# Fault Diagnosis of a Granulator Operating under Time-Varying Conditions Using Canonical Variate Analysis

**Elena Quatrini** [1],*[iD]**, Xiaochuan Li** [2]**, David Mba** [2][iD] **and Francesco Costantino** [1][iD]

[1] Department of Mechanical and Aerospace Engineering, Sapienza University of Rome, Via Eudossiana, 18, 00184 Rome, Italy; francesco.costantino@uniroma1.it
[2] Faculty of Computing, Engineering and Media, De Montfort University, Leicester LE1 9BH, UK; xiaochuan.li@dmu.ac.uk (X.L.); david.mba@dmu.ac.uk (D.M.)
*  Correspondence: elena.quatrini@uniroma1.it; Tel.: +39-0644585260

**Abstract:** Granulators play a key role in many pharmaceutical processes because they are involved in the production of tablets and capsule dosage forms. Considering the characteristics of the production processes in which a granulator is involved, proper maintenance of the latter is relevant for plant safety. During the operational phase, there is a high risk of explosion, pollution, and contamination. The nature of this process also requires an in-depth examination of the time-dependence of the process variables. This study proposes the application of canonical variate analysis (CVA) to perform fault detection in a granulation process that operates under time-varying conditions. Beyond this, a different approach to the management of process non-linearities is proposed. The novelty of the study is in the application of CVA in this kind of process, because it is possible to state that the actual literature on the theme shows some limitations of CVA in such processes. The aim was to increase the applicability of CVA in variable contexts, with simple management of non-linearities. The results, considering process data from a pharmaceutical granulator, showed that the proposed approach could detect faults and manage non-linearities, exhibiting future scenarios for more performing and automatic monitoring techniques of time-varying processes.

**Keywords:** condition monitoring; performance estimation; multivariate methods; pharmaceutical plant; machine learning; canonical variate analysis

## 1. Introduction

Granulators are very important machines for pharmaceutical companies because granulation is one of the most applied processes to produce tablets and capsule dosage forms, increasing the uniformity of drug distribution in the product and its physical properties. This means that the granulation process is widely implemented and widespread within pharmaceutical production plants. Granulation is a process with high-risk situations, which require exhaustive maintenance and fault management. Some of the major risks are the risk of explosions or degradation of the product, the possibility of pollution or the granule breakage, and the consequent emission of powder. In general, it is possible to assert that every day pharmaceutical plants should consider the risk of emission of a toxic material. Especially when approaching application areas such as the chemical sector, it is necessary to account for the dependence of process variables on time. In fact, sensory signals can show a strong correlation between the past and future measurements, in the presence of noise and disturbances [1]. This characteristic implies the need to use techniques that go beyond the assumption that process variables are time-independent. This assumption is the background for many widely

applied multivariate statistics techniques for fault detection, for example, such as principal component analysis (PCA), which more specifically require steady state data. Canonical variate analysis (CVA) considers time-dependence within the data and are thus more suitable for non-steady data.

Moreover, the considered granulation process can be defined as a batch process, because all materials are loaded inside the machine at the beginning of the production process and unloaded only at the end of all necessary operations [2]. These processes play a key role in pharmaceutical production processes [3]. Since this category of processes has characteristics that make monitoring difficult, it becomes essential to find effective methods for on-line monitoring and fault diagnosis, for improving product quality and making it possible to implement maintenance actions before the completion of the batch or the introduction of a new batch [4]. This type of process is widely present in today's industrial environment and the CVA was successfully applied to define the state-space model of a 2-D batch process, also permitting to take into account the possible correlation between two or more batches [5]. In the same way, CVA can effectively distinguish between failures that compromise product quality and failures that leave them unchanged [6]. The importance of proper monitoring in this type of process lies in the fact that without it, even if the operators identify problems in product quality, they are not able to define in advance the causes of the problems or when they occur [4].

