

Combination of Several Control Charts Based on Dynamic Ensemble Methods

Dhouha Mejri^{1,2,*}, Mohamed Limam^{2,3}, Claus Weihs¹

¹Department of Computational Statistics, Faculty of Statistics, Technical University of Dortmund, Dortmund, Germany

²Tunis Higher Institute of Management (ISG), University of Tunis, 2000, Tunisia

³Dhofar University, Oman

Copyright ©2017 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract Combining methods from Statistical Process Control (SPC) in order to benefit from more than one method's efficiency has been recently challenged. One of the reasons is that real life problems change overtime and a small improvement can lead to a very big profit. Ensemble methods from data mining domain have recently shown their effectiveness when used with SPC. The first combined control chart based on dynamic ensemble method, called Dynamic weighted control chart, is designed especially for monitoring concept drift in online processes. This article presents a new model of combining more than two control charts based on ensemble methods as well as error rates classifications to optimize the shift identification and control. This method can be applied for offline and online processes. It is based on a three step learning model: first a preprocessing step to prepare the data for classification. Second, an ensemble method based on Dynamic Weighted Majority (DWM) is applied to aggregate the decisions of the different charts at the end of the each batch. Finally, shifts are identified based on the misclassification error rates of DWM. Dynamic Ensemble Control chart model benefits from the knowledge from classification and control to give a most precise information about the process. Experiments have shown that the latter is better than the use of individual charts and classifies the variable which is responsible for the out of control.

Keywords Shift Detection, Combined Control Charts, Ensemble Methods, Dynamic Learning

1 Introduction

Detecting concept drift and identification of the causes of the processes out of control present one of the most interesting theme in industry, manufacture and services. Thus combining control charts (CC) is one of the very interesting

methods in Statistical Process Control (SPC) for many researchers. [1] has recently proposed a new heuristics combination of CC based on an online dynamic learning methods. The method uses a diversity of CCs, removes and adds CCs based on an online learning model especially designed for online data stream processes. [2] introduced a CC that monitors both the mean vector and the covariance matrix using a variable sampling interval (VSI). They propose a VSI CC with another Variable Sampling Rate (VSR) CC based on sequential sampling. This method aims to detect shifts in the mean and the variability of the process based on a combined multivariate EWMA (MEWMA)-type chart which is proved to be more performant than traditional CCs. [2] find it difficult to choose whether to use a Shewhart-type control to monitor large shift in the mean or a MEWMA chart to monitor small shift in the process mean vector. For that, authors considered combining multivariate Shewhart and MEWMA chart to control both the mean vector and covariance matrix. In fact, when monitoring a process that has multivariate normal variables, the Shewhart type CC traditionally used for monitoring the process mean vector is effective for detecting large shifts. However, for detecting small shifts, it is more effective to use the MEWMA CC. Based on a simulation of different combinations of MEWMA and Shewhart chart, it has been proved by [3] that combining two MEWMA CCs is better than combining a MEWMA and Shewhart chart.

Additionally, authors prove that using three combined CCs is better than two-combined CCs under some conditions. However [3] prove that the proposed three CCs combination is sometimes worse than a two-combined CCs for detecting shift in the variability. To determine the coefficients of the combined model, [4] used linear regression where predictors are the forecasters and the actual value is the dependent variable. [5] proposed a combination of Shewhart chart with a square regression for disease bio-surveillance. [6] presented a combination of Moving average (MA) and autoregressive (AR) model.

To treat the above-mentioned issues, this article proposes

the Dynamic Ensemble Control (DEC) chart model. It combines different CCs based on dynamic ensemble methods for concept drift, uses all the knowledge stored in different charting statistics of each individual chart, combines their decisions and monitors both large and small process shift simultaneously. The proposed combination benefits from the online characteristic of DWM-WIN algorithm of [7] in detecting the state of the process in nonstationary environment. It consists of three steps: first transforming the task of determining the state of the process into a classification problem by treating CCs as attributes of the data where the drift has to be predicted. Second, DWM-WIN is applied as an ensemble method to combine the different CCs. Third, misclassification error rates of DWM-WIN are monitored based on the time adjusting CC for concept drift detection.

The proposed model would first benefit from all the information stored in each charting statistic. Second, it would be flexible to combine all CCs type thanks to the use of DWM algorithm that is applied as a method for aggregating the different CCs decisions. Also due to the idea of monitoring the misclassification error rates of DWM-WIN, the proposed DEC chart model would be able to learn the shift which occurred in the process, facilitate the shift detection and reduce the fault detection rate. Also, because error rates are very informative about the process behavior, some specific monitoring methods could be more suitable than others. This was resolved by using time adjusting CCs for concept drift detection to ensure the correctness of the shift identification. In this regard, we propose an appropriate ensemble chart model to detect all shift sizes under the non stationarity assumption.

This article presents a new offline CC combination method which is based on three different CCs using a dynamic ensemble method that copes with concept drifting data streams: the DWM-WIN algorithm. The proposed combination benefits from the online characteristic of DWM-WIN algorithm in detecting the state of the process when a stream of data arrives over time. It consists of two steps: first constructing the data based on the combination of the charting statistics. Second, DWM-WIN is applied as an ensemble method to combine the decisions of the different CCs. A normal distribution with different shift values is used to simulate the combined CC.

The article is outlined as follows: Section 2 introduces dynamic system modeling. Section 3 and 4 present an overview of the different individual CCs and their shortcomings. In Section 5, the problem of shift classification is discussed. Section 6 introduces the DEC model. Section 7 and 8 detail the experiments and the comparisons conducted for this research. Section 9 concludes this article.

2 Dynamic systems modeling

In SPC domain, CCs are usually designed to detect constant shift levels. However, in many industrial applications, the changes are time varying. Thus, it would be of interest to explore the dynamic nature of the shift and to think about

new methods to get a suitable identification of time varying process changes.

[8] defined dynamic processes as the manner in which process variables perform, react and influence each other. [9] showed that in a dynamic process it is necessary to define a time interval called "process transition period" rather than immediately respond to the abrupt change in the process. During this process transition from one magnitude to another, a dynamic modeling system should be applied.

Dynamic systems find their applications in mechanical engineering [10] or industrial [9] processes. An adaptive forecast based monitoring approach was proposed by [9] to model a dynamic process for plastic extrusion. Their approach is based on fitting an ARIMA time series model to process data and then monitor the one step ahead forecast errors with traditional charts, such as EWMA, CUSUM or others. Several dynamic modeling systems were discussed in [11].

3 Detecting a change using control charts

3.1 EWMA chart

EWMA chart developed by [12] is a type of CC used to monitor either variables or attributes-type data using the monitored business or the industrial process entire history of output. Two parameters have to be defined in EWMA chart: $\lambda \in [0, 1]$ describing the weight given to the most recent subgroup mean and L denoting the rational subgroup standard deviation in the control limits (CLs) and it is generally set at 3. An illustration of EWMA chart in a simulated dataset is given in Figure (1). Table (2) details formulas about the charting statistics and the CLs as well as the output decision.

3.2 XBAR chart

The Shewhart chart was introduced by [13]. It monitors the process over time based on the mean of the series of instances called subgroups. It is a time based chart which stores the process history over time. Shewhart-XBAR charts are efficient in detecting large shifts but not small ones. Figure (1) provides an illustration and detailed formulas of XBAR chart. More discussion of these charts will be provided in the next Section.

3.3 CUSUM chart

The CUSUM chart was first introduced by [14]. Then, [15], [16] and [17] developed its mathematical principles. CUSUM is an efficient alternative to Shewhart procedures. It was constructed to overcome shortcomings of Shewhart CCs. It is well suited for small and moderate mean shift detection thanks to its cumulative sum function. Equations and illustrative plots are given in Table (1) and Figure (1) respectively. We mention that S_t , E_t and C_t mentioned in Table (1) are the decision rules used to identify the out of control situations in XBAR, EWMA and CUSUM charts.

Table 1. Properties of three control charts.

