Computers & Industrial Engineering 142 (2020) 106325

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

Implementation of statistical process control framework with machine learning on waveform profiles with no gold standard reference

Shih-Hsiung Chou[a,*], Shing Chang[b], Tzong-Ru Tsai[c], Dennis K.J. Lin[d], Yunfei Xia[e], Yu-Siang Lin[f]

[a] Center for Outcomes Research and Evaluation (CORE), Atrium Health, Charlotte, NC, USA
[b] Quality Engineering Laboratory, Department of Industrial and Manufacturing Systems Engineering, Kansas State University, Manhattan, KS, USA
[c] Department of Statistics, Tamkang University, New Taipei City, Taiwan
[d] Department of Statistics, Pennsylvania State University, State College, PA, USA
[e] Department of Mathematics and Statistics, University of North Carolina Charlotte, Charlotte, NC, USA
[f] Department of Industrial Management, National Taiwan University of Science and Technology, Taipei, Taiwan

ARTICLE INFO

Keywords:
Individual control chart
PAM clustering method
Support vector machine

ABSTRACT

Condensation water temperature profiles are collected from a curing process for high-pressure hose products. The shape of those profiles resembles sine waves with diminishing amplitudes. A gold standard wave profile does not exist. Instead some wave profiles with various frequency and amplitudes are deemed normal for the water release operation. To the best of our knowledge, the current practice and research on SPC do not provide a solution for monitoring wave profiles of this kind. We leveraged existing methods, tools, algorithms that can be found in open source or commercial software for quick response to this type of problem. The proposed SPC implementation framework first converts waveform profiles from the time domain to the frequency domain. Then a set of phase I *IX* control charts is constructed based on a Partition Around Medoids (PAM) clustering method. A Support Vector Machine (SVM) classifier is then used to label a new profile to its associated group for phase II monitoring so that the *IX* chart associated with a homogeneous group can provide better process monitoring. Overall 146 water temperature profiles were collected in phase I process, while 39 profiles were captured in phase II process. Out of those 39 profiles, 6 of which were recognized as abnormal waveform profiles by quality engineers and our judgements. The proposed framework with machine learning and SPC implementation in the frequency domain works well during phase I control charting with low false alarm rates. The proposed framework also outperforms the other profile analysis methods in phase II control charting in term of high detection rate of abnormal profiles.

1. Introduction/Background

High-pressure hoses are made from rubber material with layers of mental wraps. Rubber products often require a curing process called vulcanization as the final step to activate cross-linking reaction so that the tensile strength of finished rubber is stronger (Hoster, Jaunich, & Stark, 2009). In a curing process, reels of un-vulcanized hose are loaded into a sealed autoclave or vulcanizer whose in-chamber temperature and various valves are controlled by a programmable logical controller (PLC). Once all reels are loaded, the heat-up stage of curing process is enabled. The PLC monitors the chamber temperature and controls the steam valve to heat up the chamber until the temperature is reaching to a fixed target temperature, says, 500 °F. Once the target temperature is reached, the PLC will activate the curing stage to maintain the target temperature for a fixed time unit, e.g., 480. Due to the confidential agreement with the process owner, the curing recipe (i.e., 500 °F and 480 time units) is a process setting for illustration purpose only. The final stage of the curing process is called cool-down stage to decrease the chamber temperature. For more details of curing process please see Fig. 1 and Chang, Tsai, Lin, Chou, and Lin (2012).

This study focuses on an additional control of water valve during the curing stage (the second stage shown in Fig. 1 of a curing process. During the curing stage, condensation water accumulates at the bottom of the chamber. A water valve is installed at the end of the water releasing pipe for releasing condensation water. If this valve fails to open, the water would keep accumulating at the bottom of a vulcanizer and cause cosmetic or functional damage of hoses. On the other hand, if the water valve fails to close, the chamber temperature would be difficult to maintain and result in energy waste.

Fig. 2 shows a schematic diagram of a vulcanizer with accumulated

* Corresponding author.
E-mail address: shih-hsiung.chou@atriumhealth.org (S.-H. Chou).

https://doi.org/10.1016/j.cie.2020.106325
Received 23 April 2019; Received in revised form 26 December 2019; Accepted 25 January 2020
Available online 28 January 2020
0360-8352/

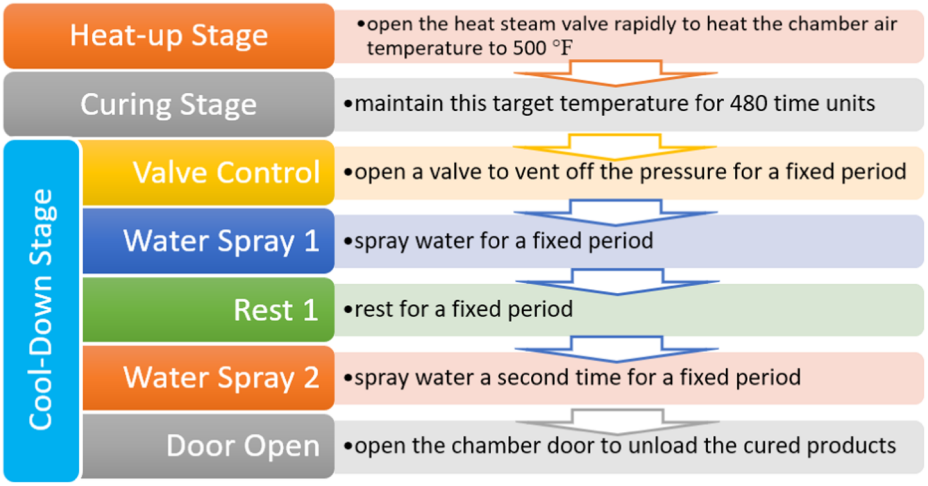

Fig. 1. Steps of curing process for high-pressure hose products.

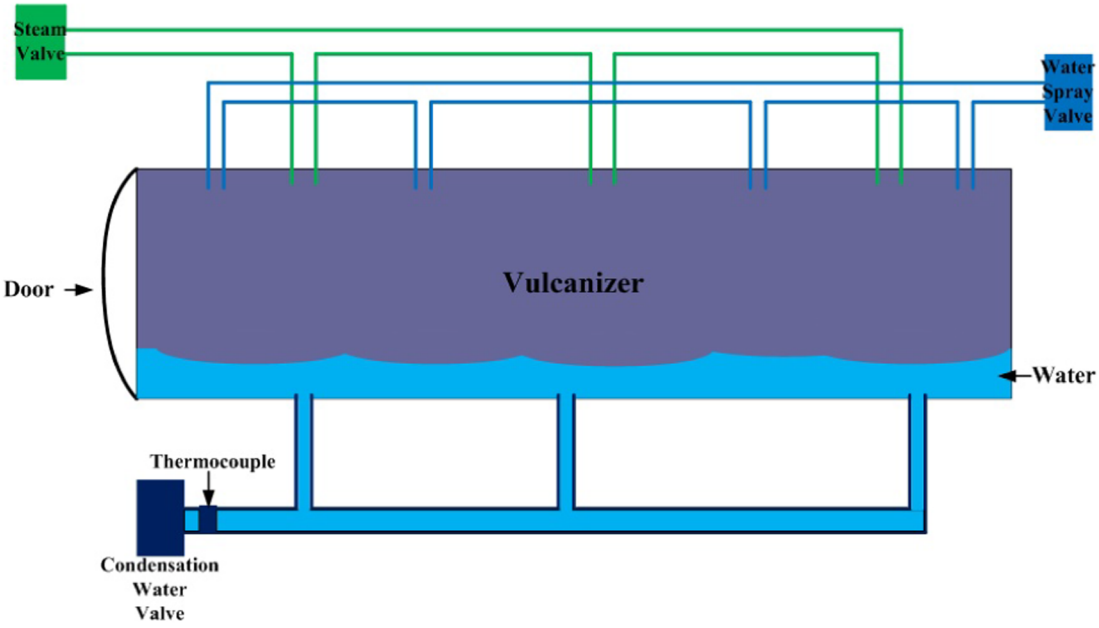

Fig. 2. A schematic diagram of vulcanizer.

condensation water and the location of valves. A thermocouple is mounted at the end of condensation water releasing pipe (before the condensation water valve) to read the condensation water temperature. One typical condensation water temperature profile is shown in Fig. 3. Note that once the condensation water temperature is decreased to a certain degree, the PLC will open the condensation water valve for a period to release water, and then the water temperature is climbing rapidly until the condensation water valve is closed. This mechanism results in the waveform shape of the condensation water profile. Since the company produces different kind of high-pressure hose products, they load different reels of hose products into the vulcanizer at the same time according to production orders. Reels in the vulcanizer may contain different amount of rubber material or different layers of metal wraps on hoses. It is not a surprise that the condensation water temperature profiles do not have a gold standard.

The condensation water temperature is recorded by a thermocouple located close to the water valve as shown in Fig. 2. It is also a suitable indicator of the operation of condensation water valve because it can be used to detect abnormal situations, such as, malfunction condensation water valve or thermocouple. However, the current detection method relies on visual inspection, i.e., through a quick glance of the water-temperature print out. If a profile contains enough number of waves, it is deemed as a normal profile. For example, the condensation water temperature profile shown in Fig. 3 is considered good because it contains 22 waves. Yet, other acceptable profiles may not contain the same number of waves. It is a challenge to develop an objective, systematic process control strategy for inexperienced engineers or operators. Fig. 4 shows three other examples of in-control waveform profiles.