Another important aspect of the examined process is its being operational under time-varying conditions, i.e., with different process conditions. The granulation process implemented consists of seven different process phases, each with its specific characteristics. This characteristic is exacerbated by the impossibility of identifying and separating the phases during the production process. All phases, in fact, are performed by the same machine, that is, the analyzed granulator, without there being evidence of the end and beginning of the phases. A further cause of time-varying conditions of the process is the non-continuity of production processing, whereby maintenance work or sensor changes could be included. This type of process has developed an important interest in the field of fault detection and maintenance, in general, in recent years, since its widespread diffusion [7–10]. Considering what was said earlier about the process under consideration, i.e., that it could be defined as a batch process, one can assert that these processes are characterized by typical changing conditions, generating nonstationary profiles for different process variables [10]. In fact, one of the greatest difficulties in monitoring and, therefore, maintaining such varying processes is that fault features are disturbed by this variability [9] and most of the existing methods, which are based on a stationary assumption, are only suitable for fault detection under constant operating conditions [11]. Rotating machines, for example, and their extremely variable speed levels, are an issue that is being addressed, due to the widespread use of rotating machines in the industrial landscape [7]. Likewise, induction motors operate mostly in time-varying conditions, with the effect that the bandwidth of the fault frequency components varies with the speed variation [8]. The articles cited are intended to show how the characteristics of the process examined are not rare to encounter, even in different industrial contexts, thus emphasizing the importance of improvement in their monitoring.

It can be said that the literature on the subject of maintenance, in processes with significant variability over time is limited [9]. Based on this, the need for action to ensure that CVA can achieve positive results in monitoring this type of process is often highlighted. In this type of process, the risk in the application of multivariate statistics techniques is that they assume that the underlying processes are linear and static. This entails the risk that normal changes in operating conditions are labelled as faults, which would result in high false-positive rates [12]. The literature shows some limitations in the application of CVA in processes of this type, with a consequent need, for example, to implement changes to traditional CVA [12,13]. The combination of CVA with other techniques, such as Fisher Discriminant Analysis, was also proposed to increase the fault detection capabilities of a model [14]. Another interesting study proposes to combine CVA and Fisher Discriminant Analysis, to handle the dynamic data with highly serial correlations, integrating it with the locality preserving projection, to further improve the performance of fault classification [15]. Another interesting research was carried out for applying CVA in such processes, developing an effective recursive CVA [16]. Nevertheless,

relatively little has yet been developed for the application of CVA in non-linear processes, i.e., extremely variable over time [17].

In general, however, CVA has achieved excellent results in many maintenance applications, either in its original form or in its modified form [18–21], even in the context of slowly evolving faults [22]. An important characteristic of CVA, when compared to other methods, is that CVA can guarantee better stability and it uses fewer parameters to represent dynamic systems [19]. The adaptive kernel canonical variable analysis, together with Bayesian fusion proved to be effective in fault management for quality-related multiple fault, because they can fully mine the local information inside the process and detect the quality-related multiple faults from plant-wide-level to subprocess-level [18]. Returning to the studies carried out on rotating machines, it is interesting to note that CVA was successfully applied in this context, even with the development of a new health indicator, compared to those traditionally considered [20]. In other cases, the standard CVA-based approach was extended by using canonical correlation features, resulting in an improved fault detection model [23]. CVA also proved to be extremely suitable for monitoring chemical processes, which, as asserted before, show a relevant correlation between the past and future instances [1]. The fundamental characteristic of CVA, therefore, is that it calculates linear combinations of the past values of the system inputs and outputs, which were most highly correlated with the linear combinations of the future values of the outputs of the process [24]. Moreover, considering, for example PCA and CVA, it is possible to assert that dimensionality reduction techniques could better generalize to new process data than representations that used the entire dimensionality [21]. Beyond this, it is necessary to take into account the computational complexity of a model, which is closely related to the size of the dataset itself. This means that it is often necessary to intervene to reduce the number of variables considered, by applying dimensionality reduction techniques, to properly manage computational time [25].