	XBAR	EWMA
Statistic	$Shewhart_t = \bar{X}_t = \sum_{i=1}^{i=n} X_i/n$ [18]	$EWMA_t = \lambda \cdot Y_t + (1-\lambda) \cdot EWMA_{t-1}$ [18]
UCL	$\bar{X} + L \cdot \hat{\sigma} / \sqrt{n}$	$EWMA_0 + L \cdot (S/\sqrt{n})\sqrt{\lambda/(2-\lambda)}[1 - (1-\lambda)^{2i}]$
LCL	$\bar{X} - L \cdot \hat{\sigma} / \sqrt{n}$	$EWMA_0 - L \cdot (S/\sqrt{n})\sqrt{\lambda/(2-\lambda)}[1 - (1-\lambda)^{2i}]$
Dec.	$S_t = \text{if } [Shewhart_t > \text{UCL}]$ or $[Shewhart_t < \text{LCL}]$	$E_t = \text{if } [EWMA_t > \text{UCL}]$ or $[EWMA_t < \text{LCL}]$
CUSUM		
Statistic	$C_t^+ = \max(0, C_{t-1}^+ + (z_t - k))$ for $t = 1, \dots, n$ $C_t^- = \max(0, C_{t-1}^- - (z_t + k))$ for $t = 1, \dots, n$ where k is the reference value and n is the sample size [31]	
UCL	h, where h is the decision boundary.	
LCL	-h	
Dec.	$C_t = \text{if } [C_t^+ > h]$ or $[C_t^- < -h]$	

4 Problem of individual charts application and motivation

Generally CCs have two important aims. First, they are used as a tool to maintain process stability and control. Thus practitioners have to identify the best CC for any monitoring situation. Second, CCs represent a data analysis tool and learning from the process history behavior over time. Despite their performance, individual CCs suffer from some weaknesses. In this paper, we are interested in improving the performances of CUSUM, EWMA and XBAR charts. The performance and the gain obtained by each CC depends on the shift size. To evaluate the CC's performance, first simulated results are based on a normal distribution with mean $\mu = 1$ and a standard deviation $\sigma = 1$ and we have simulated different shift ranges. Figure (2) provides the best selected measure of performance of each CC. **We base our analysis in terms of different evaluation metrics defined as follows:**

True Positive (TP) : It happens when a test signals an alarm in the process when it is not there (true detection).

True Negative(TN) : It happens when a test signals an alarm when it is there.

False Positive (FP) : It happens when a test signals an alarm in the process when it is not there (false detection): type I errors.

False Negative(FN) : It happens when a test does not signal an alarm when it is there (misdetection): type II errors.

Recall : True Positive Rate.

Accuracy : is computed as follows:

$$Accuracy = \frac{(TP + TN)}{N}, \tag{1}$$

where N is the total of the population.

Precision : is the positive predicted rate and it is computed as follows:

$$Precision = \frac{TP}{(TP + FP)} \tag{2}$$

F-measure : Is the harmonic mean of precision, and Recall and is computed as follows:

$$F = 2 \frac{precision \cdot recall}{(precision + recall)} \tag{3}$$

In fact, each CC's performance is highlighted in terms of one or more measures but not all. CUSUM's performance is highlighted in terms of Accuracy and False Positives (FPs). In EWMA, Accuracy and False Negatives (FNs) are the best performance features whereas Recall and FN's are the best performance features for XBAR Chart.

First of all it is obvious that each CC has some advantages and some weaknesses. CUSUM chart has higher accuracy and smaller FP for small values of shift levels showing a good performance for small shift detection. Nonetheless, it has low performance in detecting moderate and large shifts. This is explained by high FN and small accuracy for moderate and large shifts.

The out of control scenario is modelled as follows: $\mu_1 = \mu_0 + \delta \sigma_0$, where μ_1 is the new mean after the shift has occurred, μ_0 is the initial mean of the distribution, σ_0 is the initial variance of the distribution and δ is the shift size. As compared to XBAR and EWMA charts, CUSUM has better ability of small shift detection. For example, as shown in Figure (2) the ability of CUSUM chart in detecting a shift $\delta = 0.5$ increases in terms of accuracy compared to a shift $\delta = 3$ from 0.12

to 0.96, respectively. Also, the ability to reduce the FPs in CUSUM chart while detecting a small shift is more pronounced. For a shift of 1, CUSUM has a FP rate of 0.12 versus a FP of 0.93 for a shift of 4. This is explained by the fact that the charting statistic of CUSUM chart is based on the cumulative sum which makes the detection of small shifts faster. Additional to the weakness of large shift detection, CUSUM has another drawback related to the difficulty to analyze point patterns when all points are highly correlated. Recall and FN measures do not show the effect of CUSUM chart in performing detection of small shifts, that's why we only present the most relevant features of performance.

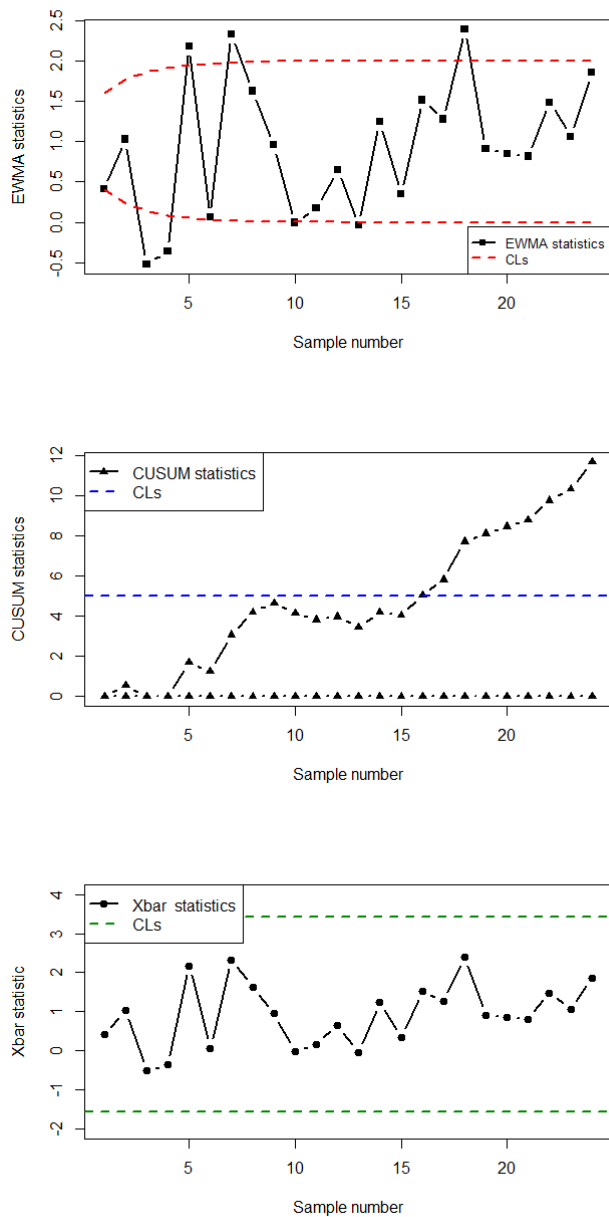


Figure 1. Illustrative plots of individual EWMA, CUSUM and Xbar charts monitoring a $N(1,1)$ with a shift in the mean of 0.5.

shifts. This effect is shown in terms of recall and FNs, where XBAR has a small recall measure when detecting a small shift. However this measure increases for moderate and large shifts. FP also reflects this positive reaction to moderate and large shifts. As an example, when detecting a shift of 1, XBAR registers a recall measure of 0.2 and an FN of 0.8 compared to 0.99 and 1 when detecting a shift of 3.

