

Ariza Guerrero, Angellys P.; Barreto Pombo, Rister; Herrera Acosta, Roberto J.

## Article

# Application of robust multivariate control chart with Winsorized Mean: A case study

Journal of Industrial Engineering International

## Provided in Cooperation with:

Islamic Azad University (IAU), Tehran

*Suggested Citation:* Ariza Guerrero, Angellys P.; Barreto Pombo, Rister; Herrera Acosta, Roberto J. (2019) : Application of robust multivariate control chart with Winsorized Mean: A case study, Journal of Industrial Engineering International, ISSN 2251-712X, Springer, Heidelberg, Vol. 15, Iss. S1, pp. 309-318,  
<https://doi.org/10.1007/s40092-019-0319-5>

This Version is available at:

<https://hdl.handle.net/10419/267652>

### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

### Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by/4.0/>



# Application of robust multivariate control chart with Winsorized Mean: a case study

Angellys P. Ariza Guerrero<sup>1</sup> · Rister Barreto Pombo<sup>1</sup> · Roberto J. Herrera Acosta<sup>1</sup>

Received: 25 November 2018 / Accepted: 13 July 2019 / Published online: 19 July 2019  
© The Author(s) 2019

## Abstract

Water pH and active ingredient concentration are two of the most important variables to consider in the manufacturing process of fungicides. If these variables do not meet the required standards, the quality of the product may be compromised and lead to poor fungicide performance when water is used as the application carrier, which is in most cases. Given the correlation between the variables, these kinds of manufacturing processes must be analyzed in multivariate settings. Thus, this paper analyzes the variables involved in the process using the multivariate control chart  $|S|$  introduced by J. A. Vargas. In the original chart, the arithmetic mean is used as the mean vector estimator. However, in this investigation the arithmetic mean was replaced by the Winsorized Mean for the purpose of evaluating the chart performance with a robust estimator. The results show that using the new estimator, the control chart is able to detect shifts in the variation of the mean vector that the traditional estimator did not. Furthermore, different subgroup sizes for the data were studied in order to examine the performance of the chart in each case. It was found that the proposed control chart is more sensible to changes when the subgroups consist of less observations, since it is able to better identify the outliers in the sample.

**Keywords** Fungicide · Variability · Determinant · Outliers

## Introduction

The manufacturing process of a fungicide, or any pesticide, consists mainly of three phases. First, the active ingredient must be synthesized in a laboratory or chemical factory. Then, it is formulated either by mixing it with a carrier if it is to be prepared in liquid form or with a dry fertilizer if the intended result is a dust pesticide. Finally, a farmer or a certified applicator dilutes the pesticide before its application. In some cases, the final product might be stored for a short amount of time before it is sent to the farmer, who must then dilute the emulsified concentration before applying it to the field (Secrest n.d.).

The majority of large pesticide manufacturers test their products for important characteristics that influence their performance, such as potency, emulsification, density, color, pH and particle size or suspension depending on the formulation process used (Secrest n.d.). Since most pesticide formulations are designed to be diluted using water as a carrier, the ideal pH range of the product is one of the most important variables for both manufacturers and farmers. A high level of pH creates alkaline conditions in which the pesticide will undergo a chemical process called hydrolysis (Deer and Beard 2001). During this process, the active ingredient breaks down into a less powerful formula and, consequently, the performance of the pesticide will be less effective.

The hydrolysis rate varies for each pesticide, depending on the active ingredient used. Generally, for each point increase in the optimal pH levels for a given pesticide, the hydrolysis rate will increase by approximately ten times (Deer and Beard 2001). However, there are other factors that play a role in the degradation of the pesticide such as its susceptibility, the temperature of the water and the amount of time it is in alkaline conditions. When using any pesticide in the industry of crop production, all of these agents must

---

✉ Rister Barreto Pombo  
rjbarreto@mail.uniatlantico.edu.co

Angellys P. Ariza Guerrero  
apaolaariza@mail.uniatlantico.edu.co

Roberto J. Herrera Acosta  
robertojoeherrera@gmail.com

<sup>1</sup> Universidad del Atlántico, Km 7, Puerto, Colombia

be adequately handled to increase the chances of a successful pesticide application.

It is the responsibility of the manufacturer to provide useful information of the product on the labels, so that it can be applied correctly. In order to achieve this goal, the variability of the factors that affect this process must be controlled prior to the application, in the different phases of the manufacturing process. Since there are many factors to consider when manufacturing pesticides, the quality control technique must be conducted in a multivariate setting. In this paper, two of the key variables involved in the manufacturing process of a pesticide, more specifically a fungicide, will be examined using a multivariate control chart.

Agrochemical companies and other economic sectors, must study and control the variables that affect their manufacturing processes to achieve total quality assurance and customer satisfaction, since it plays a critical role in the success of one manufacturing enterprise in today's globally competitive marketplace (Du et al. 2012). An influential factor in manufacturing process quality is variability, which can affect product features and make them different from each other, even when they belong to the same production line. Process variability may have assignable or non-assignable causes. The latter does not have a dominant influence on the process and is subject to changes over time. On the other hand, assignable causes affect the process in either a systematic or a sporadic way and can be easily distinguished, isolated and eventually eliminated (Rogalewicz 2012).

