the soft-DTW calculation is the most time consuming, iterations are not suggested since it could make the calculation time worse. An alternative to the iteration is to preset one centroid as the initial centroid rather than randomly choose one time series.

Two types of centroids are used and discussed for initial centroids, the arithmetic median and the DBA method. The arithmetic median method is used since it is fast to compute and it behaves the best among the three discussed arithmetic centroid methods. The DBA method is introduced since currently it is the most accurate method in terms of the mean DTW distance in the experiment. The soft-DTW-based centroid method will be evaluated to determine whether it outperforms the DBA method when the DBA centroid is preset as the initial centroid.

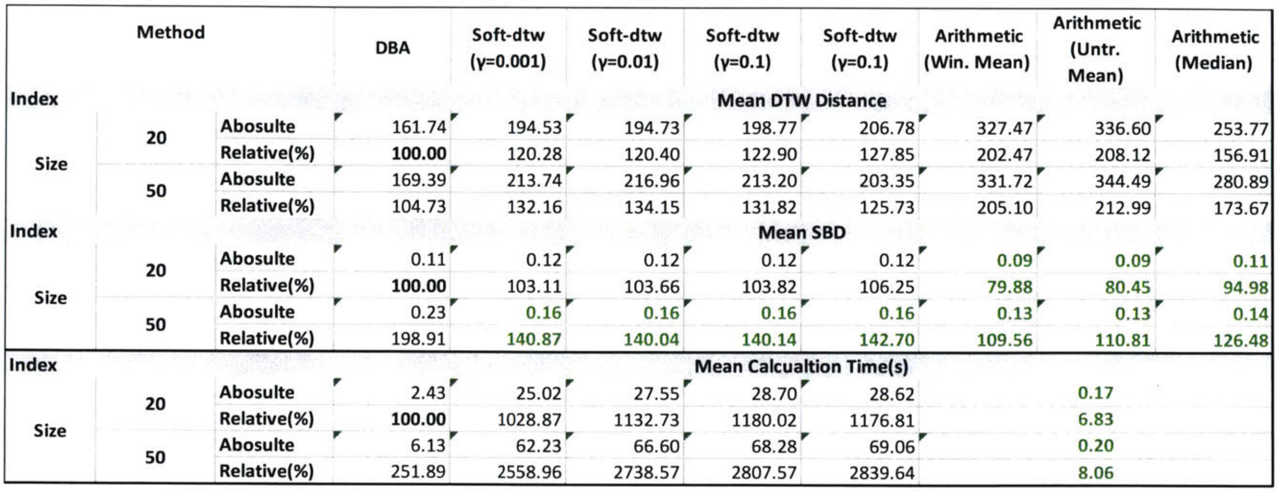
Results are listed in Table 3 through Table 6, with those preset with the arithmetic median methods presented first. Results are also listed in Table A.3 through A.6 in the Appendix for more details. The green texts show where the method outperforms the DBA method for the same parameter and the same sample size.

Table 3. Average Performance of Centroid Methods on the Parameter 'ProChm_EndPt_ChanC', with Arithmetic Median Centroid as Preset Centroid



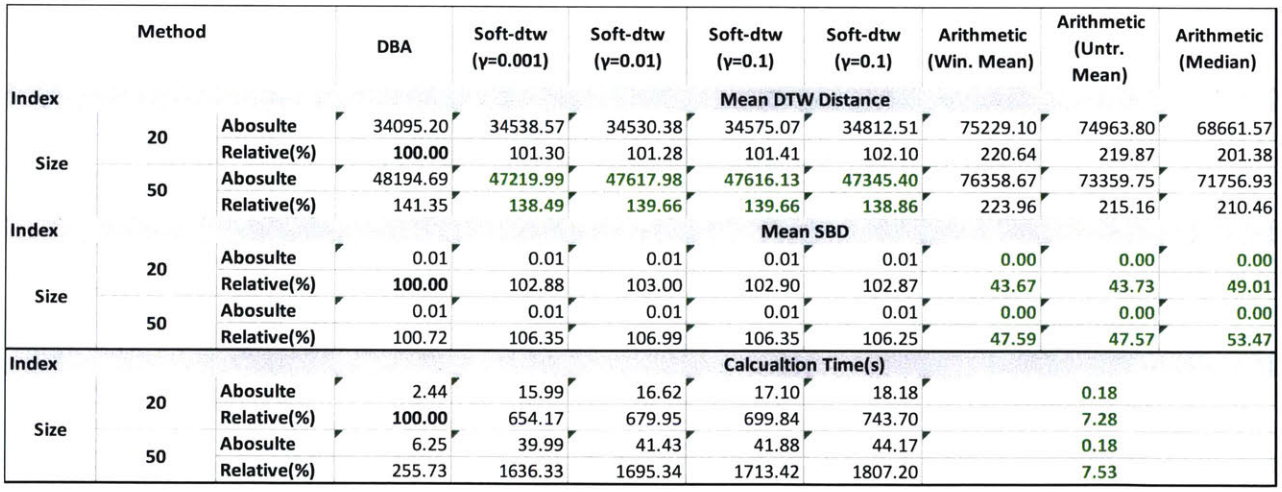
Method			DBA	Soft-dtw (γ=0.001)	Soft-dtw (γ=0.01)	Soft-dtw (γ=0.1)	Soft-dtw (γ=0.1)	Arithmetic (Win. Mean)	Arithmetic (Untr. Mean)	Arithmetic (Median)
Index						Mean DTW Distance				
Size	20	Abosulte	34394.85	48489.84	44266.94	46893.02	44948.09	75229.10	74963.80	68661.57
		Relative(%)	**100.00**	140.98	128.70	136.34	130.68	218.72	217.95	199.63
	50	Abosulte	47017.30	55424.00	54920.19	54318.26	54372.64	76358.67	73359.75	71756.93
		Relative(%)	136.70	161.14	159.68	157.93	158.08	222.01	213.29	208.63
Index						Mean SBD				
Size	20	Abosulte	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00
		Relative(%)	**100.00**	86.93	88.08	92.15	92.42	69.97	70.06	78.54
	50	Abosulte	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00
		Relative(%)	108.40	100.33	105.25	102.82	103.67	76.25	76.22	85.67
Index						Mean Calcualtion Time(s)				
Size	20	Abosulte	2.44	15.83	16.60	17.05	17.53		0.17	
		Relative(%)	**100.00**	650.00	681.36	700.08	719.46		7.14	
	50	Abosulte	6.42	41.37	43.70	42.90	45.01		0.20	
		Relative(%)	263.46	1698.36	1794.09	1760.92	1847.62		8.05	

Table 4. Average Performance of Centroid Methods on the Parameter ‘Bot_RF_RevPwr’, with Arithmetic Median Centroid as Preset Centroid



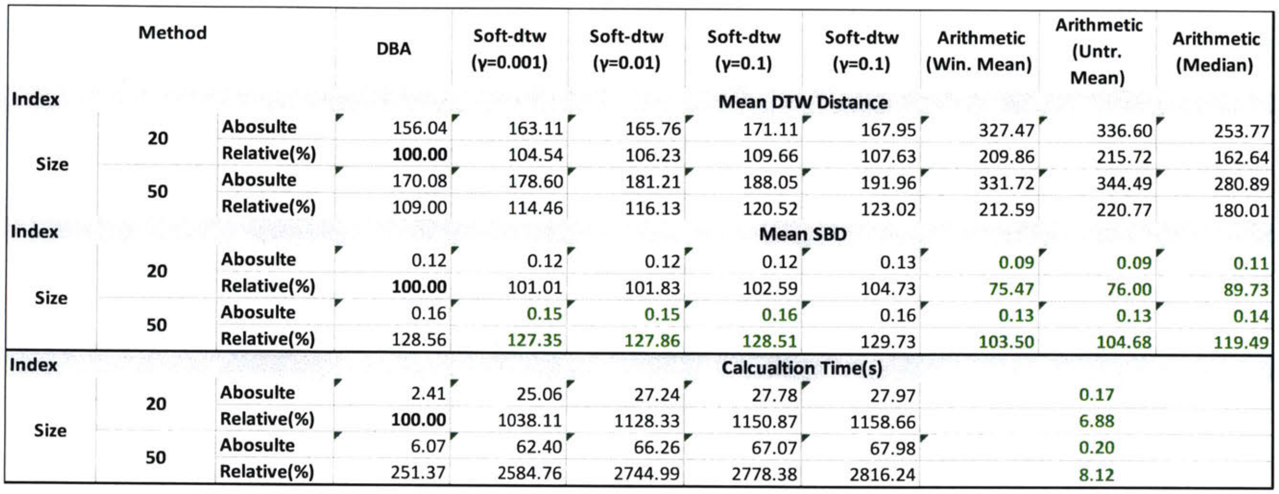
Method			DBA	Soft-dtw (γ=0.001)	Soft-dtw (γ=0.01)	Soft-dtw (γ=0.1)	Soft-dtw (γ=0.1)	Arithmetic (Win. Mean)	Arithmetic (Untr. Mean)	Arithmetic (Median)
Index			Mean DTW Distance							
Size	20	Abosulte	161.74	194.53	194.73	198.77	206.78	327.47	336.60	253.77
		Relative(%)	**100.00**	120.28	120.40	122.90	127.85	202.47	208.12	156.91
	50	Abosulte	169.39	213.74	216.96	213.20	203.35	331.72	344.49	280.89
		Relative(%)	104.73	132.16	134.15	131.82	125.73	205.10	212.99	173.67
Index			Mean SBD							
Size	20	Abosulte	0.11	0.12	0.12	0.12	0.12	0.09	0.09	0.11
		Relative(%)	**100.00**	103.11	103.66	103.82	106.25	79.88	80.45	94.98
	50	Abosulte	0.23	0.16	0.16	0.16	0.16	0.13	0.13	0.14
		Relative(%)	198.91	140.87	140.04	140.14	142.70	109.56	110.81	126.48
Index			Mean Calcualtion Time(s)							
Size	20	Abosulte	2.43	25.02	27.55	28.70	28.62	0.17		
		Relative(%)	**100.00**	1028.87	1132.73	1180.02	1176.81	6.83		
	50	Abosulte	6.13	62.23	66.60	68.28	69.06	0.20		
		Relative(%)	251.89	2558.96	2738.57	2807.57	2839.64	8.06		

Table 5. Average Performance of Centroid Methods on the Parameter ‘ProChm_EndPt_ChanC’, with the DBA as Preset Centroid



Method			DBA	Soft-dtw (γ=0.001)	Soft-dtw (γ=0.01)	Soft-dtw (γ=0.1)	Soft-dtw (γ=0.1)	Arithmetic (Win. Mean)	Arithmetic (Untr. Mean)	Arithmetic (Median)
Index			Mean DTW Distance							
Size	20	Abosulte	34095.20	34538.57	34530.38	34575.07	34812.51	75229.10	74963.80	68661.57
		Relative(%)	**100.00**	101.30	101.28	101.41	102.10	220.64	219.87	201.38
	50	Abosulte	48194.69	47219.99	47617.98	47616.13	47345.40	76358.67	73359.75	71756.93
		Relative(%)	141.35	138.49	139.66	139.66	138.86	223.96	215.16	210.46
Index			Mean SBD							
Size	20	Abosulte	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00
		Relative(%)	**100.00**	102.88	103.00	102.90	102.87	43.67	43.73	49.01
	50	Abosulte	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00
		Relative(%)	100.72	106.35	106.99	106.35	106.25	47.59	47.57	53.47
Index			Calcualtion Time(s)							
Size	20	Abosulte	2.44	15.99	16.62	17.10	18.18	0.18		
		Relative(%)	**100.00**	654.17	679.95	699.84	743.70	7.28		
	50	Abosulte	6.25	39.99	41.43	41.88	44.17	0.18		
		Relative(%)	255.73	1636.33	1695.34	1713.42	1807.20	7.53		