Based on the considerations presented and the literature on the subject, this study proposes the application of CVA for the identification of faults in a granulator, within a pharmaceutical plant. Specifically, the study proposes an approach for the management of non-linearities and discontinuities in the process, to overcome the limitations of traditional CVA, in processes that operate under time-varying conditions, without modifying the model itself. As can be seen from the literature on the subject, the non-linearities of the process tend to be managed with changes to the very structure of the CVA. The authors instead propose a different approach, deriving from the structure of the process itself, in which the CVA remains in its original form and the discontinuities are managed before its application, but not in the application phase. The lack of continuity in the delivery of the process raises doubts and questions to answer, which could be summarized as follows:

- It is necessary to understand whether a process interruption should be considered as the end of a process with certain characteristics, and thus, the beginning of a process with new characteristics. This would result in the loss of tested and performing monitoring techniques, over a period of time, with the consequent need to treat each subperiod of the same process as a separate process. Any maintenance work, sensor change, or simply a new set up, could generate specific process characteristics, in some way, to be managed ad hoc.

The sensitivity of the two indices to failures is also a point of interest in the paper. The aim was as follows:

- Make a critical assessment of the results achieved and that can be achieved, by looking individually at the health indices taken into consideration, in order to better understand how they can be analyzed together, to increase the overall performance of the model.

Finally, CVA was never applied in such contexts and the structure of the process itself incorporates many points of innovation in the presented study:

- The granulation process is a current and relevant process in the pharmaceutical sector, making it important that it is properly monitored and maintained, also considering the economic value of the

products processed with it. Beyond this, granulation is a discontinuous and multiphase process; this is detailed in the next section of the paper. These characteristics make it peculiar. However, at present, there is no study in the literature that proposes a similar approach to ours for these processes, and a critical evaluation of how the division into production periods and multiphase are managed with a technique like CVA. The evidence of the running production phase is not evident in processes such as granulation and their recognition requires a pre-processing phase with an associated risk of error, but above all with a lengthening of monitoring times. For this reason, it is interesting to evaluate how CVA manages to correctly monitor the process, without having the evidence of the production phase currently running.

For completeness and reproducibility, the authors want to specify that the methods presented in this study were implemented in MATLAB 9.8 (R2020a).

## 2. Material and Methods

### 2.1. CVA

CVA is a dimensionality reduction technique that can be used to monitor machine operation by converting process data into a one-dimensional health indicator [22]. The data extracted under normal operating conditions of the machinery were used during the training phase of the model, to define the thresholds of normality of the state of health of the machinery. This enabled the identification of non-normal states of machinery operation, when the health index value exceeded the normal threshold. The main aim of CVA was to maximize the correlation between the two datasets [21]. In fault detection application, the considered datasets were the past and the future data measurements. CVA application for fault detection was proposed in 2014 [26].

CVA can capture the process dynamics, so it can take into account, not only the correlation between the different process variables but also the time correlation within each variable. For this reason, this method was demonstrated as more suitable for time-varying process than other dynamic approaches for fault detection like PCA. Given that the industrial process investigated in this study involved various operating conditions, a dynamic process-monitoring tool like CVA would suit the data very well. The drawback of CVA is that its effectiveness is much dependent on the proper selection of a failure threshold. To solve this problem, a kernel density estimation (KDE) approach that fully considers the process nonlinearities was employed in this study for the determination of the threshold. KDE estimates control limits, according to the distributions of data, and features very few tuning parameters. Another potential shortfall of CVA is that when the deviations between different operating conditions are relatively large, operating conditions with "large variations" might lead to high false alarms, and operating conditions with "small variations" might cause low detectability. To solve this problem, a just-in-time-learning scheme was suggested to be incorporated into the CVA approach, and this would be future research direction.

The mathematical procedure for the application of CVA was described in previous studies, such as [27]. Following these previous in-depth analyses, in this study it is possible to present only the fundamental steps for the application of CVA. These steps following what was explained in [28].

The first step is to generate two data matrices from the measured data $y_t \in R^n$, with $n$ representing the process variables and $t$ being the sampling time. It was possible to expand each sampling, including $p$ past samples and $f$ future samples. After the definition of these two parameters, which should have the same value, the following step was the construction of the future and past samples vectors, $y_{f,t} \in R^{fn}$ (1) and $y_{p,t} \in R^{pn}$ (2).

$$y_{f,t} = \begin{bmatrix} y_{t+1}, & y_{t+2}, & \cdots, & y_{t+f-1} \end{bmatrix}^T \tag{1}$$

$$y_{p,t} = \begin{bmatrix} y_{t-1}, & y_{t-2}, & \cdots, & y_{t-p} \end{bmatrix}^T \tag{2}$$

