Constructing a Control Chart Using Functional Data

Miguel Flores [1], **Salvador Naya [2]**, **Rubén Fernández-Casal [3]**, **Sonia Zaragoza [4]**,
Paula Raña [3] and **Javier Tarrío-Saavedra [5,*]**

[1] MODES Group, Department of Mathematics, Escuela Politécnica Nacional, 170517 Quito, Ecuador;
miguel.flores@epn.edu.ec

[2] MODES Group, CITIC, ITMATI, Department of Mathematics, Escola Politécnica Superior, Campus
Industrial, Universidade da Coruña, Mendizábal s/n, 15403 Ferrol, Spain; salva@udc.es

[3] MODES Group, CITIC, Department of Mathematics, Faculty of Computer Science, Campus de Elviña,
Universidade da Coruña, 15008 A Coruña, Spain; ruben.fcasal@udc.es (R.F.-C.); paula.rana@udc.es (P.R.)

[4] PROTERM Group, Department of Naval and Industrial Engineering, Campus Industrial, Universidade da
Coruña, Mendizábal s/n, 15403 Ferrol, Spain; szaragoza@udc.es

[5] MODES Group, CITIC, Department of Mathematics, Escola Politécnica Superior, Campus Industrial,
Universidade da Coruña, Mendizábal s/n, 15403 Ferrol, Spain

* Correspondence: javier.tarrio@udc.es

check for
updates

Abstract: This study proposes a control chart based on functional data to detect anomalies and estimate the normal output of industrial processes and services such as those related to the energy efficiency domain. Companies providing statistical consultancy services in the fields of energy efficiency; heating, ventilation and air conditioning (HVAC); installation and control; and big data for buildings, have been striving to solve the problem of automatic anomaly detection in buildings controlled by sensors. Given the functional nature of the critical to quality (CTQ) variables, this study proposed a new functional data analysis (FDA) control chart method based on the concept of data depth. Specifically, it developed a control methodology, including the Phase I and II control charts. It is based on the calculation of the depth of functional data, the identification of outliers by smooth bootstrap resampling and the customization of nonparametric rank control charts. A comprehensive simulation study, comprising scenarios defined with different degrees of dependence between curves, was conducted to evaluate the control procedure. The proposed statistical process control procedure was also applied to detect energy efficiency anomalies in the stores of a textile company in the Panama City. In this case, energy consumption has been defined as the CTQ variable of the HVAC system. Briefly, the proposed methodology, which combines FDA and multivariate techniques, adapts the concept of the control chart based on a specific case of functional data and thereby presents a novel alternative for controlling facilities in which the data are obtained by continuous monitoring, as is the case with a great deal of process in the framework of Industry 4.0.

Keywords: functional data analysis; statistical process control; control chart; data depth; nonparametric control chart; energy efficiency

1. Introduction

Generally, univariate and multivariate control charts are applied to identify anomalies in the industry and control the quality of products and services. However, the specific characteristics of data obtained by continuous monitoring (which has become the main trend due to advancements in sensoring and communications in the framework of Industry 4.0), require the use of increasingly sophisticated tools that take into account the presence of autocorrelation in critical to quality (CTQ) variables for the process under study.

Recently, diverse solutions have been provided that propose the modification and application of exponentially weighted moving average chart (EWMA) control charts [1]; control charts based on the fitting of autoregressive integrated moving average (ARIMA) models [2]; and the use of control charts for profiles, which is understood as the control of the parameters that define the relationship between two different CTQ variables [3–6]. Additionally, there are solutions that propose the use of techniques based on the control chart concept for anomaly detection such as machine learning techniques (neural networks and support vector machines, among others) and time series [7–14]. The increasingly common use of these tools is attributed to the fact that they adapt very well to the new paradigm of data (in the framework of Industry 4.0) defined by the continuous monitoring of numerous variables. Traditional control charts cannot be applied to many of the new cases due to their non-compliance with their starting hypotheses. In the domain of statistical quality control, the use of the concept of latent variables along with multivariate control charts is one of the most popular and useful alternatives to solve the problem of process control in the Industry 4.0. domain [15–17].

Many of these new data, usually curves, can be studied as functional data. This is the case of data in the fields of energy consumption, indoor and machine temperatures, relative humidity, amount of CO_2, among other variables, measured in all types of buildings. These new data, which currently characterize industrial processes, require innovative solutions, developed by researchers in the field of statistical quality control, based on the application of functional data analysis (FDA) techniques. Until now, few works have been published (in relation to the importance of the topic and the frequent monitoring of this type of data) on controlling the quality of processes when the data to be monitored are functional CTQ variables. Among the most outstanding research in this context, Reference [18] compared the performances of different control charts used for monitoring functional data; these charts are often identified as profiles in the statistical quality control literature. According to References [19,20], in the case of the monitoring of profiles, a set of statistical techniques (usually multivariate) is applied for controlling processes when they are defined by the functional relationship between two variables. However, a control chart for monitoring functional data, which is based on bootstrap resampling, is also proposed in Reference [21], whereas, in References [22,23], two complete monographs are proposed in which the concepts of control charts for profiles are adapted to the context of the functional data. Moreover, recently, new alternatives for outlier detection based on interlaboratory studies and those with application in industrial anomaly detection have been performed in the FDA framework [24–26].

This study intends to provide an alternative solution to the detection of out-of-control anomalies or states in the area of energy efficiency in buildings, specifically in commercial areas such as different stores of textile companies. Therefore, an important aim of this work is to solve the problem of such companies, specifically the ones providing lighting facilities; office automation; and heating, ventilation and air conditioning (HVAC). One such company is Fridama SL. The group comprises Fridama's facilities, Σqus (web platforms for big data management) and Nerxus (statistical consultancy for data analysis in the energy sector); the solution to the problem raised by this company is defined by continuously monitored data over time that can be treated as profiles or functional data. Based on this need, the present study proposes methodologies to build control charts that allow us to control the aforementioned processes. in order to test the applied statistical methodologies for anomaly detection properly, these companies have provided a database on a real HVAC installation, whose anomalies and assignable causes are identified by its maintenance personnel.

These energy efficiency data are non-Gaussian and autocorrelated, as the main component of continuously monitored data in the Industry 4.0 framework. Thus, the present proposal provides alternative control charts for Phases I and II of the statistical process control that can deal with non-Gaussian and autocorrelated data. As is well-known, using control charts, the process can be controlled in the following two phases: Phase I involves the stabilization or the calibration of the process and Phase II focuses on process monitoring. For Phase I, a control chart is proposed based on the depth measurement of a functional datum and the idea of atypical detection; in Phase II, a nonparametric range control chart is proposed, based on the calculation of the functional data depth,

to monitor the process of interest. The procedure of building control charts corresponding to Phases I and II are explained in the next section; it also shows the results of their performance through a simulation study including diverse scenarios. The proposed methodology has been developed and programmed in R through different functions integrated in the "Quality Control Review" package, qcr, which can be freely and easily accessed by the practitioners.

Additionally, it provides a way to visualize the control charts for functional data, including the original data and the curves corresponding to the estimated control limits. This visual tool allows users and maintenance managers to relate each anomaly to an intuitively assignable cause. The results of the simulation study and its application to the real data show the usefulness of this control chart methodology in detecting anomalies when the process is defined by functional data, specifically the daily curves of energy consumption in commercial areas.

1.1. Alternatives of the Statistical Quality Control When the Basic Assumptions of the Control Charts Are Not Met

This control chart approximation for functional data is based on nonparametric multivariate control charts, which are useful when the assumptions of Gaussianity are not met. Therefore, a brief introduction of nonparametric control charts and statistical tools for dealing with autocorrelated data are presented in the following section.

Recently, analyses of the robustness of different control charts against the non-compliance of Gaussianity hypothesis have been developed. Particularly, several nonparametric control charts have been proposed. In this domain, we should highlight the works of Regina Liu [27–29], who has developed the r, Q and S control charts based on the data depth and rank concepts and the rank or ranges. In this regard, it is important to highlight the strong influence of Liu's works; it must be noted that several nonparametric alternatives for control charts are based on the concepts of data depth and rank. Moreover, it should be emphasized that one of the most important lines of research of the SQC, the profiles' control charts, are based, in many cases, on the application of nonparametric or semiparametric regression models [3,4]. Additionally, it is also important to highlight the use of resampling techniques for calculating natural control limits for different types of control charts [30,31]. The work in Reference [32] constitutes a complete monograph on the current trends for the development of control charts.

Despite their important advantages (no assumptions on the probability distribution of CTQ variables), nonparametric control charts are not predominantly used in the industrial and business framework. As noted in Reference [20], this can be attributed to several factors such as the lack of specific software, both commercial and free; the lack of general training in nonparametric statistics that generates insecurity and distrust in users and the lack of contrasted reference texts for the application of nonparametric methods in SQC. However, the research activity in this field is growing, as illustrated by the work in Reference [33].

Another starting hypothesis of control charts is the independence of the observations. The data continuously monitored over time by different sensors usually show a variable level of autocorrelation (the greater the correlation the closer the observations are in time). The application of standard techniques in the case of the violation of the independence hypothesis often results in the detection of an unacceptable number of false alarms [34]. Therefore, the development and analysis of techniques that remove the sample autocorrelation is fully justified. Within these techniques, the most widespread is the application of time series models (e.g., autoregressive moving average (ARMA) and ARIMA) to remove the correlation between observations and the subsequent monitoring of the error variable (difference between actual values and those estimated by the model) using control charts [34–36]. Moreover, References [37,38] propose the combination of control charts with adjustment algorithms. Finally, the monographs of [39,40] describe the most relevant lines of research for controlling and monitoring autocorrelated data.

It is important to note that this type of data is related to the control charts for functional data. The use of FDA techniques allows us to consider the autocorrelation of the data as well as, by means of resampling techniques, to circumvent parametric assumptions about the trend and dependence.

1.2. Control Charts for Phases I and II

The studied process is stabilized in the framework of the Phase I control charts, that is, the process is left under control [41]. This implies that there are no other assignable causes in the process, except for those present due to randomness itself (non assignable causes). This is equivalent to stating that the process remains stable—the parameters of the probability distribution of the CTQ characteristic remain unchanged. This allows us to estimate the natural control limits for the variable that describes the quality of the process. The natural control limits are estimated using a preliminary or calibration sample. Thus, in this stage, the main assignable causes of variation are identified and rectified through corrective actions. After the variations are removed, we estimate the natural control limits of the CTQ variable corresponding to the process under control. While most of the literature in the field has referred to Phase II control charts, within the new data paradigm for Industry 4.0, it would be essential to develop methodologies for Phase I [19]. In this regard, Reference [42] presents an overview of the recent contributions to the development of Phase I control charts. Among the latest proposals for control charts, the work of Grasso et al. [43] is especially interesting taking into account that is a Phase I control chart proposal for profiles based on the use of functional data depth concept (as well as the case of the present proposal). In fact, they proposed a Phase I control chart methodology for profiles belonging to the "multi-modelling framework", that includes the following stages: (1) a classification stage of new profiles into different operating modes or profile patterns using functional classification techniques from functional data depth measures (maximum depth approach based on Mode depth); (2) even the identification of a novel operation mode is included; (3) Once the operation mode corresponding to a new profile is identified, this is assigned to the corresponding control chart and consequently a suitable control charting method is applied to determine if the process remained in control over the period of time where those data were collected.

In this study, the atypical detection procedure presented in Reference [44] is adapted to develop a control chart for Phase I; this approach allows us to obtain a calibration sample from an under-control process that can be monitored through new measurements using a Phase II rank control chart. In fact, in Phase II, it is assumed that the process is under control [45]; additionally, in each new sample (monitored sample), the statistical rank is obtained and it is represented in the control chart with the estimated lower control limit.

2. Methodology

In this section, methodologies of control charts for functional random variables, **X**, which take values in a functional space $\mathcal{E} = L^2(T)$, with $T \subset \mathbb{R}$, are developed.

Based on observations of the functional variable **X**, we obtain a sample each of calibration and monitoring; they are functional datasets of sizes n and m, respectively, which allow us to build control charts for Phase I (in the case of the calibration sample) and Phase II (from the monitoring sample).

In the case of designing the control charts, any unstable or out-of-control process is refereed to the assignable causes of variation emerging from unusual and avoidable events that interrupt the process, that is, when they cause a change in the parameters of the underlying model of the profile or functional data [46]; these variations can be eliminated from the data by identifying and acting on the cause; this approach will avoid such variations in the future [46].

Concerning a method for building quality control charts, the probability of process instability (level of significance) presents a measure of its performance. This probability, provided that it is within the H_0, allows the derivation of at least one measure (observed value of the statistic) outside the control limits [18]. Phase I involves the development of a method and the estimation of the level of

significance; however, in the case of Phase II, is the process is assumed to be under control, that is, the level of significance is fixed.

2.1. Procedure for Building a Control Chart for Phase I (Stabilization)

As mentioned above, in Phase I, the retrospective data corresponding to a calibration sample of size n is analyzed for evaluating the stability of the process whose quality is ascertained over time and for estimating the parameters of the control chart [18].