Table 6. Average Performance of Centroid Methods on the Parameter ‘Bot_RF_RevPwr’, with Arithmetic the DBA as Preset Centroid



Method			DBA	Soft-dtw (γ=0.001)	Soft-dtw (γ=0.01)	Soft-dtw (γ=0.1)	Soft-dtw (γ=0.1)	Arithmetic (Win. Mean)	Arithmetic (Untr. Mean)	Arithmetic (Median)
Index						Mean DTW Distance				
Size	20	Abosulte	156.04	163.11	165.76	171.11	167.95	327.47	336.60	253.77
		Relative(%)	**100.00**	104.54	106.23	109.66	107.63	209.86	215.72	162.64
	50	Abosulte	170.08	178.60	181.21	188.05	191.96	331.72	344.49	280.89
		Relative(%)	109.00	114.46	116.13	120.52	123.02	212.59	220.77	180.01
Index						Mean SBD				
Size	20	Abosulte	0.12	0.12	0.12	0.12	0.13	0.09	0.09	0.11
		Relative(%)	**100.00**	101.01	101.83	102.59	104.73	75.47	76.00	89.73
	50	Abosulte	0.16	0.15	0.15	0.16	0.16	0.13	0.13	0.14
		Relative(%)	128.56	127.35	127.86	128.51	129.73	103.50	104.68	119.49
Index						Calcualtion Time(s)				
Size	20	Abosulte	2.41	25.06	27.24	27.78	27.97		0.17	
		Relative(%)	**100.00**	1038.11	1128.33	1150.87	1158.66		6.88	
	50	Abosulte	6.07	62.40	66.26	67.07	67.98		0.20	
		Relative(%)	251.37	2584.76	2744.99	2778.38	2816.24		8.12	

Initially improvements with the use of preset centroid is checked. Figure 20 and 21 indicate the improvement of mean DTW distance and calculation time. In the abbreviated form, the legend in each figure shows the following information: ‘P5’ and ‘P19’ indicates the parameter name, with the ‘P5’ for the ‘Bot_RF_RevPwr’ and the ‘P19’ for the ‘ProChm_EndPt_ChanC’; ‘20/50’ indicates the sample size; ‘Median’ indicates the preset centroid is built on arithmetic median method, and ‘DBA’ indicates it is built on the DBA.

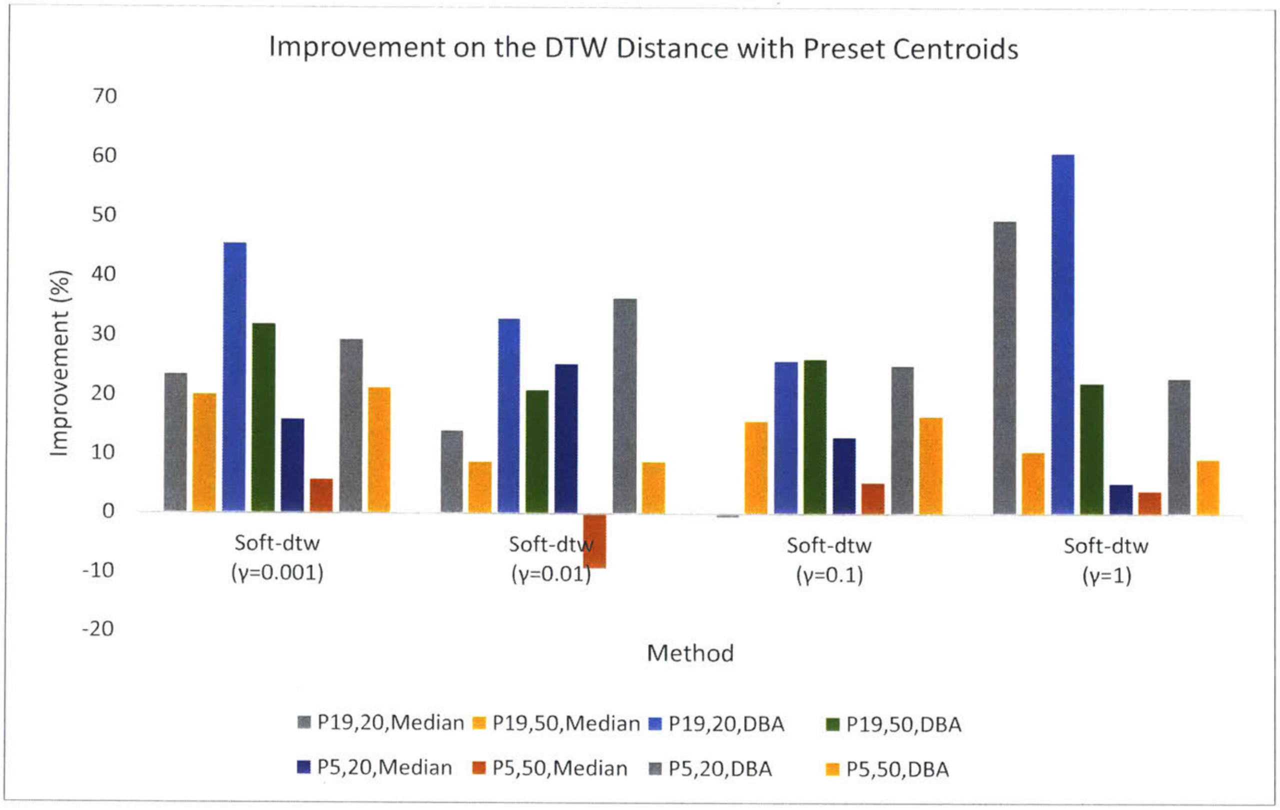


Figure 20. Improvement on the DTW Distance with Preset Centroids for Soft-DTW

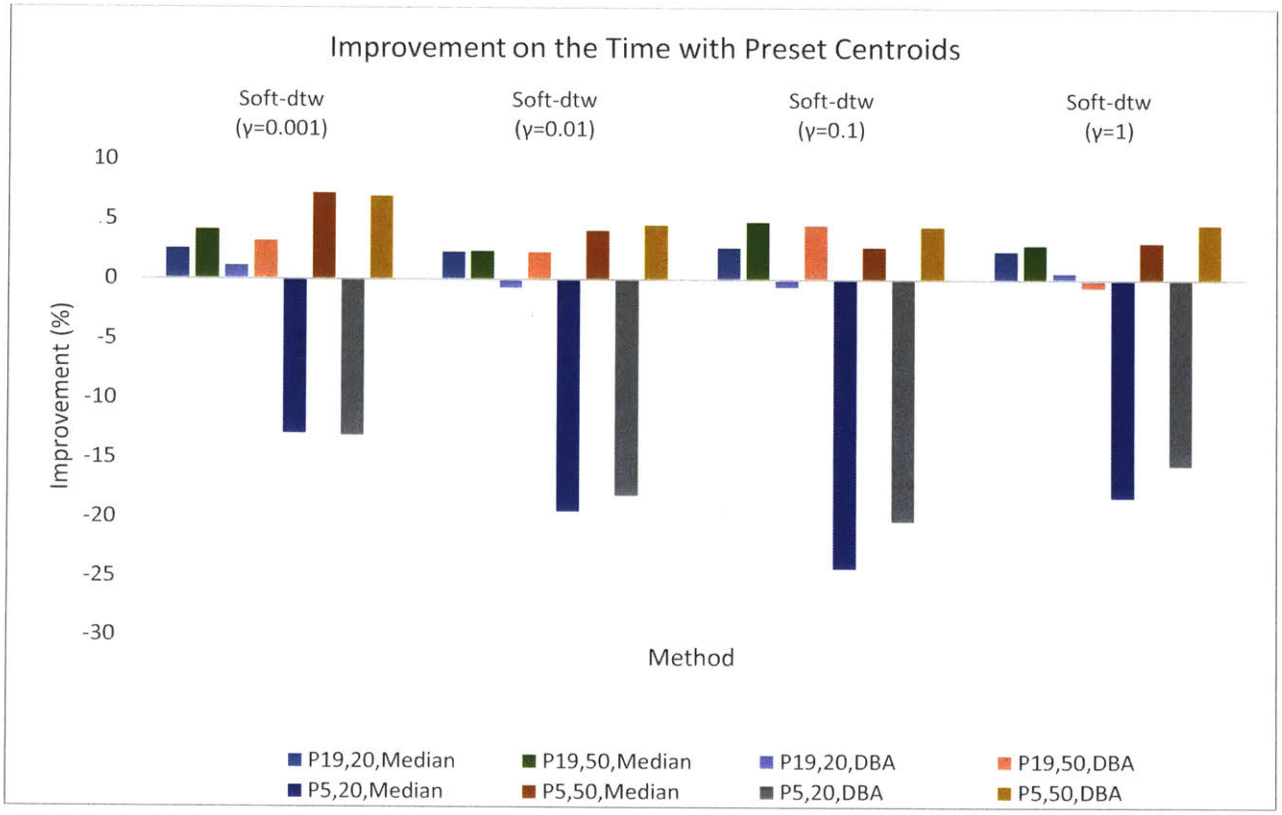


Figure 21. Improvement on the Time with Preset Centroids for Soft-DTW

In most cases, the introduction of a preset centroid helps improve the soft-DTW-based centroid method's quality, in terms of both the mean DTW distance and the calculation time. Overall the DBA preset centroid delivers shorter DTW distances than the preset arithmetic median centroid, thus making the soft-DTW-based centroid more accurate. It makes the soft-DTW method outperform the DBA method in terms of the mean DTW distance for sample size of 50 in both parameters, except for the parameter 'Bot_RF_RevPwr' with γ of 1, as shown in Table 5 and 6. While the preset centroid generally shortens soft-DTW-based centroid mehtod's calculation time by around 5%, the total calculation time is still quite long compared to that of the DBA. The extreme time elongation for the case of P5 and sample size of 20 are related to the overloaded computer used at that time and hints at the performance instability of the preset method for small sample sizes. Overall, the mean DTW improvement is within 5%, which is quite limited and expensive considering the much longer calculation time that the soft-DTW method needs.