To avoid variables with bigger absolute values that influence fault detection, $y_{f,t}$ and $y_{p,t}$, were normalized to mean zero vectors, $\hat{y}_{f,t}$ and $\hat{y}_{p,t}$. The latter, at different sampling times, were then rearranged for creating the future observation matrix $\hat{Y}_f \in R^{fnXN}$ (3), and the past observation matrix $\hat{Y}_p \in R^{pnXN}$ (4).

$$\hat{Y}_f = \left[ \hat{y}_{f,\,t+1},\ \hat{y}_{f,\,t+2},\ \cdots,\ \hat{y}_{f,\,t+N} \right] \tag{3}$$

$$\hat{Y}_p = \left[ \hat{y}_{p,\,t+1},\ \hat{y}_{p,\,t+2},\ \cdots,\ \hat{y}_{p,\,t+N} \right] \tag{4}$$

In (3) and (4), $N = m - p - f + 1$, with m representing the number of total samples in $y_t$. According to [29], the following step was the application of the Cholesky decomposition to $\hat{Y}_f$ and $\hat{Y}_p$, to configure a Hankel matrix $H$, to create a correlation matrix of $\hat{Y}_f$ and $\hat{Y}_p$, and to transfer the latter in a space with reduced dimensionality. Then, using then the singular value decomposition (5) to decompose the truncated Hankel matrix $H$, it is possible to find the linear combination that maximizes the correlation between $\hat{Y}_f$ and $\hat{Y}_p$.

$$H = \sum\nolimits_{f,f}^{-1/2} \sum\nolimits_{f,p} \sum\nolimits_{p,p}^{-1/2} = U\Sigma V^T \tag{5}$$

In (5), $\Sigma_{f,f}$ and $\Sigma_{p,p}$ are the covariance matrices of $\hat{Y}_f$ and $\hat{Y}_p$ and $\Sigma_{f,p}$ is their cross-covariance matrix. Considering $r$ as the order of $H$:

$$U = [u_1,\ u_2,\ \ldots,\ u_r] \in R^{npXr} \tag{6}$$

$$V = [v_1,\ v_2,\ \ldots,\ v_r] \in R^{nfXr} \tag{7}$$

$$\Sigma = \begin{pmatrix} d_1 & . & . & . & 0 \\ & . & . & & \\ . & & . & & \\ . & & & . & \\ 0 & & & & d_r \end{pmatrix} \in R^{rXr} \tag{8}$$

The columns of $U$ and the columns of $V$ are called the left-singular and right-singular vectors of $H$. $\Sigma$ is a diagonal matrix, and its diagonal elements are called singular values. They represent the degree of correlation between the corresponding left-singular and right-singular vectors. The largest $q$ singular values permit to truncate the matrix $V$ in $V_q$ (9).

$$V_q = [v_1,\ v_2,\ \ldots,\ v_q] \in R^{npXq} \tag{9}$$

$V_q$ allow converting $\hat{Y}_p$ in a reduced $q$-dimensional matrix $\phi \in R^{qXn}$ (10).

$$\phi = [z_{t=1},\ z_{t=2},\ \ldots,\ z_{t=N}] = K \cdot \hat{Y}_p \tag{10}$$

$$K = V_q^T \sum\nolimits_{p,p}^{-1/2} \in R^{qXnp} \tag{11}$$

$z_t$ is the canonical state space. The residual space $\psi$ is computed as (12).

$$\psi = [\varepsilon_{t=1},\ \varepsilon_{t=2},\ \ldots,\ \varepsilon_{t=N}] = G \cdot \hat{Y}_p \tag{12}$$

$$G = I - V_q V_q^T \sum\nolimits_{p,p}^{-1/2} \in R^{npXnp} \tag{13}$$

The canonical variates $z_t$ and residual variates $\varepsilon_t$ are used to construct the health indicators. The first one represents the projection of the measurement matrix into the $q$-dimensional space, the second one represents system variations not considered by the state space.