Current profile analysis techniques can be simply characterized into two categories, linear and nonlinear profiles regarding to profile shape structure. Many studies monitored parameters of the linear regression model, such as, intercept or slope parameter. For example, Kang and Albin (2000) monitored slope and intercept with the Hotelling's T^2 control chart as the first method to detect abnormal profiles, and they also proposed the second method, which monitored average residuals between sample profiles and reference profile followed by EWMA and R chart. Kim, Mahmoud, and Woodall (2003) proposed three univariate EWMA charts which monitoring the slope, intercept, and the variance of deviation between samples and regression line simultaneously. Hosseinifard, Hosseinifard, Abdollahian, and Zeephongsekul (2011)

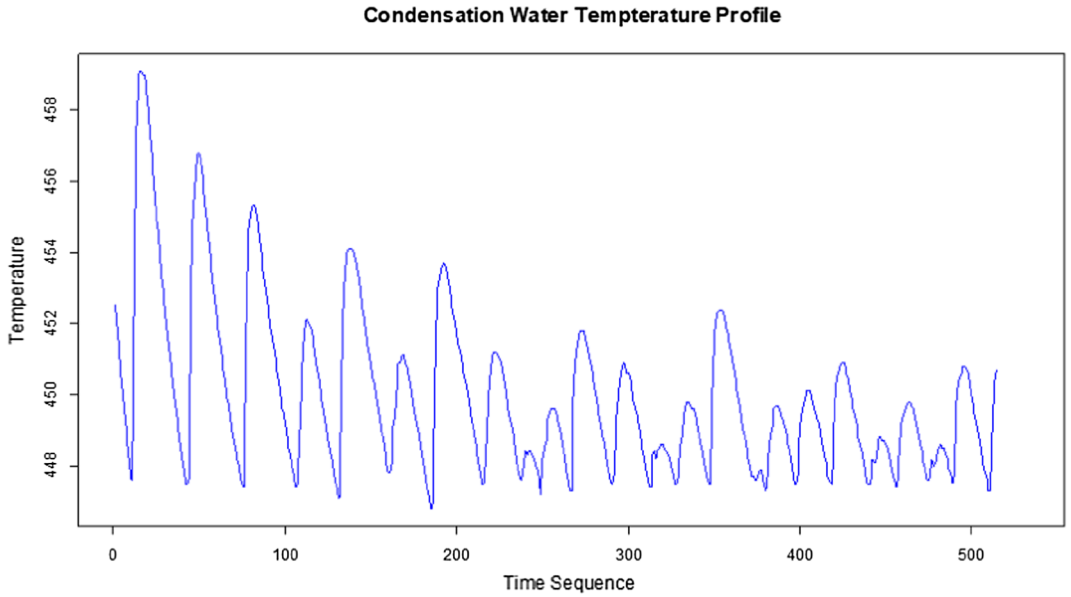

Fig. 3. An example of condensation water temperature profile.

used perceptron neural networks to monitor linear profiles shifted on y-intercept, slope, or residual variance. Abbas, Qian, Ahmad, and Riaz (2017) used three univariate Bayesian double exponentially weighted moving average control charts to monitor the Y intercepts, slopes, and error variances of linear profiles.

Many techniques in nonlinear profile analysis can be found in Woodall (2007) and Noorossana, Saghaei, and Amiri (2011). Woodall (2007) classified nonlinear profile analysis into four categories—applying multiple and polynomial regression (Kazemzadeh, Noorossana, & Amiri, 2008; Mahmoud, 2008; Zou, Tsung, & Wang, 2007), applying nonlinear regression models (Ding, Zeng, & Zhou, 2006; Williams, Woodall, & Birch, 2007; Shiau, Huang, Lin, & Tsai, 2009; Chang & Yadama, 2010; Chen & Nembhard, 2011; Wu, Liu, & Zhou, 2016; Yang, Zou, & Wang, 2017, and Awad, AlHamaydeh, & Faris, 2018), the use of mixed models (Abdel-Salam, Birch, & Jensen, 2013; Gomaa & Birch, 2019; Jensen & Birch, 2009; Jensen, Birch, & Woodall, 2008; Paynabar & Jin, 2011; Qiu, Zou, & Wang, 2010), and the use of wavelets (Chicken, Pignatiello, & Simpson, 2009; Reis & Saraiva, 2006; Zhou, Sun, & Shi, 2006).

Although the above-mentioned methods are successful in dealing with the situations in their problem domains, those techniques deal with profiles that contain specific shape with a standard or model profile that can be predefined or estimated. In addition, none of these methods are tested for waveform shape profile and most of these methods were not software/library/toolbox-readily for quality engineers to use. Since the quality engineers are interested in if water-releasing cycle is under a statistical control, this study develops a general SPC implementation framework with machine learning that can be found in existing open source/commercial software to monitor waveform shape profiles when no gold standard profile can be established. In the next section, we will describe how the wave profiles are collected.

2. Data collection

In this case study, 146 condensation water temperature profiles were deemed good, i.e., in control because of their frequent temperature oscillations. Most importantly, these profiles were collected from the production batches where high-pressure hose products, the condensation water valve, and the condensation water thermocouples were

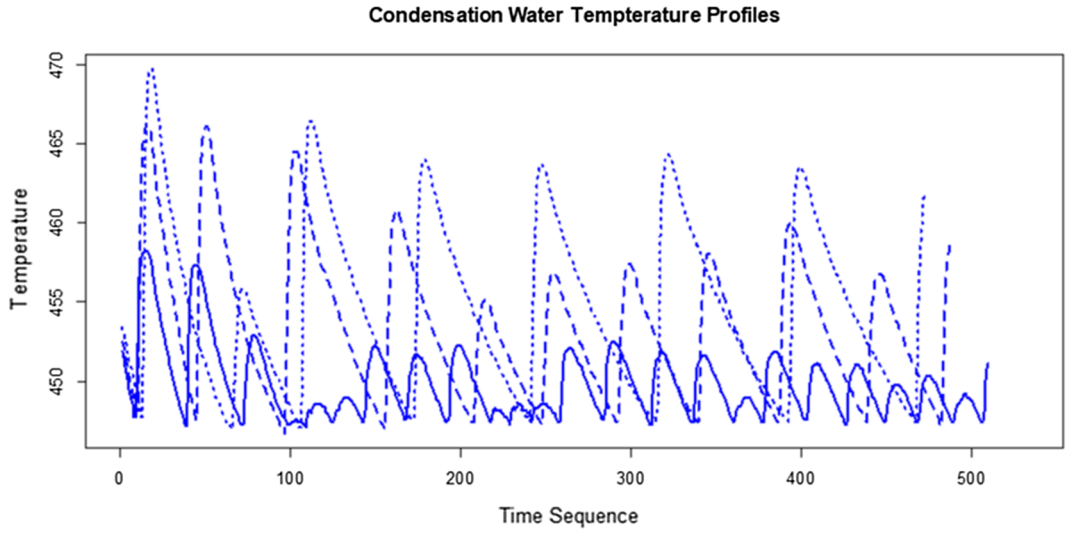

Fig. 4. Example of three in-control condensation water profiles.

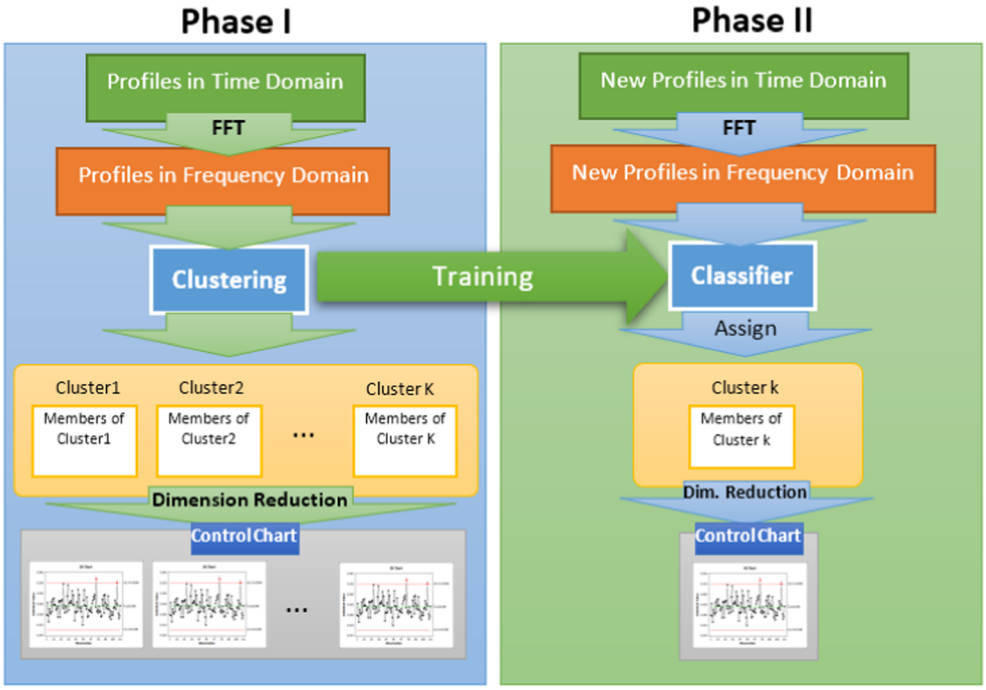

Fig. 5. The proposed framework for condensation water temperature monitoring.

all in good conditions. No cosmetic or malfunction hose was found in those production batches. Each wave profile can be presented as vector as shown in Eq. (1) where n is number of profiles and p_i is number of data points within i^{th} profile. The number of data points within a profile were collected by sensors of the chamber. These profiles (n = 146 in this case) will be used to construct phase I control charts.

$$W_{ji} = [w_{11}, w_{12}, ..., w_{ji}, ..., w_{pi,n}],\ i = 1, 2, 3, ..., n,\ j = 1, 2, 3, ..., p_i \quad (1)$$

The proposed framework is robust in that it strives to achieve a balance between the lowest false alarm rate and highest accuracy rate. An experiment is also conducted to examine various well-known clustering methods and dimension-reduction methods that are often applied to profile analysis with promising results. Then, the best method is chosen based on the criteria of false alarm rates and accuracy rates. In this case study, 39 new condensation water temperature profiles were collected as phase II data to test the proposed framework.