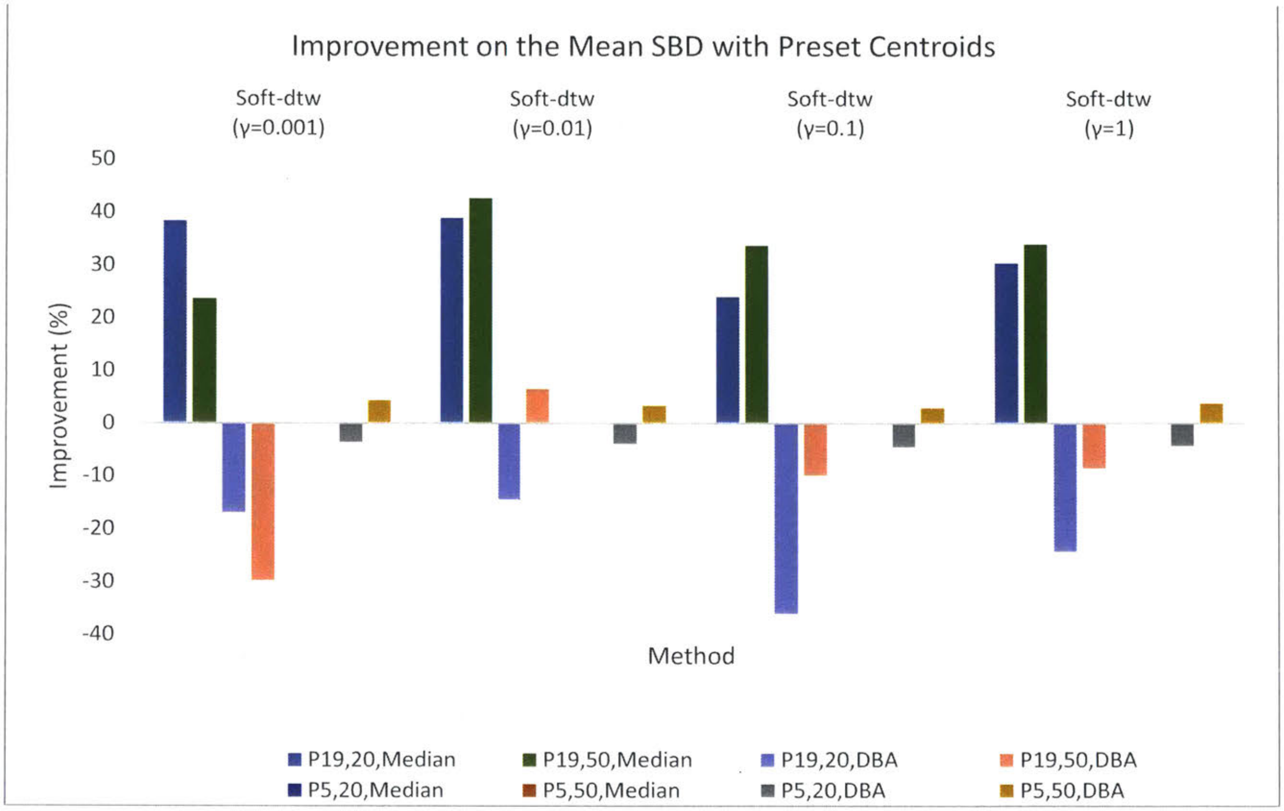


Figure 22. Improvement on the Mean SBD with Preset Centroids for Soft-DTW

In addition, whether the preset centroid helps improve the mean SBD for the soft-DTW-based centroid is tested, with results shown in Figure 22. It shows no strong relation between the improvement in the SBD and the preset method used.

As a result, the soft-DTW-based centroid does not outperform other centroid building methods for the plasma etching process considered in the thesis in a meaningful way. In view of the advantageous mean DTW distance using the DBA, ways to further upgrade the DBA performance and reduce its calculation time are further discussed in the following section.

4.2 Improvements on the DBA Method

In Section 4.1 the DBA method is found to be the best behaved centroid building method after comprehensive evaluations. Further improvements are introduced in this section to further improve it. Section 4.2.1 discusses the possibility of applying preset arithmetic median method to optimize its accuracy and calculation time. Section 4.2.2 discusses whether the window constraint reduces calculation time substantially with acceptable losses on accuracy. Finally, Section 4.2.3 considers the situation where new time series are added into the sample and discusses whether the innovative recalculation method by combining previous centroids and new series can reduce recalculation time without great losses on accuracy.

4.2.1 Preset Arithmetic Median Centroid

The preset arithmetic median centroid can be one of the possible solutions to improve the DBA method. Since the arithmetic median method uses much less calculation time than the DBA method, it could be a meaningful improvement if the preset centroid effectively reduces the DBA centroid's mean DTW distance from the time series it builds upon, while requiring negligible extra calculation time. Therefore related tests are made similar to what has been done on the soft-DTW-based methods, and the results are shown in Figures 23 and 24.

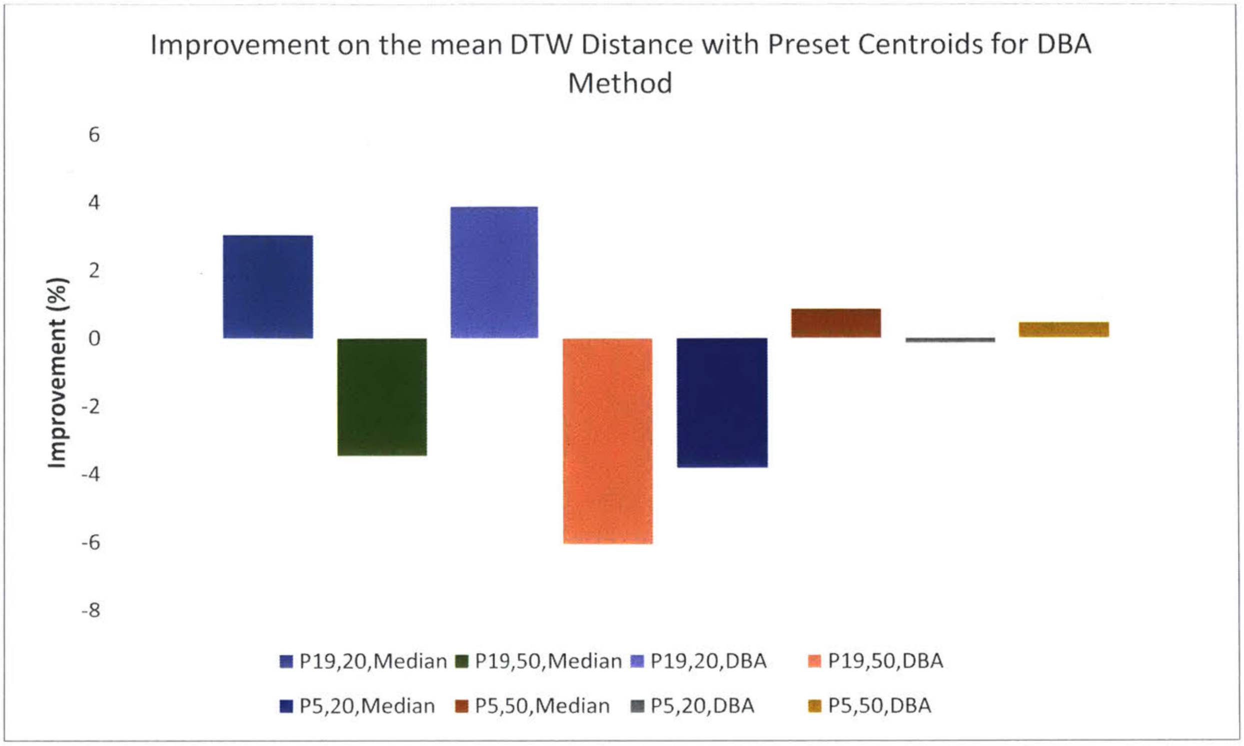


Figure 23. Improvement on the DTW Distance with Preset Centroids for the DBA Method

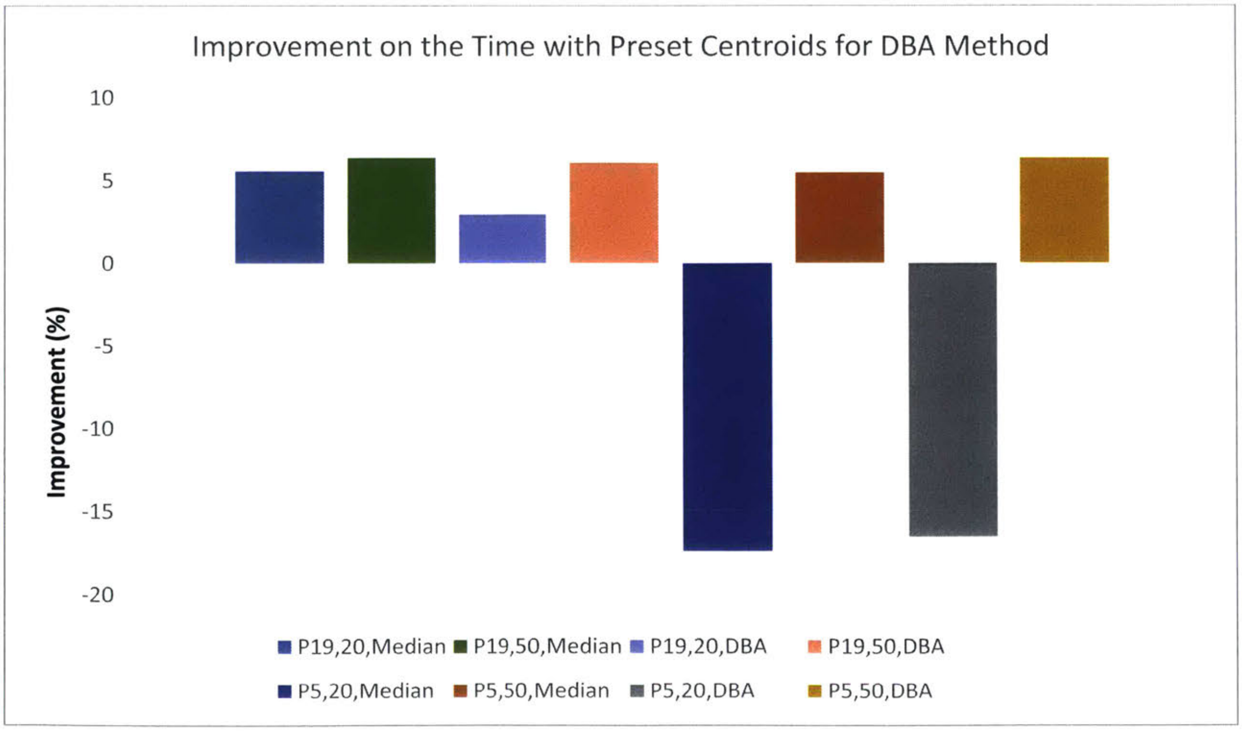


Figure 24. Improvement on the Time with Preset Centroids for the DBA Method

Results point out that there is no strong correlation between improvement on the mean DTW distance and a preset arithmetic median centroid. Overall, the preset median centroid reduces the total calculation time that the DBA method spends slightly (less than 5%). The two negative improvement data points in Figure 24 are not caused by the preset centroid. Similar to the longer time of soft-DTW calculation shown in Figure 22, the longer calculation time is related to the DBA calculation delays because of the overloaded computer. With limited improvement in calculation time and uncertain influence on the mean DTW distance, a preset arithmetic median centroid does not bring significant improvement on the DBA performance.

4.2.2 Window Constraint

As discussed in Section 2.1, the DTW creates a warping path between the query and inference time series to minimize the distance. Without any constraint, an *m*-by-*n* distance matrix is created for the two time series with lengths of *m* and *n* respectively and then an optimized path is created considering all the possibilities of matching the two time series. However, this creates high computational complexity. However, in most cases, not all elements in the distance matrix need to be used. In this way a global DTW warping constraint, the window constraint, is introduced that constrains the warping path to be considered for a DTW. Figure 25 is an example of the window-constraint DTW. The warping path does not consider red elements and the entire warping path is within $[(i, j-w), (i, j+w)]$ for all (i,j) points along the diagonal. Here w is the window size, normally set as 10% of the series length. Sometimes a smaller window size even produces a better result [22].