Statistical process control and specifically control charts are presented as a tool that helps monitor variability and process behavior. Moreover, its tools play an important role in controlling the quality of the products, which ultimately ensures that they will meet client needs and general requirements (Haridy and Wu 2009). In this sense, control charts allow manufacturers to determine whether the behavior of their processes is appropriate or non-assignable causes that affect its performance can be found. Manufacturing industries were a pioneer in control charts application, more exactly univariate control charts that were used to improve quality of products, particularly in the automobiles, technology, chemical and pharmaceutical sectors; nevertheless, service companies such as electrical energy, public transportation, banking, retailing and health care sector also make use of these charts to seek satisfaction and customer loyalty (Montgomery 2009).

A process is considered statistically unstable or out of control when the control chart shows one or more points that are higher (or lower) than the established control limits. This can also be said if the behavioral pattern of the chart is not random (Du et al. 2013). However, most processes have more than one variables causing disturbances in their performance, leading to difficulties in the identification of main causes for variability. Furthermore, manufacturing processes

are often highly correlated and univariate control charts are unable to display their interrelation (Stefatos and Hamza 2009), since they can only be used to analyze isolated variables. Applying individual control charts to each variable independently has been proven an inefficient solution that can lead to distortions.

By applying multivariate control charts, out-of-control signals can be detected and this allows manufacturers to diagnose and identify the factors in the process that affect the behavior of these variables, as well as establish corrective actions (Costa et al. 2015). There is a large number of published studies describing the role of control charts in industrial processes. Therefore, it is possible to find different approaches that can be more or less efficient in any given situation. In some cases, the effectiveness of any statistical analysis technique is influenced by the traditional assumptions about the distribution of the sample data.

When the normality and homoscedasticity (i.e., homogeneity of variance) of the sample data are assumed, the presence of outlying observations that challenge these assumptions may result in different types of errors (Goodwyn 2012). The justification for such assumptions is that if they cause insignificant errors in the model, the resulting errors in the conclusion will also be small, but this is not always the case. In the past decades, it has been observed that the most common statistical procedures are extremely sensitive to minor deviations originated from the traditional assumptions and many robust procedures have been proposed in order to obtain more accurate measures (Huber 1981).

Robust statistical methods are less influenced by outlying scores in the sample data (Thompson 2006). An outlier is an observation that lie downs an abnormal distance from other values in a dataset (Al-Khazaleh et al. 2015). Such observations can lead to inaccurate results in the calculation of statistical measures, more specifically for this case study, central tendency measures, which look for a single value to represent the typical individual in a dataset (Wilcox 2001). In this sense, when the observations in a dataset are not homogenous (i.e., it contains extreme values), this measure tends to follow these outliers and the resulting single value does not represent adequately the sample data.

In the  $|S|$  control chart, the arithmetic mean is used as the central tendency measure, assuming the normality of the data distribution. In order to reduce the impact of the outlying observations in the sample, the usage of the Winsorized Mean will be presented as an alternative for a more robust estimation. The Winsorized Mean is the mathematical average of the Winsorized distribution (Goodwyn 2012) and when it is applied on normal data, it does not produce a large loss in efficiency (Rivest 1994). Since simply eliminating the outliers in the sample is not considered an estimation robust method and can cause the results of the analysis to be distorted, the Winsorized Mean is a better option because

it, rather than trimming the extreme values, replaces a percentage of the data with other values of the same set that are closer to its actual mean.

In this paper, the Winsorized Mean will be applied to the construction of the  $|S|$  control chart, more specifically, in the calculation of the variance–covariance matrix. The main objective of this modification is to reduce the influence of the largest observations in the dataset obtained from a fungicide manufacturing process. The rest of this paper is organized as follows. Firstly, the Problem Formulation will explain the effects of pH water and active ingredient concentration variation on the manufacturing process of a fungicide and the problems that the lack of monitoring can cause during the application. In the Building section, the proposed chart and the methodology of experimentation will be presented. In the Results section, the traditional and proposed chart construction methods will be compared, as well as the results for different subgroup sizes. Finally, the last section contains the conclusion and future work.

## Problem formulation

Fungicide is a type of pesticide that specifically kills or inhibits fungi, thus protecting plants against fungal diseases. Managing fungal diseases is a very important part of crop production because of the impact they have on yield and quality (McGrath 2009). The application of a fungicide consists not only of the interaction between fungicide and fungus, but also between the plant and the fungicide (Smart 2003). In other words, some environmental conditions can have an impact on the effectiveness of the fungicide such as temperature, type of plant and method of cultivation.

The majority of pesticides are commercialized in a concentrated form, which must be dissolved in water in order to activate its components. This is also the case for fungicides, since they are often sold in powder and granulated form. Considering the use of water in the application of fungicides, it is important to examine the way that the active ingredient reacts to its properties, such as hardness and pH level. If the pH of the water is higher than 7.5, it is considered alkaline and can affect the performance of some pesticides. The alkalinity of the water can degrade the active ingredient of the fungicide, thereby making it non-toxic and ineffective against fungal diseases (Schilder 2008).

The half-life is a measurement that indicates how long it takes for alkaline water to degrade the pesticide. For some active ingredients, a pH level of 5–7 is optimal and they become unstable if it is above 9 and below 4. In other instances, fungicides can be stable at pH = 10 and 11, depending on the product and the active ingredient used. Overall, the half-life indicates the percentage of the active ingredient that is degraded in a certain amount of

time. Dimethoate, for example, is reduced up to 50% in 48 min at a pH = 9. This results in the application of half the acquired amount of the active ingredient and, therefore, poor pesticide performance.