3. The proposed SPC framework

The proposed framework is shown in Fig. 5 that demonstrates the monitoring the condensation water temperature profiles during the curing stage of vulcanization process. In phase I, the original waveform profiles in time domain is transformed to frequency domain using the Fast Fourier Transformation (FFT), which is a logical choice for dealing with wave signals. Then, a clustering method is applied to the frequency domain profiles. This step groups frequency-domain signals into homogeneous clusters. Once all profiles have been clustered to their associated group, all profiles are processed by a dimension reduction method so that the reduced outcome for each profile can be plotted into a control chart. A control chart for each cluster is then to be constructed based on members of the cluster. Note that the information of the clusters and their associated membership will be utilized as a dataset for training a classifier model in phase II. Once the classifier is trained, it is ready for the phase II process to determine which group a new profile in frequency domain belongs to. For example, if a new waveform profile is assigned to cluster k by the classifier, then the parameters constructed in phase I, such as, mean, standard deviation, and control limits, can be obtained for phase II process monitoring. If the control chart indicates that an out-of-control signal occurs, the profile contributed to this out-of-control signal is then considered as an abnormal profile. More details of the classifier component will be introduced in the later section.

There are many candidate methods that may be appropriate for the clustering and dimension reduction functions shown in Fig. 5. We will not introduce the detailed mechanism of all techniques we examined in this study. Instead, we will only introduce the methods recommended in this study. We will also discuss why these methods are applied and how we expect the results to be. The proposed framework will be still valid when new or more advanced clustering techniques and dimension reduction methods are introduced in the future. Quality engineers may also choose the other methods that they are more familiar with due to the availability of the computational software and package available to them.

3.1. Fast Fourier transformation component

Fast Fourier Transformation (FFT) developed by Cooley and Tukey (1965) is one of the most well-known algorithms to calculate discrete Fourier Transformation (DFT) for converting a signal from time domain to frequency. We used Matlab's fft function to convert profiles. The output vector of FFT can be defined as Eq. (2) where p_f is the size of the FFT-transformed vector.

$$W_{ji}^{f} = [w_{11}^{f}, w_{12}^{f}, ..., w_{ji}^{f}, ..., w_{pf,n}^{f}],\ i = 1, 2, 3, ..., n,\ j = 1, 2, 3, ..., p_f \quad (2)$$

We convert the condensation water temperature profiles from the time domain to the frequency domain in that the original waveform profiles are too complicated to be directly applied to the existing SPC profile monitoring methods as reviewed earlier. Fig. 6 shows an example of the in-control profile 15 and its Fourier transformed profile. Although we can observe that the majority frequency of the profile 15 is 0.00091 Hz, there are other frequencies between 0 Hz and 0.01 Hz that cannot be ignored. After the FFT transformation, we can now treat the FFT transformed profiles as the other regular profile analysis problems except that we still need to address the issue of diverse frequency domain profiles in the next section.

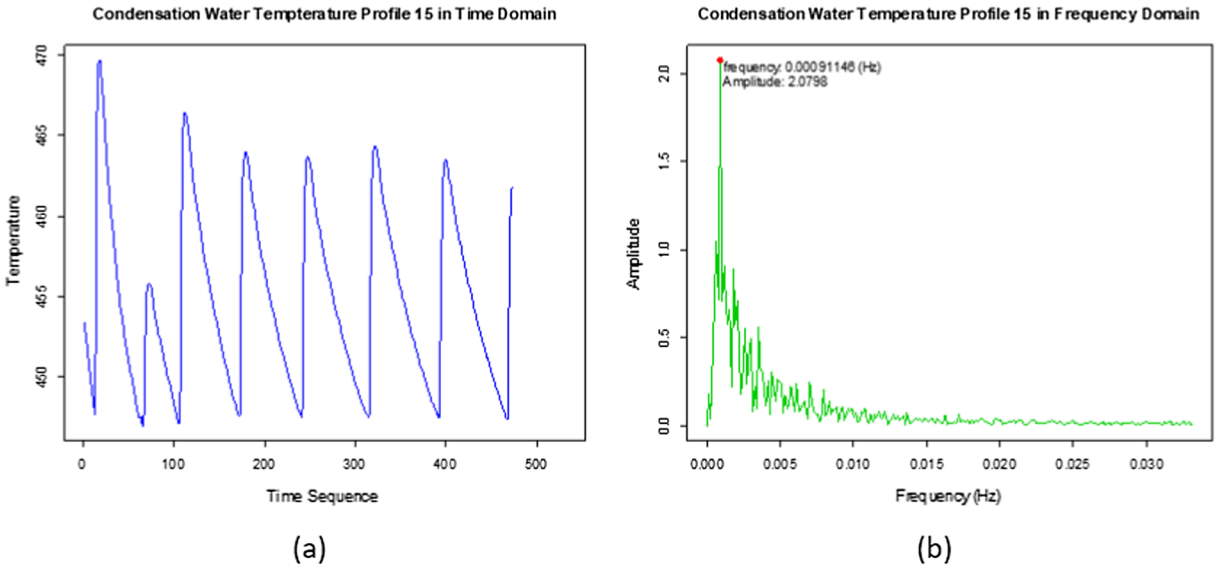

Fig. 6. Condensation water temperature profile 15 in (a) time domain and (b) frequency domain.

3.2. Clustering component

Even though the in-control FFT transformed profiles are potential representations of the original time domain profiles, they are not homogeneous. For example, the in-control profile 5 and 15 in frequency domain has different profile shape in frequency domain as shown in Fig. 7. Profile 5 has two majority frequencies while profile 15 only has one peak with gradient decline between 0.0025 Hz and 0.01 Hz. Therefore, we proposed to cluster the in-control FFT transformed profiles into homogeneous groups.

The purpose of a clustering method is to group a set of frequency domain members as similar as possible within a group, and those members can be distinguished from the other groups. In this study, we examine several widely used clustering analysis methods to the FFT transformed profiles such that members within the same group can be more homogeneous for subsequent profile analysis.

Table 1 shows the selected clustering methods examined in this study. Those well-developed methods including hierarchical clustering, mclust, K-means, Partition Around Medoids, fuzzy clustering, and fuzzy C-means are candidates for performance comparison (Leisch & Gruen, 2013). These clustering methods were used to group 146 in-control FFT transformed profiles as defined in Eq. (2). Then, the variance ratio criterion (VRC) is used as the clustering methods evaluation criterion (Mooi & Sarstedt, 2011) as shown in Eq. (3).

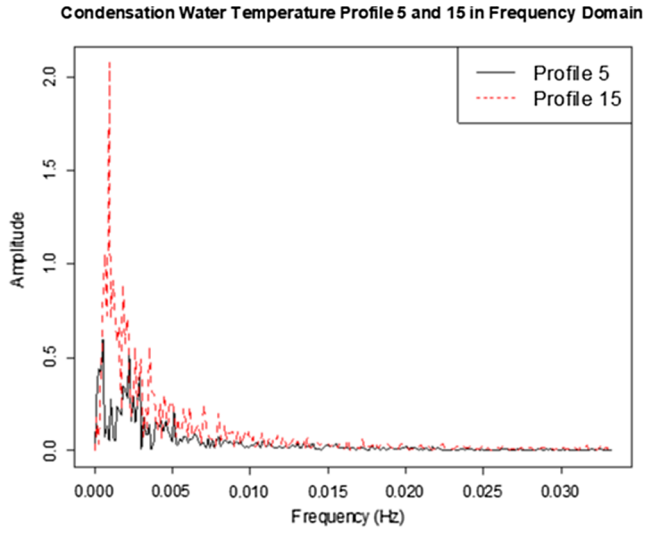

Fig. 7. Profile 5 and 15 in frequency domain.

Table 1
Clustering methods examined in this study.

Clustering Method	R Package	Function	SS_W	SS_B	*VRC*
Hierarchical Clustering	stat or cluster	hclust()	171.949	71.604	29.774
mclust	mclust	Mclust()	179.856	63.698	25.323
K-means	stats	kmeans()	156.309	87.245	39.909
PAM	cluster	pam()	164.266	101.880	**44.345**
Fuzzy clustering	cluster	fanny()	166.989	76.564	32.782
Fuzzy C-Means	e1071	cmeans()	158.9717	37.172	16.719

$$VRC = \frac{SS_B/(K-1)}{SS_W/(n-K)} \tag{3}$$

where K is the number of clusters, n is total number of profiles to be clustered, SS_B is the sum of the squares between clusters, and SS_W is the sum of the squares within the clusters. Note that, the larger the value of VRC, the better the performance of the clustering method.