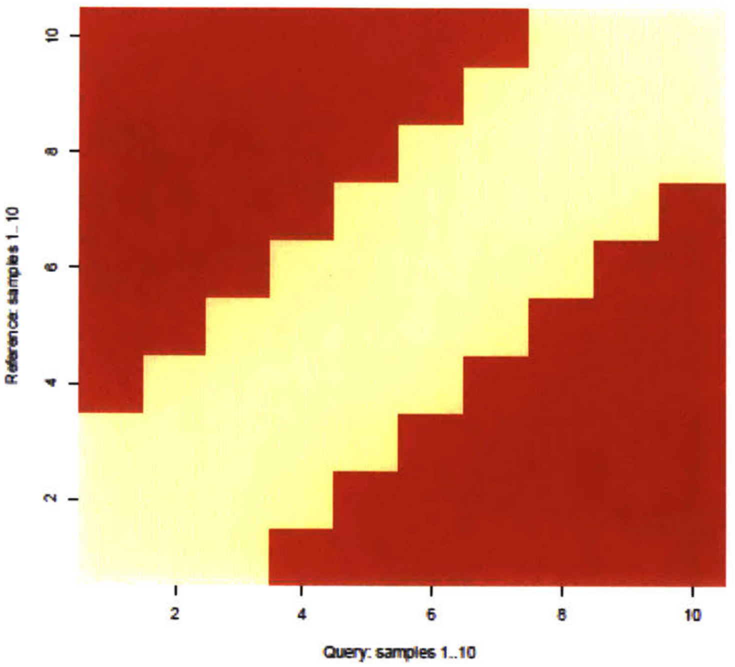


Figure 25. The Window-Constrained Warping Path, with Red Elements not Considered in the DTW [7]

A test of the effect of the window size on the DBA centroid's mean DTW distance and calculation time is made on time series of the parameter 'Bot_RF_RevPwr', with sample sizes of 20 and 50 and window sizes of 5%, 10%, 20% and 30% of the series lengths. The results are compared with those from the unconstrained DBA, as shown in Figure 26 and 27. The original data is listed in Appendix Table A.7, and the average performance is shown in Table 7. Again, the green texts indicate that the method outperforms the unconstrained DBA method.

It is seen that overall with a 20% window size of the series length, calculation time is reduced by over 15%, and the increase in the mean DTW distance is within 5%. The reduction on the centroid accuracy is small compared with the significant benefit of calculation time savings. Therefore a choice of 20% window size is the suggested setting for calculation time saving.

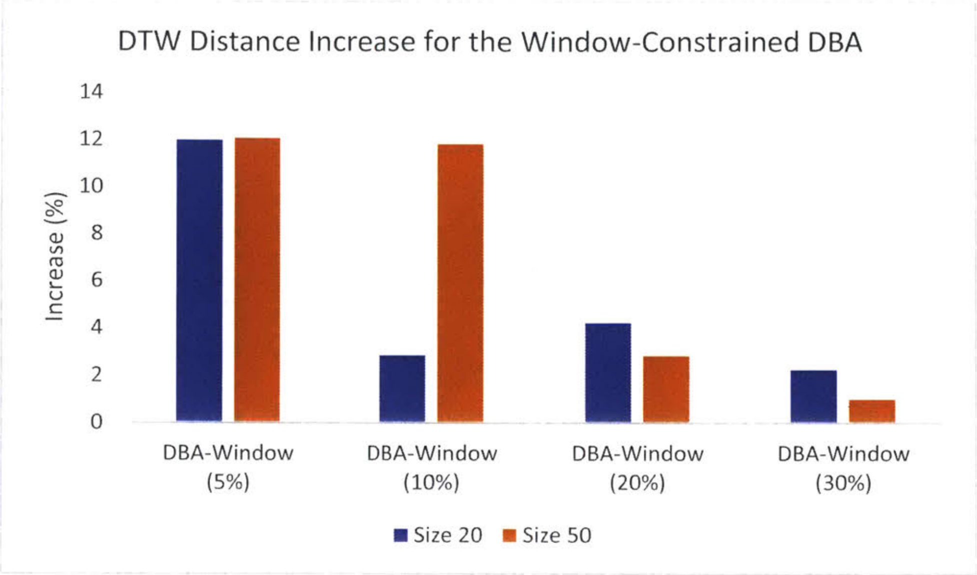


Figure 26. Mean DTW Distance Increase for the Window-Constrained DBA

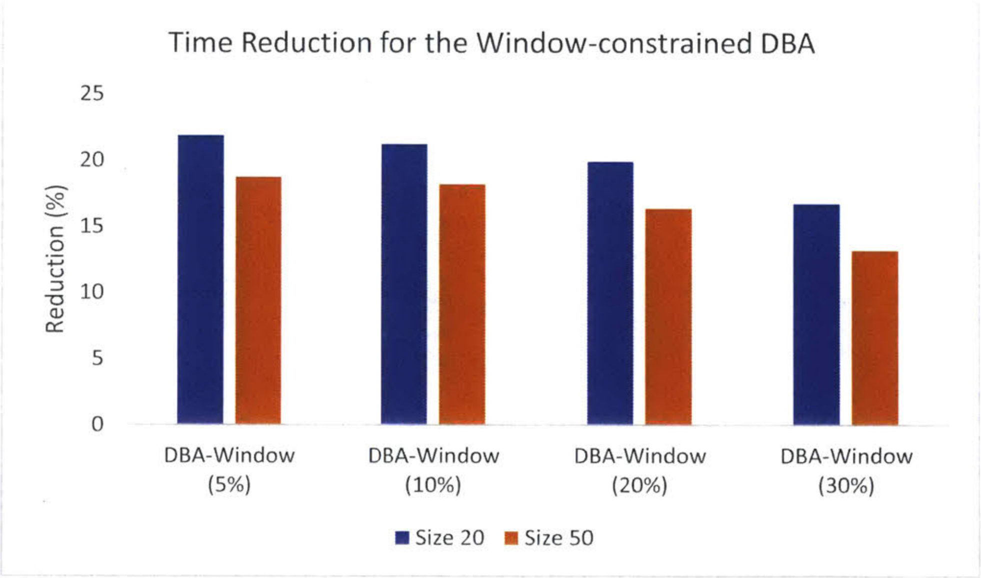
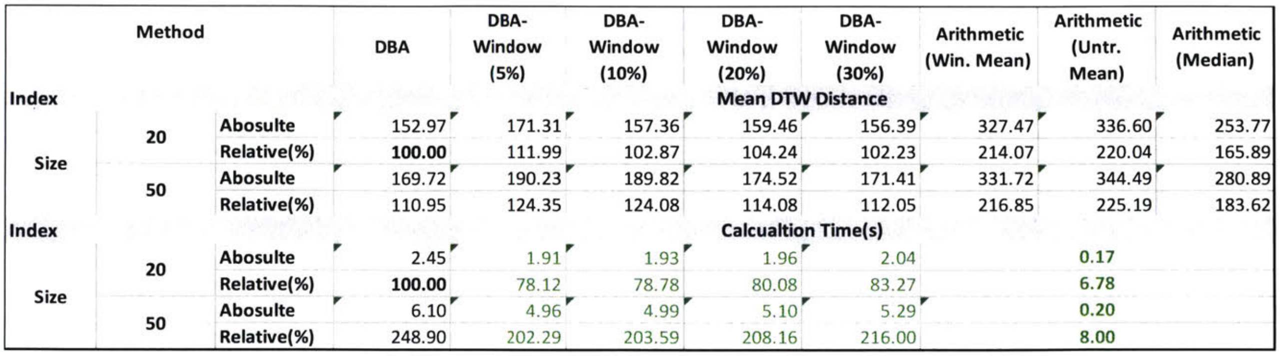


Figure 27. Time Reduction for the Window-constrained DBA

Table 7. Average Performance of Window-constrained DBA on the Parameter 'Bot_RF_RevPwr'



	Method		DBA	DBA-Window (5%)	DBA-Window (10%)	DBA-Window (20%)	DBA-Window (30%)	Arithmetic (Win. Mean)	Arithmetic (Untr. Mean)	Arithmetic (Median)
Index						Mean DTW Distance				
Size	20	Abosulte	152.97	171.31	157.36	159.46	156.39	327.47	336.60	253.77
		Relative(%)	**100.00**	111.99	102.87	104.24	102.23	214.07	220.04	165.89
	50	Abosulte	169.72	190.23	189.82	174.52	171.41	331.72	344.49	280.89
		Relative(%)	110.95	124.35	124.08	114.08	112.05	216.85	225.19	183.62
Index						Calcualtion Time(s)				
Size	20	Abosulte	2.45	1.91	1.93	1.96	2.04		**0.17**	
		Relative(%)	**100.00**	78.12	78.78	80.08	83.27		**6.78**	
	50	Abosulte	6.10	4.96	4.99	5.10	5.29		**0.20**	
		Relative(%)	248.90	202.29	203.59	208.16	216.00		**8.00**	

4.2.3 Recalculation with New Data

During an ongoing manufacturing process, it is normal to add new time series data and seek to update centroids. Since the centroid building method is computationally expensive, it is worth checking the feasibility of updating the centroid by combining the previous centroid and the new time series data.

Given a set of randomly picked 200 time series datasets for the parameter 'Bot_RF_RevPwr', 150 datasets randomly picked and classified as the "previous datasets" and others as the "new datasets." A centroid is initially calculated based upon all 200 individual sets of data. The method is defined as the "complete recalculation." Then the centroid of these 200 datasets is calculated by a combination of centroid and individual datasets, where the centroid is calculated based upon the previous datasets and individual datasets are those 50 new datasets. Each individual dataset is given a weight of 1 and the centroid is given a weight equal to the quantity of datasets it is built up with. This method is defined as the "reconstructed calculation."

Table 8. Comparison between Recalculation Methods

Method	Complete Recalculation	Reconstructed Calculation
Mean DTW	253.57	240.59
Time (s)	19.83	11.97

The average result is shown in Table 8 with comparisons of mean DTW distance between the centroid and the individual time series, as well as the calculation time. On average, the calculation time is reduced by 39.62% and the mean DTW distance is also reduced by 5.12%. The result points out that the reconstructed calculation method not only reduces calculation time substantially, but also improves accuracy and representativeness of the centroid. Therefore, when adding new time series data, the reconstructed calculation method that uses weighted previous centroid is preferred.

4.3 Conclusion on Centroid Building

Based upon complete comparisons among the arithmetic, DBA, and soft-DTW-based centroid building methods, the DBA method is believed to be the most appropriate for the plasma etching case in the thesis. It keeps the mean DTW distance much lower than the arithmetic methods, beneficial to clustering and other operations. It has appropriate and acceptable calculation time compared with the soft-DTW-based centroid building method. Although the soft-DTW centroid is claimed to be smoother with better shape quality, the mean SBD result indicates that it is not obviously better than the DBA. Preset centroids cannot help it improve effectively, either. With the help of window size (preferably 20% of the series length) and efficient recalculation method combining the latest centroid and new datasets, the DBA's efficiency is further improved, and thus the DBA method is suggested for use in building centroids for both online and offline references.

4.4 k-shape Clustering

k-shape clustering is applied to the plasma etching case and compared with the currently mainstream partitional clustering method, k-means. They have similar algorithms: both iteratively calculate the distance, assign the time series to the closet centroid and then recalculate the centroid based upon the time series it is assigned to. The difference is that the k-shape uses the SBD and the shape extraction to calculate distances and build centroids, while the k-means uses the DTW distance and the arithmetic mean method instead. The data is a mixture of 341 'ProChm_EndPt_ChanC' time series data, as shown in Figure 28. The good series are listed in Figure 3 in Chapter 1. The k-shape clustering is tested whether it can extract all of the 60 good series effectively and efficiently.