Though farmers and certified applicators can fix the problem that hydrolysis supposes, by lowering and controlling the pH levels of the water with buffering agents (Deer and Beard 2001), the susceptibility to degradation of a fungicide is a factor that manufacturers must be concerned with. It is their responsibility to offer quality products with standard characteristics that will respond to buffering agents utilized by the customer. The active ingredient concentration and optimal pH water play an important role in the performance of the fungicide during the application and therefore the homogeneity and correlation between these variables are key factors to control during production.

Generally, manufacturers provide the information necessary for the user to correctly apply the fungicide. This includes the level of pH at which the product performs best and the rate at which the fungicide hydrolyzes, among others (McKie and Johnson 2002). In order to accurately present this information, both the active ingredient concentration and the water pH level that will cause the product to degrade must be controlled during the manufacturing process. If these characteristics are too variable, different units of the same product will perform differently for each customer, which makes the fungicide unreliable.

The active ingredient concentration in a fungicide can influence its susceptibility to alkaline hydrolysis, since the latter is usually determined by the characteristics of the first. As mentioned before, depending on the active ingredient used, a fungicide may be resistant to pH levels above 7.0 for up to 10 days but this may vary in a product in this ingredient concentration is inconsistent throughout the production process. Thus, it is important for manufacturers to control the levels of both variables during production as to determine whether they are liable to change for each unit.

In this case, the application of a multivariate control chart, more specifically the  $|S|$  control chart, allows the monitoring of the optimal pH and the concentration of the active ingredient at which the product performs best. For the purpose of analyzing the interaction between these variables and the impact of variations in the fungicide manufacturing process, the  $|S|$  control chart uses the determinant of the variance covariance matrix as a measure of variability (Morales and Vargas 2008). However, the usage of the arithmetic mean as the measure of central tendency leads to inaccurate parameter calculations, especially in the presence of outlying observations. For this reason, robust estimates for the mean vector were necessary to enhance the performance of the control chart.

### Building section

This section is dedicated to a brief overview of existing studies about mean vector estimators and original multivariate control chart proposals. These investigations support the applied methodology in this paper.

Vargas (2003) stated that the usual estimators are very sensitive to outliers in the sample. The  $T^2$  chart and other similar multivariate control charts are based on sensible estimators. In these cases, the statistic plotted performs poorly and is ineffective in the detection of shifts in the mean vector. Under such circumstances, the  $|S|$  chart control limits become unreliable and the statistic may not represent the actual interaction between the variables of interest for a given process. With the objective of eliminating these limitations, Vargas proposed using robust estimators that provide more statistical power than the traditional ones.

Morales and Vargas (2008) claimed that the EWMA control charts are very efficient when monitoring small changes in the variability. In their research, they studied different multivariate control charts, including EWMA, CUSUM and Shewhart-type control charts, whose measure of global variability is the generalized variance. This measure is defined as the determinant of the variance covariance matrix. The  $|S|$  chart, which is one of the Shewhart-type control charts, will be modified to include a robust estimator and will be applied to the manufacturing process of a fungicide.

Moreover, Dixon (1960) reviewed different approaches to the use of robust estimators of mean and standard deviation. In his study, Dixon explained different methods for eliminating or censoring observations with magnitudes that are more extreme than the other observations in the sample. The approach called symmetrical censoring requires that the observations be organized in descending order to facilitate the replacement of outliers from each extreme of the sample. This way, the extreme values are not eliminated but rather the mean estimator is adjusted, thereby making it more reliable. Dixon concluded that for  $k$  observations censored at each extreme this estimator is

$$m_w = \frac{[(k + 1)x_{k+1} + x_{k+2} + \dots + x_{N-k-1} + (k + 1)x_{N-k}]}{N} \tag{1}$$

In the original  $|S|$  chart, for  $i$  random observation samples, let  $S_i$  be the variance–covariance matrix of the  $i$ th sample and  $|S_i|$  its determinant. According to Lee and Khoo (2016), where

$$|S_i| = \left| \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})' \right| \tag{2}$$

With the purpose of reducing the influence of the outliers in the dataset, the Winsorized Mean will be defined as the

measure of central tendency for the proposed chart. Thus, let  $S_{i_{mw}}$  be the variance–covariance matrix of the  $i$ th sample and  $|S_{i_{mw}}|$  its determinant, where

$$|S_{i_{mw}}| = \left| \frac{1}{n-1} \sum_{k=1}^n (X_k - m_w)(X_k - m_w)' \right| \tag{3}$$

When  $|\Sigma_0|$  is unknown, it is necessary to use  $|\bar{S}_i|$  as an estimator, where  $|\bar{S}_i|$  is the mean vector for the  $|S_i|$  determinants obtained in each sampling point. Under these conditions, the control limits for the  $|S|$  chart proposed by Vargas (2006) are defined as:

$$LCL = \max\left(0, |\bar{S}_i| \left(1 - 3 \frac{\sqrt{b_2}}{b_1}\right)\right) \tag{4}$$

$$CL = |\bar{S}_i| \tag{5}$$

$$UCL = |\bar{S}_i| \left(1 + 3 \frac{\sqrt{b_2}}{b_1}\right) \tag{6}$$

Nevertheless, using the  $|\overline{S_{i_{mw}}}|$  as the new mean estimator, which corresponds to the Winsorized Mean of the determinants calculated from the variance–covariance matrices of each subgroup, the control limits for the proposed model are:

$$LCL_{mw} = \max\left(0, |\overline{S_{i_{mw}}}| \left(1 - 3 \frac{\sqrt{b_2}}{b_1}\right)\right) \tag{7}$$

$$CL_{mw} = |\overline{S_{i_{mw}}}| \tag{8}$$

$$UCL_{mw} = |\overline{S_{i_{mw}}}| \left(1 + 3 \frac{\sqrt{b_2}}{b_1}\right) \tag{9}$$