According to Table 1, the largest VRC among all clustering methods is PAM. Note that, PAM requires the prior knowledge of the number of clusters in advance. Users can either determine the number of clusters by using Hierarchical Clustering (hclust) or using mclust method, in which the prior knowledge of the number of clusters is not required. For more information regarding to clustering analysis methods in R please refer to Leisch and Gruen (2013).

3.2.1. Partition around medoids

The PAMs clustering algorithm proposed by Kaufman and Rousseeuw (2005) is the first known algorithm of k-medoids clustering method (Han, Kamber, & Pei, 2006). Unlike a k-means algorithm that calculates the mean value of the cluster (centroid) as representative object of the cluster, a k-medoids algorithm uses the actual data point to represent the cluster. The objective of PAM is to minimize the cost function, i.e., sum of dissimilarities between given data points and the medoids as shown in Eq. (4). Note that, the measurement of dissimilarity between objects can be calculated by Euclidean distance or Manhattan distance as described in Kaufman and Rousseeuw (2005). A PAM cost function F is defined as:

Randomly select K FFT-Transformed profiles as medoids of the clusters
Repeat
  Assign each non-medoid FFT-Transformed profile to the nearest medoid using the dissimilarity function
  For each medoid m
    For each non-medoid FFT-Transformed profile W^f
      Swap m and W^f then calculate F
  Determine the new set of K medoids with the lowest F
Until no change in the medoid.

Fig. 8. PAM algorithm.

$$F = \sum_{j}^{K} \sum_{W^f \in C_j} d(W^f, m_j) \tag{4}$$

where K is the number of clusters, W^f defined in Eq. (2) is a FFT-transformed profile assigned in cluster C_j, m_j is the medoid of C_j, which is also a vector of size p_f, and $d(W^f,m_j)$ is the function of dissimilarity between FFT-transformed profile W^f and medoid m_j. The algorithm of PAM is shown in Fig. 8. The output of PAM is the FFT-transformed profile with its class label.

3.3. Classifier component

In data mining, classification analysis constructs a model or classifier to predict the categorical labels. There are two steps in the classification analysis, training phase (or learning phase) and classification step. Training phase builds a classifier by learning from a given training dataset with class labels predetermined, while the classification step classifies the new data without a given class label to one of the associated classes (Han et al., 2006). In Fig. 5, a classifier is needed to identify a new profile's membership. However, it needs to be trained first during the phase I process. The training dataset is generated from a clustering method, whose output contains the attributes of FFT-transformed profiles along with class labels. Then, the proposed classifier is constructed by learning from the trained data. Once the classifier is trained, it assigns a profile to an appropriate cluster for phase II process monitoring. The classifier used in this study is the SVM classifier.

The SVM classification method has become an indispensable classifier in machine learning and pattern recognition field. SVM is adopted as the proposed phase II classifier in that it is one of the most competitive classification methods. Meyer, Leisch, and Hornik (2003) confirmed that the SVM is the best classifier among 16 popular classifiers. Details of the theory and application of SVM may be found in Cortes and Vapnik (1995) and Campbell and Ying (2011).

The main idea of SVM classifier is mapping input variables into higher dimensional space using a kernel function to distinguish nonlinearly separable datasets. The choice of a kernel function may result in different accuracy rates in the same problem domain. Many kernel functions have been embedded in packages of R. For example, linear, polynomial, Gaussian radial basis function kernel (RBF), and sigmoid kernel function. We will apply the svm() function in the e1071 R package to various training and test datasets. We will also examine the following kernel functions for the best performance in term of the accuracy rate in phase II process:

$$\text{Linear Kernel:} \quad K(u, v) = u'v \tag{5}$$

$$\text{Polynomial Kernel of degree } h\text{:} \quad K(u, v) = (u'v + 1)^h \tag{6}$$

$$\text{Gaussian Radial Basis Function Kernel (RBF):} \quad K(u, v) = e^{-\|u-v\|^2/2\theta^2} \tag{7}$$

$$\text{Sigmoid Kernel:} \quad K(u, v) = \tanh(\kappa u'v - c) \tag{8}$$

Note that u and v are both p-dimensional vectors, and h, θ, κ, and c are all parameters that determined by users. For more details in adjusting kernel function, please refer to Karatzoglou, Meyer, and Hornik (2006).

3.4. Dimension reduction component

Montgomery (2009) suggested that the maximum number of dimension for a multivariate control chart should be smaller than ten. When the quality characteristic is defined as all observations in a profile, the dimensions may be in hundreds or thousands. In this case study, the FFT-transformed profile still contains 256 data points. No multivariate control chart can handle such a large dimension effectively. Therefore, various traditional profile analysis techniques have leveraged dimension reduction methods before applied multivariate control chart to the quality characteristics. They include: (1) wavelet transformation (Reis & Saraiva, 2006; Zhou et al., 2006; Chicken et al., 2009); (2) Principal Component Analysis (Ding et al., 2006; Noorossana, Amiri, & Soleimani, 2008; Shiau et al., 2009); (3) B-Spline Fitting (Chang & Yadama, 2010; Walker & Wright, 2002; Williams et al., 2007).

Choosing an appropriate dimension reduction method, we need to consider the balance between computational cost and performance. The Euclidian distance (ED) method seems to be a good candidate. Not only this approach reduces the number of dimensions to one, but also the steps of calculating the ED between two profiles are simpler than those of the other dimension reduction methods. In this study, we will compare the proposed method to the other methods. All methods studied and their associated R packages used in this study are listed as follows: (1) discrete wavelet transformation: dwt() in wavelets package; (2) principal component analysis: prcomp() in stats package; (3) cubic B-spline (BS): ns() in splines package; (4) ED: dist() in stats package. Since a multivariate control chart, e.g., Hotelling's T^2 or MEWMA, is only effective in handling the number of dimension of less than ten, the number of dimensions reduced by wavelet transformation is specified to eight (or less), while the maximum number of dimensions that reduced by PCA is no larger than ten.

3.5. Control chart component

In the proposed SPC framework, a control chart plays the role of decision making. Both multivariate and univariate control charts with individual observation can be used. If the dimensionality of the dimension is reduced to one, individual X control chart (*IX* chart) can be obtained. On the other hand, when the number of dimensions is larger than one after running the dimensional reduction method, the Hotelling's T^2 control chart with individual observations can then be used. We can substitute the *IX* chart with either a EWMA or a CUSUM chart, and the Hotelling's T^2 control chart with a MEWMA chart depending on the magnitude of the expected shifts. Montgomery (2009) provides details of both control charts.

To establish the individual control chart based on the independent observations x_i, $i = 1, 2, \ldots n$, the I*X* control chart parameters are formulated as follows:

$$\text{UCL} = \bar{x} + 3\overline{\text{MR}}/d_2, \tag{9}$$

$$\text{CL} = \bar{x} \tag{10}$$

$$\text{LCL} = \bar{x} - 3\overline{\text{MR}}/d_2, \tag{11}$$

where $\bar{x}$ is the average of individual observations while $\overline{MR}$ is the average of moving ranges of consecutive observations, specifically, $\overline{\text{MR}} = \sum_{i=2}^{n} |x_i - x_{i-1}|/(n-1)$, and $d_2 = 1.128$.

To construct the Hotelling's T^2 control chart with individual observations (Tracy, Young, & Mason, 1992), one can follow the formula:

$$T_i^2 = (x - \bar{x})_i' S^{-1} (x - \bar{x})_i \tag{12}$$

$$\text{UCL} = \frac{p(n+1)(n-1)}{n^2 - np} F_{\alpha,p,n-p} \tag{13}$$

$$\text{UCL} = \frac{(n-1)^2}{n} \beta_{\alpha,p/2,(n-p-1)/2} \tag{14}$$

$$\mathrm{LCL} = 0 \tag{15}$$

$$\mathrm{S} = \frac{V'V}{2(n-1)} \tag{16}$$

$$V = [v_1'v \cdots v_{n-1}']' \tag{17}$$

$$v_i = x_{i+1} - x_i \tag{18}$$

The statistics T^2 of the i^{th} observation is shown in Eq. (12), where S is calculated by using the Eq. (16). Moreover, S estimates the variance-covariance matrix better than that of using the conventional approach if there was no trend, cycle, etc., in the process. If the process was totally random, the variance-covariance structure determined by Eq. (16) and the conventional approach would have no difference (Holmes & Mergen, 1993). The UCL in Eq. (13) should be chosen when the control chart monitors phase II process, while Eq. (14) shows the UCL in the phase I process. Note that, n is number of profiles, p is number of dimensions, and α is the confidence level in Eqs. (13) and (14). LCL in both phase I and phase II is zero.