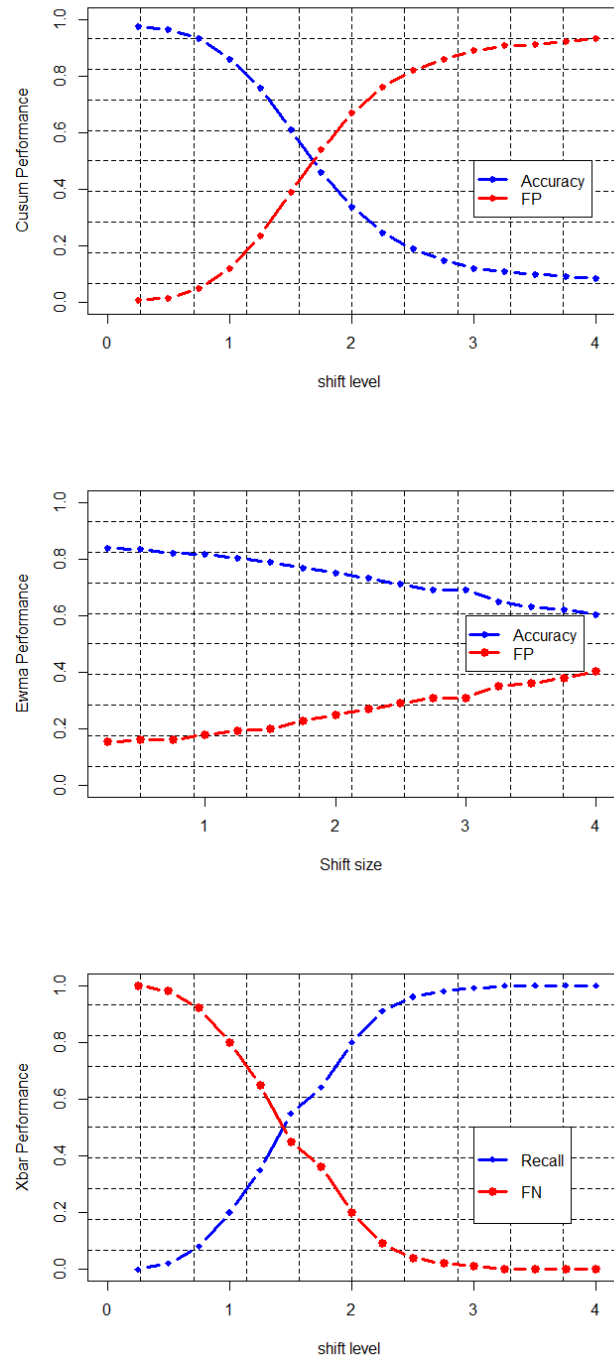


Figure 2. Individual charts's shortcomings in terms of different performance measures where (a) CUSUM chart, (b) EWMA chart and (c) Xbar chart based on $N(1, 1)$ using 400 observations.

On the other hand, XBAR chart is unable to detect small

Besides its weakness in detecting small shifts, XBAR has

another disadvantage that each new point in the monitored process depends only on one subgroup of the data without considering the process history. Moreover, EWMA chart is better in detecting small and moderate shifts than large ones. This is shown in Figure 2 in terms of accuracy and FP. A small decrease in the accuracy and an increase in the FP are shown for large shifts. For example, an accuracy of 0.815 when a shift is 1 decreases by more than 0.2 when detecting a shift of 4. Also, the FN rate increases from 0.15 to 0.36 for a shift of 0.25 and 3.5, respectively.

While being relatively good in detecting different shift values, CUSUM is still better than EWMA and XBAR in detecting small shifts whereas XBAR is better than EWMA and CUSUM in detecting large shifts. The reason of that is that CUSUM chart depends on the entire history of the process making the small shift detection easier. However, Xbar depends only on the last subgroup of the data process improving its ability to detect large shifts, while EWMA is a weighted sum of the recent history. Given these conclusions, we need to find a combination of these different CCs. Because each CC has specific advantage and knowledge. Our aim is to benefit from all these advantages and competencies in detecting several shift ranges. Also, another aim of our approach is to ease the analysis of monitoring a process through the use of dynamic learning control model.

5 Classification of the time varying shift for control

Time varying shift detection presents a serious issue in SPC because it requires a dynamic learning monitoring model as well as a re-designing identification system. Although many machine learning methods were integrated into SPC charts, in order to facilitate the shift identification and to assign a new competence of shift cause identification, the detection of time varying shift in nonstationary environment is still a common issue in SPC. In fact, many techniques use machine learning to improve the identification of special causes of process shift are proposed in the literature.

[19] and [20] have proposed the use of an hybrid ensemble learning model to monitor both the mean and variance simultaneously. A combination of the prediction of several artificial neural networks was proposed by [21] to improve the wind speed and power forecast in the context of a weather prediction application.

Machine learning algorithms were proposed by [22] through using a neural network and a support vector machine to detect a shift in the mean and to determine variable causes of the control. One of the shortcomings of these methods is that learning models were separately applied. Also, the used learning method is not adapted to the process with time varying shifts being based on a fixed learning approach. Many other works have handled SPC shift monitoring problems but without considering the time varying shift and the dynamic nature of the required learning approach to be applied.

Accordingly, all previous methods did not use a dynamic learning model for re-learning and re-designing the time va-

rying shift level. Also, ensemble methods which are designed to cope with shift and concept drift detection were never applied to SPC. That's why, our aim is to propose a method that has at least three advantages: (1) to use dynamic classification methods that can learn the shift and adapt the model to the time varying magnitude of the process change, (2) to aggregate decision from several CCs used simultaneously using an ensemble method for non stationary environment (3) to propose a method that is able to detect both small and large shifts.

In fact, this research is based on a general set up that works with little number of required information. The most important and only information needed is online data with zeros and ones as class labels, where 0 indicates that the batch or the instance has been correctly classified whereas 1 indicates that the batch or the instance has been wrongly classified. The properties we have to comply with are described in the following points: (a) the data can be in forms of instances or batches, (b) the required information is the classification error rates, (c) the change to be detected can be sudden or gradual, (d) specific or ensemble method can be used and (e) attributes charts or variable charts can be used.

6 Dynamic ensemble CC model

6.1 Data set assessment of the ensemble method

In order to minimize the risk of choosing a non adequate CC, as well as to enhance the detection of concept drift, an ensemble of CCs is used instead of choosing the best one. The idea of using adaptive techniques to detect concept drift is inspired from machine learning techniques such as STAGGER of [23], DWM of [24] and [25], DWM-WIN of [7], SVM of [26] and IFCS of [27]. An adaptive combined chart is developed based on a weighted majority vote over the different classifiers' decisions of the CCs for each instance of the batch. We use the DWM of [24] and DWM-WIN of [7] to get this class prediction about the learned shift. Algorithm (1) defines these steps.

In Step 1, the dataset of the DEC chart model is based on the combination of the charting statistics of k individual CCs. The proposed method is suitable to be applied to all CCs types. Each charting statistic is presented in a vector X_k . Then, the different vectors are concatenated and stored in a matrix called X_k for $k = 1, \dots, p$ given by

$$X_k = \begin{pmatrix} x_{1k} \\ x_{2k} \\ \vdots \\ x_{pk} \end{pmatrix} \quad (4)$$

where p is the number of CCs, X_k is the charting statistics of each CC and k is the range of the CC. For three CCs, the vectors are:

Algorithm 1: Combined chart model based on Dynamic weighted majority algorithm for drift detection.

- 1: **Inputs** p : Initialize number of CCs to monitor the process.
 β : parameter for decreasing classifier's weight when state decision is correct.
 η : parameter for increasing classifier's weight when state decision is incorrect.
 θ : threshold for removing classifiers from the ensemble
 - 2: **Step1: Data Pre-processing**
 - 1.1: Apply p chart for each batch of data
 - 1.2: Store the charting statistics of each chart in vectors X_1, X_2, \dots, X_p .
 - 1.3: Concatenate vectors X_n in a matrix of p columns: this represents the combined data D .
 - 3: **Step2: Apply DWM-WIN algorithm to combine chart's decisions.**
 - 2.1 Initialize classifier's weights.
 - 2.2 Apply DWM to the new generated data D
 - 2.3 Compute the error rates of DWM-WIN based on D
-

$$X_{Shewhart} = \begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{k1} \end{pmatrix}, X_{CUSum} = \begin{pmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{k2} \end{pmatrix}, X_{ewma} = \begin{pmatrix} x_{13} \\ x_{23} \\ \vdots \\ x_{k3} \end{pmatrix} \quad (5)$$

Equations (4)-(9) describe the different sub-steps of the pre-processing data phase.