In Phase I, a control chart is used to test the hypothesis that there is no change in the distribution of observations of the variable ordered with respect to time $\{\mathcal{X}_1(t), \mathcal{X}_2(t), \ldots, \mathcal{X}_n(t)\}$.

These changes can be punctual (freaks or bunches) or they can be related to a change in process that is evaluated (observable through patterns of sudden or gradual change in the mean of the process). Concerning isolated changes, it refers to the occurrence of at least one observation of the observed variable that deviates from the distribution of the other observations [47]. The hypothesis tested in Phase I is:

$$
\begin{aligned}
&H_0 : \mathcal{X}_i(t) \stackrel{d}{=} \mathcal{X}_j(t), \forall i, j \in \{1, \ldots, n\} \\
&H_a : \mathcal{X}_i(t) \stackrel{d}{\neq} \mathcal{X}_j(t), \text{for some } i, j \in \{1, \ldots, n\}.
\end{aligned}
\tag{1}
$$

The stabilization phase of a process consists of applying an iterative method that allows the detection and elimination of those observations (in this context curves) that have a deviation with respect to the shape or magnitude of most of the observed curves. In other words, a curve is an atypical value if it has been generated by a different stochastic process or there is a change in the trend or variability of the stochastic process with respect to that corresponding to the remaining data [48]. The quantity of outliers is assumed to be unknown, although small.

In this study, the proposed control charts for Phase I are based on the adaptation of outlier detection methodologies from functional data, based on the data depth calculations [44].

The method of the detection of outliers for functional data [44] considers an atypical curve if its depth is less than a specific quantile of the distribution of depths estimated by bootstrap. In other words, an atypical curve will have a significantly low depth.

This procedure can be used with different types of functional depths. In the library `fda.usc` [49], the following alternatives are offered: Fraiman and Muniz (FM) [50,51], mode depth [52] and random projections depth (RP) [52,53]. In this work, the outlier detection procedure proposed by Febrero et al. `fda.usc` has been adapted to estimate a specific quantile of the depth distribution that plays the role of the lower control limit (LCL) for a Phase I control chart.

The control chart proposed for Phase I is estimated and plotted from a depth measurement (FM, RP or mode) and only the lower control limit (LCL) is considered to detect if the process is out-of-control (the depth of a curve is less than the LCL). In addition to this representation, an additional chart showing the original curves is proposed to provide an intuitive idea about the cause behind the identified anomaly (by checking its shape or magnitude) and thus to identify assignable causes. For instance, in the case of the HVAC energy consumption in buildings, anomalies may include stopping air handler, failure of the counter or sensor, a change in the regulation of the machines and adverse weather conditions.

In Phase I, we consider the functional random variable $\mathcal{X}$, from which a random sample is drawn—$\{\mathcal{X}_1(t), \mathcal{X}_2(t), \ldots, \mathcal{X}_n(t)\}$. These data are used to formulate the following steps to build the control chart for Phase I:

1. Calculate the depth corresponding to each observation of the dataset, $D(\mathcal{X}_i)_{i=1}^n$ and make a control chart based on the depth of each datum with respect to time.
2. Choose the parameter LCL according to the significance level of the control chart, that is, the percentage of false alarms (observations under control but erroneously detected as

out-of-control) is small (for example, $\alpha = 1\%$). The following procedures are used to estimate the LCL.

- Bootstrap procedure based on Trimmed:

  - Reorder the curves according to their depths in a decreasing way. $\mathcal{X}_{(1)}, \ldots, \mathcal{X}_{(n)}$.
  - It is assumed that, at most, $\alpha\%$ of the sample can be considered outliers.
  - B samples are obtained by a smoothed bootstrap procedure from the dataset resulting from discarding the $\alpha\%$ of the less depth curves. Let $\mathcal{X}_i^{*b}, i = 1, \ldots, n, b = 1, \ldots, B$ these bootstrap samples. To obtain each bootstrap resample:

    * A uniform sampling is done, i^* of $1, \ldots, [n(1 - \alpha)]$.
    * Z_{i*} is generated as a Gaussian process with zero mean and variance and covariance matrix.

      $\gamma \Sigma_{\mathcal{X}}$ with $\gamma \in [0, 1]$, where $\Sigma_{\mathcal{X}}$ is the variance and covariance matrix of the observations $\mathcal{X}_{(1)}, \ldots, \mathcal{X}_{([n(1-\alpha)])}$.
    * Finally, you get $\mathcal{X}_i^{*b} = \mathcal{X}_{(i*)} + Z_{i*}$.

  - For each $b = 1, \ldots, B$, we obtain C^b, which is the $\alpha\%$ empirical quantile corresponding to the distribution of the depths, $D(\mathcal{X}_i^{*b})$. The output value, C = LCL, is the median of the values $C^b, b = 1, \ldots, B$.

- Bootstrap procedure based on weighing:

  - The depths of the $\mathcal{X}_1, \ldots, \mathcal{X}_n$ curves are obtained.
  - B samples are obtained through a smoothed bootstrap from the original dataset weighted by their depths. Let $\mathcal{X}_i^{*b}, i = 1, \ldots, n, b = 1, \ldots, B$ these bootstrap samples. These replicas would be obtained:

    * Weighted sampling is performed, with i^* of $1, \ldots, n$ and with probability proportional to $D(\mathcal{X}_1), \ldots, D(\mathcal{X}_n)$.
    * Z_{i*} is generated as a Gaussian process with zero mean and variance-covariance matrix $\gamma \Sigma_{\mathcal{X}}$, with $\gamma [0, 1]$, where $\Sigma_{\mathcal{X}}$ is the variance and covariance matrix of the observations $\mathcal{X}_1, \ldots, \mathcal{X}_n$.
    * Finally, we get $\mathcal{X}_i^{*b} = \mathcal{X}_{i*} + Z_{i*}$

  - For each $b = 1, \ldots, B$, we get C^b, which is the empirical quantile corresponding to the $\alpha\%$ of the distribution of the depths, $D(\mathcal{X}_i^{*b})$. The final value C = LCL is the quantile β of the values C^b, with $b = 1, \ldots, B$.

3. If there is any curve such that $D(\mathcal{X}_i) \leq$ LCL for a given LCL, then it would be considered atypical and the process would be out-of-control.

4. Additionally, a control chart including the original curves and the functional envelope obtained from 99% of the deeper bootstrap replicas is also developed.

Moreover, once the atypical curves are detected, they are removed and the procedure is repeated until the process becomes stable (under control), namely, defined by a total absence of atypical data.

2.2. Procedure for Building a Control Chart for Process Monitoring (Phase II)

Phase II deals with process monitoring; it involves quick detection of changes from the calibrated sample stabilized in Phase I [48]. For the scalar and multivariate cases, the process is monitored by taking the estimated control limits in Phase I [46] as a reference. In this phase, the average run length (ARL) is used to evaluate the performance of the control charts [35].

In this context, we test if there are deviations between the data obtained in Phase II, also called monitoring sample, $\{\mathcal{X}_{n+1}(t), \mathcal{X}_{n+2}(t), \ldots, \mathcal{X}_m(t)\}$ and the reference data $\{\mathcal{X}_1(t), \mathcal{X}_2(t), \ldots, \mathcal{X}_n(t)\}$ or calibration sample, taking into account its distribution.

In Phase II, in the univariate case, an *F* distribution for the under-control process is estimated from a calibration sample or reference data. It is assumed that *F* is the distribution of the CTQ variable of an under-control process (Phase I). This distribution is used to establish control limits that will be used to monitor the process in Phase II. The limits comprise an interval that will cover new observations of the process with a high probability, assuming that the process is under control. In Phase II, a sample of the *G* distribution is monitored. Therefore, in this stage, the methods for constructing control charts are based on contrasting the hypothesis:

$$
\begin{array}{ll}
H_0: & F = G \\
H_1: & F \neq G.
\end{array}
\tag{2}
$$

In the FDA context, we do not have a density function for a functional random variable $\mathcal{X}$ that allows us to perform different tests corresponding to Phases I and II. Alternatively, we can estimate the distribution of the depth corresponding to the curves that belong to a sample of functional data. Thus, for Phase II, the use of the rank control charts [27] is proposed in an FDA context. The calculation of depths for functional data is proposed, which facilitate the calculation of ranks. These form the basis for developing *r* control charts, called rank charts.

The adaptation of *r* control charts involves the calculation of the rank statistic from functional depth measurements. The *r* chart plots the rank statistic as a function of time. The central control line CL = 0.5 serves as a reference point for observing possible patterns or trends. The lower limit is LCL = α, where α is the false alarm rate.

For the practical application of this proposal, the `qcr` [54] and `fda.usc` [49] R packages are used. The `fda.usc` package provides tools for the calculation of functional data depth, while the rank control chart, among other nonparametric charts proposed in Reference [27], is applied using the `qcr` package, which was developed by the authors.

As mentioned above, in Phase II, the curves corresponding to the calibration sample of Phase I, $\{\mathcal{X}_1(t), \mathcal{X}_2(t), \ldots, \mathcal{X}_n(t)\}$, are used for detecting changes or deviations with respect to the behavior of the process described in Phase I. The curves of the monitoring sample, $\{\mathcal{X}_{n+1}(t), \mathcal{X}_{n+2}(t), \ldots, \mathcal{X}_{n+m}(t)\}$, are collected; additionally, we test the hypothesis of each new curve belonging to the same distribution that corresponds to the calibration sample.

The procedure for estimating control charts for Phase II follows the same scheme presented in Reference [27]. Particularly, we assume that the rank statistic follows a uniform asymptotic distribution. This result is applicable to the functional case because of the way the rank corresponding to each observation is calculated (percentage of less deep curves than the observed ones). This fact provides a computational advantage in the monitoring of continuous processes since it eliminates the need to estimate the LCL. However, it is set as the quantile of a uniform distribution at a significance level α.

The procedure to develop the rank control chart for the functional univariate case (process defined by one functional variable) is detailed below, which can be easily generalized to the functional multivariate case (process defined by more than one functional variable).

1. From the reference sample $\{\mathcal{X}_1(t), \mathcal{X}_2(t), \ldots, \mathcal{X}_n(t)\}$, get the depths of the dataset, $D(\mathcal{X}_i)_{i=1}^n$ and the depths of the curves that make up the monitoring sample, $D(\mathcal{X}_j)_{j=n+1}^{n+m}$. The depth of each curve corresponding to the monitoring sample is calculated from the *n* curves of the calibration sample, that is, with respect to $n + 1$ cases.

2. The rank statistic is calculated for each curve of the monitoring sample, $r_G(\mathcal{X}_{n+1}), \ldots, r_G(\mathcal{X}_{m+n})$, considering the calibration sample $\{\mathcal{X}_1(t), \mathcal{X}_2(t), \ldots, \mathcal{X}_n(t)\}$ as sample of reference.

$$
r_G(\mathcal{X}) = \frac{\#\{\mathcal{X}_i | D(\mathcal{X}_i) \leq D(\mathcal{X}), i = 1, \ldots, n\}}{n}.
$$

3. The values of the rank statistic, the lower control limit LCL = α and central line CL = 0.5 (the expected value of the rank statistic) are plotted, thereby generating the control chart.
4. Proceed to monitor the process. If at least the rank of a curve, $\mathcal{X}$, is such that $r_G(\mathcal{X}) \leq$ LCL, then the process is considered out-of-control.
5. A functional control chart is developed. This is a graphical tool that allows us to identify the possible assignable causes of the out-of-control states. The original curves are included, those correspond to the reference and monitoring samples, in addition to the functional envelope obtained from $(1 - \alpha)\%$ of the deepest curves of the calibration sample.

3. Data Collection: Case Study of HVAC Installations in Commercial Areas

Here, the case study of HVAC installations' control are considered for a clothing store of a commercial area in the Panama City [24]. The data stream has been obtained by using the Σqus energy web platform. Sixteen CTQ variables are measured taking into account their ability to provide information about the energy efficiency, air quality and the thermal comfort of the store environment—indoor temperatures, overall energy consumption, HVAC energy consumption, CO_2 content in the air (ppm), relative humidity (%), temperatures of impulsion and return temperatures of the chillers in different areas of the store (see Figure 1).
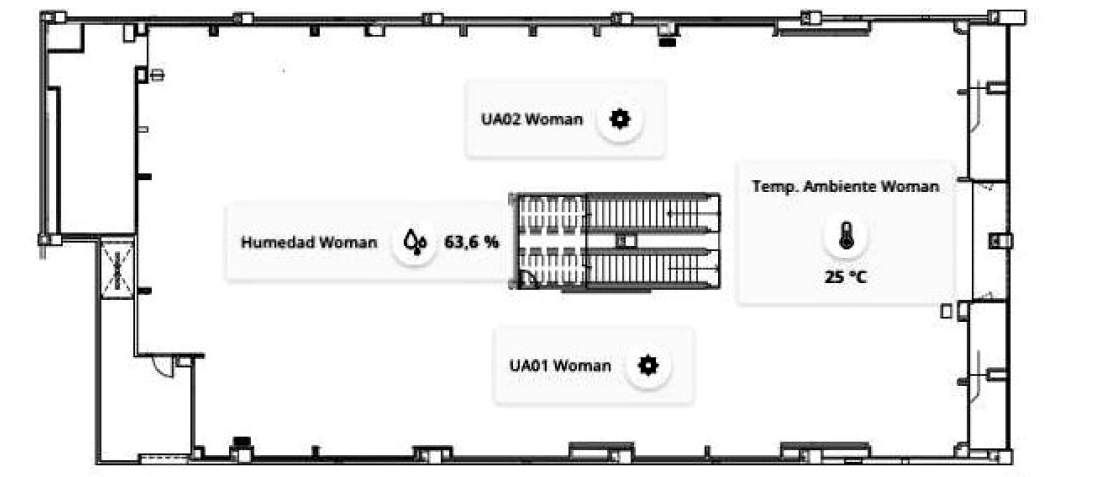

Figure 1. Plan of the case study store located in the Panama City.