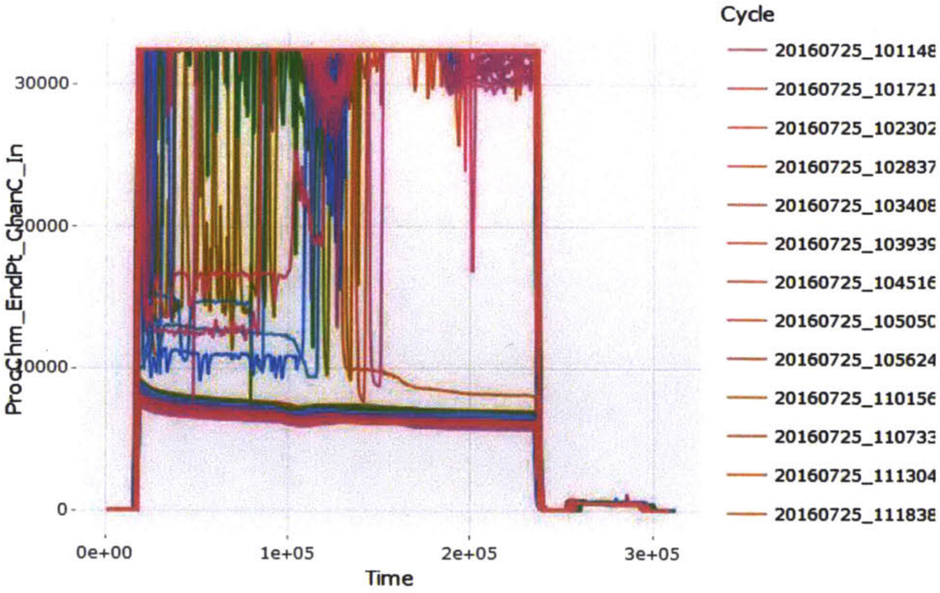


Figure 28. Mixture of 341 'ProChm_EndPt_ChanC' Time Series Data

Too few or too many clusters cannot distinguish the time series homogenously and well- separated. Since a preset number of clusters is required for the partitional method, an equation is added into the code to automatically optimize the quantity of the cluster and make the entire clustering algorithm more automatic. Within the clustering method, the summed square SBD between the time series and their centroids keeps dropping with the increased quantity of clusters. At some time the summed square SBD will be stable within a limited range. When first approaching this range, the second derivative of the summed square SBD with respect to the quantity of clusters turns positive and this "elbow point" is chosen as the optimized cluster quantity. The elbow point is marked in Figure 29 and the optimized quantity of clusters should be 4 for this case. The derivatives are also shown in Table 9.

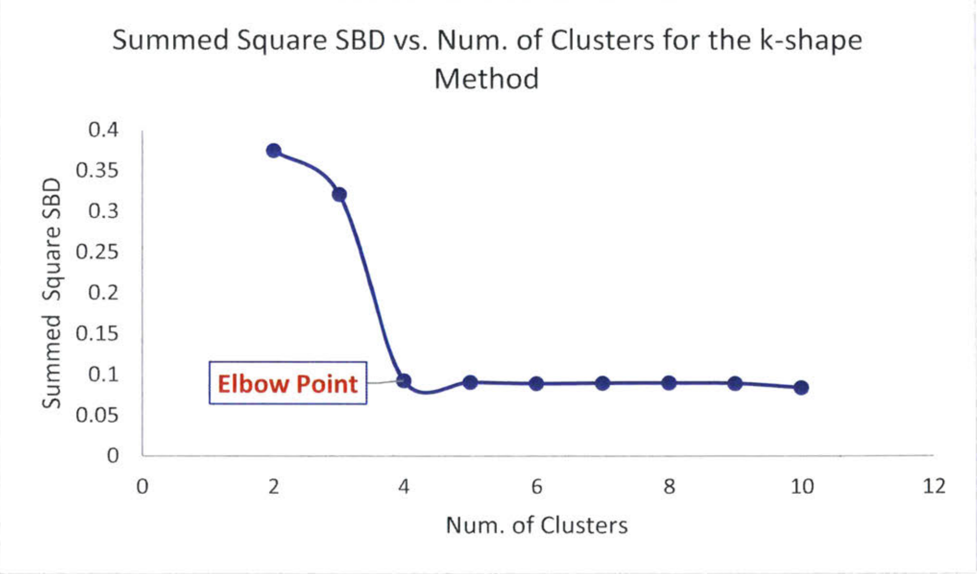


Figure 29. Summed Square SBD vs. Quantity of Clusters for the k-shape Clustering

Table 9. Derivatives of the Summed Square SBD vs. Quantity of Clusters for the k-shape Clustering

Num. of Clusters	2	3	4	5	6
Summed Square SBD	0.375	0.321	0.093	0.091	0.089
1st Derivative	-0.054	-0.141	-0.115	-0.002	N/A
2nd Derivative	N/A	-0.031	0.070	N/A	N/A

The operation is performed from 2 to 6 clusters to identity the elbow point, with a total calculation time of 523.63s. The result is shown in Figure 30, where the vertical axis unit is the normalized unit since the datasets need to be z-normalized before clustering. Black dashed and red thinned lines are the centroids for each cluster. The cluster sizes are 60, 100, 173 and 8, respectively.

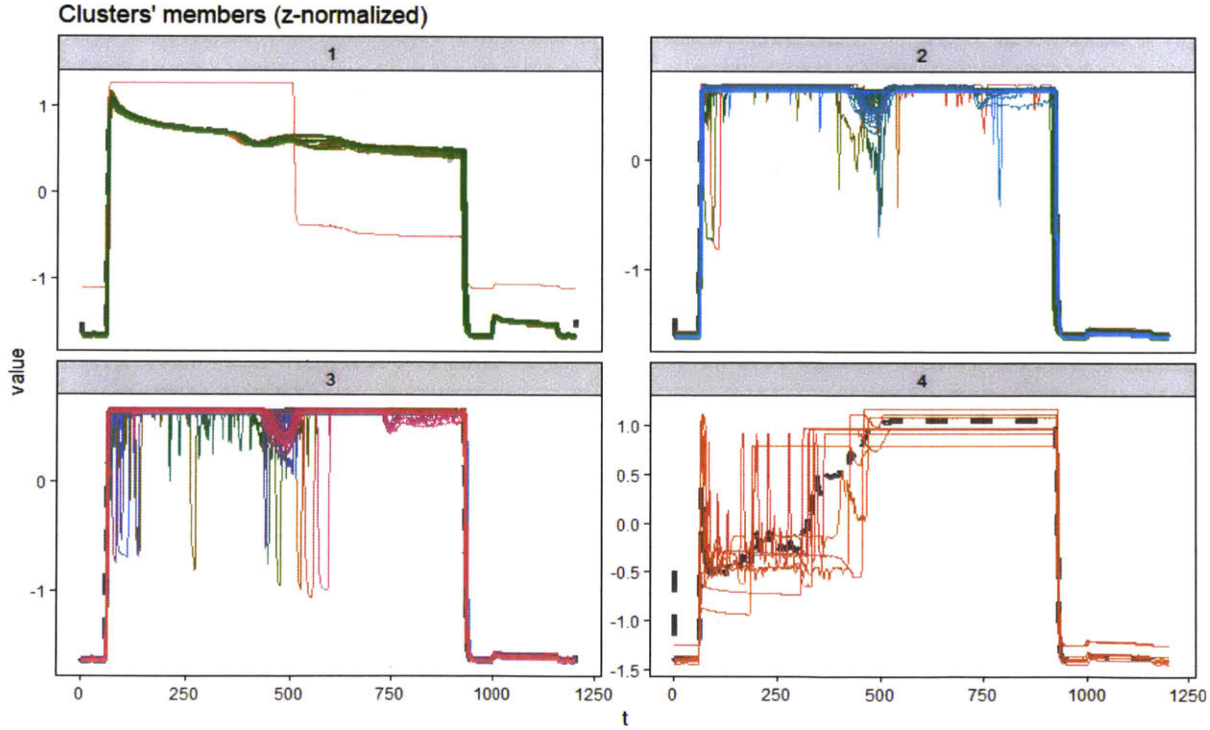


Figure 30. Result of the 4-cluster k-shape Clustering on the 341 'ProChm_EndPt_ChanC' Time Series

The normal shape is shown in Figure 3 in Chapter 1. It is seen that clearly all 60 normal cycles are clustered into the first cluster, and others are filtered into three other clusters. In comparison, the k-means clustering is applied. However, due to the extremely long process time, the test is terminated with only the results of 2 and 3 clusters. The 2-cluster clustering requires 143.97s and 3-cluster clustering requires as long as 7065.41s. The 3-cluster works better, with 59 normal cycles distinguished, and results are shown in Figure 31. Quantitative comparisons are shown in Tables 10 through 12.

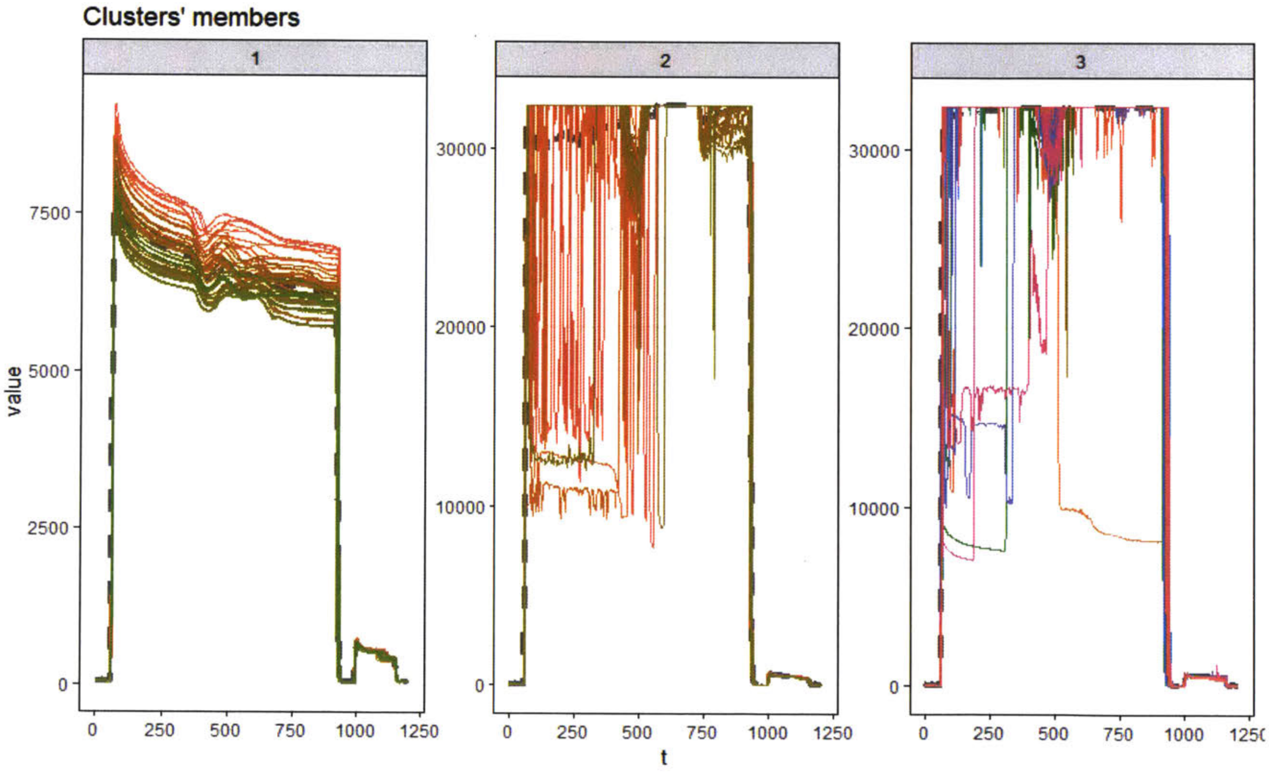


Figure 31. Result of the 3-cluster k-means Clustering on the 341 'ProChm_EndPt_ChanC' Time Series

Table 10. Confusion Matrix of k-shape on the 341 'ProChm_EndPt_ChanC' Time Series

		Actual	
		Anomaly	Non-anomaly
Predicted	Anomaly	60	0
	Non-anomaly	0	281

Table 11. Confusion Matrix of 3-Cluster k-means on the 341 'ProChm_EndPt_ChanC' Time Series

		Actual	
		Anomaly	Non-anomaly
Predicted	Anomaly	59	0
	Non-anomaly	1	281

Table 12. Comparison of k-shape and k-means Performance on the 341 'ProChm_EndPt_ChanC' Time Series

Method	k-shape	k-means
Precision	1	1
Recall	1	0.983
F_1	1	0.992
Calculation Time(s)	523.63	>7209.38

Results show that the k-shape clustering method gets more accurate results with substantially shorter calculation time in the plasma etching case. The shorter time should be related to the time savings from using the SBD calculation rather than the DTW distance. However, since time series are locally z-normalized before clustering for the k-shape method, the method is effective for data not volatile in scales, and further tests should be made to test the applicability of the k-shape method for other semiconductor production monitoring cases.