4. Experimental design

To optimize and test the proposed SPC implementation framework, we conduct an experiment as shown in Fig. 9. In this experiment, we use 146 in-control condensation water temperature profiles that collected within two months in 2011 from the hose production manufacturing process to construct phase I process. Note that those 146 in-control profiles will be the training data for SVM classifier for phase II process if the clustering method is applied to the phase I process. We also collected extra 39 profiles in different months in 2011 in which 6 of those were abnormal profiles identified by quality engineers' and our judgments. Those abnormal profiles are profile 16, 24, 26, 27, 37, and 39 as shown in Fig. 10, which are superimposed on overall 146 in-control phase I profiles. According to Fig. 10, the abnormal profile 16, 24, and 27 have unusual period that the condensation water valve was kept closed after 200-time units. Profile 26 indicated that the valve was kept closed between 200- and 390-time units. The profile 37 is an obvious failure run of the hoses manufacturing process while the profile 39 indicates that the water valve opened and closed too frequently comparing to a normal pattern.

Moreover, since the proposed framework consists of machine learning methods, we investigate the performance in terms of accuracy rate with and without applying clustering/classification. Note that, all methods that without applying machine learning methods are categorized as traditional profiles analysis methods, such as, DWT + Hotelling's T^2, PCA + Hotelling's T^2, BS + IX, and ED + IX. The clustering method used in the proposed framework is PAM due to its highest VRC value based on the phase I data described in the above section. PAM is used to determine the clusters and their memberships which are then used as the training data for a SVM classifier. Because different kernel function in SVM classifier may result in different accuracy rates, therefore, four popular kernel functions specified in Fig. 9 will be investigated in this experiment so that we can optimize the proposed framework. The performance criteria used in this experiment, false alarm rates and accuracy rate, are introduced in the next section.

5. Performance comparison

We use false alarm rates to evaluate the proposed framework in a phase I process, and the accuracy rate in a phase II process in this study. The accuracy rate can be calculated from the information listed in a confusion matrix that is often used in the machine learning field. The true positives (TP) are the number of in-control observations assigned to the in-control group, while the true negatives (TN) are out-of-control observations classified as the out-of-control group. If in-control

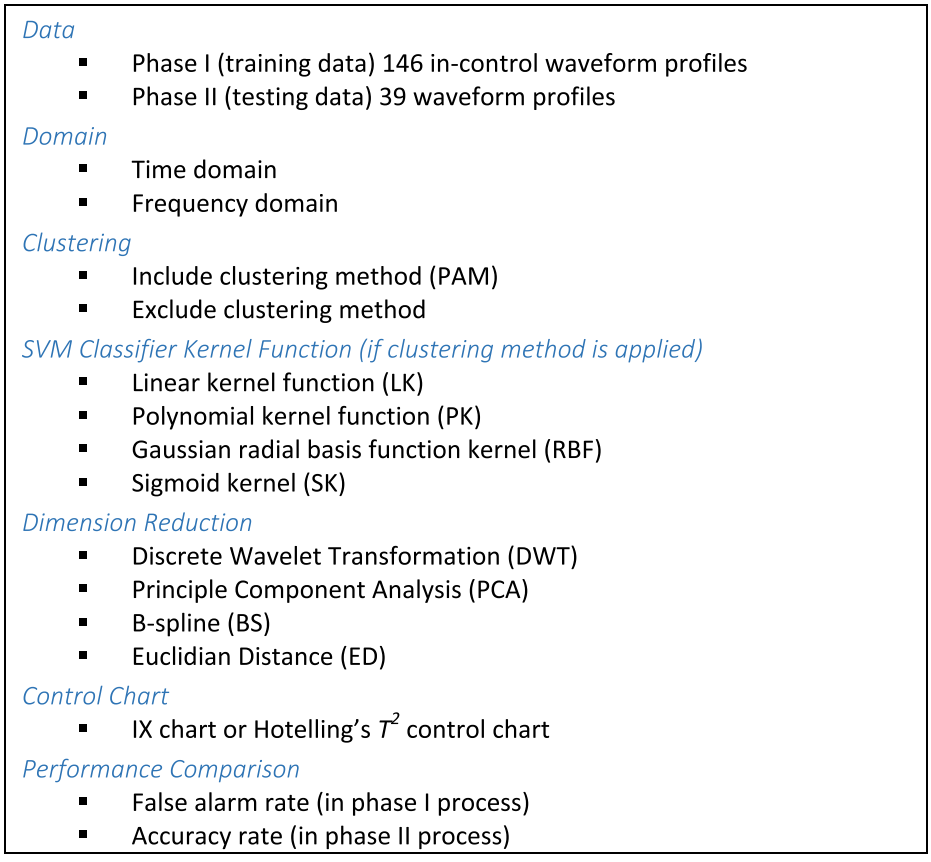

Fig. 9. The experimental design of the proposed SPC implementation framework for the condensation water temperature profiles.

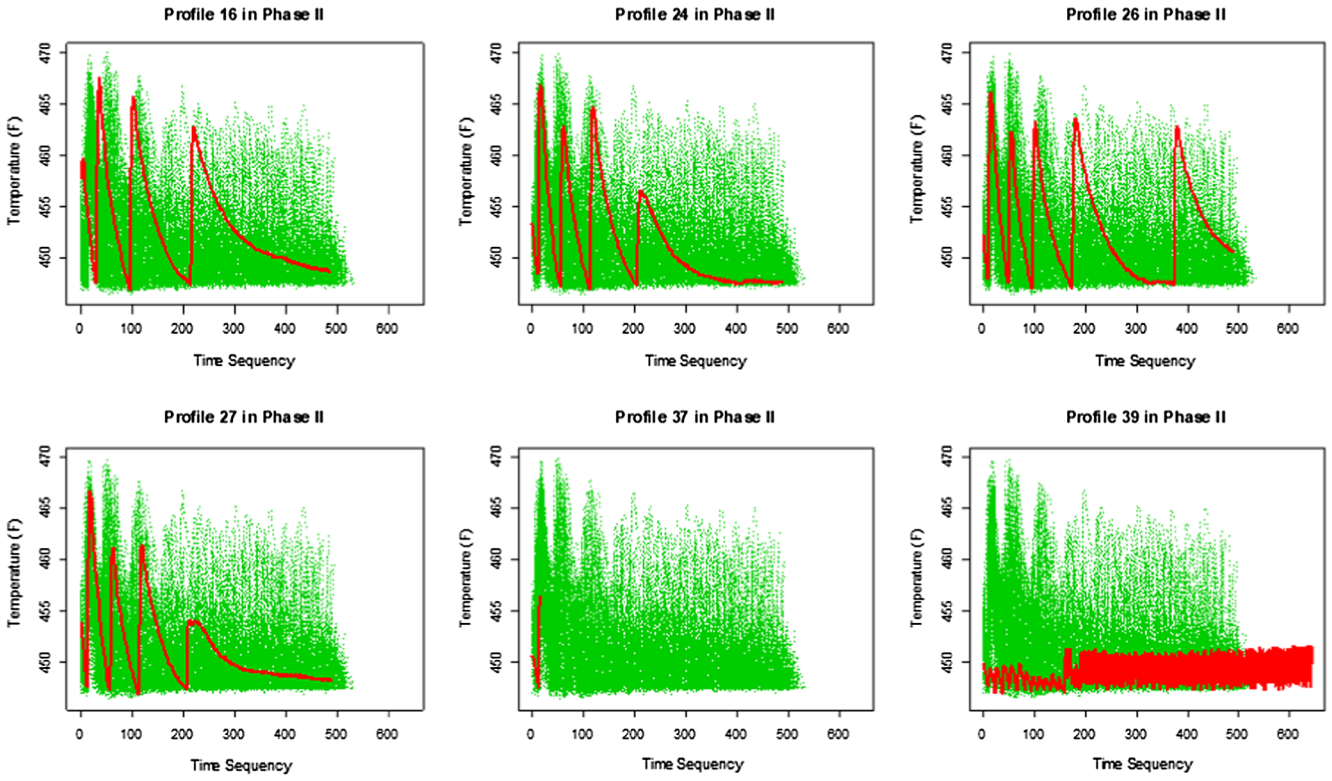

Fig. 10. Six abnormal condensation water temperature profiles in phase II process (thick solid line: abnormal profiles; thin dot lines: overall 146 in-control profiles in phase I).

Table 2
Performance results of the experiment for phase I data.

	Domain	dim. Reduction	# of false alarm	False alarm rate
The proposed framework	FFT	BS	1	**0.0068**
		ED	2	**0.0137**
		DWT	19	0.1301
		PCA	28	0.1918
Traditional profile analysis tools	Time	BS	3	0.0205
		ED	5	0.0342
		DWT	10	0.0685
		PCA	15	0.1027
	FFT	BS	0	**0.0000**
		ED	5	0.0342
		DWT	11	0.0753
		PCA	16	0.1096

w/: with clustering method; w/o: without clustering method; WT: wavelet transformation; PCA: principle component analysis; BS: cubic b-spline; ED: Euclidian distance.

observations are assigned to the out-of-control group, they are called false negatives (FN). On the other hand, when out-of-control observations are classified to the in-control group, they are called false positives (FP). Moreover, according Han et al. (2006), the accuracy rate can be defined in Eq. (19). The accuracy is a function of sensitivity and specificity defined in Eqs. (20) and (21), respectively. Sensitivity and specificity are known as true positive rate and true negative rate, i.e., the proportion of positive tuples and negative tuples are all correctly identified. The accuracy rate is a good indicator of optimizing the proposed framework because it provides overall performance criteria, such as, sensitivity and specificity, in one value.

$$accuracy = sensitivity\frac{(TP + FN)}{(TP + FN + FP + TN)} + specificity\frac{(FP + TN)}{(TP + FN + FP + TN)} \tag{19}$$

$$sensitivity = \frac{TP}{(TP + FN)} \tag{20}$$

$$specificity = \frac{TN}{(FP + TN)} \tag{21}$$

6. Analyses and interpretations

In this section, the performance results of the experiment expressed in above section will be shown and discussed. We will discuss the results in the two phases of process monitoring, i.e., phase I and phase II process.