For a combination of 3 CCs, the new data will be

$$X^3 = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ \vdots & \vdots & \vdots \\ x_{k1} & x_{k2} & x_{k3} \end{pmatrix} \quad (6)$$

More generally, the dataset representing the combination of p charts is:

$$X^p = \begin{pmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1p} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2p} \\ x_{m1} & x_{m2} & x_{m3} & \dots & x_{mp} \\ x_{m1+1} & x_{m2+1} & x_{m3+1} & \dots & x_{mp+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{k1} & x_{k2} & x_{k3} & \dots & x_{kp} \end{pmatrix} \quad (7)$$

The next pre-processing step consists on applying DWM algorithm to the generated data in order to monitor its classification error rates and to detect the shift in the data. This is done through a simultaneous use of more specific knowledge about the data by using the charting statistics information and the DWM as a technique of concept drift detection. We simultaneously base our experiments on data with different shifts in the mean aiming to prove the detection ability of the combined CC versus single charts in detecting concept drift during the classification process. To apply DWM, class

labels have to be defined. Thus, before the shift occurs in the process we assume that the class label is 0. Then, when a shift occurs, the class label is 1. For p variables, if the shift occurs in observation $\frac{k}{2}+1$ the class label will be as follows:

$$Classlabel = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad (8)$$

Finally, the new data is:

$$X = \begin{pmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1p} & 0 \\ x_{21} & x_{22} & x_{23} & \dots & x_{2p} & 0 \\ x_{m1} & x_{m2} & x_{m3} & \dots & x_{mp} & 0 \\ x_{m1+1} & x_{m2+1} & x_{m3+1} & \dots & x_{mp+1} & 1 \\ x_{m1+2} & x_{m2+2} & x_{m3+2} & \dots & x_{mp+2} & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{k1} & x_{k2} & x_{k3} & \dots & x_{kp} & 1 \end{pmatrix} \quad (9)$$

DWM is called to detect the concept drift. More precisely, each classifier will initially have a weight of one. The weight of classifiers will be decreased if the classifier predicts incorrectly the state of the process based on the ensemble of chart's statistics and increased if the classifier's global prediction is correct. This step will be applied for the three CCs until having three final weights. If the weight of each CC is under the threshold θ , the classifier will be removed from the ensemble, else it will be maintained. Then, a global prediction vector will be computed from the remaining ensemble of classifiers. When this global prediction of classifiers based on the CC decisions is wrong, a new classifier is added. The same reasoning would continue until finding the best combination of classifiers which detects if the process is in control or out of control for the different batches of the process. The size of the batch affects the detection of the shift via the misclassification error rates. A shift could either be detected or absorbed by the system. In fact, for large batch sizes, the shift is detected and considered while updating the weights inside one batch. However, for small batch sizes, the shift induces a high error which is expected to be more obvious with the monitoring method.

7 Experiments

7.1 Simultaneous small and large shift detection

DEC is a shift detection method based on dynamic classification error rates analysis. In this section, we further explain this method and illustrate it with a small experiment to show the robustness of DEC model in detecting a shift and updating itself. This method is at least as good as or better than the best single charts. Two illustrative examples based

on two datasets, with a small shift in the mean of 0.25 and a large shift of 3.5, are simulated to illustrate the proposed model. The dataset consists of 30 observations where the shift in the mean occurs in observation 11. Therefore, the first 10 observations are simulated based on $N(0.5, 0.5^2)$ in both datasets and the second 20 observations are simulated based on a $N(0.75, 0.5^2)$ and $N(4, 0.5^2)$ in the first and the second example respectively. First data is denoted D_1 and is shown in Table (2) and second data is denoted D_2 and is shown in Table (3). The charting statistics of EWMA, CUSUM, XBAR and DEC charts are X_{1i} , X_{2i} , X_{3i} and Z_i respectively. The upper and lower CLs of EWMA, CUSUM, XBAR and DEC charts are (LCL_1, UCL_1) , (LCL_2, UCL_2) , (LCL_3, UCL_3) and (LCL_z, UCL_z) respectively. Tables (2) and (3) summarize the results for the different methods using the same random generation numbers denoted "seed".

For the first dataset, EWMA and CUSUM chart can detect the shift that occurred in observation 11 at time t_{12} and t_{10} . However, XBAR detects the small shift in D_1 in observation 20. One reason of XBAR's delay to detect the small shift is that its charting statistic is based on an independent subgroup of the dataset which decreases the ability of a quick small shift identification. For DEC chart, the shift is detected by the increase in the misclassification error rates. Then, the DWM algorithm needs some time to update the internal knowledge after the shift detection by removing and adding of experts as well as weight adjusting and other steps. All these steps require some time before reaching a certain stability which can be shown when the error goes back to 0. This explains how the shifts are detected and how the errors are failing back to 0 after a while. Furthermore, thanks to its prediction ability our proposed method has an immediate detection ability explained by the use of classification methods which eases the shift prediction from observation 10. Accordingly, DEC chart is as good as CUSUM chart but much faster than EWMA and XBAR in detecting small shift range.

For the second dataset with large shift in the mean, a shift in the last 20 observations is also simulated. According to Table (3), DEC model is as good as Xbar chart to identify such shift. It predicts, identifies the shift at observation 10 thanks to the learning classification methods aggregating the decisions of the ensemble of charts' decision, EWMA detects it at t_{11} , CUSUM at t_{11} and XBAR at t_{10} . As expected, XBAR has an early detection ability compared to EWMA and CUSUM to detect such shift in D_2 .

This advantage is explained by the use of the misclassification error rates as charting statistics of the monitoring process in DEC chart.

7.2 Improvement of the misclassification error rates

In this section we first concentrate on evaluating the control charts in terms of error of classification which is measured with the misclassification error rates (or type II errors). In the next section, we will compare the CCs based on other measures of performance such as FPs (error type I), accuracy and recall measures. The reason of that

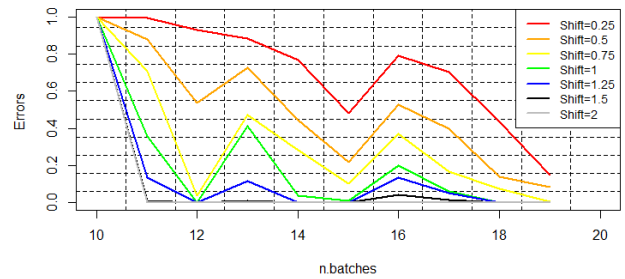


Figure 3. Shift comparison based on ewma misclassification error rates.

is that the misclassification error rates are very informative about the shift during processes and thus we studies the type II error rates first.

We focus on a shift simulation analysis based on several values of shift parameters. In Figure (3), the error rates of EWMA chart for shift values parameters are plotted based on a $N(\mu = 0, \sigma = 1)$ of 400 observations with a shift in the middle. It shows that the change in the error rates of EWMA chart is clearly impacted by the change of the shift level, except the change point of the shift detection which is practically the same. The relation between the speed of the adaptation of the error process after the shift detection and the shift level is clear. Changes in batches 13, 15 and 16 are random change points which are detected with different levels of errors impacted by the level of the shift.