Hourly measurements are obtained from 1 August 2017 to 31 October 2018. The operations of the HVAC facilities of the store start at 9:00 a.m. or 10:00 a.m. At start-up, the energy consumption peaks due to the characteristics of the HVAC installation. From 12:00 p.m., the consumption remains relatively constant the store closure at 20:00 p.m., 21:00 p.m. or 22:00 p.m. The shutdown takes about 1 or 2 h, with consumption falling at a constant rate of change. The resulting data can be considered functional data and thus FDA techniques can be applied. It is also important to note that this case study is a controlled study in which the anomalies and their assignable causes have been previously detected for the maintenance staff.

The data were obtained in the framework of a controlled environment where anomalies were identified by the maintenance staff. They are briefly described as follows:

-   On 11 September, there was a decline in the air conditioning consumption at about midday.
-   On 21, 22 and 30 September, the shopping center was closed and thus there was no energy consumption and the temperatures remained high.
-   On 27 September, several maintenance tests were applied to the store's HVAC installations.
-   On 29 September, the store's HVAC installations were stopped one hour earlier.

- Additionally, on 19 September, the air conditioning was stopped half an hour before the usual time. Particularly, there was a regulation change in the HVAC system.
- In mid-October, there was a leak in the air conditioning circuit. From that moment, energy consumption began to rise.
- On 1 November, repairing activities were performed. Consequently, the consumption decreased and the start-up consumption peak was removed.
- Between 17 and 20 November, the energy consumption in HVAC was again increased.

It is important to note that only working days have been studied in this work to evaluate the performance of the proposed control charts.

4. Application to Real Data

This section shows the usefulness and performance of the new graphical methodology for quality control using functional data, which is evaluated in the case study on the detection of energy efficiency anomalies of an HVAC installation. Specifically, the case study considers a commercial area of a well-known Galician apparel brand located in the Panama City. In this controlled case study, the anomalies and their assignable causes were detected by the maintenance personnel.

The following section shows the need to develop and apply FDA methodologies for control charts, considering the observations of the data of the present case study, particularly those corresponding to August. As mentioned above, in August, no event destabilized the process due to assignable causes. However, by using a methodology for scalar data (ignoring the autocorrelation between the variables), an unacceptable number of false alarms could be detected.

In the scalar case, boxplot [55] is commonly used to detect anomalous or atypical data. Figure 2 shows a traditional scalar approach for detecting outliers using boxplot. The left panel shows the boxplot for each variable of energy consumption in HVAC systems per hour, while the right panel shows the curves of daily energy consumption in HVAC systems, highlighting curves detected as outliers by the descriptive procedure based on the application of boxplots to each hourly consumption. In the usual procedure, atypical curves are those in which at least one point has been detected as an outlier in some boxplot; however, the drawback of this approach is that it increases the probability of type I error. It detects 12 daily energy consumption curves as outliers.
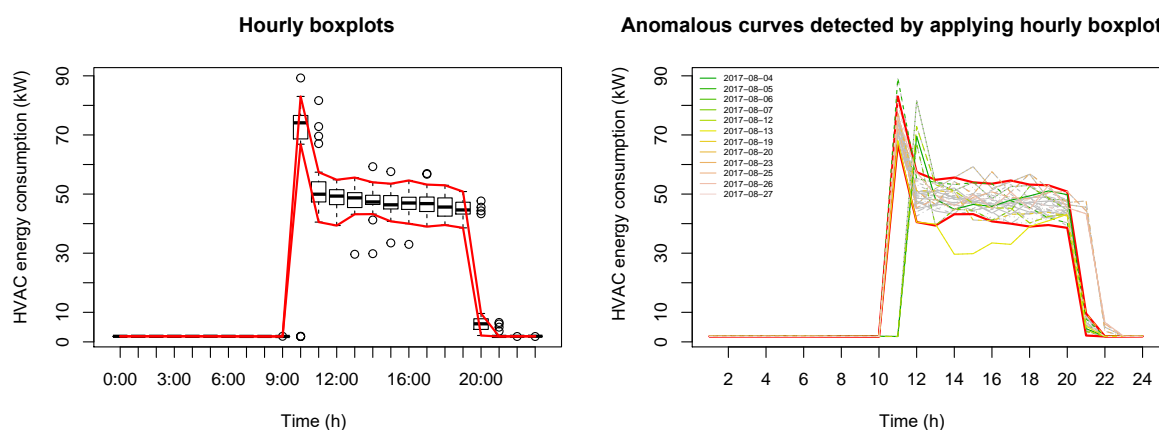

Figure 2. Detecting outliers in the heating, ventilation and air conditioning (HVAC) energy consumption by developing a boxplot for energy consumption per h.

Based on the information described in the previous section, first, we apply the data depth control chart for Phase I and, subsequently, the rank control chart to monitor the process during Phase II. The application of these two statistical techniques, together with the contribution of an intuitive graphic tool (to facilitate the detection of assignable causes for the anomalies), constitutes the new

proposed procedure of control charts for functional data. Generally, the procedure can be summarized as follows:

- The energy consumption curves in HVAC during August and September accounts for the calibration sample. A broad temporal range has been taken to estimate the natural control limits of energy consumption in HVAC accurately.
- For Phase I, the reference or calibration sample is stabilized by using the control chart developed from the depth measures of the curves. Particularly, curves detected as atypical by FDA outlier data detection methods are removed from the calculations to estimate the natural control limits.
- Data corresponding to October and November are monitored to confirm that there is no deviation in the HVAC processes. If there is a change, that is, there are days with out-of-control HVAC energy consumption, then it was recommended that the possible assignable causes should be sought, which, when detected, should be removed in order to remove the corresponding process variations.

In Figure 3, the black curves correspond to August (23 curves), whereas the gray curves account for the HVAC energy consumption in September (21 curves). The days from Monday to Friday are used in this study, taking into account that the work schedule is different on Saturday and Sunday. In September, the actual anomalies were detected (see Figure 3, curves in red) using the Phase I control chart based on the functional data depth outlier detection methods. Particularly, the anomalies corresponding to 11, 21, 22, 27 and 29 September were identified (refer to Section 3 for more information on the assignable causes of these anomalies).

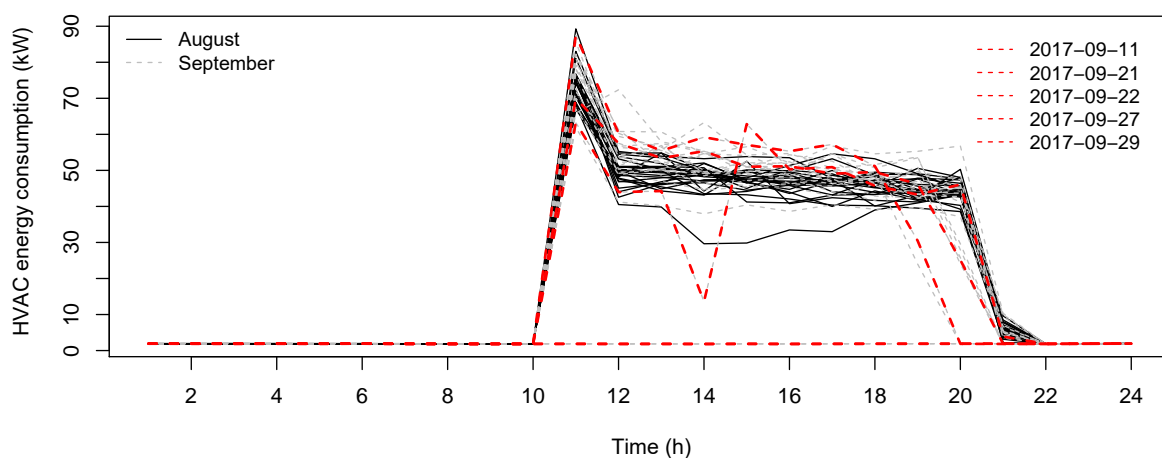**Curves belonging to the calibration sample**

Figure 3. Daily curves of energy consumption in HVAC facilities of the Panama City's store. The natural control limits are estimated from the curves belonging to the calibration sample, which have been shown which (Phase I).

From this reference sample and based on the simulation study corresponding to the control chart based on the functional depths (see Section 5), the mode functional depth and the weighed method for outlier detection are used. Additionally, a significance level of $\alpha = 0.025$ is used for the estimation of the LCL from $B = 500$ bootstrap resamples, a smoothing coefficient $\gamma = 0.8$, and a percentage of Trimming $trim = 0.05$, which allows obtaining an envelope of 95% of the deepest curves. The advantage of this procedure of detecting atypical curves in the Phase I control chart based on functional data is its flexibility to adapt to a wide variety of real problems, by regulating its parameters (γ and $trim$).

The left panel of the Figure 4 shows the original gray curves, the estimate of the median (blue curve), the functional trimmed mean (red curve) and the envelope corresponding to the 95% of the

deepest curves, which is plotted in red. The curves detected as anomalous are shown in gray and dotted in black. After the mentioned atypical energy consumption curves are identified, they are removed from any calculation related to the estimation of the natural control limits, given that their assignable causes have been identified (causes apart from the randomness of the data). Subsequently, the process is repeated according to an iterative scheme until atypical curves associated with assignable causes are not identified.

Figure 4 shows the functional approximation of the rank control chart, where each datum or point accounts for a daily curve of energy consumption in HVAC. The first 23 points represent the HVAC energy consumption curves for August, whereas the next 21 points represent the energy consumption curves of September.
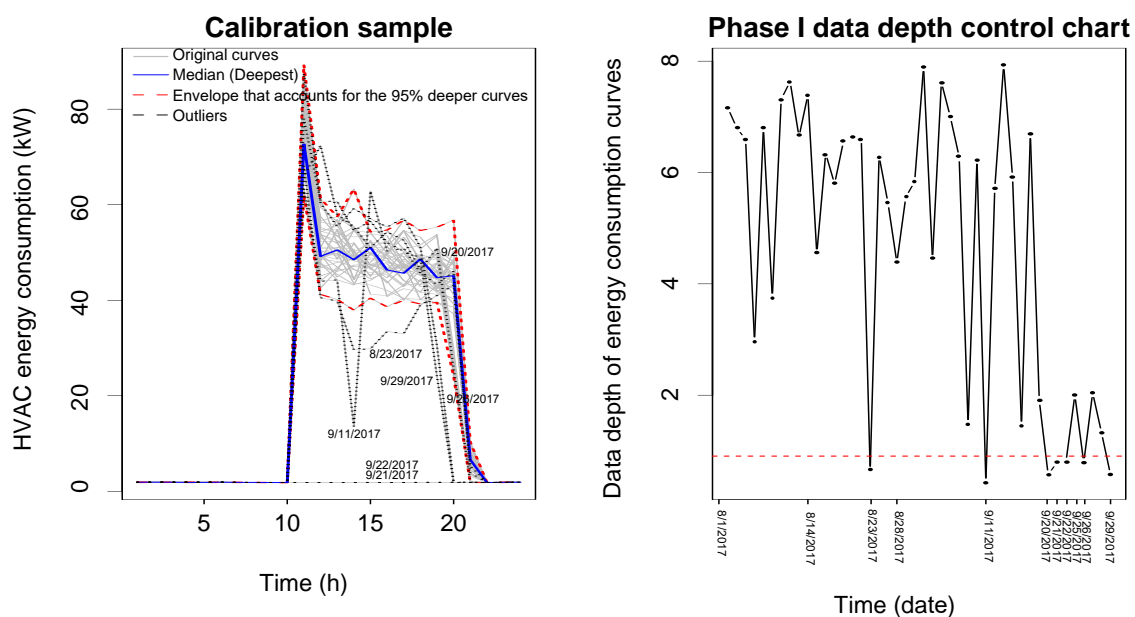

Figure 4. Left panel: Curves of daily energy consumption in HVAC facilities of the Panama City's store. The figure identifies curves detected as anomalies as well as functional position measurements. **Right** panel: Control chart corresponds to Phase I; they facilitate the detection of anomalous curves, thus eliminating the assignable causes of variation. The depths of each of the daily consumption curves and the natural control limit are shown.