6.1. Phase I process

In the phase I process, we examined the performance of four dimension-reduction methods on frequency domain with and without applying a clustering method based on 146 in-control condensation water temperature profiles collected from the hoses manufacturing process. The performance results of the experiment for phase I process are shown in Table 2 in ascending order of the false alarm rate. From Table 2, the analysis in the frequency domain dominates the top three spots. Specifically, the cubic BS method shows the lowest false alarm rate as well as smallest number of false alarms with and without applying clustering method among all dimension reduction methods. Moreover, according to the bar chart of the four dimension- reduction methods shown in Fig. 11, BS and ED method performs better (i.e. with smaller average false alarm rates) than other methods in phase I process. In fact, both BS and ED methods show very competitive results in frequency domain with or without applying the clustering method to the waveform profiles in phase I process as shown in Table 2. We examine the top three methods that result the first three lowest false alarm rate in phase I data, i.e., FFT + BS, FFT + PAM + BS, and FFT + PAM + ED. Note that, the BS method we used in the experiment

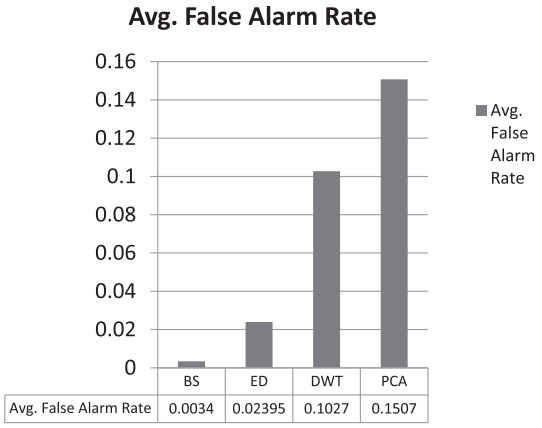

Fig. 11. Bar chart of the four dimension reduction methods and their average false alarm rates.

is modified from one segment method of Chang and Yadama (2010). Specifically, the output of the BS is calculated by sum of absolute value of deviation between the cubic BS fitting curve of the FFT-transformed profile and the mean profile of the cluster. The mean profile of each cluster is also generated by cubic-BS fitting. Therefore, the dimensionality of the problem is reduced from 256 to one.

Fig. 12, Fig. 13, and Fig. 14 show the *IX* charts for these three methods. Note that, although FFT + BS' false alarm rate is zero, the control limits calculated in IX chart maybe misleading. Considering there are three distributions within the 146 profiles, it is not proper directly applying an IX chart to the data. In our case, those 146 data points didn't pass the normality test (Shapiro-Wilk Normality Test, P-Value < 0.01), therefore the control limited in IX chart may not reflect the proper range for those data points and cause higher type II error. And we do not suggest using FFT + BS in this case. On the other hand, we examined the normality (Shapiro-Wilk) for those three clusters determined by PAM, the P-Value of cluster 1, 2, and 3 are 0.092, 0.742, and 0.81. In other words, we cannot reject those three groups are three normal distributions, thus, IX charts are rational to use on those three groups respectively.

Although Fig. 13 and Fig. 14 shows three out-of-control points in total in the IX chart, it is a reasonable step to remove those out-of-control points from the *IX* chart during constructing phase I process control chart. After those three points were removed from the cluster 1 in Fig. 13 and Fig. 14, the results of *IX* charts of cluster 1 by using FFT + PAM + BS and FFT + PAM + ED show that both methods are ready for phase II monitoring since no point is outside the control limits as shown in Fig. 15. Note that FFT + BS in Fig. 12 only uses one control chart for phase II monitoring because no clustering method is used to assigned profiles into homogeneous groups.

6.2. Phase II process

Phase II data contains 39 condensation water temperature profiles, and 6 of which are recognized as abnormal waveform profiles by quality engineers and our judgments as shown in Fig. 10. The performance results of the experiment for phase II data are shown in Table 3 in the descending order of the accuracy rate defined in Eq. (19). In Table 3, the method that combines the SVM classifier with RBF kernel function and ED has the best performance due to its largest accuracy rate among all methods studied. Moreover, the first two highest accuracy rates that were applied SVM classifier with ED dimension reduction method on frequency domain to the phase II data provide sensitivity rate of 1. In other words, they can identify all in-control profiles correctly, that is, no profile has been falsely detected as out of control in the phase II process using those methods. Even though the specificity rates that provided by SVM_{RBF} + ED and $SVM_{sigmoid}$ + ED are not the highest rate, their overall performance scores in terms of accuracy rates are higher than the other methods.

Although the FFT + BS shows the lowest false alarm rate in phase I process, its specificity rate is 0.33 and the overall performance in terms of accuracy rate is 0.87. In other words, the proposed SPC implementation framework with machine learning that applies FFT and clustering/classification to the waveform profile analysis can not only provide a solid phase I process control chart but also construct competitive phase II process control chart than other methods examined in this study. Therefore, based on the sensitivity rate of 1, specificity rate of 0.83, and the overall performance in terms of accuracy rate of 0.97, we recommend the combination of the FFT, the clustering (PAM)/classification (SVM with RBF kernel function) method, and the dimension reduction approach (ED).

7. Instruction for using the proposed framework

The proposed SPC with machine learning methods use wide varieties of signal processing and statistical techniques. Bridging the model development and practical implantation, we offer a step-by-step

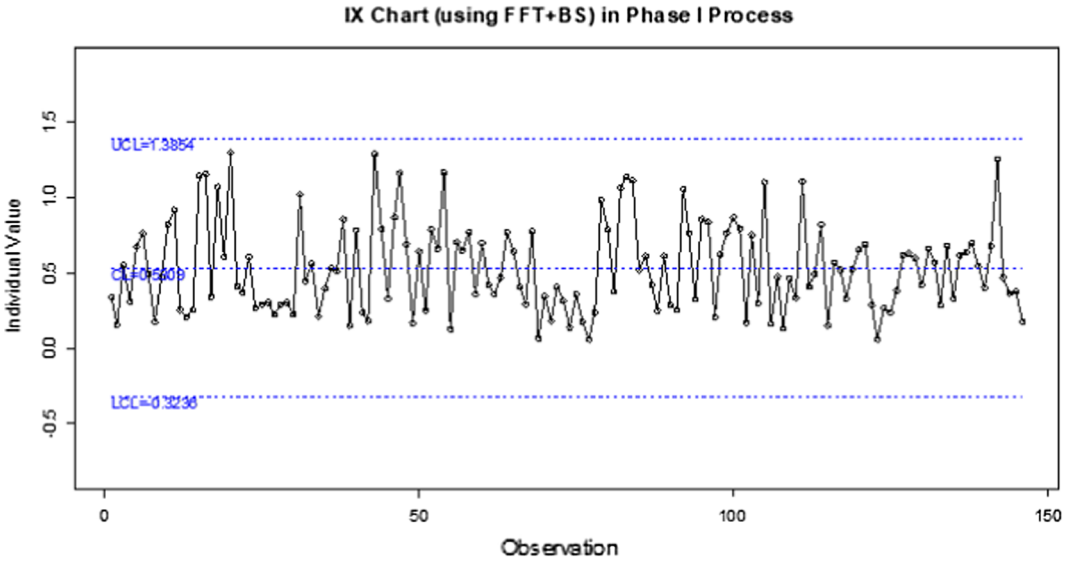

Fig. 12. The IX chart of 146 in-control waveform profiles in phase I process using FFT + BS method.

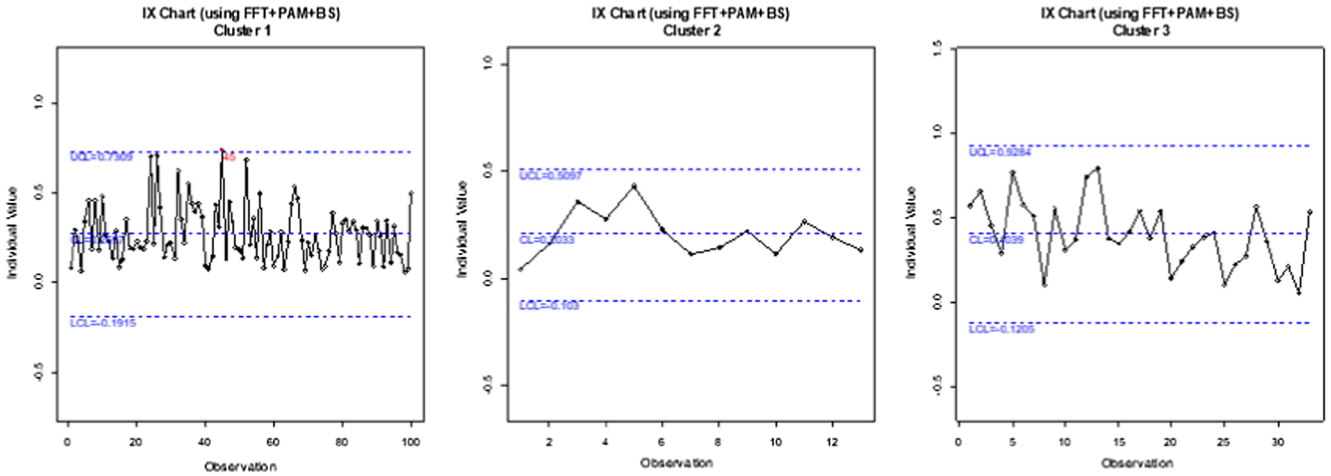

Fig. 13. The IX charts of 146 in-control waveform profiles in phase I process for each cluster using FFT + PAM + BS method.

instruction to implement the proposed methods.