Because the reaction of the chart to the shift is highly informative about the robustness of the method to detect the changes and to adapt itself after the shift detection, we conduct a comparative study between the different variants of DEC chart and the individual charts. Thus, we compare results obtained by the different variants of DEC chart denoted: (1) EWMA-CUSUM, (2) EWMA-Xbar, (3) CUSUM-Xbar and (4) XBAR-CUSUM-EWMA with the application of individual CCs: EWMA, CUSUM and Xbar. The combined CCs were implemented based on 400 observations with a mean shift in the middle using a number of batches of 5, 10, 20, 30 and 40. The procedure used by DWM algorithm, to divide the data in different batches, train on the $n - 1$ recent batches and test on the last batch, is the same procedure used for generating a k fold cross validation. The latter consists on portioning the data into n subsets, performing the training on one subset and test on the other subsets. So, the number of batches is equal to the number k used to generate k fold cross validations. Indeed, we compute the mean over 10 runs for each method. By using this procedure, we abstain sampling biases. Furthermore, we use other measures of performance such as F-measure, recall, FPs ad FNs and statistically compare the difference between the methods.

DEC chart is configured with the following parameters: $\beta = 0.449$, $\eta = 1.13$ and $\theta = 0.01$. These values were selected after applying Genetic Algorithm (GA) with a population size = 50. For each method 10 runs were performed for each method using each time a different random seed and providing

Table 2. Illustrative examples of EWMA, CUSUM, XBAR and ensemble chart model for small shift in the mean $\delta = 0.25$ in $N(0.5, 0.5^2)$. The first 10 observations follow a $N(0.5, 0.5^2)$ and the last 20 observations follow a $N(\text{shift} + 0.5, 0.5^2)$ using the same set.seed (random number generation). For EWMA, $\lambda = 0.2$ and $L = 3$ are used and for CUSUM, $k = 1$ and $h = 3$.

n	EWMA			CUSUM			XBAR			DEC		
	X_{1i}	LCL	UCL	X_{2i}	LCL	UCL	X_{3i}	LCL	UCL	Z_i	LCL	UCL
1	0.472	0.160	0.609	0.660	-2	2	-	0.432	0.787	-	-	-
2	0.462	0.090	0.672	0.260	-2	2	-	0.432	0.787	0	0	0
3	0.603	0.063	0.706	1.850	-2	2	-	0.432	0.787	-	-	-
4	0.518	0.040	0.720	0.790	-2	2	-	0.430	0.787 2	0	0	0
5	0.423	0.031	0.730	0	-2	2	0.605	0.432	0.787	-	-	-
6	0.364	0.241	0.740	0	-2	2	-	0.432	0.787	0	0	0
7	0.278	0.0190	0.750	0	-2	2	-	0.432	0.787	-	-	-
8	0.289	0.016	0.750	0	-2	2	-	0.432	0.787	0	0	0
9	0.203	0.014	0.750	0	-2	2	-	0.432	0.787	-	-	-
10	0.476	0.013	0.750	2.670	-2	2	0.480	0.432	0.787	1	0	0
11	0.598	0.012	0.750	4.040	-2	2	-	0.432	0.787	-	-	-
12	0.780	0.011	0.750	6.660	-2	2	-	0.432	0.787	1	-0.841	1.840
13	0.720	0.011	0.750	8.361	-2	2	-	0.432	0.787	-	-	-
14	0.920	0.011	0.750	7.210	-2	2	-	0.432	0.787	1	-0.840	1.84
15	0.94	0.011	0.750	10.220	-2	2	0.692	0.432	0.787	-	-	-
16	0.74	0.011	0.750	11.450	-2	2	-	0.432	0.787	1	-0.840	1.840
17	0.804	0.011	0.750	9.750	-2	2	-	0.432	0.787	-	-	-
18	0.709	0.0110	0.750	11.040	-2	2	-	0.432	0.787	1	-0.840	1.84
19	0.590	0.011	0.750	10.380	-2	2	-	0.432	0.787	-	-	-
20	0.690	0.011	0.750	9.180	-2	2	0.880	0.432	0.78	0	-0.840	1.840
21	0.790	0.011	0.750	10.560	-2	2	-	0.432	0.787	-	-	-
22	0.680	0.011	0.750	12.230	-2	2	-	0.432	0.787	0	-0.840	1.84
23	0.570	0.011	0.750	11.390	-2	2	-	0.432	0.787	-	-	-
24	0.600	0.011	0.750	10.190	-2	2	-	0.432	0.787	0	-0.840	1.84
25	0.680	0.011	0.750	10.630	-2	2	0.420	0.432	0.787	-	-	-
26	0.480	0.011	0.750	11.82	-2	2	-	0.432	0.787	0	-0.840	1.840
27	0.469	0.011	0.750	9.370	-2	2	-	0.432	0.787	-	-	-
28	0.532	0.011	0.750	8.960	-2	2	-	0.432	0.787	0	-0.840	1.840
29	0.503	0.011	0.750	9.530	-2	2	-	0.432	0.787	-	-	-
30	0.508	0.011	0.750	9.030	-2	2	0.649	0.432	0.787	0	-0.840	1.840

results by the mean of these iterations. The task of DEC chart is to mine the charting statistics of the different CCs. To treat this problem, DEC proposes a classification task solution. It combines many decisions at each time step based on DWM-WIN algorithm to increase the probability of correctly classifying a change point.

Here, we track the degree of diversity of the CC ensemble. Thus, we analyze the capacity of the combined CC model based on a normal distribution with $\mu = 0$ and $\sigma = 1$ by comparing it with the individual chart model in terms of misclassification error rate of DWM-WIN. For the EWMA chart, many values of the parameter λ are used in the literature. In general, values of this parameter are between 0 and 1 as stated by [26]. In this experiment, $\lambda = 0.2$ and $L = 3$ are used with the notation that these parameters can be optimized. For CUSUM chart we use the decision boundary $h = 4$, the reference value $k = 1$ and the target value representing the actual process mean. $T = 1$. For XBAR chart used in this simulation, we use the Xbar'one (xbar) chart function from the quality control chart (qcc) R package which computes statistics required by the xbar chart for one at-time data. The reason of this choice is that we need to combine vectors of

charting statistics of equal size. This can only be obtained by the Xbar'one chart where the number of subgroups is equal to the dataset size as in CUSUM and EWMA charts.

In Figure (4), the horizontal axis represents the shift level and the vertical axis the error rates of the classification ensemble method used to combine the different chart models. We base our analysis on 6 shift levels. The analysis is as follows:

Shift = 0.25: The differences between methods are more pronounced for small shift values. Although all other individual chart models as well as the combined ones: EWMA-CUSUM, EWMA-XBAR and XBAR-EWMA charts begin with relatively high errors, EWMA-CUSUM-XBAR starts with low error rates by perfectly coping with the shift level 0.25. In fact, the ensemble chart model begins with misclassification error rates between 0.055 and 0.12 for all different number of batches situations, however it is greater than 0.4 for individual charts. Thus, our proposed CC outperforms other CCs and is able to track small shifts when considering the statistics information of more than one chart.

Shift = 0.5: When the shift level is 0.5, errors of the combined three charts increase but they are still less than 0.22.