The constants  $b_1$  and  $b_2$  in Eqs. (4), (6), (7) and (9) have been defined by several authors. According to Levinson (2011), they are calculated as follows:

$$b_1 = \frac{1}{(k-1)^p} \prod_{i=1}^p (k-1) \tag{10}$$

And

$$b_2 = \frac{1}{(k-1)^{2p}} \prod_{i=1}^p (k-1) \left\{ \prod_{i=1}^p (k-j+2) - \prod_{i=1}^p (k-j) \right\} \tag{11}$$

In Eqs. (10) and (11),  $k$  and  $p$  represent the sample size for each subgroup  $i$  and the number of variables to study, respectively.



In the fungicide manufacturing process, the active ingredient concentration and the optimal pH of the water are the most important quality characteristics. In this process, these variables are highly correlated, which requires that the analysis be conducted in a multivariate setting, using the  $|S|$  control chart. For the purpose of evaluating chart performance, the  $|S|$  chart proposed by Vargas (2006) using the mean vector and the model with Winsorized Mean were compared and analyzed. The analysis was organized in the following way.

A total of  $i$  subgroups were defined out of a hundred observations that followed a normal distribution. The subgroups were conformed by  $k$  observations for each of the  $p$  quality characteristics. Firstly, the arithmetic mean of each subgroup and the differences between each observation and the estimated mean were calculated. The latter were used to calculate the variance–covariance matrix of each subgroup and its corresponding  $|S_i|$  determinant as shown in Eq. (2). Since  $|\Sigma_0|$  is unknown, Eq. (5) was used to find the central limit of this control chart, which corresponds to  $|\bar{S}_i|$ . Upper and lower control limits are calculated with Eqs. (4) and (6), respectively.

On the other hand, Eq. (1) was used to calculate the Winsorized Mean of every subgroup, censoring 10% of each extreme by replacing the outliers with the magnitude of their lower or higher neighbor. Secondly, the differences between each observation and the new estimated mean were used to calculate the variance–covariance matrix and its corresponding  $|S_{i_{mw}}|$  determinant as shown in Eq. (3).

Furthermore, the central limit of this control chart was defined by the Winsorized Mean of the determinants, as shown in Eq. (8), while constants  $b_1$  and  $b_2$  are calculated with Eqs. (10) and (11), respectively. These values are used in Eqs. (4), (6), (7) and (9) to find the upper and lower control limits. Finally, the values of the determinants  $|S_{i_{mw}}|$  were plotted to evaluate the performance of the process.

## Case study

The object of study is a product manufactured by an agrochemical plant, specialized in crop protection, with a product portfolio designed to increase productivity and help producers protect their fields against weeds, insects and diseases. Primary data were collected from the manufacturing process of a fungicide whose active ingredient is propiconazole. This active ingredient protects the plant from within and is effective against a wide spectrum of diseases that affect banana, coffee and rice crops.

Let the variables  $p_1$  and  $p_2$  be the optimal water pH and the propiconazole concentration, respectively. The two building models described in the previous section were executed in order to evaluate their performance and capability

of shift detection in the mean vector. Additionally, three different subgroup sizes were studied with the proposed control chart, using the Winsorized Mean as the central tendency measure.

The analysis was conducted as follows. First, the set of a hundred observations was divided into  $i = 10$  subgroups with  $k = 10$  observations each. Both the traditional and the proposed control charts were constructed with these parameters. Then, the performance of the proposed control chart was analyzed with different values for these parameters:  $i = 5$  subgroups with  $k = 20$  observations and  $i = 20$  subgroups with  $k = 5$  observations each. Note that only when  $k = 10$  was the traditional chart constructed; the variations in the subgroup sizes were only considered for the proposed chart to evaluate the results of the robust estimator for different construction methods. The graphics were compared to assess the sensibility of the chart for each case.

Table 1 shows the determinants of the variance covariance matrices for both the traditional and the proposed control chart. These results correspond to the  $i = 10$  subgroups, and it is possible to observe the differences in the calculations when the traditional and robust estimators are used.

For the proposed control chart, constants  $b_1$  and  $b_2$  calculated using Eqs. (10) and (11) are 0.8889 and 0.4170, respectively. The three values for upper, central and lower control limits are 0.1263, 0.0397 and 0.0000. The central limit is the Winsorized Mean of determinants set  $|S_{i_{mw}}|$ , and Eq. (3) was used to calculate this value. The values for constants  $b_1$  and  $b_2$  are the same for the traditional chart, while the corresponding upper, central a lower control limits, using the mathematical average, are 0.1371, 0.0431 and 0.000 and are obtained using Eqs. (4), (5) and (6), respectively.