7.1. Construction of the phase I process:

1. Screen out all abnormal wave profiles and treat the rest of profiles as phase I data.
2. Convert all phase I data from time domain to frequency domain using FFT.
3. Apply the PAM cluster method to all FFT-transformed profiles. If users want to use another cluster method, performance of the new cluster method should be compared with PAM in terms of VRC as shown in Eq. (3).
4. Once all profiles have been clustered to their associated groups, the users should—
   4.1. Use information of the clusters and their associated memberships to train a classifier model (SVM) for the phase II process.
   4.2. Utilize a dimension-reduction method (ED) for all profiles within a homogenous group. If users would like to try new dimension-reduction method, and accuracy rate should be utilized for performance comparison.
5. A control chart is then constructed for each cluster. If we have K clusters, then we will have K control charts. All information, such as mean, variance, and control limits of each cluster, will be used in the phase II process.
6. If batch combinations are changed so that new acceptable clusters emerge, the phase I process should be repeated again. Specifically, quality engineers can repeat steps 1 to 3 for data collected in a future production period. If additional clusters are generated, the proposed phase I process should be repeated to establish new control charts.

7.2. Construction for phase II process

1. Convert a new wave profile from time domain to frequency domain using FFT.
2. Apply a trained classifier (SVM) to the FFT-transformed profile and determine which cluster the new profile belongs.
3. Apply the dimension-reduction method (ED) to the new FFT-transformed profile to calculate the plotting statistic.
4. Plot the statistic in step 3 to control chart k with corresponding mean, variance, and control limits. If the statistic exceeds control limits, the new wave profile is deemed an abnormal profile.

8. Conclusions and recommendations

Profile monitoring using SPC has been studied in manufacturing process in recent years. Many researches provided successful approaches in their problem domain. However, no study has been found in the monitoring of process stability when the profile shape is a waveform without a gold standard. Since waveform profiles generated from a manufacturing process may consist of various magnitudes and frequencies, the process monitoring problem becomes more challenging

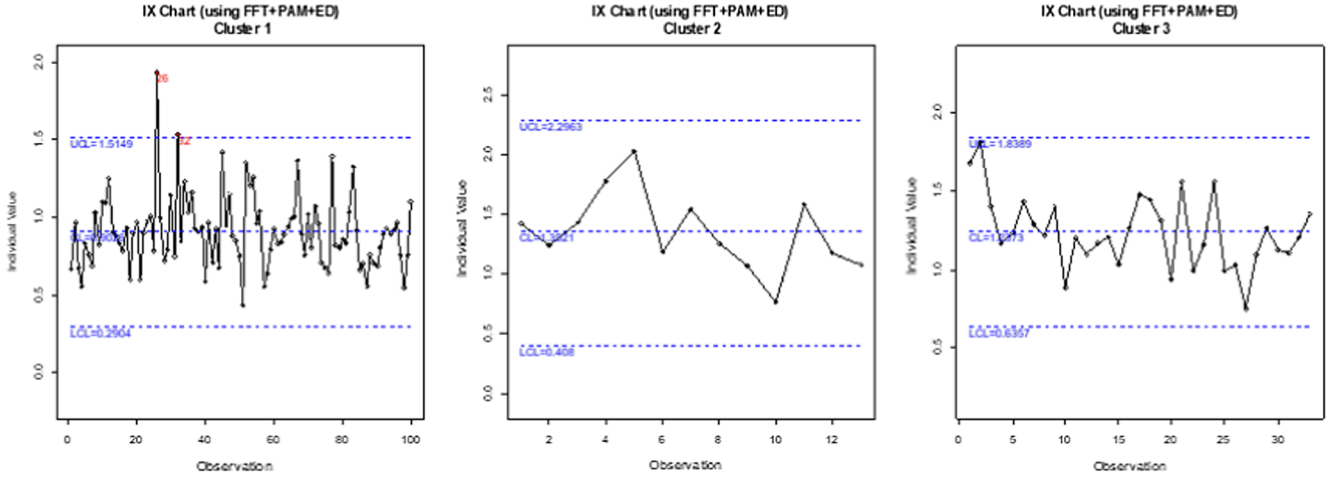

Fig. 14. The IX charts of 146 in-control waveform profiles in phase I process for each cluster using FFT + PAM + ED method.

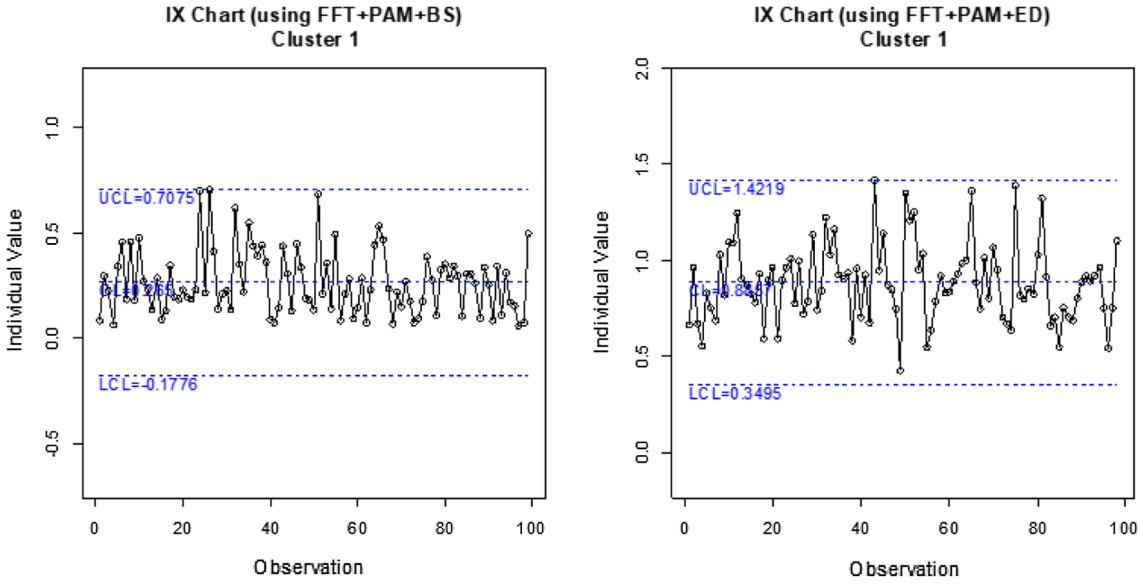

Fig. 15. The IX charts of Cluster 1 using FFT + PAM + BS and FFT + PAM + ED with out-of-control points removed in constructing phase I process.

Table 3
Performance results of the experiment for phase II data.

	Domain	Dim. Reduct.	kernel	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy
The proposed framework	FFT	ED	RBF	33	5	1	0	1.0000	0.8333	**0.9744**
			Sigmoid	33	4	2	0	1.0000	0.6667	**0.9487**
			Linear	33	2	4	0	1.0000	0.3333	0.8974
			Poly	25	5	1	8	0.7576	0.8333	0.7692
		BS	RBF	26	6	0	7	0.7879	1.0000	0.8205
			Sigmoid	26	6	0	7	0.7879	1.0000	0.8205
			Linear	26	6	0	7	0.7879	1.0000	0.8205
			Poly	24	6	0	9	0.7273	1.0000	0.7692
		DWT	RBF	23	1	5	10	0.6970	0.1667	0.6154
			Sigmoid	23	1	5	10	0.6970	0.1667	0.6154
			Linear	23	0	6	10	0.6970	0.0000	0.5897
			Poly	24	1	5	9	0.7273	0.1667	0.6410
		PCA	RBF	30	5	1	3	0.9091	0.8333	0.8974
			Sigmoid	30	5	1	3	0.9091	0.8333	0.8974
			Linear	30	3	3	3	0.9091	0.5000	0.8462
			Poly	30	3	3	3	0.9091	0.5000	0.8462
Traditional profile analysis tools	Time	ED	NaN	33	1	5	0	1.0000	0.1667	0.8718
		BS		33	1	5	0	1.0000	0.1667	0.8718
		DWT		26	5	1	7	0.7879	0.8333	0.7949
		PCA		33	1	5	0	1.0000	0.1667	0.8718
	FFT	ED		33	0	6	0	1.0000	0	0.8462
		BS		32	2	4	1	0.9697	0.3333	0.8718
		DWT		31	5	1	2	0.9394	0.8333	**0.9231**
		PCA		32	1	5	1	0.9697	0.1667	0.8462

w/: with clustering method; w/o: without clustering method; WT: wavelet transformation; PCA: principle component analysis; BS: cubic b-spline; ED: Euclidian distance; linear: linear kernel function; RBF: RBF kernel function; poly: Polynomial kernel function; sigmoid: sigmoid kernel function.

due to the homogeneity issue. In this study, we propose a SPC implementation framework that consists of FFT and the clustering/classification method as well as the dimension reduction approach to monitor waveform profiles. The proposed framework can identify abnormal wave profiles with minimal false alarms based on both phase I and phase II data sets.