Table 3. Illustrative examples of EWMA, CUSUM, XBAR and ensemble chart model for large shift in the mean of 3.5 in $N(1, 1)$. The first 10 observations follow a $N(0.5, 0.5^2)$ and the last 20 observations follow a $N(0.5 + \text{shift}, 0.5^2)$ using the same set.seed (random number generation). For EWMA, $\lambda = 0.2$ and $L = 3$ are used and for CUSUM, $k = 1$ and $h = 3$.

n	EWMA			CUSUM			XBAR			DEC		
	X_{1i}	LCL	UCL	X_{2i}	LCL	UCL	X_{3i}	LCL	UCL	Z_i	LCL	UCL
1	0.472	0.160	0.609	0.084	-2	2	-	2.450	3.101	-	-	-
2	0.460	0.090	0.672	0	-2	2	-	2.450	3.101	0	0	0
3	0.603	0.063	0.706	0	-2	2	-	2.450	3.101	-	-	-
4	0.518	0.043	0.726	0	-2	2	-	2.450	3.101	0	0	0
5	0.420	0.031	0.738	0	-2	2	2.550	2.450	3.101	-	-	-
6	0.360	0.024	0.740	0	-2	2	-	2.450	3.101	0	0	0
7	0.270	0.019	0.746	0	-2	2	-	2.450	3.101	-	-	-
8	0.289	0.0160	0.751	0	-2	2	-	2.450	3.101	0	0	0
9	0.203	0.016	0.754	0	-2	2	-	2.450	3.101	-	-	-
10	0.470	0.014	0.756	1.580	-2	2	2.350	2.450	3.101	1	0	0
11	1.240	0.0132	0.757	8.700	-2	2	-	2.450	3.101	-	-	-
12	1.950	0.012	0.750	16.740	-2	2	-	2.450	3.101	1	-0.090	1.690
13	2.450	0.011	0.758	24.090	-2	2	-	2.450	3.101	-	-	-
14	2.640	0.011	0.758	29.320	-2	2	-	2.450	3.101	1	-0.090	1.690
15	3.106	0.011	0.759	37.650	-2	2	2.642	2.450	3.101	-	-	-
16	3.340	0.011	0.759	44.660	-2	2	-	2.450	3.101	1	-0.090	1.690
17	3.310	0.011	0.759	49.470	-2	2	-	2.450	3.101	-	-	-
18	3.509	0.011	0.759	56.520	-2	2	-	2.450	3.101	1	-0.090	1.690
19	3.52	0.011	0.759	62.122	-2	2	-	2.450	3.101	-	-	-
20	3.49	0.011	0.759	67.305	-2	2	2.830	2.450	3.101	1	-0.090	1.690
21	3.66	0.011	0.759	74.420	-2	2	-	2.450	3.101	-	-	-
22	3.810	0.011	0.759	81.750	-2	2	-	2.450	3.101	1	-0.090	1.690
23	3.75	0.011	0.759	87.210	-2	2	-	2.450	3.101	-	-	-
24	3.680	0.011	0.759	92.390	-2	2	-	2.450	3.101	0	-0.090	1.690
25	3.740	0.011	0.759	98.810	-2	2	3.022	2.450	3.101	-	-	-
26	3.840	0.011	0.759	105.790	-2	2	-	2.450	3.101	0	-0.090	1.690
27	3.650	0.011	0.759	110.040	-2	2	-	2.450	3.101	-	-	-
28	3.660	0.011	0.759	115.820	-2	2	-	2.45	3.101	0	-0.090	1.690
29	3.730	0.011	0.759	122.330	-2	2	-	2.450	3.101	-	-	-
30	3.710	0.011	0.759	128.304	-2	2	-	2.450	3.101	0	-0.090	1.690

On the other hand, all combined two chart models: EWMA-CUSUM, EWMA-XBAR and CUSUM-XBAR perform better than individual charts for all batch size situations. For example, for $N = 5$, these methods greatly outperform individual charts. This is due first to the advantages of the ensemble method technique over the individual ones. Second, this performance is also due to the fact that statistics used for each CC are very informative about the quality control of process behavior. Also, the DWM-WIN used in the combined chart models allows this technique to easily detect shifts in the process and to update the internal knowledge announced in the CC about this shift.

Shift = 1: For this shift level, errors of individual charts decrease compared to smaller shift level situations because all these CCs are more sensitive to detect large shifts, except the CUSUM which is based on accumulating changes in one direction and hence is more sensitive to small shifts. Combined models are still more robust to react to the shift better than individual charts. XBAR, CUSUM and EWMA do not show this effect in detecting shifts in processes because the internal information does not allow it. More rigorously, the combined chart model performs better than other individual charts, in particular XBAR-CUSUM and EWMA-CUSUM

are the best ones for $N = 10$ and $N = 30$. This is due to the fact that CUSUM chart is more sensitive to small shifts and EWMA and XBAR perform better with large shift cases as explained earlier. Thus, combining CUSUM with XBAR and with EWMA outperforms the decision of only one chart model in detecting a shift of 1.

Shift = 1.5: All CC error rates decrease when the shift level increases to 1.5. In particular, EWMA-CUSUM-XBAR chart errors are between 0.05 and 0.12 whereas other charts reach error rates of 0.2 and even more for $N = 5$. An error rate of 0.28 is obtained with XBAR chart. More precisely, XBAR-CUSUM shows a good reaction capacity to shift detection for all batch size situations. This is due to the combined effect of the two charts and the diversity of the internal knowledge given by the two different charts, one is sensitive to very small shifts and the other is more sensitive to large shifts.

Shift = 2 Combined EWMA-CUSUM-XBAR chart outperforms other CCs for moderate and large shifts. In fact, it has a maximum error of 0.12, individual charts reach 0.28, two combined models reach 0.18 for $N = 30$.

Shift = 3: Similar results are obtained as for shift level = 2.

Experimental results show that combined EWMA-

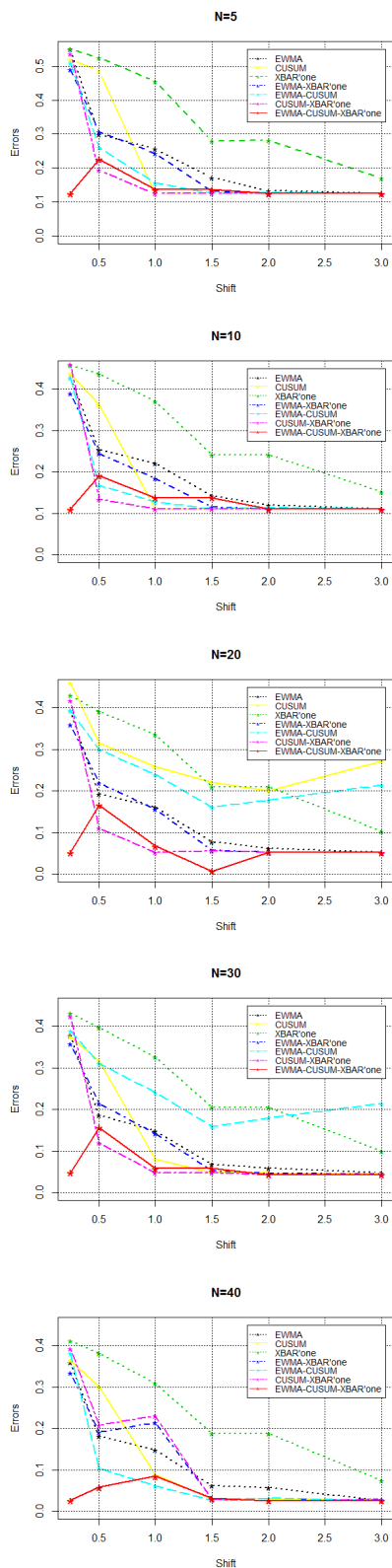


Figure 4. Comparison of the misclassification error rates of DWM-WIN monitoring the different combined control chart models versus the individual charts for a variety of shift levels.

CUSUM-XBAR chart outperforms individual CCs. When the diversity of combined CCs increases, the misclassifica-

tion error rates decrease accordingly. When combining three CCs based on a dynamic ensemble method for concept drift, the accuracy of the CC improves. As shown in Figure 4, EWMA-CUSUM-XBAR chart depicts smaller values of misclassification error rates better than the combined two chart model. Interestingly, the diversity in combining CCs presents an important factor to improve the classification accuracy rate of the proposed chart in both monitoring small and large shifts in the mean.

8 Comparison of DEC chart with combined SFEWMA-X chart

In this section we focus on comparing the proposed dynamic chart DEC with one of the combined charts from the literature called Single Featured EWMA chart (SFEWMA-X) of [29].