In this first iteration, the previously identified curves are detected as anomalous. Their structure corresponds to an assignable cause of variation (see Section 3). They correspond to 11, 21, 22 and 29 September. However, in this iteration, the anomalous curve corresponding to 27 September has been not detected. Additionally, the curves corresponding to 23 August, 20 and 26 September are detected. Moreover, the data depth corresponding to the curves from 19 September are very small—they are outside the lower control limit and thus they are identified as out-of-control states. The assignable cause corresponding to this behavior is that the air conditioning was being shut off half an hour earlier than usual.

In the next iteration, the following daily consumption curves are detected as out-of-control or anomalous: 4 August, 15, 19, 25 and 27 September. In this iteration, the HVAC energy consumption curve of 27 September and those corresponding to the last days of September are identified; the timing of HVAC shutdown was changed for the latter part of September. Finally, energy consumption curves are not detected as anomalous after the second iteration (see Figure 5).
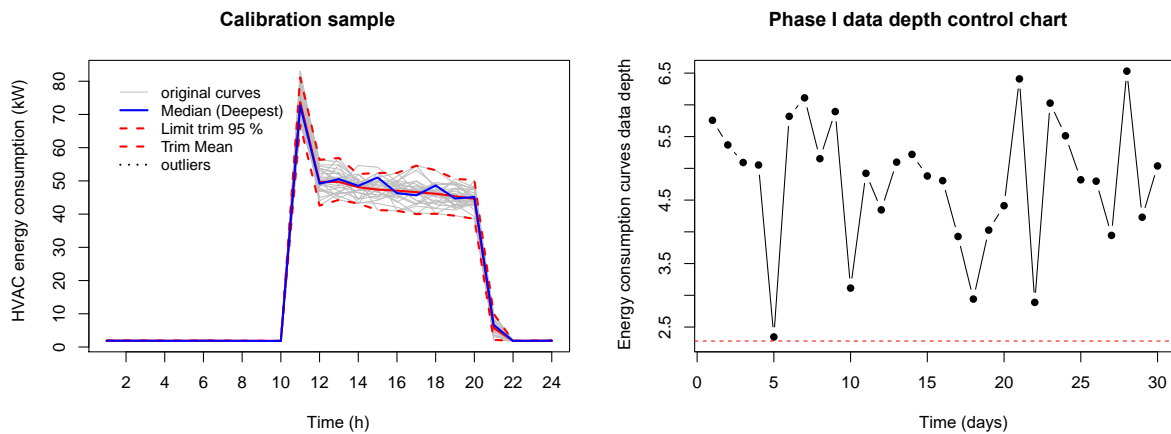
Figure 5. Under-control process (HVAC energy consumption is stabilized). **Left** panel: daily HVAC energy consumption curves in the Panama City store, showing the envelope corresponding to 95% of deeper curves. **Right** panel: control chart for Phase I based on functional data depth.

The process began with 44 curves, of which 9 (in the first iteration) and 5 (in the second iteration) curves were identified as anomalous. Of these 14 energy consumption curves, 2 and 12 curves correspond to August and September, respectively. Hence, the calibration sample comprises days in August and September, until 18 September. The current performance of energy efficiency of HVAC installation has been characterized through this reference or calibration sample. The next step is to detect changes (from the reference HVAC performance) in the HVAC system. This task belongs to Phase II control charts.

After Phase I, the Phase II control charts based on functional data, also called monitoring phase, is performed. Therefore, the sample to be monitored comprises days corresponding to October in which the HVAC facilities of the apparel store were operated at regular times. This is in consideration of the fact that a leak in the air conditioning circuit was recorded in mid-October, especially after the HVAC consumption began to rise. This behavior can be observed in Figure 6; it plots the monitored sample as well as the curves of the calibration sample and their 95% envelope, which is estimated in Phase I.

Figure 6 shows that the HVAC energy consumption, the CTQ variable for the energy efficiency of HVAC installations, is out of control. The assignable cause is the leak in the air conditioning system. In order to rectify the anomalous operation of the HVAC installations, on 1 November, a provisional repair was performed. These actions produced a decrease in the HVAC energy consumption and attenuated peak energy consumption corresponding to the start-up (see the energy consumption curves before 11:00). Consequently, the consumption of the HVAC facilities in November is slightly different from that of August and September, which account for the calibration sample. However, this energy consumption resumes its rise between 17 and 20 October.

These changes can be observed in Figure 7 wherein it is observed that the HVAC energy consumption corresponding to the monitored sample is characterized by a greater variability than that corresponding to the calibration or reference sample. The results of the application of the rank control chart show that the process is out-of-control both for the monitoring samples of October and November. In October, the assignable cause is the fault in HVAC facilities, whereas the assignable cause in November is the change in the process due to repair activities that deals on a decreasing in HVAC energy consumption.
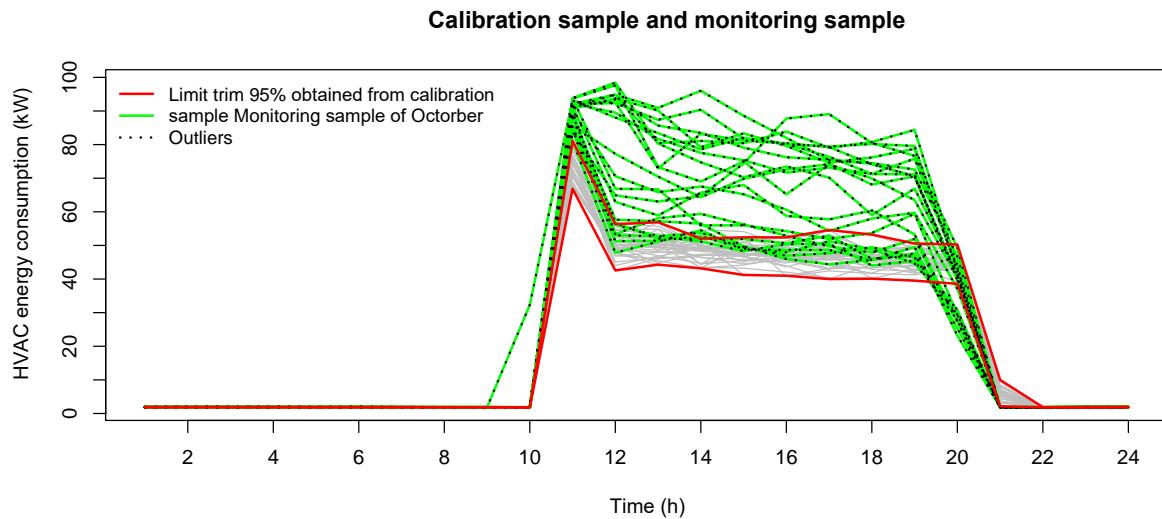
Calibration sample and monitoring sample

Figure 6. Reference of calibration and monitored samples corresponding to October.

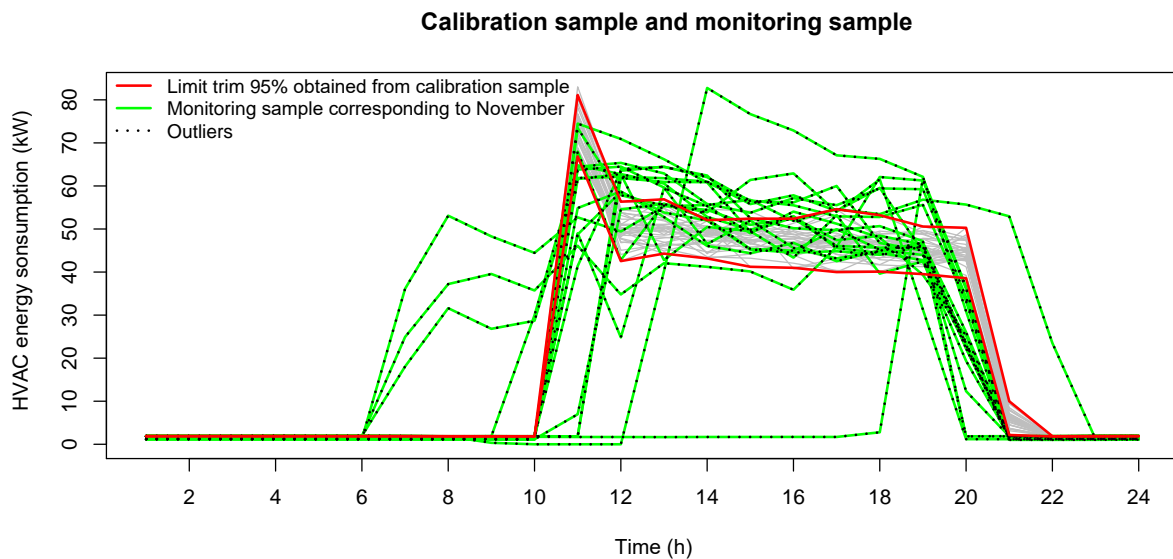**Calibration sample and monitoring sample**

Figure 7. Reference (August and September) and monitoring samples corresponding to November.

Given that the process has changed, a new reference sample should be obtained and studied. The stabilization and monitoring process should begin for the next year, starting December. Figure 8 shows the rank control chart for the calibration (August and September) and monitoring samples of November. In practice, monitoring can be efficiently performed using the proposed control charts because you do not need to estimate the LCL through resampling procedures. Additionally, the rank control chart can be adapted for monitoring multivariate functional data, that is, when there are different types of curves that define the quality of a specific system or process simultaneously. For example, a direct case of application considers their first derivative along with the original curves.

The application of rank control charts also allows us to monitor the process by studying more than one functional variable. Particularly, it not only allows us to study the energy consumption curves but also the curves of daily temperature, daily relative humidity, daily CO_2 concentration, among others; they completely characterize the energy efficiency of the facilities and their thermal comfort and ambient air quality. Precisely, this procedure would allow us to perform a functional multivariate monitoring.
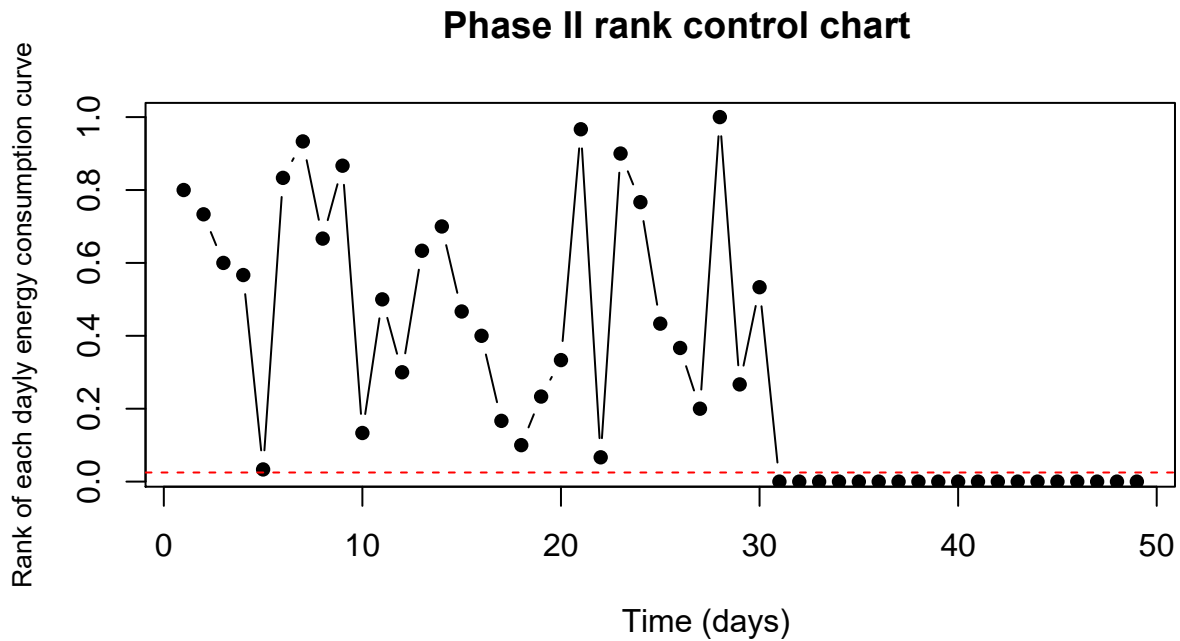
Phase II rank control chart

Figure 8. Rank control chart for Phase II of quality control. Both reference (corresponding to August and September) and monitoring (corresponding to November) samples are shown.

5. Simulation Study

The control chart performance is many times measured by observing its power $(1 - \beta)$. It is defined as the probability of identifying an out-of-control state when the process state is actually out-of-control [35]. Moreover, assuming that the process is actually under control, the type I error (α) or false alarm rate will be defined as the probability of detecting an out-of-control signal [35].

In the case of the process is under control, the probability of identifying an out-of-control observation should be small enough to prevent an unacceptable number of false alarms. Otherwise, if the process is effectively out of control, the power should be high enough to detect the process change as quickly as possible [35].

Another common index for measuring the performance of a control chart is Average Run Length (ARL), which is defined as the average number of observations plotted before a signal is out-of-control. The ARL is equal to $\dfrac{1}{p}$ (if we can assume that the signals are independent, that the run length distribution is geometric), where p is the probability of having an out-of-control signal [35].