5. Conclusion

Comprehensive conclusions are made in this chapter from the perspective of the centroid building method and the clustering methods themselves, the value brought to ADI, and finally recommendations for future experiments and improvements.

5.1 Suggested Centroid Building Method

The three centroid building methods, arithmetic, DBA, and soft-DTW, are compared in terms of accuracy and calculation time. The accuracy is evaluated through two indices, the mean DTW distance between the centroid and the time series it is built on, as well as the mean SBD. The former is a quantitative distance-measurement parameter and the latter indicates the average similarity between the centroid and the time series it is built upon. Overall, the DBA outperforms the arithmetic method with a much better mean DTW distance, and outperforms the soft-DTW with a much shorter calculation time and a shorter mean DTW distance. The arithmetic method outperforms the other two methods on the mean SBD, and all three methods perform well in terms of the shape similarity. Considering the overall performance, the DBA method is chosen as the recommended centroid building method.

Further optimizations are made on the DBA method. With the window size set to 20% of the time series length, fewer warped paths are considered so that calculation time is substantially shortened, with a modest mean DTW distance loss within 5%. Considering the case when new datasets are added, two recalculation methods are compared, the complete recalculation and the reconstructed recalculation. The former updates the centroid based upon all datasets individually , while the latter reconstructs the centroid by combining the previous centroid and new datasets. The result prefers the reconstructed method, since it reduces calculation time by 39.62% with a mean DTW distance improvement of 5.12% relative to the "complete recalculation" method.

5.2 K-shape Clustering

A novel clustering method, k-shape clustering, is explored in this thesis. It is a partitional method based upon the SBD and the shape extraction to iteratively reassign datasets and rebuild centroids. Datasets are z-normalized locally first before clustering. The clustering of the mixture of 341 'ProChm_EndPt_ChanC' datasets are performed with the k-shape clustering method and compared with the k-means method. The result indicates that a number of 4 clusters create the optimal result for the k-shape clustering method. On top of that, the k-shape clustering method outperforms the k-means method, particularly with respect to the calculation time.

5.3 Value to ADI

An accurate and representative centroid plays an important role in ADI's online and offline anomaly detections. For online detections where anomalies are detected automatically, the centroid is the basis of clustering that determines clustering accuracy. With more accurate centroids, both possibilities of false alarms and missed anomalies are reduced. With an assumption that the majority of datasets are normal, centroid buildings and following anomaly detections can be operated automatically.

For the offline detection where process engineers extract the data and make related analysis individually or collaboratively, currently centroids are mainly made by rules of thumb. It is hard to build a centroid considering the phase dislocations between time series. Such manually built centroids reduce not only judgment accuracy and consistency, but also efficiency, particularly for the case where engineers discuss issues with diverse centroids in their minds. A quantitively and qualitatively accurate centroid built by the optimized DBA method introduced in this thesis facilitates engineers making accurate and consistent judgements efficiently. The improvements on the DBA centroid building with the help of window size and reconstructed recalculation further improves the anomaly detection efficiency both offline and online.

The k-shape clustering method is an accurate and efficient method which could be practical in anomaly detection and classifications for scale-stable univariate single-recipe process analysis. With time series data updated and primary parameters to consider reselected, the centroid building and clustering methods discussed in this thesis could be applied to other recipes and processes flexibly.

5.4 Recommendations

Further tests and extensions can be made to further improve the method quality. Tests can be made on larger sample sizes, and on different parameters, recipes and processes, in order to make related adjustments and optimizations. The current demonstration is mainly based on the univariate analysis. Since the DBA supports multivariate analysis, further evaluations can be made on multiple process parameters of the time series to consider take interdependent influence of parameters into account. The k-shape clustering method can be further tested and evaluated on univariate analysis, especially for the case where scales differ volatilely. On top of that, the influence of optimized centroids on the clustering and neural network models that T. Chen and O. Makhlouk work on [3, 4] can be further tested and discussed.

Reference

[1] Analog Devices Inc. (2015) "Analog Devices Inc. Form 10-K." Derived from http://files.shareholder.com/downloads/ADI/2383156919x0x872073/9B336071-EF60-43AF-9E98-A424EEF6634C/2015_AnalogDevices_AR.FINAL_for_Posting.pdf

[2] Analog Devices Inc. (2018) "About ADI." Derived from http://www.analog.com/en/about-adi.html

[3] T. Chen. (2018) "Anomaly Detection in Semiconductor Manufacturing through Time Series Forecasting using Neural Networks" M.Eng. Thesis, Massachusetts Institute of Technology, Cambridge, MA.

[4] O. Makhlouk. (2018) "Time Series Data Analytics: Clustering-based Anomaly Detection Techniques for Quality Control in Semiconductor Manufacturing" M.Eng. Thesis, Massachusetts Institute of Technology, Cambridge, MA.

[5] Oxford Instruments. (2010) "Plasma Etch-Plasma Etching." Derived from http://www.oxnist.com

[6] RPI SCOREC. (2018) "Physically-Based Models of Reactive Ion Etching." Derived from https://scorec.rpi.edu/research_plasmaetchmodeling.php

[7] A. Sard´a-Espinosa. (2018) "Comparing Time-Series Clustering Algorithms in R Using the dtwclust Package." Derived from https://cran.rproject.org/web/packages/dtwclust/vignettes/dtwclust.pdf

[8] R. Wicklin. (2017, February 08) "Winsorization: the Good, the Bad, and the Ugly." Derived from https://blogs.sas.com/content/iml/2017/02/08/winsorization-good-bad-and-ugly.html

[9] V. Niennattrakul, D. Srisai, C.A. Ratanamahatana. (2011) "Shape-based Template Matching for Time Series Data"

[10] D. Sankoff, J. Kruskal. (1983) "The Symmetric Time-warping Problem: from Continuous to Discrete." In "Time Warps, String Edits and Macromolecules: The Theory and Practice of Sequence Comparison", pp. 125–161, 1983.

[11] J. Aach, G.M. Church. (2001) "Aligning Gene Expression Time Series with Time Warping Algorithms." Bioinformatics, pp. 495-508, 2001.

[12] Z. Bar-Joseph, G. Gerber, D.K. Gifford, T.S. Jaakkola, I. Simon. (2002) "A New Approach to Analyzing Gene Expression Time Series Data." RECOMB: Proceedings of the Sixth Annual International Conference on Computational Biology, pp. 39–48, 2002.

[13] D.M. Gavrila, L.S. Davis. (1995) "Towards 3-D Model-based Tracking and Recognition of Human Movement: a Multi-view Approach." IEEE International Workshop on Automatic Face- and Gesture-Recognition, pp. 272–277, 1995.

[14] T. Rath, R. Manmatha. (2003) "Word Image Matching Using Dynamic Time Warping." IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 521–527, 2003.

[15] M. Cuturi, M. Blondel. (2017) "Soft-DTW: a Differentiable Loss Function for Time-Series." In 34th International Conference on Machine Learning, vol. 70, pp. 894–903,2017. Derived from http://proceedings.mlr.press/ v70/cuturi17a.html

[16] H. Sakoe, S. Chiba. (1978) "Dynamic programming algorithm optimization for spoken word recognition." IEEE Trans. on Acoustics, Speech, and Sig. Proc., vol. 26, pp. 43–49, 1978.

[17] M. Cuturi, J-P. Vert, O. Birkenes, and T. Matsui. (2007) "A kernel for time series based on global alignments." 2007 IEEE International Conference on Acoustics, Speech and Signal Processing ICASSP'07, vol. 2, pp. II–413, 2007.

[18] Fréchet, Maurice. (1948) "Les éléments aléatoires de nature quelconque dans un espace distancié." Annales de l'institut Henri Poincaré, vol. 10, no. 4, pp. 215-310, 1948.

[19] J. Paparrizos, L. Gravano. (2015) "k-shape: Efficient and Accurate Clustering of Time Series"

[20] J. Paparrizos, L. Gravano. (2017) "Fast and Accurate Time Series Clustering." ACM Transactions on Database Systems, vol. 42, no. 2, article 8, June 2017.

[21] NIST. (2017) "Minkowski Distance." Derived from https://www.itl.nist.gov/div898/software/dataplot/refman2/auxillar/minkdist.htm

[22] CA. Ratanamahatana, E. Keogh. (2004) "Everything You Know About Dynamic Time Warping is Wrong." Third Workshop on Mining Temporal and Sequential Data.

Appendix

Table A.1. Centroid Performance of the 'ProChm_EndPt_ChanC', without Preset Centroids

Parameter			ProChm_EndPt_ChanC							
Method			DBA	Soft-dtw (γ=0.001)	Soft-dtw (γ=0.01)	Soft-dtw (γ=0.1)	Soft-dtw (γ=0.1)	Arithmetic (Win. Mean)	Arithmetic (Untr. Mean)	Arithmetic (Median)
Index			Mean DTW Distance							
Sample Size(20)	Group	1	58615.56	131328.48	80674.87	80068.23	100106.48	95697.09	86107.37	94395.96
		2	31014.48	39811.14	35560.97	36687.23	37833.05	65952.12	76312.55	63519.5
		3	48820.57	87318.62	83713.12	75510.86	268807.56	145882.78	142461.7	134257.02
		4	24097.12	43312.76	41302.31	25759.39	24193.1	48528.2	48177.83	33726.34
		5	14829.94	15169.96	16175.38	15026.58	14945.7	20085.32	21759.56	17409.04
Sample Size(50)	Group	1	67769.09	100827.69	85359.46	80977.33	79741.82	111995.32	107064.83	97626.38
		2	42666.46	55801.82	56687.34	87691.77	75042.69	68549.84	68617.38	72983.55
		3	21388.78	24143.01	27675.07	23387.52	27547.74	37306.84	37249.46	33578.82
		4	73021.7	133222.69	99570.32	101899.76	89534.97	123816	115157.89	116080.94
		5	22375.65	32577.46	31464.61	28049.31	31778.82	40125.34	38709.19	38514.97
Index			Mean SBD							
Sample Size(20)	Group	1	5.353E-03	5.829E-03	5.728E-03	4.197E-03	4.081E-03	3.062E-03	3.061E-03	3.357E-03
		2	8.911E-03	6.447E-03	9.209E-03	6.686E-03	7.093E-03	4.720E-03	4.741E-03	5.610E-03
		3	9.790E-03	6.351E-03	5.456E-03	7.722E-03	4.478E-03	2.421E-03	2.442E-03	2.568E-03
		4	6.312E-03	1.513E-02	1.136E-02	6.944E-03	1.067E-02	5.136E-03	5.132E-03	5.639E-03
		5	7.139E-03	4.407E-03	7.262E-03	7.272E-03	9.590E-03	3.605E-03	3.591E-03	4.087E-03
Sample Size(50)	Group	1	1.419E-02	8.493E-03	6.308E-03	7.410E-03	1.276E-02	4.500E-03	4.505E-03	5.108E-03
		2	1.273E-02	7.106E-03	1.379E-02	1.329E-02	7.339E-03	4.571E-03	4.568E-03	5.139E-03
		3	4.778E-03	5.966E-03	6.938E-03	5.431E-03	5.799E-03	3.842E-03	3.832E-03	4.398E-03
		4	1.511E-02	7.730E-03	1.360E-02	7.991E-03	6.023E-03	3.764E-03	3.762E-03	4.021E-03
		5	6.506E-03	6.267E-03	9.010E-03	7.860E-03	1.053E-02	3.966E-03	3.968E-03	4.526E-03
Index			Calcualtion Time(s)							
Sample Size(10)	Group	1	2.53	15.89	16.1	17.12	17.45	0.17		
		2	2.56	15.89	16.2	17.77	18.34	0.17		
		3	2.53	16.56	17.72	17.68	17.42	0.17		
		4	2.6	17.4	17.6	18.13	18.87	0.18		
		5	2.79	16.86	17.41	18.92	18.19	0.19		
Sample Size(20)	Group	1	6.31	39.2	41.08	41.16	42.01	0.19		
		2	6.31	39.86	41.46	44.53	42.25	0.19		
		3	6.3	39.37	41.25	41.75	42.94	0.19		
		4	6.17	39.28	40.43	40.75	41.97	0.18		
		5	6.52	45.08	43.92	44.09	45.77	0.18		