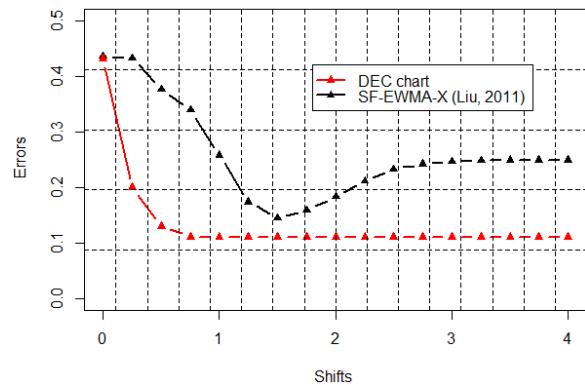


Figure 5. Performance evaluation of proposed DEC chart model versus combined SFEWMA-X chart based on $N(\mu = 1, \sigma^2 = 1)$ using 400 observations.

Like DEC chart, the latter was proposed with the aim to both monitor small and large shifts in the process. SFEWMA chart was proposed to handle the problem of combining EWMA-X chart of using two sets of statistics and CLs by proposing only one set of statistic and CLs. This method aims to identify both small and large shifts. The basic idea of SFEWMA-X is to transform the EWMA statistic into a new one in order to have the same scale for Shewhart-X and EWMA charts. This is done by performing a variable transformation called "multiplier" to rescale the statistic of EWMA into the same range of the statistic of Shewhart-X chart and thus make the application of the combined chart easier. The main reason of choosing this method among others is because it is similar to our method in the idea of unifying the statistics of CCs. In their article, [29] compared SFEWMA-X only to individual charts. Here, we compare SFEWMA-X to individual charts and also the proposed DEC chart. Comparisons were conducted in terms of errors and in terms of different performance measures. Figure (5) compares error rates of DEC chart and SFEWMA-X chart for the different shift sizes.

The two methods begin with equal performance when the dataset doesn't contain any shift. Then, as the shift increa-

ses, DEC's error decreases until achieving a stable level of error of 0.11 for shifts ranging from 0.75 to 4. Concerning SFEWA-XBAR chart, the error is decreasing until a shift of 1.5 achieving 0.13 of error. Then, it increases again to achieve 0.25 for shift 2.5 until 4. Thus, the performance of dynamically combining three CCs based on DWM algorithm outperforms the SFEWMA-X chart. The reason for that is that first, DEC chart uses more statistical information than SFEWMA-X chart. Second, the dynamic combination through the use of DWM algorithm is much more informative about the shift detection points than the use of the rescaled parameter proposed to combine EWMA and XBAR in SFEWMA-X. Indeed we note that the decrease in the error is impacted by the shift level. This is explained by the relation between the knowledge about the shift level and the classification errors. In fact, a clear knowledge about the shift, obtained by higher levels, impacts the classification errors. Thus, the clear is the knowledge representation of the shift obtained, the smaller is the classification error. This reasoning is confirmed by [30] when they show the high relation between the knowledge-based levels and the error of classification.

To go further in our analysis, we want to understand the cause of high errors in SFEWMA-X compared to DEC chart. Thus, we concentrated on the behavior of the two charts for shifts smaller than 1. Figure (6) shows the high differences in errors between the two methods based on the mean over the Type I and Type II errors.

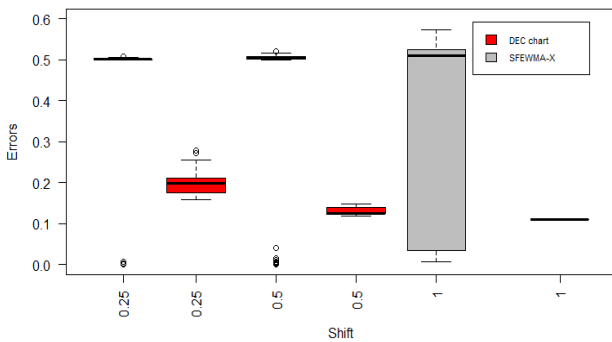


Figure 6. Boxplot comparing the mean over Type I and Type II errors of DEC chart and SFEWMA-X charts based on 100 runs.

As noticed, the variance in SFEWMA-X is higher than in DEC chart for shift of 1. Compared to SFEWMA-X chart, the DEC chart is more stable in results and procures smaller error variance.

To go further in our analysis, we compare DEC chart with SFEWMA-X as well as individual CUSUM, XBAR and EMWA charts in terms of accuracy, recall and FP measures. Figures (7) to (9) show the comparisons of performances of the different methods for small-moderate and large shifts.

The performance of each method could be highlighted based on one measure and not the others. Each measure is sensitive to evaluate one or more methods but not all of them.

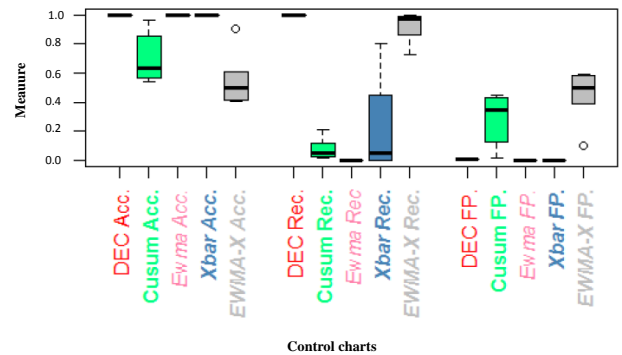


Figure 7. Performance evaluation of DEC model compared to individual EWMA, CUSUM and Xbar in terms of Accuracy, Recall, FP and FN rates: small-moderate shifts.

That's why, instead of error rates, we use accuracy, FPs and FNs. For small and moderate shifts as shown in Figure (7), DEC chart outperforms SFEWMA-X and individual charts in terms of accuracy, recall and FP rates. SFEWMA-X is competitive with respect to DEC chart and outperforms individual EWMA and XBAR in terms of recall. The performance of CUSUM chart is not highlighted here because of the low performance of this type of chart in detecting the moderate shift which is included in the computations done in the boxplots.

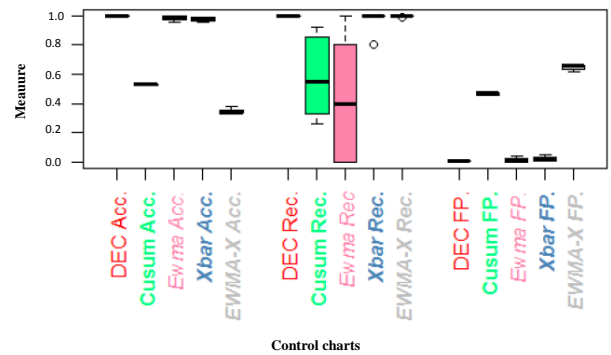


Figure 8. Performance evaluation of DEC model compared to individual EWMA, CUSUM and XBAR in terms of Accuracy, Recall, FP and FN rates: large shifts.

For large shifts, the performance of DEC chart is also maintained. It optimizes the different measures compared to other methods. The performance of SFEWMA-X chart is highlighted in terms of recall measure when a competitive recall measure is achieved. Concerning the comparison between the individual charts in detecting a large shift range, results show that as expected, XBAR chart outperforms CUSUM and EWMA charts in terms of Accuracy, Recall and FPs. The performance of CUSUM is not shown because of

the moderate shift included in the computations done in Figure (8). DEC chart's performance is noticed through the different measures and a high robustness and stability to the shift detection than the combined SFEWMA-X chart as well as when compared to other individual charts.

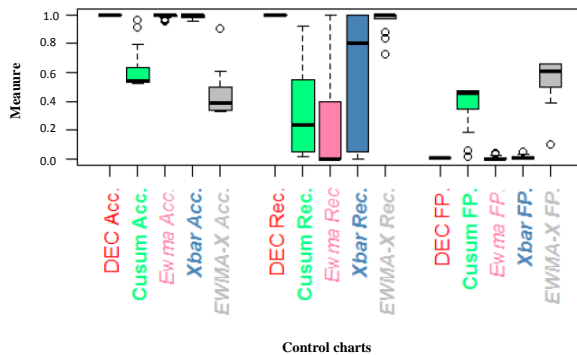


Figure 9. Performance evaluation of DEC model compared to individual EWMA, CUSUM and Xbar in terms of Accuracy, Recall, FP and FN rates: all shifts.