To evaluate the performance of a control chart in Phase II, the ARL_0 and ARL_1 are often used, which are the average number of observations until the first out of control is detected, in cases where the process is really under control ($ARL_0 = \frac{1}{\alpha}$) or actually is not ($ARL_1 = \frac{1}{1-\beta}$) [35]. The ARL_1 should be at its lowest to increase the probability of quickly identifying events $(1 - \beta)$, power of a test that lead to the process being out-of-control [35].

Given that the F distribution is not known, a Monte Carlo simulation is designed to calculate the control charts power. The simulated scenarios allow us to estimate and compare the power of the control charts for different data depth measurements and for the case of independent and dependent functional data.

In this section, the performance of the control graphics proposed for Phases I and II will be evaluated. First, the simulation scheme designed in Febrero et al. [44] will be used to evaluate the performance of the control chart proposal for Phase I. Realizations of a Gaussian stochastic process have been proposed following the expression below [44,48]:

$$\mathcal{X}(t) = \mu(t) + \sigma(t) \cdot \epsilon(t), \tag{3}$$

whereby $\sigma^2(t) = 0.5$ and

$$\mu(t) = \mathbf{E}\left(\mathcal{X}(t)\right) = 30t(1-t)^{3/2}, \tag{4}$$

whereas $\epsilon(t)$ is a Gaussian process $\epsilon(t) \sim GP(\mathbf{0}, \Sigma)$ with $\mathbf{0}$ mean and variance-covariance matrix equal to

$$\mathbf{E}\left[\epsilon(t_i) \times \epsilon(t_j)\right] = e^{-\frac{|t_i - t_j|}{0.3}}.$$

Additionally, Reference [44] developed an alternative model to generate atypical curves, $\mu(t) = 30t^{3/2}(1-t)$. In Figure 9a, the two functional means are presented. The black curve accounts for the process mean without atypical curves, while the red curve is the mean of the process that generates the atypical curves.

The control charts proposed in this work have been designed to monitor the functional mean to detect two events—change in the mean of the process in terms of the magnitude and shape—which reveal that the process is not under control. For designing control charts for Phase II, it is assumed that the process is under control, that is, outliers are not detected. To generate simulation scenarios for each of these events, the following functional means have been considered:

- Mean of the model with a change in the magnitude:

$$\mu(t) = 30t(1-t)^{3/2} + \delta, \tag{5}$$

   by which δ denotes the change that goes from 0.4 to 2 in steps of 0.4.
- Mean of the model with a change in form:

$$\mu(t) = (1-\eta) \cdot 30t(1-t)^{3/2} + \eta \cdot 30t^{3/2}(1-t), \tag{6}$$

   where η is the change from 0.2 to 1 in steps of 0.2.

In Figure 9b, the green curve accounts for the functional mean of a process when there is a change in the magnitude ($\delta = 0.7$), while the curve of blue denotes the mean of a process when there is a change in the shape ($\eta = 0.3$).

In Febrero et al. [44], the functional data $\mathcal{X}_1, \ldots, \mathcal{X}_n$ denote realizations of a stochastic process $X(\cdot)$, assuming continuous trajectories in the $[a, b] = [0, 1]$ period and independence between the curves. However, simulation scenarios in which the simulated curves are defined by a variable degree of dependence have also been considered. This is because several practical applications of this type of chart are related to continuously monitored data with respect to time, forming functional time series, such as the curves of daily energy consumption in commercial areas. In this way, dependent curves are generated from the model $\tilde{Y}_i(t) = \mu(t) + \sigma(t) \cdot \tilde{\epsilon}(t)$, with $\tilde{\epsilon}(t) = \rho \cdot \tilde{\epsilon}_{i-1}(t) + (1-\rho) \cdot \epsilon_i(t)$, where ρ is the correlation measure between curves and $\sigma(t) = 0.5$ and both $\epsilon(t)$ and $\tilde{\epsilon}(t)$ are Gaussian processes [48].

In order to compare the results of the simulations in the scenarios defined by independence and dependence between the curves, the variance of ϵ is rescaled (we define the variance of the error $\tilde{\epsilon}$ to be one). Specifically, considering $\sigma_{\tilde{\epsilon}}^2 = \frac{(1-\rho^2)}{(1-\rho)^2} = \frac{(1+\rho)}{(1-\rho)}$, you have $\sigma_{\tilde{\epsilon}}^2 = 1$.

In Figure 10, different scenarios are presented considering the changes in the functional mean of the process, in the shape and magnitude, in the cases of independence and dependence between curves. The gray curves show the realizations of the process when it is under control (whose mean is the Equation (4)). However, the red curve in each graph accounts for the scenarios in which the presence of events that destabilize the process is considered, that is, the process is not under control. In Figure 10a,b, the cases of independence between curves and the presence of events defined by changes in the functional mean in terms of the magnitude and in shape, respectively, are shown.

However, Figure 10c,d show two cases defined by the presence of dependence between curves, including changes in the magnitude of the mean (panel c), with respect to its shape (panel d).
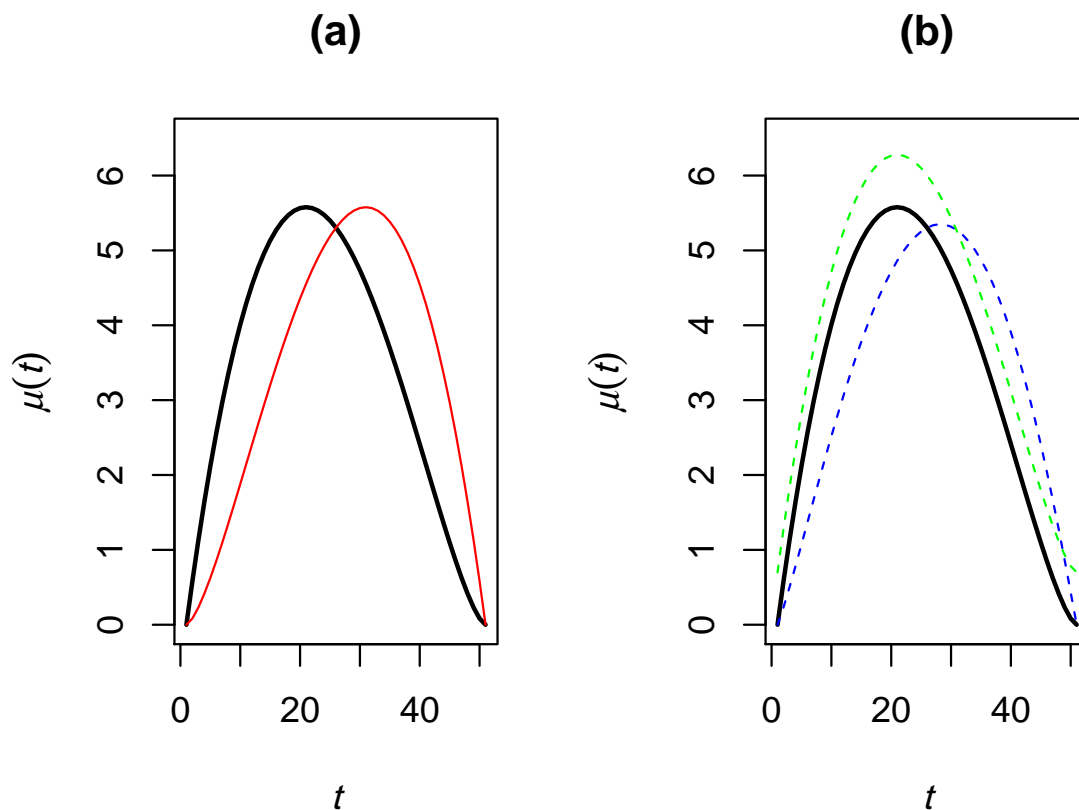
(a)　　　　　　　　　　　　　　　　**(b)**

Figure 9. (**a**) Functional means and (**b**) changes in the shape and the magnitude in the mean of the process.

　　　In the building energy efficiency domain, the out of control signals of energy consumption, temperature, CO_2 proportion and humidity, among others, can be defined by a change of shape and/or magnitude with respect to the under control signals analogous to those analyzed in the simulation study. The change in magnitude is related with a change in the scale of the studied process, for example, the increasing of energy consumption, temperature, humidity or concentration of CO_2 in all the hours corresponding of a specific day with respect to the otherwise normal pattern. This type of change is accounted by the addition of the δ term to the Equation (4) in order to obtain the Equation (5). Thus, the amount of that change is controlled by δ parameter. The green curve of Figure 9 accounts for an example of magnitude change. If we compare with the black one, we could realize that the two curves have the same shape and the only difference is the scale. On the other hand, changes in the shape of the curves are introduced by modifying the Equation (4) by the η parameter resulting in the Equation (6). In the energy efficiency domain, this type of change can be related to a change in the HVAC facilities programing (changes in temperature regulation, changes in the time schedule), in a failure of the HVAC in just one interval of the day, an extremely high or low level of occupation in a building and extreme changes in the weather, among other causes. The proposed simulation study has performed taking into account the specific domain where the FDA control chart approach is applied, namely energy efficiency in buildings. Thus, the shape of these types of profiles is similar to that corresponding to CO_2, temperature, energy consumption and humidity curves in buildings. In order to measure the performance of the proposed control charts for very different types of profiles, new studies may be necessary.

In the following section, a simulation study is performed to determine conditions under which it can be verified that the smooth bootstrap procedure works when there is independence and dependence between curves.
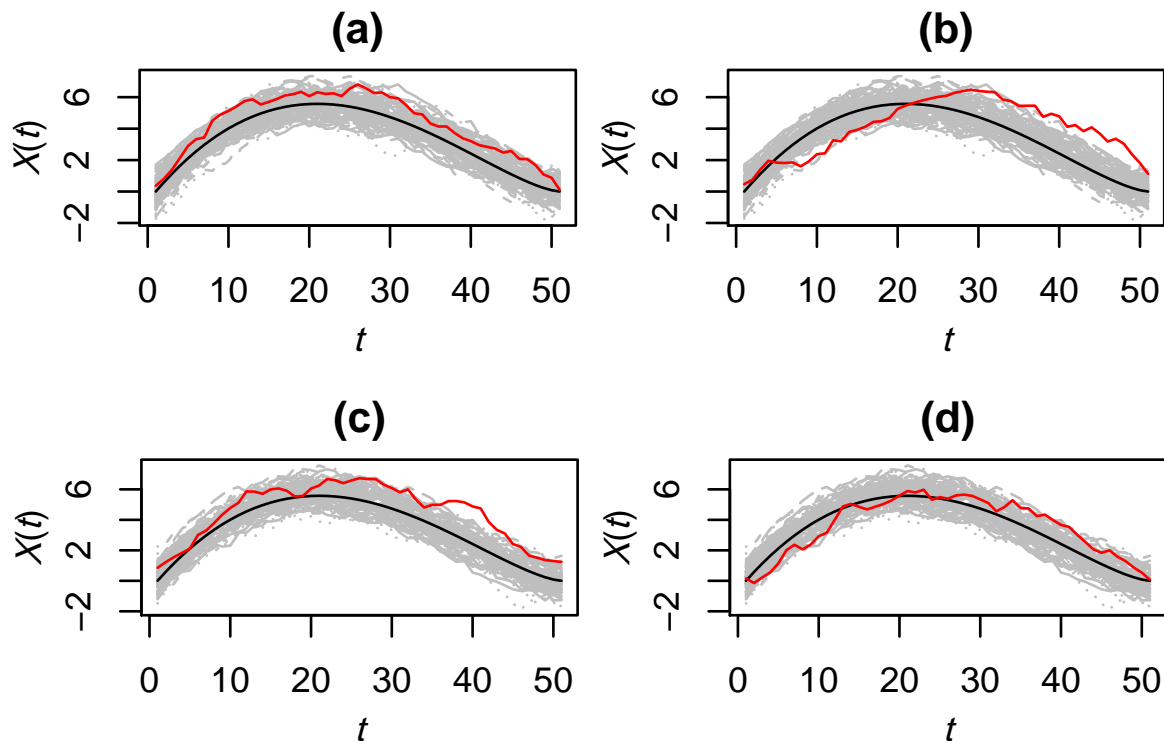

Figure 10. Scenarios in which independence between curves is studied. Changes in the functional mean with respect to its magnitude (**a**) and shape (**b**) are shown. In the case of dependence, panels (**c**) and (**d**) show the simulation scenarios in which changes in the magnitude and shape, respectively, are observed in the functional mean.

5.1. Measurement and Comparison of the Performance of the Control Chart Proposed for Phase I

The performance of the control chart is estimated and compared from the generation of calibrated samples of size 50 and 100 (curves). For each sample, different functional depth measurements described in Section 2 are calculated and the outlier detection robust procedures (weighted and trimmed) are applied for the estimation of type I error when the process is under control and the power of the test when the process is out-of-control. For the estimation of type I error, each scenario is replicated 1000 times ($n = 50, 100$, assuming independence and dependence between curves). When the power of the test is estimated, in each scenario (assuming independence and dependence), a curve within the alternative hypothesis is generated; this procedure is also repeated 1000 times.