In Fig. 1, it can be observed that determinant  $|S_{6_{mw}}|$  is an out-of-control point. This value is higher than the upper control limit with a difference between both values of 0.0108. Moreover, the chart shows high variability with determinants that are considerably higher or lower than the central

**Table 1** Determinants of  $i = 10$  subgroups

Subgroup	$ S $ control chart with $m_w$	$ S $ control chart with $\bar{x}$
1	0.0401	0.0401
2	0.0271	0.0281
3	0.0810	0.0802
4	0.0506	0.0508
5	0.0397	0.0353
6	0.1371	0.1251
7	0.0066	0.0042
8	0.0531	0.0481
9	0.0053	0.0053
10	0.0140	0.0141

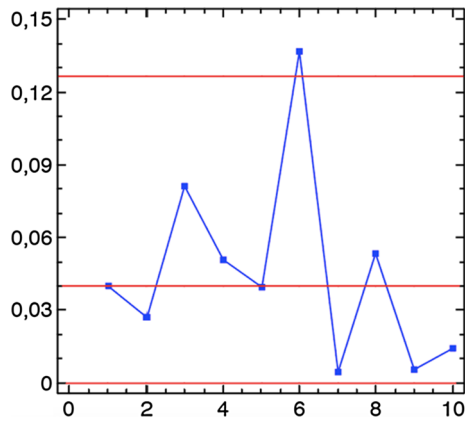


Fig. 1 Proposed  $|S_{mw}|$  chart for  $i = 10$

line value but do not represent out-of-control signals. These signals, combined with the non-random pattern observed in the behavior of the control chart, are indicators of an alarm sign that describes a statistically unstable process.

In order to analyze the out-of-control signal shown in Fig. 1, it was necessary to study the composition of the subgroup that caused  $|S_{6_{mw}}|$  to be higher than the UCL of the chart. Such analysis helps determine whether the alarm signal is a result of high process variability or it can be considered a false alarm. The standard deviation for the  $i = 10$  subgroups of  $|S_{mw}|$  chart is illustrated in Fig. 2. This graphic shows that  $|S_{6_{mw}}|$  has the second highest deviation in the active ingredient concentration values, which is an indicator of heterogeneity in this variable. However, that is also the case for many of the other subgroups in this chart configuration. The determining factor for the out-of-control signal in subgroup six is the high deviation value in the pH variable of the subgroup.

High variability in both of the studied variables results in greater differences between the observations of any given subgroup and the robust mean estimator used. This is also an

indicator that the manufacturing process is failing to maintain the specification values and their variability within the tolerance interval and, therefore, can lead to uneven production batches and affect the reliability in the product. Note that other subgroups have higher deviation values in one of the variables but they do not represent an out-of-control signal because the other variable is more stable. Since the  $|S|$  control chart focuses on the correlation between the  $p$  variables included, an alarm signal is presented when this correlation indicates disperse characteristic values that fall outside the control limits of the manufacturing process, thereby making it easier to identify efficiency problems in the production system.

Figure 3, in contrast, corroborates that the traditional method of construction, i.e., using the arithmetic mean, does not detect out-of-control signals, since all the points in the graphic stay within the limits of the control chart. In this case, the highest determinant value is also  $|S_6|$  but the upper control limit is higher by 0.012 points. Although both charts show a similar behavioral pattern, the traditional method is proven to be less sensitive to the largest observations in subgroup six. The results obtain from both charts, albeit fairly similar, provide different information about the same process. In fact, the absence of out-of-control signals in the traditional chart may be enough for manufacturers to ignore the variability in their processes.

For the second scenario analyzed, the observations were divided into  $i = 5$  subgroups of  $k = 20$  observations each. Table 2 shows the determinants for each variance–covariance matrix calculated from the sample. The control limits of the chart were found using the corresponding equations and are equal to 0.1163, 0.0478 and 0.000 for the upper, central and lower control limit, respectively. In this scenario, constant  $b_1 = 0.9474$  and constant  $b_2 = 0.2047$ ; note that these values are different from those obtained earlier with a smaller subgroup size because the parameter  $k$  changed in Eqs. (10) and (11).

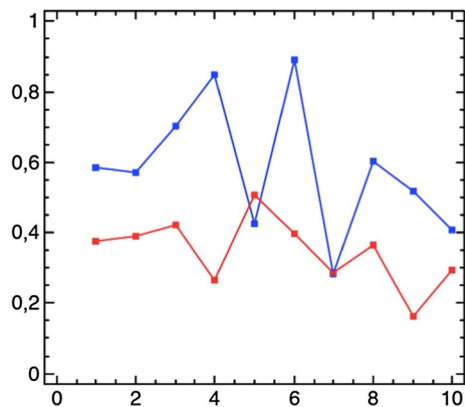


Fig. 2 Standard deviation for each subgroup when  $i = 10$

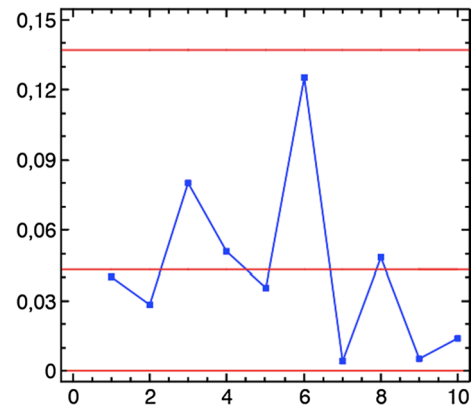
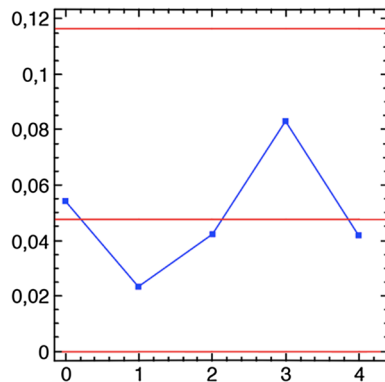