Figure (9) illustrates a summary for more general results over the different shift sizes. SFEWMA-X outperforms individual EWMA and XBAR and is as good as DEC chart in terms of Recall. However, this effect is not shown in terms of accuracy and FP rates. Being already very performant to shift detection, this method assumes that the class labels are unknown. In the next chapter, we propose a new combined chart without the assumption of known classes in order to make the application in real word data possible.

9 Conclusions

The proposed CC does not only exhibit superior robustness to individual EWMA, CUSUM and XBAR for some performance measures but also presents a new heuristic for shift learning and monitoring in nonstationary environment. DEC chart presents the second CC proposed by aggregating the decisions of three different CCs based on DWM method after DWMCC. Results show that the use of the ensemble techniques in CCs outperforms the use of single CCs as well as other combined charts when several shift levels have to be monitored in terms of different measures thanks to the fact that the use of DWM in CCs allows an identification of the shift learning process. In order to enable the application of combined charts to real world data, a new model based on both DWM-WIN as well as TAACL chart was proposed. In fact, DWM was used in learning while monitoring with CCs by applying exactly the same mechanism dealing with classifiers in the DWM to CCs. Using this new heuristic of treating CCs as classifiers by adding, removing and dynamic weighting of CCs based on their performance leads to a better

shift identification ability. Results show that the DWM chart model leads to the best model to deal with the non stationarity of the data compared to models from the most successful individual SPC techniques. **This article presents the application of the method on different simulated scenarios of non-stationarity processes and on time varying processes. An interesting extension of this article would be to apply this method to a real dataset and to study the relationship between the process variability and the type of charts to be combined.**

Another possible perspective to this research is that instead of applying many CCs to one dataset and combining the decision over the different charts, one can apply one CC to different features of the data, then combine the decision over the different features. This would represent a new multivariate application of ensemble methods in SPC.

REFERENCES

- [1] D. Mejri, M. Limam, and C. Weihs, (2016), "A new dynamic weighted Majority Control chart for data streams", *Journal of Soft Computing*, DOI: 10.1007/s00500-016-2351-3, pp:1-12.
- [2] M. R. R., Jr and G.-Y., Chob, (2011). Multivariate control charts for monitoring the mean vector and covariance matrix with variable sampling intervals. *Sequential Analysis: Design Methods and Applications*, 30(1):140.
- [3] M. R. S. Z. G., Reynolds, JR., (2008). Combinations of multivariate shewhart and mewma control charts for monitoring the mean vector and covariance matrix. *Quality Technology*, 40(4):381393.
- [4] D. Zeng, I. Gotham, K. Komatsu, C. Lynch, M. Thurmond, D. Madigan, B. Lober, J. Kvach, and H. Chen, (2007). *Intelligence and Security Informatics : Biosurveillance*. Second NSF Workshop, New Brunswick, NJ, USA, Proceedings, vol. 4506 of Lecture Notes in Computer Science. Springer.
- [5] J. C. Brillman, T. Burr, D. Forslund, E. Joyce, R. Picard, and E. Umland, (2005). Modeling emergency department visit patterns for infectious disease complaints: results and application to disease surveillance. *BMC medical informatics and decision making*, 5:4.
- [6] B. Y. Reis, and K. D. Mandl, (2003). Time series modeling for syndromic surveillance. *BMC medical informatics and decision making*, 3:111.
- [7] D. Mejri, R. Khanchel, and M. Limam, (2013). An ensemble method for concept drift in nonstationary environment. *Journal of Statistical Computation and Simulation*, 83:11151128.

- [8] H. Wang, W. Fan, P. S. Yu, and J. Han, (2003). Mining concept-drifting data streams using ensemble classifiers. In Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 03, pages 226235, New York, NY, USA. ACM.
- [9] H. B. Nembhard, and M. S. Kao, (2003). Adaptive forecast based monitoring for dynamic systems. *Technometrics*, 45(3):208219.
- [10] X. Zhang, I. Elishakoff, and R. Zhang, (1991). A Stochastic Linearization Technique Based on Minimum Mean Square Deviation of Potential Energies. Springer Verlag.
- [11] M.C. Georgiadis, J. R. Banga, and E. N. Pistikopoulos, (2010). Dynamic process modeling: Combining models and experimental data to solve industrial problems. *Process Systems Engineering: Dynamic Process Modeling*, 7:208219.
- [12] S. W. Robert, (1959). Control chart test based on geometric moving average. *Technometrics*, 42(1):97102.
- [13] W. A. Shewhart, (1931). *Economic Control of Quality of Manufactured Product*. New York: D. Van Nostrand Company.
- [14] E. S. Page, (1954). Continuous inspection scheme. *Biometrika*, 41(1/2):100115.
- [15] W. D. Ewan, (1963). When and how to use cusum charts. *Technometrics*, 5(1):122.
- [16] N. L. Johnson, (1961). A simple theoretical approach to cumulative sum control chart. *Journal of the American Statistical Association*, 56(296):835840.
- [17] N. L. Johnson, and F. C. Leone, (1962). Cumulative sum control charts: Mathematical principles applied to their construction and use, part i. *Industrial Quality Control*, 18(12):1521.
- [18] D. Montgomery, (2005). *Introduction to Statistical Quality Control*. 5th edition, New York: John Wiley and Sons.
- [19] W. Yang, Y. G. and W. Liao, (2015). A hybrid learning based model for simultaneous monitoring of process mean and variance. *Quality Reliability Engineering*, 31:445463
- [20] B. Wu, and J.B. Yu, (2010). A neural network ensemble model for online monitoring of process mean and variance shifts in correlated processes. *Expert Systems with Applications*, 37(6):4058–4065.
- [21] Z. Men, E. Yee, F.S. Lien, Z. Yang, and Y. Liu, (2014). Ensemble nonlinear autoregressive exogenous artificial neural networks for short term wind speed and power forecasting. *International Scholarly Research Notices*, ID 972580.
- [22] C. S. Cheng, and H. P. Cheng, (2008). Identifying the source of variance shifts in the multivariate process using neural networks and support vector machines. *Expert Systems with Applications*, 35(1-2):198206.
- [23] J. C. Schlimmer, and R. H. Granger, (1986). Beyond incremental processing: Tracking concept drift. In Proceedings of the Fifth National Conference on Artificial Intelligence, pages 502507. AAAI Press, Menlo Park, CA.
- [24] J. Z. Kolter, and M. A. Maloof, (2007). Dynamic weighted majority: An ensemble method for drifting concepts. *Journal of Machine Learning Research*, 8:2755 2790.
- [25] J. Z. Kolter, and M. A. Maloof, M. A. (2005a). Dynamic weighted majority: A new ensemble method for tracking concept drift. In *International Conference on Data Mining (ICDM)*, pages 123130. IEEE.
- [26] S. N. Ahmed, L. Huan, and S. K. Kay, (1999). Handling concept drifts in incremental learning with support vector machines. In Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 99, pages 317 321, New York, NY, USA. ACM.
- [27] J. Schlimmer, and R. Granger, (2011). Fuzzy classification in dynamic environments. *Soft Computation*, 15(5):10091022.
- [28] S. H. Steiner, (1999). Ewma control charts with time-varying control limits and fast initial response. *Journal of Quality Technology*, 31(1):7586.
- [29] C. S. Liu, and F. C. Tien, (2010). Design of single featured ewma x control chart for process mean shift detection. In Proceedings of the 2nd International conference on Applied Operational research, pages 301314. *Lecture Notes in Management Science*.
- [30] H. Pasman, (2015). *Risk Analysis and Control for Industrial Processes - Gas, Oil and Chemicals. A System Perspective for Assessing and Avoiding Low-Probability, High-Consequence Events*. Elsevier Science.
- [31] J. M. Lucas, (1982). Combined shewhart cusum quality control schemes. *Journal of Quality Technology*, 14(2):5159.