Table A.2. Centroid Performance of the 'Bot_RF_RevPwr', without Preset Centroids

Parameter			Bot_RF_RevPwr							
Method			DBA	Soft-dtw (γ=0.001)	Soft-dtw (γ=0.01)	Soft-dtw (γ=0.1)	Soft-dtw (γ=0.1)	Arithmetic (Win. Mean)	Arithmetic (Untr. Mean)	Arithmetic (Median)
Index			Mean DTW Distance							
Sample Size(20)	Group	1	167.99	316.76	359.79	336.12	318.45	386.44	389.95	311.81
		2	120.51	130.66	130.82	188.81	125.97	262.75	276.59	199.75
		3	221.06	366.21	452.66	319.28	351.90	472.74	483.28	362.61
		4	147.52	201.32	223.40	166.74	169.56	321.65	327.46	250.95
		5	121.95	140.12	135.23	129.95	123.51	193.76	205.73	143.75
Sample Size(50)	Group	1	181.79	270.35	219.35	278.55	278.69	363.86	376.48	296.87
		2	181.36	220.56	200.87	246.44	227.52	360.01	374.89	319.15
		3	170.90	262.31	207.40	192.29	162.24	329.53	354.88	265.19
		4	158.69	188.49	161.97	192.25	187.62	270.65	267.62	241.86
		5	161.55	191.69	203.64	215.94	201.43	334.56	348.55	281.38
Index			SBD							
Sample Size(20)	Group	1	0.122	0.139	0.138	0.139	0.137	0.075	0.075	0.088
		2	0.117	0.131	0.132	0.131	0.145	0.093	0.095	0.108
		3	0.140	0.137	0.137	0.137	0.142	0.107	0.106	0.146
		4	0.126	0.116	0.117	0.117	0.111	0.117	0.118	0.137
		5	0.066	0.066	0.068	0.068	0.072	0.064	0.065	0.064
Sample Size(50)	Group	1	0.148	0.205	0.207	0.211	0.219	0.121	0.122	0.134
		2	0.164	0.155	0.152	0.151	0.155	0.126	0.127	0.146
		3	0.593	0.211	0.208	0.207	0.209	0.172	0.175	0.209
		4	0.081	0.082	0.083	0.082	0.082	0.077	0.078	0.082
		5	0.149	0.151	0.151	0.149	0.150	0.130	0.132	0.153
Index			Calcualtion Time(s)							
Sample Size(20)	Group	1	1.4	15.38	14.28	14.33	14.72	0.15		
		2	1.14	12.81	13.11	13.25	14.72	0.17		
		3	2.47	25.35	27.76	27.36	28.03	0.19		
		4	2.69	30.03	30.41	30.37	30.55	0.17		
		5	2.66	27.17	29.72	30.15	32.91	0.15		
Sample Size(50)	Group	1	6.41	64.64	68.44	68.08	69.79	0.2		
		2	6.69	69.89	69.79	71.91	72.81	0.18		
		3	6.4	71.07	70.28	74.75	71.22	0.21		
		4	6.52	65	71.22	68.18	70.61	0.2		
		5	6.39	64.86	67.53	68.09	71.99	0.19		

Table A.3. Centroid Performance of the 'ProChm_EndPt_ChanC', with the Arithmetic Median as the Preset Centroid

Parameter			ProChm_EndPt_ChanC							
Method			DBA	Soft-dtw (γ=0.001)	Soft-dtw (γ=0.01)	Soft-dtw (γ=0.1)	Soft-dtw (γ=0.1)	Arithmetic (Win. Mean)	Arithmetic (Untr. Mean)	Arithmetic (Median)
Index			Mean DTW Distance							
Sample Size(20)	Group	1	64929.66	78368.31	71707.35	86256.39	84085.95	95697.09	86107.37	94395.96
		2	30910.87	39148.87	39551.56	41903.54	37277.87	65952.12	76312.55	63519.5
		3	42646.49	84792.87	73022.8	68538.12	65240.5	145882.78	142461.7	134257.02
		4	19271.59	25235.83	22178.2	22972.06	23081.35	48528.2	48177.83	33726.34
		5	14215.63	14903.3	14874.81	14795	15054.76	20085.32	21759.56	17409.04
Sample Size(50)	Group	1	65909.54	80881.56	73344.89	72323.35	70788.57	111995.32	107064.83	97626.38
		2	53224.23	59924.12	59321.99	59295.12	58523.28	68549.84	68617.38	72983.55
		3	21472.35	23324.89	23566.05	23486.34	23248.87	37306.84	37249.46	33578.82
		4	72864.8	86848.14	91161.15	88077.85	92620.12	123816	115157.89	116080.94
		5	21615.57	26141.3	27206.87	28408.64	26682.37	40125.34	38709.19	38514.97
Index			Mean SBD							
Sample Size(20)	Group	1	4.055E-03	3.554E-03	3.450E-03	4.318E-03	3.974E-03	3.062E-03	3.061E-03	3.357E-03
		2	6.760E-03	6.202E-03	6.341E-03	6.421E-03	6.396E-03	4.720E-03	4.741E-03	5.610E-03
		3	5.875E-03	3.902E-03	4.179E-03	4.311E-03	4.760E-03	2.421E-03	2.442E-03	2.568E-03
		4	6.163E-03	5.743E-03	5.740E-03	5.767E-03	5.770E-03	5.136E-03	5.132E-03	5.639E-03
		5	4.219E-03	4.131E-03	4.135E-03	4.130E-03	4.118E-03	3.605E-03	3.591E-03	4.087E-03
Sample Size(50)	Group	1	8.574E-03	7.274E-03	8.089E-03	7.925E-03	7.610E-03	4.500E-03	4.505E-03	5.108E-03
		2	5.356E-03	5.426E-03	5.348E-03	5.372E-03	5.355E-03	4.571E-03	4.568E-03	5.139E-03
		3	4.630E-03	4.673E-03	4.669E-03	4.676E-03	4.663E-03	3.842E-03	3.832E-03	4.398E-03
		4	5.932E-03	5.219E-03	5.839E-03	5.280E-03	5.914E-03	3.764E-03	3.762E-03	4.021E-03
		5	4.856E-03	4.568E-03	4.549E-03	4.581E-03	4.524E-03	3.966E-03	3.968E-03	4.526E-03
Index			Calcualtion Time(s)							
Sample Size(20)	Group	1	2.49	15.78	16.11	16.52	17.34	0.17		
		2	2.32	15.35	15.98	16.66	17.36	0.18		
		3	2.38	15.59	16.38	17	17.2	0.17		
		4	2.54	16.63	18.29	18.04	18.09	0.18		
		5	2.45	15.82	16.23	17.05	17.64	0.17		
Sample Size(50)	Group	1	6.26	40.9	42.79	42.79	44.25	0.19		
		2	6.57	42.68	44.66	42.47	43.59	0.2		
		3	6.54	41.39	45.15	44.11	44.85	0.19		
		4	6.47	41.65	44.19	43.5	44.83	0.19		
		5	6.25	40.24	41.73	41.61	47.52	0.21		

Table A.4. Centroid Performance of the ‘Bot_RF_RevPwr’, with the Arithmetic Median as the Preset Centroid

	Parameter		Bot_RF_RevPwr							
	Method		DBA	Soft-dtw (γ=0.001)	Soft-dtw (γ=0.01)	Soft-dtw (γ=0.1)	Soft-dtw (γ=0.1)	Arithmetic (Win. Mean)	Arithmetic (Untr. Mean)	Arithmetic (Median)
	Index		Mean DTW Distance							
Sample Size(20)	Group	1	168.1317	225.5783	231.2395	240.4125	252.67	386.4423	389.9541	311.8083
		2	130.288	142.8848	146.4394	144.0133	154.7286	262.7486	276.5872	199.7517
		3	202.9497	259.4248	261.9103	245.582	274.4279	472.7362	483.2808	362.6097
		4	145.9281	158.223	156.2988	173.0567	167.7131	321.6545	327.4646	250.9507
		5	161.3786	186.557	177.7784	190.7695	184.3533	193.7552	205.7322	143.7536
Sample Size(50)	Group	1	188.8292	240.3366	235.3826	239.017	245.2013	363.8622	376.4825	296.8668
		2	177.7011	238.4747	245.4393	239.7266	218.4103	360.0056	374.8913	319.1509
		3	167.5353	207.3496	213.8141	219.066	181.5471	329.5268	354.8758	265.1896
		4	153.9625	175.1082	179.0958	181.8727	187.5194	270.6501	267.623	241.8604
		5	158.8994	207.4555	211.0785	186.2927	184.0695	334.5623	348.5527	281.3806
	Index		Mean SBD							
Sample Size(20)	Group	1	0.122	0.139	0.138	0.139	0.137	0.075	0.075	0.088
		2	0.117	0.131	0.132	0.131	0.145	0.093	0.095	0.108
		3	0.140	0.137	0.137	0.137	0.142	0.107	0.106	0.146
		4	0.126	0.116	0.117	0.117	0.111	0.117	0.118	0.137
		5	0.066	0.066	0.068	0.068	0.072	0.064	0.065	0.064
Sample Size(50)	Group	1	0.148	0.205	0.207	0.211	0.219	0.121	0.122	0.134
		2	0.164	0.155	0.152	0.151	0.155	0.126	0.127	0.146
		3	0.593	0.211	0.208	0.207	0.209	0.172	0.175	0.209
		4	0.081	0.082	0.083	0.082	0.082	0.077	0.078	0.082
		5	0.149	0.151	0.151	0.149	0.150	0.130	0.132	0.153
	Index		Calcualtion Time(s)							
Sample Size(20)	Group	1	2.42	24.43	28.04	31.16	28.78		0.15	
		2	2.38	24.69	27.19	27.5	27.47		0.17	
		3	2.39	25.67	27.31	27.53	28.5		0.19	
		4	2.5	26.24	29	29.08	29.13		0.17	
		5	2.47	24.08	26.2	28.22	29.22		0.15	
Sample Size(50)	Group	1	6.04	61.01	66.72	67.34	67.68		0.2	
		2	6.26	61.24	67.17	68.11	72.06		0.18	
		3	6.17	63.08	67.67	71.05	68.15		0.21	
		4	6.1	61.96	65.4	68.79	70.67		0.2	
		5	6.06	63.88	66.05	66.11	66.74		0.19	