Following the scheme described in Reference [44], curves observed at equidistant points are considered; the number of points that define each curve is 51 in the interval $[0, 1]$. From 1000 resamples ($B = 1000$) and with a 2.5% trimming procedure (removing less deep curves), a smoothing bootstrap procedure defined by a smoothing factor $\gamma = 0.05$ is applied to estimate the $C = 0.01$ quantile representing the LCL.

First, a simulation study is performed to estimate and compare the type I error ($\alpha = 0.01$ is fixed) of the proposed control chart, assuming scenarios with independence and dependence between curves. Subsequently, a similar study is carried out to estimate and evaluate the power of the control chart to detect out-of-control signals in different situations (independence, dependence, different sample size and a change in the shape or magnitude).

In Table 1, the results of the estimation of the false alarm rate (type I error) in the independence scenario are shown. The average of the percentage of false out-of-control signals (type I error) detected by the procedure shown above are very close to the nominal 1% for the two considered sample sizes. Furthermore, it can be observed that, when n increases, the type I error percentages are closer to the nominal level. In general, for a sample of size $n = 100$, the results of applying the weighted method are closer to α, especially when using the mode depth measurement. The results obtained in the simulations are similar to those presented in Reference [44].

Table 1. Estimation of the false alarm rate (%) for the case of independence between curves using a nominal type I error of 1%.

n	Method	FM	RP	Mode
50	Weighted	1.94	1.89	1.49
	Trimmed	1.34	1.95	1.36
100	Weighted	1.55	1.75	1.25
	Trimmed	1.67	2.33	1.76

In any process, type I error increases the production cost. Hence, it is essential not to overestimate this error rate when managing the quality.

Table 2 shows the results of the simulation to evaluate the ability of the control chart to detect a change in the shape or magnitude of the functional mean of the process through the estimation of its power $(1 - \beta)$. The percentage of out-of-control signals (outliers) correctly detect when the population defined by the Equation (4) is contaminated with curves belonging to the M_1 model (Equation (5)) and M_2 model (Equation (6)); it is denoted by p_c; however, the percentage of false alarms (false states out of control) is p_f. These parameters have been estimated, in all the scenarios assuming the independence of curves, using the average of the corresponding empirical values, $\hat{p}_c$ and $\hat{p}_f$.

Table 2 shows that a better performance is achieved when the curves of model M_1 (where changes in the magnitude of the proposed control chart are simulated) are studied. Precisely, $\hat{p}_f$ and $\hat{p}_c$ are closer to the nominal α and $(1 - \beta)$. When identifying changes in the shape of the process average, M_2, the mode depth provides the highest percentages of correctly detected out-of-control signals. However, in the case of the M_1 model, the use of RP depth provides percentages of the detection of the true out-of-control states lower than those corresponding to the use of FM and mode depths. With respect to a robust method for outlier detection, the performance is similar in all the scenarios. However, an exception is the case wherein the RP depth is used; it reveals the low performance of the control chart in detecting observations corresponding to actual out-of-control states.

Briefly, the detection rate of false out-of-control signals for the independence scenario is close to 1%. However, when using the trimmed method, the detection rate of false out-of-control signals is overestimated but this percentage decreases when the sample size increases.

The results of the false alarm rate (type I error) for scenarios defined by dependence between curves are shown in Table 3. It is important to note that, for different values of ρ, very similar results with respect to the independence scenarios have been obtained. Precisely, the average of the percentages of false out-of-control signals are close to the nominal 1% in the two studied sample sizes. Additionally, when n increases, the type I error percentages are closer to the nominal level. However, some differences are observed when the RP data depth measure is used to develop the control chart. In this case, there is an overestimation of the percentage of false out-of-control signals.

Table 2. Percentages of $\hat{p}_c$ and $\hat{p}_f$ for the cases of curves simulated with M_1 (Equation (5)) and M_2 (Equation (6)) models, assuming independence between curves.

Stage	Displays	Depth	Method	δ 0.4 $\hat{p}_f$	$\hat{p}_c$	0.8 $\hat{p}_f$	$\hat{p}_c$	1.2 $\hat{p}_f$	$\hat{p}_c$	1.6 $\hat{p}_f$	$\hat{p}_c$	2 $\hat{p}_f$	$\hat{p}_c$
M_1	50	FM	Weighted	1.80	8.30	1.63	27.80	1.47	56.50	1.48	85.30	1.60	95.70
			Trimmed	1.25	7.70	1.19	26.80	1.19	56.00	1.22	83.00	1.20	95.50
		RP	Weighted	1.88	6.80	1.68	20.10	1.49	45.10	1.42	71.90	1.42	87.70
			Trimmed	1.88	7.00	1.89	23.20	2.08	49.30	2.14	75.70	2.20	91.90
		Mode	Weighted	1.41	5.70	1.27	19.90	1.09	45.10	1.08	76.70	1.17	94.30
			Trimmed	1.33	6.80	1.32	22.20	1.33	48.60	1.38	77.50	1.39	95.20
	100	FM	Weighted	1.51	5.70	1.40	23.30	1.34	53.50	1.34	82.00	1.38	96.20
			Trimmed	1.63	7.40	1.51	25.90	1.32	56.10	1.13	83.60	1.00	96.70
		RP	Weighted	1.68	5.00	1.57	18.60	1.50	45.40	1.45	71.40	1.52	90.10
			Trimmed	2.28	7.40	2.22	23.70	2.23	54.20	2.13	78.70	2.08	93.50
		Mode	Weighted	1.22	4.20	1.15	17.90	1.08	46.80	1.06	78.90	1.12	94.70
			Trimmed	1.70	5.80	1.62	22.90	1.46	54.20	1.28	83.50	1.13	95.90

Stage	Displays	Depth	Method	η 0.2 $\hat{p}_f$	$\hat{p}_c$	0.4 $\hat{p}_f$	$\hat{p}_c$	0.6 $\hat{p}_f$	$\hat{p}_c$	0.8 $\hat{p}_f$	$\hat{p}_c$	1 $\hat{p}_f$	$\hat{p}_c$
M_2	50	FM	Weighed	1.91	1.50	1.85	3.40	1.70	11.00	1.46	26.30	1.25	55.00
			Trimmed	1.28	1.90	1.23	4.10	1.19	11.60	1.13	26.70	1.18	50.60
		RP	Weighted	1.94	2.40	1.90	3.30	1.75	8.00	1.64	16.90	1.42	32.20
			Trimmed	1.89	2.30	1.83	4.60	1.73	10.10	1.84	22.50	1.92	41.40
		Mode	Weighted	1.43	2.80	1.35	10.70	1.19	31.00	1.06	65.30	1.12	91.70
			Trimmed	1.33	4.00	1.31	12.30	1.31	33.70	1.42	68.00	1.48	92.10
	100	FM	Weighted	1.53	2.00	1.51	2.70	1.45	12.20	1.35	27.90	1.25	53.00
			Trimmed	1.66	2.50	1.63	4.30	1.57	14.40	1.51	33.40	1.43	58.00
		RP	Weighted	1.69	2.70	1.66	4.10	1.63	8.20	1.55	17.30	1.48	34.70
			Trimmed	2.25	2.90	2.19	5.60	2.15	13.60	2.15	25.40	2.21	45.50
		Mode	Weighted	1.22	3.30	1.18	11.90	1.12	34.10	1.07	66.80	1.09	89.60
			Trimmed	1.72	4.60	1.65	15.70	1.57	41.70	1.41	72.70	1.22	92.90

Table 3. Results for the scenarios with dependence between curves. The estimates of false alarm rate (type I error) for values of ρ between 0.3 and 0.7 are shown.

ρ	n	Method	FM	RP	Mode
0.3	50	Weighted	1.85	1.88	1.41
		Trimmed	1.31	1.90	1.33
	100	Weighted	1.54	1.71	1.20
		Trimmed	1.65	2.23	1.71
0.5	50	Weighted	1.80	1.81	1.28
		Trimmed	1.26	1.82	1.29
	100	Weighted	1.46	1.68	1.15
		Trimmed	1.61	2.21	1.65
0.7	50	Weighted	1.66	1.73	0.93
		Trimmed	1.25	1.74	1.06
	100	Weighted	1.42	1.61	0.99
		Trimmed	1.56	2.15	1.57

Tables 4–6 show the results of the empirical estimation of p_c and p_f, assuming different values of ρ (from 0.3 to 0.7). The power (estimated by $\hat{p}_c$) of the control chart proposed for the model M_1 (Equation (5)) performs better when the weighted method is applied and if the sample size is increased. It is also observed that the performance of the control chart tends to be the same, independent of the type of data depth measurement used. Certainly, the performance of control charts in detecting real changes in the process, related to differences in the shape and mean, is better when the mode depth is used.

Table 4. Empirical values of $\hat{p}_f$ and $\hat{p}_c$, with $\rho = 0.3$ (assuming dependence between curves).

Model	Sample Size	Depth	Method	δ 0.4 $\hat{p}_f$	$\hat{p}_c$	0.8 $\hat{p}_f$	$\hat{p}_c$	1.2 $\hat{p}_f$	$\hat{p}_c$	1.6 $\hat{p}_f$	$\hat{p}_c$	2 $\hat{p}_f$	$\hat{p}_c$
M_1	50	FM	Weighted	1.58	8.90	1.22	33.60	0.98	69.80	1.07	91.70	1.19	99.30
			Trimming	1.12	9.20	0.75	26.60	0.37	47.40	0.09	54.30	0.02	49.20
		RP	Weighted	1.75	7.10	1.38	26.30	0.97	53.80	0.74	77.60	0.67	90.70
			Trimmed	1.72	7.60	1.42	25.90	0.89	45.90	0.49	60.90	0.20	68.70
		Mode	Weighted	1.21	5.80	0.90	24.90	0.67	59.80	0.70	88.30	0.86	98.50
			Trimmed	1.21	6.40	0.89	24.00	0.50	49.40	0.19	73.50	0.04	89.80
	100	FM	Weighted	1.47	7.30	1.23	29.60	1.07	67.80	1.13	92.30	1.24	99.20
			Trimmed	1.53	7.60	1.29	31.20	1.04	68.70	0.96	91.30	0.92	99.10
		RP	Weighted	1.69	5.50	14.67	24.40	1.21	58.00	1.06	83.90	1.02	95.90
			Trimmed	2.19	8.80	1.99	27.40	1.88	64.80	1.92	89.30	1.85	97.30
		Mode	Weighted	1.17	4.60	0.98	23.00	0.86	62.40	0.89	91.00	1.01	98.50
			Trimming	1.62	6.90	1.38	27.90	1.15	67.10	1.07	91.70	1.06	98.90
				η 0.2		0.4		0.6		0.8		1	
M_2	50	FM	Weighted	1.75	1.30	1.61	1.95	1.31	7.70	0.94	16.85	0.64	26.05
			Trimmed	1.25	1.40	1.15	2.35	0.97	7.70	0.73	15.95	0.51	20.50
		RP	Weighted	1.86	1.20	1.75	3.00	1.43	5.50	1.24	10.85	0.96	16.55
			Trimmed	1.79	1.65	1.68	3.40	1.58	7.15	1.46	13.10	1.16	19.40
		Mode	Weighted	1.31	3.70	1.08	16.00	0.78	44.90	0.62	79.30	0.81	96.90
			Trimmed	1.22	4.70	1.11	17.50	0.74	42.20	0.29	65.00	0.07	82.50
	100	FM	Weighted	1.53	0.90	1.46	2.05	1.30	7.45	1.07	19.50	0.92	32.35
			Trimmed	1.65	1.20	1.57	2.70	1.44	9.40	1.22	22.75	1.12	35.35
		RP	Weighted	1.76	0.60	1.73	2.20	1.57	5.10	1.41	12.10	1.27	21.80
			Trimmed	2.23	1.10	2.18	3.30	2.09	8.25	1.91	17.45	1.97	29.35
		Mode	Weighted	1.17	3.55	1.07	14.70	0.91	47.20	0.86	83.10	0.98	97.50
			Trimmed	1.64	4.70	1.48	20.50	1.25	52.60	1.13	85.10	1.12	98.20

Table 5. Empirical values of $\hat{p}_f$ and $\hat{p}_c$, with $\rho = 0.5$ (assuming dependence between curves).