Fig. 3 Traditional  $|S|$  chart for  $i = 10$

**Table 2** Determinants of  $i = 5$  subgroups

Subgroup	Determinants for $ S $ control chart with $m_w$
1	0.0551
2	0.0243
3	0.0431
4	0.0841
5	0.0429



**Fig. 4**  $|S_{mv}|$  chart for  $i = 5$

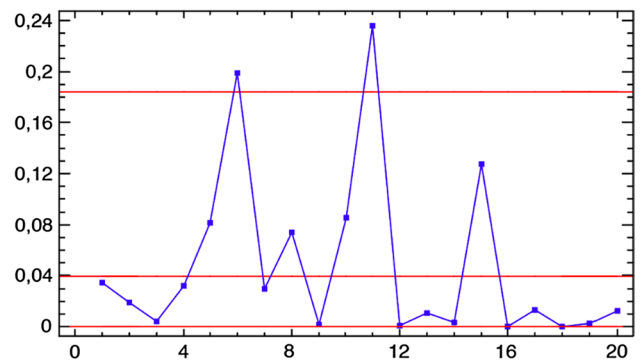
Figure 4 shows that the multivariate control chart does not detect out-of-control signals. The determinant of subgroup  $|S_{4,mw}|$  is the highest plotted point in the chart. However, the difference between this point and the UCL has a magnitude of 0.0679, which is very significant considering the performance of the control chart when each subgroup had fewer observations. Furthermore, all the determinants are close to the value of the process mean vector indicating that under these new conditions, the control chart is far more centered than the previous scenario analyzed, as shown in Table 3.

Lastly, a third scenario was analyzed, where  $k = 5$  observations per subgroup. Using Eq. (1), the Winsorized Mean was calculated for each one of the 20 subgroups defined. Only one of the top and bottom values of the observations were censored, which means that three out of the five original observations in each subgroup were included in the calculation of its mean value. Since the parameter  $k$  changed in Eqs. (10) and (11), the resulting  $b_1$  and  $b_2$  were 0.7500 and 0.8437, respectively. The values for the control limits, obtained using the corresponding Eqs. (7), (8) and (9), are  $UCL_{mw} = 0.1840$ ;  $CL_{mw} = 0.0394$ ; and  $LCL_{mw} = 0.000$ .

In Fig. 5, it can be observed that the control chart detected two out-of-control signals in the process,  $|S_{6,mw}|$  and  $|S_{11,mw}|$ . Although this chart was constructed using different parameters, its behavior is congruent with the results of the other

**Table 3** Determinants for  $i = 20$  subgroups

Subgroup	Determinants for $ S $ control chart with $m_w$
1	0.0342
2	0.0190
3	0.0044
4	0.0323
5	0.0813
6	0.1985
7	0.0293
8	0.0740
9	0.0014
10	0.0851
11	0.2358
12	0.0007
13	0.0105
14	0.0033
15	0.1274
16	0.0001
17	0.0131
18	0.0001
19	0.0029
20	0.0122



**Fig. 5**  $|S_{mv}|$  chart for  $i = 20$

methods.  $|S_{11,mw}|$  is the highest point in the chart, with a difference of 0.0518 between its value and the upper control limit; this point and the determinant for subgroup six when  $k = 10$  are analogous, which substantiates that the robust estimator in the construction method helps identify the largest observation in the dataset by indicating alarm signals in the subgroups in which they are included. In this case, there is a second out-of-control signals that the other charts had not detected. However, this point is consistently the second highest in every other chart, underlining the fact that the effect of the robust estimator intensifies as the subgroups



size decreases. Besides the out-of-control signals, the graphic when  $k = 20$  shows the lowest points found in the analysis across the cases, with  $|S_{16_{mv}}|$  and  $|S_{18_{mv}}|$  being nearly the same value as the lowest control limit.

## Conclusions

This paper focused on the analysis of two important variables involved in the manufacturing process of a fungicide, using a multivariate control chart with a robust mean estimator. Three separate cases were created from the same case study by changing the parameters established for structuring the chart. The charts were presented in order to test their accuracy in performance and effectiveness. In addition, the original control chart introduced by Vargas was compared to one of the cases of the proposed chart to determine whether or not differences can be found when using robust estimators in the mean calculation.

The information available in the control chart shown in Fig. 1 indicates that the process is mostly unstable (i.e., the majority of the points are not centered) but the statistic plotted stays within the control limits. There is, however, an out-of-control sign that represents existing problems in the process. In terms of fungicide manufacturing, this means that the final product does not meet the requirements and could be subject to hydrolysis when mixed into water. On the other hand, Fig. 3 presents a control chart obtained by using the traditional construction method, which does not detect out-of-control signals in the process.

Notwithstanding the similarities between the proposed and traditional methods, it was found that if the latter is used, the resulting chart is less sensitive to small shifts in the mean vector. This serves as evidence that the usage of robust estimators in the construction method of the chart, such as the Winsorized Mean, leads to a more effective process variability diagnosis. Furthermore, there are great differences between the control limits for the proposed and traditional chart, namely the central control limit. Since the CL is defined by the central tendency measure used, it can be expected that these values influence the behavior of the plotted statistic. In this instance, the value of the central limit for the traditional chart caused the upper control limit to be higher and, consequently, no points move outside the limits established.