We compared the proposed method to a few widely used dimension reduction techniques in profile analysis, such as, wavelet transformation, principle component analysis, and BS transformation. We also considered applying clustering/classification method on frequency domain. The phase I and phase II datasets in the experiment are from the condensation water temperature profiles that collected from the curing process of the high-pressure hoses. According to Table 2, the proposed framework in phase I (FFT + PAM + ED) constructs a solid phase I process control chart with competitive performance in terms of false alarm rates after removing abnormal data points. In addition, as shown in Table 3, the proposed framework in phase II (FFT + PAM + SVM_{RBF} + ED) dominates other profile analysis techniques with respect to the accuracy rate. In summary, we recommend the use of FFT + PAM + ED in constructing phase I process, and FFT + SVM_{RBF} + ED in phase II. To leverage the proposed framework when new or more advanced clustering, classification, and control charting techniques are introduced in the future, quality engineers may follow the procedure to choose the other methods that provide better performance, or they are more familiar with due to the availability of the computational software and package available to them.

Although the proposed method provides robust results than the other profile analysis techniques mentioned above in this problem domain, new fitting/control charting techniques may provide better performance in the future. We encourage more researchers conduct study

when no gold standard reference exists in a process.

CRediT authorship contribution statement

Shih-Hsiung Chou: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing - original draft, Data Scientist. **Shing Chang:** Supervision, Conceptualization, Methodology, Writing - review & editing. **Tzong-Ru Tsai:** Conceptualization, Methodology, Writing - review & editing. **Dennis K.J. Lin:** Conceptualization, Methodology, Writing - review & editing. **Yunfei Xia:** Software, Validation. **Yu-Siang Lin:** Data curation.

References

Abbas, T., Qian, Z., Ahmad, S., & Riaz, M. (2017). Bayesian monitoring of linear profile monitoring using DEWMA charts. *Quality and Reliability Engineering International. 33*, 1783–1812.

Abdel-Salam, A.-S. G., Birch, J. B., & Jensen, W. A. (2013). A semiparametric mixed model approach to Phase I profile monitoring. *Quality and Reliability Engineering International, 29*(4), 555–569.

Awad, M. I., AlHamaydeh, M., & Faris, A. (2018). Fault detection via nonlinear profile monitoring using artificial neural networks. *Quality and Reliability Engineering International. 34*, 1195–1210.

Campbell, C., & Ying, Y. (2011). *Learning with Support Vector Machines.* Morgan & Claypool.

Chang, S. I., Tsai, T. R., Lin, D. K. J., Chou, S. H., & Lin, Y. S. (2012). Statistical process control for monitoring nonlinear profiles: A six sigma project on curing process. *Quality Engineering, 24*, 251–263.

Chang, S. I., & Yadama, S. (2010). Statistical process control for monitoring non-linear profiles using wavelet filtering and B-spline approximation. *International Journal of Production Research, 48*(4), 1049–1068.

Chicken, E., Pignatiello, J., Jr., & Simpson, J. R. (2009). Statistical process monitoring of nonlinear profiles using wavelets. *Journal of Quality Technology, 41*(2), 198–212.

Chen, S., & Nembhard, H. B. (2011). A high-dimensional control chart for profile monitoring. *Quality and Reliability Engineering International, 27*(4), 451–464.

Cooley, J. W., & Tukey, J. W. (1965). An algorithm for the machine calculation of complex fourier series. *Mathematics Computation, 19*, 297–301.

Cortes, C., & Vapnik, V. (1995). Support-vector network. *Machine Learning, 20*, 273–297.

Ding, Y., Zeng, L., & Zhou, S. (2006). Phase I analysis for monitoring nonlinear profiles in manufacturing processes. *Journal of Quality Technology, 38*(3), 199–216.

Gomaa, A. S., & Birch, J. B. (2019). A semiparametric nonlinear mixed model approach to phase I profile monitoring. *Communications in Statistics-Simulation and Computation, 48*(6), 1677–1693.

Han, J., Kamber, M., & Pei, J. (2006). *Data mining, second edition: concepts and techniques.* Elsevier Science.

Holmes, D. S., & Mergen, A. E. (1993). Improving the performance of the T2 control chart. *Quality Engineering, 5*(4), 619–625.

Hosseinifard, S. Z., Hosseinifard, S. Z., Abdollahian, M., & Zeephongsekul, P. (2011). Application of artificial neural networks in linear profile monitoring Elsevier. https://doi.org/10.1016/j.eswa.2010.09.160.

Hoster, B., Jaunich, M., & Stark, W. (2009). Monitoring of the vulcanisation process by ultrasound during injection moulding. *ndt.net, 14*(9), 1–9.

Jensen, W. A., & Birch, J. B. (2009). Profile monitoring via nonlinear mixed models. *Journal of Quality Technology, 41*(1), 18–34.

Jensen, W. A., Birch, J. B., & Woodall, W. H. (2008). Monitoring correlation within linear profiles using mixed models. *Journal of Quality Technology, 40*(2), 167–183.

Karatzoglou, A., Meyer, D., & Hornik, K. (2006). Support vector machines in R. *Journal of Statistical Software, 15*(9), 1–28.

Kaufman, L., & Rousseeuw, P. J. (2005). *Finding groups in data: An introduction to cluster analysis.* New York: Wiley.

Kazemzadeh, R. B., Noorossana, R., & Amiri, A. (2008). Phase I monitoring of polynomial profiles. *Communications in Statistics—Theory and Methods, 37*(10), 1671–1686.

Kang, L., & Albin, S. L. (2000). On-line monitoring when the process yields a linear profile. *Journal of Quality Technology, 32*(4), 418–426.

Kim, K., Mahmoud, M. A., & Woodall, W. H. (2003). On the monitoring of linear profiles. *Journal of Quality Technology, 35*, 317–328.

Leisch, F. and Gruen, B. (2013). Cluster Analysis & Finite Mixture Models. In CRAN Task View. Retrieved 2/28/2013, from http://cran.at.r-project.org/web/views/Cluster.html.

Mahmoud, M. A. (2008). Phase I analysis of multiple regression linear profiles. *Communications in Statistics - Simulation and Computation, 37*(10), 2106–2130.

Meyer, D., Leisch, F., & Hornik, K. (2003). The support vector machine under test. *Neurocomputing, 55*, 169–186.

Montgomery, D. C. (2009). *Introduction to statistical quality control.* New York, NY: John Wiley & Sons.

Mooi, E. A., & Sarstedt, M. (2011). *A concise guide to market research: the process, data, and methods using IBM.* SPSS Statistics: Springer.

Noorossana, R., Amiri, A., & Soleimani, P. (2008). On the monitoring of autocorrelated linear profiles. *Communications in Statistics-Theory and Methods, 37*(3), 425–442.

Noorossana, R., Saghaei, A., & Amiri, A. (2011). *Statistical analysis of profile monitoring.* Hoboken, New Jersey: John Wiley & Sons, Inc.

Paynabar, K., & Jin, J. (2011). Characterization of non-linear profiles variations using mixed-effect models and wavelets. *IIE Transactions, 43*(4), 275–290.

Qiu, P., Zou, C., & Wang, Z. (2010). Nonparametric Profile Monitoring by Mixed Effects Modeling. *Technometrics, 52*(3), 265–277.

Reis, M. S., & Saraiva, P. M. (2006). Multiscale statistical process control of paper surface profiles. *Quality Technology and Quantitative Management, 3*(3), 263–282.

Shiau, J. J. H., Huang, H. L., Lin, S. H., & Tsai, M. Y. (2009). Monitoring nonlinear profiles with random effects by nonparametric regression. *Communications in Statistics-Theory and Methods, 38*(10), 1664–1679.

Tracy, N., Young, J., & Mason, R. (1992). Multivariate Control charts for individual observations. *Journal of Quality Technology, 24*(2), 88–95.

Walker, E., & Wright, S. P. (2002). Comparing curves using additive models. *Journal of Quality Technology, 34*(1), 118–129.

Williams, J. D., Woodall, W. H., & Birch, J. B. (2007). Statistical monitoring of nonlinear product and process quality profiles. *Quality & Reliability Engineering International, 23*(7), 925–941.

Woodall, W. H. (2007). Current research on profile monitoring. *Produção, 17*(3), 420–425.

Wu, J., Liu, Y., & Zhou, S. (2016). Bayesian hierarchical linear modeling of profile data with applications to quality control of nanomanufacturing. *IIIE Transactions on Automation Science and Engineering. 13*(3), 1355–1366.

Yang, W., Zou, C., & Wang, Z. (2017). Nonparametric profile monitoring using dynamic probability control limits. *Quality and Reliability Engineering International. 33*, 1131–1142.

Zhou, S. Y., Sun, B. C., & Shi, J. J. (2006). An SPC monitoring system for cycle-based waveform signals using Haar transform. *IEEE Transactions on Automation Science and Engineering, 3*(1), 60–72.

Zou, C., Tsung, F., & Wang, Z. (2007). Monitoring general linear profiles using multivariate exponentially weighted moving average schemes. *Technometrics, 49*(4), 395–408.