Model	Sample Size	Depth	Method	δ 0.4 $\hat{p}_f$	$\hat{p}_c$	0.8 $\hat{p}_f$	$\hat{p}_c$	1.2 $\hat{p}_f$	$\hat{p}_c$	1.6 $\hat{p}_f$	$\hat{p}_c$	2 $\hat{p}_f$	$\hat{p}_c$
M_1	50	FM	Weighted	1.46	11.50	1.04	43.90	0.84	80.80	1.06	97.40	1.14	99.90
			Trimmed	1.02	10.80	0.63	33.40	0.23	50.70	0.28	53.80	0.00	45.60
		RP	Weighted	1.68	9.10	1.21	31.10	0.80	63.00	0.65	85.60	0.60	94.60
			Trimmed	1.73	9.60	1.26	30.70	0.66	52.60	0.28	66.40	0.11	68.80
		Mode	Weighted	1.10	7.50	0.76	34.40	0.55	73.00	0.69	95.20	0.82	99.90
			Trimmed	1.05	8.10	0.72	31.40	0.32	62.60	0.06	86.50	0.01	96.40
	100	FM	Weighted	1.20	10.00	1.08	39.60	1.01	79.40	1.10	97.30	1.18	99.80
			Trimmed	1.30	10.00	1.18	41.60	0.96	80.20	0.89	97.00	0.84	99.80
		RP	Weighted	1.40	10.00	1.39	32.30	1.09	69.20	0.97	90.60	0.97	97.90
			Trimmed	1.80	10.00	1.93	38.20	1.84	77.10	1.80	94.70	1.73	98.80
		Mode	Weighted	0.90	10.00	0.86	32.10	0.79	75.80	0.91	96.00	0.98	99.90
			Trimmed	1.30	10.00	1.30	38.30	1.10	79.40	1.06	97.20	1.02	99.90
				η 0.2		0.4		0.6		0.8		1	
M_2	50	FM	Weighted	1.63	1.40	1.43	3.55	1.02	10.80	0.68	21.60	0.50	33.10
			Trimmed	1.16	1.50	1.02	3.60	0.80	11.20	0.54	18.00	0.32	21.70
		RP	Weighted	1.89	1.75	1.65	3.30	1.32	7.30	0.98	14.30	0.75	22.55
			Trimmed	1.79	2.30	1.62	3.80	1.54	9.05	1.26	16.40	0.95	22.90
		Mode	Weighted	1.13	5.40	0.86	22.60	0.56	60.50	0.59	91.50	0.85	99.80
			Trimmed	1.10	6.30	0.83	23.40	0.46	54.00	0.11	77.20	0.02	93.80
	100	FM	Weighted	0.01	0.01	0.01	0.03	1.13	11.90	0.90	28.25	0.77	38.90
			Trimmed	0.02	0.01	0.01	0.04	1.28	15.25	1.11	30.90	1.04	39.90
		RP	Weighted	0.02	0.01	0.02	0.03	1.42	7.65	1.26	16.65	1.08	27.55
			Trimmed	0.03	0.02	0.02	0.05	2.02	12.60	1.91	23.25	1.96	35.00
		Mode	Weighted	0.01	0.05	0.01	0.23	0.82	63.00	0.87	92.50	0.99	99.80
			Trimmed	0.02	0.07	0.01	0.28	1.21	68.40	1.12	93.40	1.12	99.60

Table 6. Empirical values of $\hat{p}_f$ and $\hat{p}_c$, with $\rho = 0.7$ (assuming dependence between curves).

Model	Sample Size	Depth	Method	δ 0.4 $\hat{p}_f$	$\hat{p}_c$	0.8 $\hat{p}_f$	$\hat{p}_c$	1.2 $\hat{p}_f$	$\hat{p}_c$	1.6 $\hat{p}_f$	$\hat{p}_c$	2 $\hat{p}_f$	$\hat{p}_c$
M_1	50	FM	Weighted	1.32	16.00	0.87	57.70	0.85	91.40	1.01	99.30	1.01	100.00
			Trimmed	0.87	13.40	0.42	40.30	0.07	53.60	0.00	47.50	0.00	35.80
		RP	Weighted	1.41	13.00	0.97	43.40	0.61	76.70	0.56	92.00	0.53	98.00
			Trimmed	1.51	14.00	0.97	40.70	0.39	61.40	0.14	68.20	0.04	71.30
		Mode	Weighted	0.71	14.00	0.44	52.20	0.41	88.90	0.51	99.30	0.53	100.00
			Trimmed	0.82	14.30	0.43	47.40	0.09	80.20	0.01	96.00	0.00	99.90
	100	FM	Weighted	1.29	12.00	0.99	54.90	1.05	90.40	1.14	99.50	1.13	100.00
			Trimmed	1.36	13.80	1.08	56.00	0.89	90.20	0.84	99.20	0.76	100.00
		RP	Weighted	1.50	9.90	1.20	48.30	0.97	80.80	0.93	96.40	0.96	99.50
			Trimmed	2.00	13.90	1.80	53.10	1.80	87.30	1.72	98.30	1.60	99.90
		Mode	Weighted	0.90	9.80	0.68	51.00	0.69	90.00	0.81	99.40	0.83	99.90
			Trimmed	1.43	13.40	1.14	58.10	1.01	92.40	0.97	99.60	0.92	99.90

Model	Sample Size	Depth	Method	η 0.2 $\hat{p}_f$	$\hat{p}_c$	0.4 $\hat{p}_f$	$\hat{p}_c$	0.6 $\hat{p}_f$	$\hat{p}_c$	0.8 $\hat{p}_f$	$\hat{p}_c$	1 $\hat{p}_f$	$\hat{p}_c$
M_1	50	FM	Weighted	1.52	1.80	1.21	6.05	0.76	17.90	0.49	30.70	0.41	39.95
			Trimming	1.07	2.10	0.85	6.15	0.56	15.30	0.31	22.30	0.18	24.05
		RP	Weighted	1.62	2.10	1.36	5.25	0.98	11.40	0.68	19.95	0.56	28.30
			Trimmed	1.59	2.40	1.41	5.80	1.19	13.85	0.92	21.05	0.70	28.15
		Mode	Weighted	0.75	8.60	0.47	39.90	0.37	82.80	0.50	99.10	0.61	100.00
			Trimmed	0.87	10.30	0.51	38.30	0.14	73.10	0.01	93.80	0.00	99.30
	100	FM	Weighted	2.02	0.00	1.24	6.35	0.99	21.20	0.81	36.45	0.75	45.10
			Trimmed	3.03	0.00	1.36	7.80	1.16	24.30	1.05	37.95	1.03	44.55
		RP	Weighted	3.69	0.00	1.50	5.25	1.25	13.00	1.08	24.50	0.99	34.85
			Trimmed	6.28	0.00	1.98	7.20	1.84	18.50	1.89	31.80	1.89	41.60
		Mode	Weighted	1.63	0.00	0.72	37.90	0.69	82.70	0.80	98.60	0.90	100.00
			Trimmed	3.48	0.00	1.24	46.70	1.10	87.10	1.05	99.00	1.04	100.00

With respect to the false out-of-control rate $\hat{p}_f$, in the scenarios corresponding to the use of M_1 model, when the trimmed method is also used, a lower rate is obtained. In the case of the M_2 model, there are similar results on the scenarios defined by independence between curves, that is, the $\hat{p}_f$ is lower when the trimmed method for outlier detection is used.

In Reference [56], new methods for the detection of outliers were proposed for the case in which there is dependence between curves. From the simulation studies carried out in this study, at different degrees of dependence, we can say that the outlier detection method proposed in Reference [44] was relatively robust against the presence of dependence between curves. The simulation study performed in this section supports the results obtained in the work in Reference [56] and, in conclusion, justifies the use of this method within the new control charts proposed for Phase I, even in scenarios with dependence between curves.

Although the application of the weighed outlier detection method to Mode data depth has generally provided Phase I control charts with best performance, if the false alarm rate of Tables 1 and 3 are observed, the use of trimmed outlier detection method tends to provide values of $\hat{p}_f$ slightly closer to $\alpha = 1\%$ (with respect to the weighted method) when the process is actually under control, the curves are independent and the sample is relatively small ($n = 50$). In addition, if the process is out of control, the curves are independent and the outliers are generated by the Equation (5) (changes in magnitude), the trimmed method applied to FM data depth provide a $\hat{p}_f$ close to $\alpha = 1\%$ and the highest $\hat{p}_c$, as shown in Table 2. In all the remaining scenarios, the use of weighed method applied to Mode data depth tends to provide the closest to $\alpha\%$ $\hat{p}_f$ and the highest $\hat{p}_c$ (see Tables 3–6).

5.2. Measurement and Comparison of the Performance of the Control Chart Proposed for Phase II

For Phase II, the monitoring stage, the use of the rank control chart has been proposed. The application of the rank control chart allows simultaneous monitoring of changes in the mean and variability of a process. In the functional case, in order to calculate the rank statistic, the functional FM, RP and mode depths are used.

An $\mathrm{ARL}_0 = \frac{1}{\alpha = 0.025}$ (the monitoring sample is assumed under control) is assumed to evaluate the performance of the control chart. The power of the control chart is estimated and compared for an under-control process, based on the generation of a calibration sample of size $n = 50$ by a Monte Carlo procedure.

Following the simulation scheme of Phase I, curves observed at equidistant points are assumed; they are composed of 51 points at $[0, 1]$ interval. A smoothed bootstrap with a smoothing factor $\gamma = 0.05$, 1000 resamples ($B = 1000$), using a 2.5% trimmed procedure (removing the shallowest curves), is applied to estimate and compare the power of the control chart to detect out-of-control signals when a significance level of $\alpha = 0.025$ is assumed. Additionally, in the same way as in Phase I, the simulation of scenarios with independence and dependence between curves are assumed.

Table 7 shows the estimates of the power (%) of the control charts for the scenario of independence between curves, whereas Table 8 shows power of the control chart for the scenario with dependence. In both cases, the ability of the control chart to detect a change in the magnitude of the functional mean of the process is evaluated by the estimation of its power $(1 - \beta)$.

Table 7. Power of the control chart, $1 - \beta$, for the case M_1 (Equation (5)) model in the scenario of independence between curves.

δ	FM	RP	Mode
0.5	14.5	14.8	14.8
1	47.6	47.2	48.0
1.5	83.3	80.8	83.0
2	97.8	97.5	97.7

From the results of the Table 7, any depth measure can be used to detect a shift in the process mean, since the same performance, in terms of power, is obtained.

Table 8. Power of the control chart, $1 - \beta$, for the scenarios defined by the M_1 (5) model, assuming dependence between curves.

ρ	δ	FM	RP	Mode
	0.5	20.28	23.73	21.84
0.3	1	62.95	64.75	64.06
	1.5	92.98	92.33	93.64
	2	99.56	99.16	99.75
	0.5	25.59	29.41	28.41
0.5	1	73.84	75.12	76.24
	1.5	97.15	96.56	97.94
	2	99.91	99.77	99.98
	0.5	35.47	40.23	43.55
0.7	1	86.14	86.87	90.38
	1.5	99.48	99.11	99.78
	2	100.00	99.99	100.00

The results of the detection of a shift in the process mean are shown in Table 8. A similar performance is observed when using any depth measure for different values of ρ. Apparently, the control chart for Phase II is robust against the existence of dependence between curves.

As observed in the simulation study and in the analysis of the case study with real data, the present proposal of control charts for functional data, including Phase I and II control charts, can be useful to detect anomalies in diverse scenarios. In the case of its application to real data, the set of proposed techniques is being examined for implementation in the web platform Σqus and for its use by the company Nerxus for detecting false alarms in facilities in commercial areas. The present control chart methodology can be used for control tasks, monitoring, anomaly detection and continuous

improvement in diverse industrial processes, monitoring of environmental variables, chemical industry and, in general, any process involving continuous monitoring of functional data over time.

Regarding the use of our methodology in more complex case studies defined by different operation modes, the application of the multi-modelling framework methodology in combination with our proposal could be useful. Indeed, in the building energy efficiency domain, there are many different operation modes of installations, each one defined by a specific operation pattern. Namely, the HVAC installations can be operated in heating or ventilation modes (there are even different modes within ventilation or heating). The automatic classification of each profile in the corresponding profile pattern could be very useful in the building energy efficiency field and previously to the application of our control chart proposal for Phase I and Phase II. With respect to the work of Grasso et al. [43], it is also interesting to mention that the proposed profile monitoring control charting scheme is that based on functional PCA and described in Colosimo and Pacella [57].

6. Conclusions

A new alternative of control charts has been proposed when CTQ variables of the process are functional. The proposal includes alternatives to develop the the Phase I and II control charts for stabilizing and monitoring the processes, respectively. In order to develop Phase I control charts based on functional data, outlier detection methods are used based on a method of smooth bootstrap resampling and the depth calculation of functional data is proposed. However, in order to implement Phase II, the use of rank-type nonparametric control charts based on the concept of functional data depth is proposed. This Phase II control chart is directly estimated assuming that the asymptotic distribution of the rank statistic is a uniform distribution. The application of the control charts to the two-process control phases and the development of a new graphic tool for visualizing functional data (including an envelope with 95% of the deepest curves that facilitate the identification of the assignable cause of each anomaly) give rise to the proposed methodology. It has been successfully applied in real case studies belonging to the framework of anomaly detection in building energy efficiency. Additionally, a simulation study is conducted to measure the performance (as the percentage of rejection when the null hypothesis is not met) of the control charts, depending on the functional data depth used, the sample size, the presence of dependence between curves and the use of different FDA procedures for outlier detection.

In the simulation study, the use of different types of functional depths has been compared to develop Phase II of the proposed control chart. In case of the univariate functional data (single type of curves), for the three scenarios, a better performance is obtained with the mode depth measurement combined with the weighted outlier detection method and moderately large samples. Additionally, one of the final observations of the simulation study is that the control chart methodology is robust against the presence of dependence between curves. Thus, this alternative tool can be applied to the framework of continuously monitored data streams.