Although they were based on the same set of observations, Figs. 1 and 4 are drastically different. With 20 observations for each subgroup, the third control chart does not present out-of-control signals and the determinants are less dispersed. This does not necessarily imply that the process is in control but rather serves as a sensitivity measurement for the second subgroup arrangement. Figure 5 presents the

opposite scenario, with more out-of-control signals than the previously studied cases. Nevertheless, this is not enough justification to claim that the chart performs better under these conditions because the quantity of censored data with smaller subgroup sizes may sacrifice a significant part of the original collected data.

With  $k = 5$ , it is not possible to symmetrically censure less than 20% off each tail. This means that in each subgroup two values were replaced, causing the sample to have fewer original data. By using these parameter values, 40 out of 100 observations are censored, which can cause the results to be less reliable. Moreover, in cases where there is extreme variation among the data, fewer data per subgroup size will result in a control chart similar to a traditional Shewhart-type one because extreme values will remain in each subgroup and the mean estimator will still be influenced by large observations. On the other hand, when the subgroups have more observations, the control chart is unable to effectively detect the outliers in the sample, and the censored data may not be enough to obtain an effective mean estimator. Consequently, conclusions based on this chart can also be misleading and inaccurate.

Generally speaking, the outlying observations stand out more in smaller subgroup sizes when the Winsorized Mean is used. As a result, the control chart in Fig. 1 can provide more meaningful information about the fungicide manufacturing process and its performance than Fig. 3. Even though subgroups with fewer observations may be more effective for outlier detection, many of the original data are lost and the veracity of the conclusions might be questioned. The comparative results show that variations in the chart construction method affect the conclusions about process performance and out-of-control signals. Therefore, it is recommended that the collected data be analyzed prior to applying the control chart, with the objective of establishing parameter values that might be more appropriate for the situation.

In terms of the fungicide manufacturing process, out-of-control signals indicate that there is not homogeneity in the production batch. If the determinant of a subgroup causes an out-of-control point in the control chart, it means that the products included in the subgroup sample had a high variation in the propiconazole concentration and optimal pH water levels from unit to unit. As a consequence, some of the units will hydrolyze more rapidly than others because of their lower active ingredient concentration and lower pH water resistance. In this highly competitive market, manufacturers are required to not only apply statistical process control tools to their production systems but also choose correct methods to identify alarm signs that may not be detected with inadequate measures. This will directly affect the quality of their products and customer satisfaction levels.

Lastly, future studies may compare the proposed multivariate control chart with an already existing multivariate

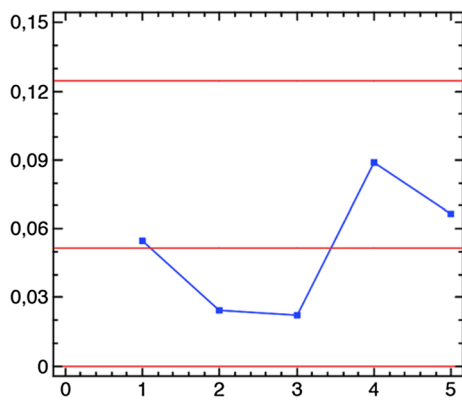


Fig. 6  $|S|$  chart for  $i = 5$

control chart, such as MEWMA in order to assess the effectiveness of each method for a given manufacturing process. Furthermore, non-symmetrical censoring methods may be applied to small subgroup sizes to decrease the number of original observations censored while still applying robust estimators for central tendency measures. This would reduce the impact seen on the control chart when the Winsorized Mean was used.

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

### Appendix: Table of parameters

Sign	Description
$b_1, b_1$	Constant values used to calculate the control limits
$i$	Index; number of subgroups into which the dataset was divided
$j$	Index
$k$	Index; number of observations per subgroup $i$
$m_w$	Winsorized Mean
$N$	Count of numbers of the Winsorized Mean
$p$	Number of variables included in the multivariate analysis
$S_i$	Variance–covariance matrix
$S_{i_{mw}}$	Variance–covariance matrix calculated using the Winsorized Mean
$ S_i $	Determinant of the variance–covariance matrix
$ S_{i_{mw}} $	Determinant of the variance–covariance matrix calculated with the Winsorized Mean

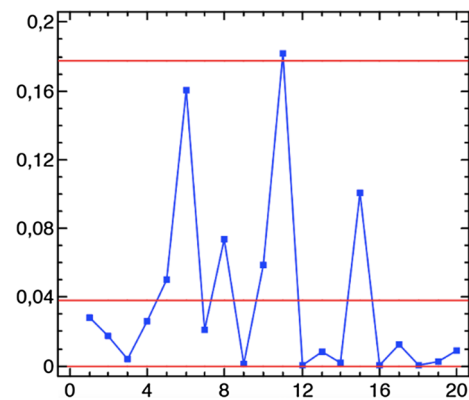


Fig. 7  $|S|$  chart for  $i = 20$

Sign	Description
UCL	Upper control limit
CL	Central control limit
LCL	Lower control limit
$UCL_{mw}$	Upper control limit calculated with the Winsorized Mean
$CL_{mw}$	Central control limit calculated with the Winsorized Mean
$LCL_{mw}$	Lower control limit calculated with the Winsorized Mean
$ \Sigma_0 $	Determinant of the variance–covariance matrix of a statistical population

Figure 6 shows the behavior of the variables when the control chart introduced by Vargas is used with  $i = 5$  subgroups. It can be observed that, in comparison with the proposed control chart, the performance follows a similar pattern since there are no out-of-control signals. However, there are variations in the values of determinants  $|S_3|$  and  $|S_5|$  between one control chart and the other, which is caused by fluctuations in the behavior of the variables involved in the analysis. For  $|S_3|$ , the value of this determinant with the original control chart is the lowest of the determinants set, while in the proposed chart with the Winsorized Mean, the value of this determinant is considerably closer to the central control limit. On the other hand, the value of determinant  $|S_5|$  is greater than the central control limit in the control chart introduced by Vargas but in the proposed chart, it is lower than this limit.