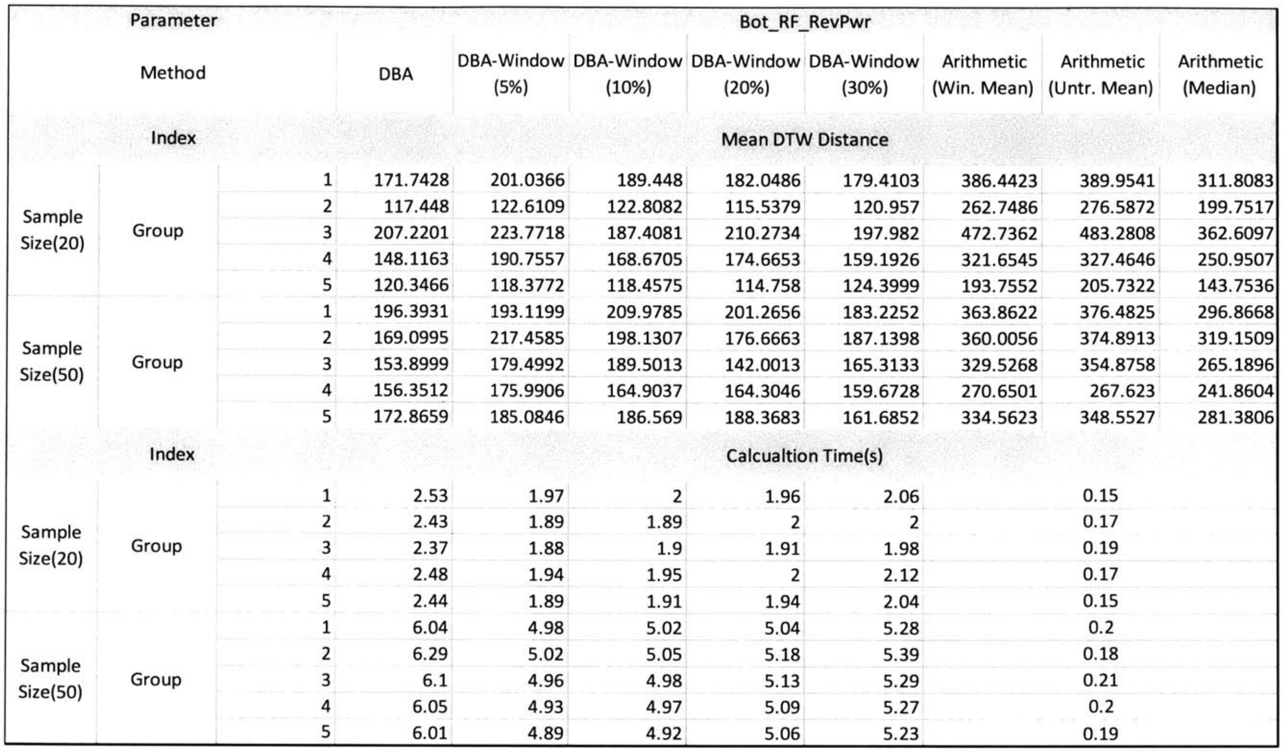
Table A.5. Centroid Performance of the 'ProChm_EndPt_ChanC', with the DBA as the Preset Centroid

Parameter			ProChm_EndPt_ChanC							
Method			DBA	Soft-dtw (γ=0.001)	Soft-dtw (γ=0.01)	Soft-dtw (γ=0.1)	Soft-dtw (γ=0.1)	Arithmetic (Win. Mean)	Arithmetic (Untr. Mean)	Arithmetic (Median)
Index			Mean DTW Distance							
Sample Size(20)	Group	1	58615.56	58980.13	58694.75	58697.93	59031.31	95697.09	86107.37	94395.96
		2	32582.24	34243.79	34800.71	34958.19	35208.78	65952.12	76312.55	63519.5
		3	39733.92	40724.39	40420.47	40469.6	40558.68	145882.78	142461.7	134257.02
		4	24854.68	24291.07	24253.59	24106.42	24716.93	48528.2	48177.83	33726.34
		5	14689.61	14453.48	14482.37	14643.23	14546.87	20085.32	21759.56	17409.04
Sample Size(50)	Group	1	73121.68	70231.38	70277.62	70369.59	69803.42	111995.32	107064.83	97626.38
		2	53702.89	53637.27	54151.67	53745.79	53703.04	68549.84	68617.38	72983.55
		3	22685.54	20844.06	20902.44	20764.96	20872.06	37306.84	37249.46	33578.82
		4	70314.47	70089.61	71594.45	71609.73	70975.63	123816	115157.89	116080.94
		5	21148.86	21297.62	21163.72	21590.56	21372.85	40125.34	38709.19	38514.97
Index			Mean SBD							
Sample Size(20)	Group	1	0.012	0.012	0.012	0.012	0.012	0.003	0.003	0.003
		2	0.008	0.009	0.009	0.009	0.009	0.005	0.005	0.006
		3	0.008	0.009	0.009	0.009	0.009	0.002	0.002	0.003
		4	0.009	0.009	0.009	0.009	0.009	0.005	0.005	0.006
		5	0.007	0.007	0.007	0.007	0.007	0.004	0.004	0.004
Sample Size(50)	Group	1	0.011	0.011	0.011	0.011	0.011	0.005	0.005	0.005
		2	0.008	0.008	0.009	0.008	0.009	0.005	0.005	0.005
		3	0.006	0.006	0.006	0.006	0.006	0.004	0.004	0.004
		4	0.012	0.014	0.014	0.014	0.014	0.004	0.004	0.004
		5	0.007	0.007	0.007	0.007	0.007	0.004	0.004	0.005
Index			Calcualtion Time(s)							
Sample Size(20)	Group	1	2.45	16.31	16.28	16.52	18.26	0.19		
		2	2.34	15.38	16.48	16.89	17.3	0.18		
		3	2.45	15.91	16.34	17.08	18.08	0.17		
		4	2.5	16.25	16.91	17.7	19.05	0.18		
		5	2.48	16.09	17.08	17.33	18.19	0.17		
Sample Size(50)	Group	1	6.25	40.64	41.42	41.86	43.81	0.19		
		2	6.35	40.64	41.419	42.14	45.99	0.18		
		3	6.32	39.67	41.58	42.06	43.84	0.17		
		4	6.21	40.24	42.01	41.57	43.2	0.19		
		5	6.12	38.77	40.741	41.75	44	0.19		

Table A.6. Centroid Performance of the 'Bot_RF_RevPwr', with the DBA as the Preset Centroid

	Parameter		Bot_RF_RevPwr							
	Method		DBA	Soft-dtw (γ=0.001)	Soft-dtw (γ=0.01)	Soft-dtw (γ=0.1)	Soft-dtw (γ=0.1)	Arithmetic (Win. Mean)	Arithmetic (Untr. Mean)	Arithmetic (Median)
	Index		Mean DTW Distance							
Sample Size(20)	Group	1	175.8923	178.0444	182.0747	194.264	201.3757	386.4423	389.9541	311.8083
		2	128.3601	142.319	143.4397	148.6287	129.3902	262.7486	276.5872	199.7517
		3	207.1175	205.893	210.5131	218.8455	220.4146	472.7362	483.2808	362.6097
		4	154.3825	159.0656	165.4264	168.1156	164.1243	321.6545	327.4646	250.9507
		5	114.4398	130.2523	127.3323	125.6956	124.4242	193.7552	205.7322	143.7536
Sample Size(50)	Group	1	167.8156	171.1952	177.5469	189.2493	202.2118	363.8622	376.4825	296.8668
		2	202.0836	227.5745	218.2022	221.5763	212.333	360.0056	374.8913	319.1509
		3	155.8517	159.5098	164.4968	169.5233	173.173	329.5268	354.8758	265.1896
		4	152.6304	150.9018	155.2226	163.9295	175.9866	270.6501	267.623	241.8604
		5	172.0321	183.8324	190.5995	195.9717	196.1019	334.5623	348.5527	281.3806
	Index		Mean SBD							
Sample Size(20)	Group	1	0.12944657	0.12940956	0.12916514	0.12891086	0.13134067	0.074803	0.07467951	0.08762304
		2	0.11702149	0.11564387	0.11609416	0.12009114	0.1313216	0.09291928	0.09547493	0.10831085
		3	0.1686562	0.1619796	0.1634693	0.1614284	0.162182	0.1070259	0.106196	0.1461103
		4	0.1184045	0.1309864	0.1355217	0.137243	0.1339568	0.1170676	0.1184105	0.1366777
		5	0.07092605	0.07256962	0.07124534	0.07243803	0.07423446	0.06434794	0.06463911	0.06367104
Sample Size(50)	Group	1	0.1673809	0.1687019	0.1701285	0.1732039	0.1764448	0.1205555	0.1221447	0.1335347
		2	0.182465	0.1808778	0.1807327	0.1823362	0.1799542	0.1260516	0.1266255	0.1455341
		3	0.1926049	0.1897609	0.1908518	0.1900552	0.1926849	0.172049	0.1745466	0.2090131
		4	0.07569581	0.07828798	0.07883781	0.07892478	0.07927705	0.07715668	0.07766332	0.08163696
		5	0.1589673	0.152168	0.1522886	0.1522715	0.1557775	0.1298136	0.1317777	0.1525269
	Index		Calcualtion Time(s)							
Sample Size(20)	Group	1	2.43	25.47	26.19	27.97	27.96	0.15		
		2	2.33	24.52	28	27.22	27.22	0.17		
		3	2.38	25.19	26.78	27.36	27.54	0.19		
		4	2.46	25.73	28.16	28.72	28.87	0.17		
		5	2.47	24.39	27.06	27.64	28.26	0.15		
Sample Size(50)	Group	1	6.06	62.69	65.7	66.83	67.58	0.2		
		2	6.17	62.84	67.06	68.56	69.17	0.18		
		3	6.11	62.34	67.05	67.63	67.85	0.21		
		4	6.03	61.72	65.85	65.7	67.85	0.2		
		5	5.97	62.39	65.66	66.63	67.47	0.19		

Table A.7. Centroid Performance of the 'Bot_RF_RevPwr' for the Window-constrained DBA Method



	Parameter		Bot_RF_RevPwr							
	Method		DBA	DBA-Window (5%)	DBA-Window (10%)	DBA-Window (20%)	DBA-Window (30%)	Arithmetic (Win. Mean)	Arithmetic (Untr. Mean)	Arithmetic (Median)
	Index		Mean DTW Distance							
Sample Size(20)	Group	1	171.7428	201.0366	189.448	182.0486	179.4103	386.4423	389.9541	311.8083
		2	117.448	122.6109	122.8082	115.5379	120.957	262.7486	276.5872	199.7517
		3	207.2201	223.7718	187.4081	210.2734	197.982	472.7362	483.2808	362.6097
		4	148.1163	190.7557	168.6705	174.6653	159.1926	321.6545	327.4646	250.9507
		5	120.3466	118.3772	118.4575	114.758	124.3999	193.7552	205.7322	143.7536
Sample Size(50)	Group	1	196.3931	193.1199	209.9785	201.2656	183.2252	363.8622	376.4825	296.8668
		2	169.0995	217.4585	198.1307	176.6663	187.1398	360.0056	374.8913	319.1509
		3	153.8999	179.4992	189.5013	142.0013	165.3133	329.5268	354.8758	265.1896
		4	156.3512	175.9906	164.9037	164.3046	159.6728	270.6501	267.623	241.8604
		5	172.8659	185.0846	186.569	188.3683	161.6852	334.5623	348.5527	281.3806
	Index		Calcualtion Time(s)							
Sample Size(20)	Group	1	2.53	1.97	2	1.96	2.06		0.15	
		2	2.43	1.89	1.89	2	2		0.17	
		3	2.37	1.88	1.9	1.91	1.98		0.19	
		4	2.48	1.94	1.95	2	2.12		0.17	
		5	2.44	1.89	1.91	1.94	2.04		0.15	
Sample Size(50)	Group	1	6.04	4.98	5.02	5.04	5.28		0.2	
		2	6.29	5.02	5.05	5.18	5.39		0.18	
		3	6.1	4.96	4.98	5.13	5.29		0.21	
		4	6.05	4.93	4.97	5.09	5.27		0.2	
		5	6.01	4.89	4.92	5.06	5.23		0.19	