Generally, the authors recommend using the weighed method and the Mode functional data depth for the case of Phase I taking into account the values of $\hat{p}_f$ and $\hat{p}_c$. Thus, when the Phase I control chart is evaluated, both weighted method and Mode data depth are generally the best options in those scenarios defined by under control assumption and even in those where out of control curves are simulated. In the latter, when the change in magnitude or shape is very small, the corresponding $\hat{p}_c$ tends to be not higher to those obtained by the use of other combinations of data depth measure and outlier detection method. Nevertheless, when the change in magnitude (δ) or shape (η) increases, the power of the combination of weighted method and Mode depth tends to be higher than those corresponding to the other combinations. Moreover, regarding the Phase II control chart and taking into account the higher values of power estimates included in Tables 7 and 8, we also recommend the use of Mode data depth for Phase II control charts.

The present proposal has been verified by its application in a real case study dealing with the detection of energy efficiency anomalies in buildings. Specifically, all the previously identified real

anomalies (by the maintenance personnel) have also been successfully identified by the application of this functional approximation of control charts for Phases I and II of process control. Additionally, the proposed graphical tool helps to intuitively identify the assignable causes corresponding to each anomaly.

This procedure can be used in different industrial and scientific domains in which the control procedures are defined by functional CTQ variables.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Lu, C.W.; Reynolds, M.R, Jr. EWMA control charts for monitoring the mean of autocorrelated processes. *J. Qual. Technol.* **1999**, *31*, 166. [CrossRef]
2. Alwan, L.C.; Roberts, H.V. Time-series modeling for statistical process control. *J. Bus. Econ. Stat.* **1988**, *6*, 87–95.
3. Qiu, P.; Zou, C.; Wang, Z. Nonparametric profile monitoring by mixed effects modeling. *Technometrics* **2010**, *52*, 265–277. [CrossRef]
4. Shiau, J.J.; Huang, H.L.; Lin, S.H.; Tsai, M.Y. Monitoring nonlinear profiles with random effects by nonparametric regression. *Commun. Stat. Methods* **2009**, *38*, 1664–1679. [CrossRef]
5. Noorossana, R.; Saghaei, A.; Amiri, A. *Statistical Analysis of Profile Monitoring*; John Wiley & Sons: Hoboken, NJ, USA, 2011; Volume 865.
6. Chipman, H.; MacKay, R.; Steiner, S. Comment on Nonparametric profile monitoring by mixed effects modeling. *Technometrics* **2010**, *52*, 280–283. [CrossRef]
7. Lavin, A.; Ahmad, S. Evaluating Real-Time Anomaly Detection Algorithms–The Numenta Anomaly Benchmark. In Proceedings of the 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA), Miami, FL, USA, 9–11 December 2015; pp. 38–44.
8. Kroll, B.; Schaffranek, D.; Schriegel, S.; Niggemann, O. System modeling based on machine learning for anomaly detection and predictive maintenance in industrial plants. In Proceedings of the 2014 IEEE Emerging Technology and Factory Automation (ETFA), Barcelona, Spain, 16–19 September 2014; pp. 1–7.
9. Ahmad, S.; Lavin, A.; Purdy, S.; Agha, Z. Unsupervised real-time anomaly detection for streaming data. *Neurocomputing* **2017**, *262*, 134–147. [CrossRef]
10. Meshram, A.; Haas, C. Anomaly detection in industrial networks using machine learning: A roadmap. In *Machine Learning for Cyber Physical Systems*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 65–72.
11. Basu, S.; Meckesheimer, M. Automatic outlier detection for time series: An application to sensor data. *Knowl. Inf. Syst.* **2007**, *11*, 137–154. [CrossRef]
12. Mosallam, A.; Medjaher, K.; Zerhouni, N. Nonparametric time series modelling for industrial prognostics and health management. *Int. J. Adv. Manuf. Technol.* **2013**, *69*, 1685–1699. [CrossRef]
13. Talagala, P.D.; Hyndman, R.J.; Smith-Miles, K.; Kandanaarachchi, S.; Muñoz, M.A. Anomaly detection in streaming nonstationary temporal data. *J. Comput. Graph. Stat.* **2019**, pp. 1–28. [CrossRef]

14. Hyndman, R.J.; Liu, X.A.; Pinson, P. Visualizing big energy data: Solutions for this crucial component of data analysis. *IEEE Power Energy Mag.* **2018**, *16*, 18–25. [CrossRef]

15. Golshan, M.; MacGregor, J.F.; Bruwer, M.J.; Mhaskar, P. Latent Variable Model Predictive Control (LV-MPC) for trajectory tracking in batch processes. *J. Process Control* **2010**, *20*, 538–550. [CrossRef]

16. Kourti, T.; MacGregor, J.F. Multivariate SPC methods for process and product monitoring. *J. Qual. Technol.* **1996**, *28*, 409–428. [CrossRef]

17. Ferrer-Riquelme, A. 1.04-Statistical Control of Measures and Processes. In *Comprehensive Chemometrics*; Elsevier: Amsterdam, The Netherlands, 2009; pp. 97–126.

18. Colosimo, B.M.; Pacella, M. A comparison study of control charts for statistical monitoring of functional data. *Int. J. Prod. Res.* **2010**, *48*, 1575–1601. [CrossRef]

19. Megahed, F.; Jones-Farmer, L. Statistical Perspectives on "Big Data". In *Frontiers in Statistical Quality Control 11*; Springer: Berlin, Germany, 2015; pp. 29–47.

20. Woodall, W.; Montgomery, D. Some current directions in the theory and application of statistical process monitoring. *J. Qual. Technol.* **2014**, *46*, 78. [CrossRef]

21. Sheu, S.H.; Ouyoung, C.W.; Hsu, T.S. Phase II statistical process control for functional data. *J. Stat. Comput. Simul.* **2013**, *83*, 2144–2159. [CrossRef]

22. García, D.R. Cartas de Control Para Datos Funcionales. Master's Thesis, Centro de Investigación en Matemáticas (CIMAT), Guanajuato, Mexico, 2011.

23. Rodrigo, O.P. Monitoreo de Datos Funcionales. Master's Thesis, Centro de Investigación en Matemáticas (CIMAT), Guanajuato, Mexico, 2013.

24. Flores, M.; Tarrio-Saavedra, J.; Fernandez-Casal, R.; Naya, S. Functional extensions of Mandel's h and k statistics for outlier detection in interlaboratory studies. *Chemom. Intell. Lab. Syst.* **2018**, *176*, 134–148. [CrossRef]

25. Flores, M.; Fernández-Casal, R.; Naya, S.; Tarrío-Saavedra, J.; Bossano, R. ILS: An R package for statistical analysis in Interlaboratory Studies. *Chemom. Intell. Lab. Syst.* **2018**, *181*, 11–20. [CrossRef]

26. Flores, M.; Naya, S.; Tarrío-Saavedra, J.; Fernández-Casal, R. Functional data analysis approach of Mandel's h and k statistics in Interlaboratory Studies. In *Functional Statistics and Related Fields*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 123–130.

27. Liu, R.Y. Control Charts for Multivariate Processes. *J. Am. Stat. Assoc.* **1995**, *90*, 1380–1387. [CrossRef]

28. Liu, R.Y.; Tang, J. Control charts for dependent and independent measurements based on bootstrap methods. *J. Am. Stat. Assoc.* **1996**, *91*, 1694–1700. [CrossRef]

29. Liu, R. On a notion of data depth based on random simplices. *Ann. Stat.* **1990**, *18*, 405–414. [CrossRef]

30. Chatterjee, S.; Qiu, P. Distribution-free cumulative sum control charts using bootstrap-based control limits. *Ann. Appl. Stat.* **2009**, *3*, 349–369. [CrossRef]

31. Jones, L.; Woodall, W. The performance of bootstrap control charts. *J. Qual. Technol.* **1998**, *30*, 362. [CrossRef]

32. Chakraborti, S. Nonparametric (Distribution-Free) Quality Control Charts. In *Encyclopedia of Statistical Sciences*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2011.

33. Qiu, P.; Li, Z. On nonparametric statistical process control of univariate processes. *Technometrics* **2011**, *53*, 390–405. [CrossRef]

34. Barbeito, I.H.; Zaragoza, S.; Tarrío-Saavedra, J.; Naya, S. Assessing thermal comfort and energy efficiency in buildings by statistical quality control for autocorrelated data. *Appl. Energy* **2017**, *190*, 1–17. [CrossRef]

35. Montgomery, D. *Introduction to Statistical Quality Control.*; John Wiley & Sons: New York, NY, USA, 2009.

36. Ledolter, J.; Bisgaard, S. Challenges in constructing time series models from process data. *Qual. Reliab. Eng. Int.* **2011**, *27*, 165–178. [CrossRef]

37. Box, G.E.; Luceño, A.; Paniagua-Quinones, M.D.C. *Statistical Control by Monitoring and Adjustment*; John Wiley & Sons: Hoboken, NJ, USA, 2011; Volume 700.

38. Del Castillo, E. Statistical process adjustment: A brief retrospective, current status, and some opportunities for further work. *Stat. Neerl.* **2006**, *60*, 309–326. [CrossRef]

39. Knoth, S.; Schmid, W. Control charts for time series: A review. *Front. Stat. Qual. Control* **2004**, *7*, 210–236.

40. Prajapati, D.; Singh, S. Control charts for monitoring the autocorrelated process parameters: a literature review. *Int. J. Prod. Qual. Manag.* **2012**, *10*, 207–249. [CrossRef]

41. Tsai, T.R.; Hsieh, Y.W. Simulated Shewhart control chart for monitoring variance components. *Int. J. Reliab. Qual. Saf. Eng.* **2009**, *16*, 1–22. [CrossRef]

42. Jones-Farmer, L.; Woodall, W.; Steiner, S.; Champ, C. An overview of phase I analysis for process improvement and monitoring. *J. Qual. Technol.* **2014**, *46*, 265. [CrossRef]

43. Grasso, M.; Colosimo, B.M.; Tsung, F. A phase I multi-modelling approach for profile monitoring of signal data. *Int. J. Prod. Res.* **2017**, *55*, 4354–4377. [CrossRef]

44. Febrero-Bande, M.; Galeano, P.; González-Manteiga, W. Outlier detection in functional data by depth measures, with application to identify abnormal NOx levels. *Environmetrics* **2007**, *19*, 331–345. [CrossRef]

45. Kazemzadeh, R.B.; Noorossana, R.; Ayoubi, M. Change point estimation of multivariate linear profiles under linear drift. *Commun. Stat.-Simul. Comput.* **2015**, *44*, 1570–1599. [CrossRef]

46. Woodall, W.H.; Spitzner, D.J.; Montgomery, D.C.; Gupta, S. Using control charts to monitor process and product quality profiles. *J. Qual. Technol.* **2004**, *36*, 309–320. [CrossRef]

47. Lanhede, D. Statistical Process Control: Evaluation and Implementation of Methods for Statistical Process Control at GE Healthcare. Master's Thesis, Umeå University, Umeå, Sweden, 2015.

48. Flores Sánchez, M. Nuevas Aportaciones del Análisis de Datos Funcionales en el Control Estadístico de Procesos. Ph.D. Thesis, Universidade da Coruña, A Coruña, Spain, 2019.

49. Febrero-Bande, M.; Oviedo de la Fuente, M. Statistical computing in functional data analysis: The R package fda.usc. *J. Stat. Softw.* **2012**, *51*, 1–28. [CrossRef]

50. Fraiman, R.; Muniz, G. Trimmed means for functional data. *Test* **2001**, *10*, 419–440. [CrossRef]

51. López-Pintado, S.; Romo, J. On the concept of depth for functional data. *J. Am. Stat. Assoc.* **2009**, *104*, 718–734. [CrossRef]

52. Cuevas, A.; Febrero, M.; Fraiman, R. Robust estimation and classification for functional data via projection-based depth notions. *Comput. Stat.* **2007**, *22*, 481–496. [CrossRef]

53. Cuesta-Albertos, J.A.; Nieto-Reyes, A. The random Tukey depth. *Comput. Stat. Data Anal.* **2008**, *52*, 4979–4988. [CrossRef]

54. Flores, M.; Fernández-Casal, R.; Naya, S.; Tarrío-Saavedra, J. qcr: Quality Control Review. R Package Version 1.0. 2016. Available online: https://cran.r-project.org/web/packages/qcr/index.html (accessed on 26 December 2019).

55. Hyndman, R.J.; Shang, H.L. Rainbow plots, bagplots, and boxplots for functional data. *J. Comput. Graph. Stat.* **2010**, *19*, 29–45. [CrossRef]

56. Raña, P. Pointwise Forecast, Confidence and Prediction Intervals in Electricity Demand and Price. Ph.D. Thesis, Universidade da Coruña, A Coruña, Spain, 2016.

57. Colosimo, B.M.; Pacella, M. On the use of principal component analysis to identify systematic patterns in roundness profiles. *Qual. Reliab. Eng. Int.* **2007**, *23*, 707–725. [CrossRef]