In Fig. 7, it can be observed that the original control chart is able to detect one out-of-control signal in the manufacturing process of fungicides. Although this value is greater than the upper control limit, it is relatively close to its value with a difference of 0.0048 between them. In comparison, the control chart proposed using the Winsorized Mean detected two out-of-control signals in the manufacturing process analyzed indicating more sensibility. Lastly, both control charts present a similar behavior and follow the same pattern.

## References

- Al-Khazaleh AMH, Al Wadi S, Ababneh F (2015) Wavelet transform asymmetric Winsorized Mean in detecting outlier values. *Far East J Math Sci (FJMS)* 96:339–351
- Costa FSL, Pedroza RHP, Porto DL, Amorim MVP, Lima KMG (2015) Multivariate control charts for simultaneous quality monitoring of isoniazid and rifampicin in a pharmaceutical formulation using a portable near infrared spectrometer. *J Braz Chem Soc* 26:64–73. <https://doi.org/10.5935/0103-5053.20140214>
- Deer HM, Beard R (2001) Effect of water pH on the chemical stability of pesticides. *Utah State University Extension* vol 14, pp 1–3. <https://extension.usu.edu/>. Accessed 31 Jan 2019
- Dixon WJ (1960) Simplified estimation from censored normal samples. *Ann Math Stat* 2:385–391
- Du S, Lv J, Xi L (2012) On-line classifying process mean shifts in multivariate control charts based on multiclass support vector machines. *Int J Prod Res* 50:6288–6310. <https://doi.org/10.1080/00207543.2011.631596>
- Du S, Huang D, Lv J (2013) Recognition of concurrent control chart patterns using wavelet transform decomposition and multiclass support vector machines. *Comput Ind Eng* 66:683–695. <https://doi.org/10.1016/j.cie.2013.09.012>
- Goodwyn F (2012) The utility of robust means in statistics. Paper presented at the annual meeting of the Southwest Educational Research Association, Texas A&M University, pp 1–13
- Haridy S, Wu Z (2009) Univariate and multivariate control charts for monitoring dynamic-behavior processes: a case study. *J Ind Eng Manag* 3:464–498. <https://doi.org/10.3926/jiem.v2n3.p464-498>
- Huber PJ (1981) *Robust statistics*. Wiley, Hoboken
- Lee MH, Khoo MBC (2016) Optimal designs of multivariate synthetic ISI control chart based on median run length. *Commun Stat Theor Methods* 4:3034–3053. <https://doi.org/10.1080/03610926.2015.1048884>
- Levinson WA (2011) *Statistical process control for real-world applications*. Taylor & Francis Group, LLC, New York
- McGrath MT (2009) What are fungicides? *Plant Health Instr*. <https://doi.org/10.1094/phi-i-2004-0825-01>
- McKie P, Johnson W (2002) Water pH and its effect on pesticide stability. University of Nevada, Reno. <https://www.unce.unr.edu/>. Accessed 01 Feb 2019
- Montgomery DC (2009) *Introduction to statistical quality control*, 6th edn. Wiley, Hoboken
- Morales V, Vargas JA (2008) EWMA chart based on the effective variance for monitoring the variability of multivariate quality control process. *Rev Colomb Estad* 31:131–143
- Rivest LP (1994) Statistical properties of Winsorized Means for skewed distributions. *Biometrika* 81:373–383
- Rogalewicz M (2012) Some notes on multivariate statistical process control. *Manag Prod Eng Rev* 4:80–86. <https://doi.org/10.2478/v10270-012-0036-7>
- Schilder A (2008) Effect of water pH on the stability of pesticides. Michigan State University Extension. <http://msue.anr.msu.edu>. Accessed 30 Aug 2018
- Secret R (n.d.) How products are made: pesticide. <http://www.madehow.com/Volume-1/Pesticide.html>. Accessed 02 Feb 2019
- Smart NA (2003) *Encyclopedia of food sciences and nutrition*. Elsevier, Maryland
- Stefatos G, Hamza AB (2009) Fault detection using robust multivariate control chart. *Expert Syst Appl* 3:5888–5894. <https://doi.org/10.1016/j.eswa.2008.07.028>
- Thompson B (2006) *Foundations of behavioral statistics: An insight-based approach*. The Guilford Press, New York
- Vargas JA (2003) Robust estimation in multivariate control charts for individual observations. *J Qual Technol* 4:367–376. <https://doi.org/10.1080/00224065.2003.11980234>
- Vargas JA (2006) *Control Estadístico de Calidad*. Universidad Nacional de Colombia, Bogotá
- Wilcox RR (2001) *Fundamentals of modern statistical methods*. Springer, New York