The health indicators adopted in this study were the Hotelling $T^2$ (14) and $Q$ (SPE) (15) statistics [30]. A machine fault occurs when the value of the health indicator goes over the computed threshold.

$$T_t^2 = \sum_{j=1}^{q} z_{t,j}^2 \tag{14}$$

$$Q_t = \sum_{j=1}^{np} \varepsilon_{t,j}^2 \tag{15}$$

### 2.2. Case Study

The case study presented here, concern the granulation process within a pharmaceutical plant. As explained before, granulation is a process that is heavily involved in many pharmaceutical productions, such as the production of tablets and capsule dosage forms. Granulation aims to augment the uniformity of drug distribution, as well as its physical properties. This process involves a high level of risk for the plant, for the characteristics of the products involved, as well as for the temperatures applied. Specifically, the high temperatures, the possible introduction of nitrogen, and the wrongly managed vacuum can generate explosions or the degradation of the product; the condensation of the evaporated liquid can cause pollution; the granule breakage results in the emission of powder. Figure 1 shows the structure of the considered granulator, with evidence of the components. The peristaltic pump, including the product's temperature sensor and nozzle, were inside the main section of the granulator.
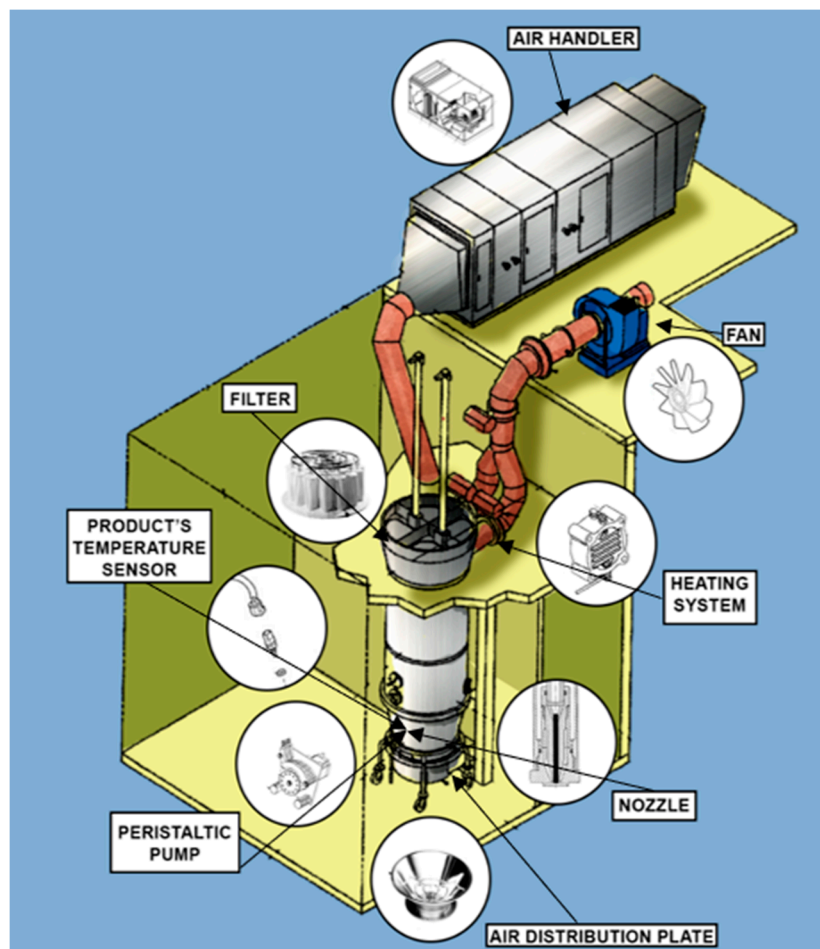


**Figure 1.** Granulator with the evidence of the components.

One aspect of some industrial processes (and particularly of chemical processes) that should not be overlooked is their strong dependence on time [1]. This requirement, as asserted before, implies the need for monitoring and subsequent data analysis with the application of techniques that are not based on an assumption of independence from the time, of the process variables.

Other features that distinguish this case study are the management and the structure of the production process itself. First, the process is not delivered continuously over the course of a year but is only implemented in specific time windows, due to managerial choices or for the performance of maintenance work. The covered periods are:

- From 2 to 12 January (A)
- From 11 to 21 February (B)
- From 4 to 15 May (C)
- From 1 to 22 June (D)
- From 29 July to 17 September (E)

The second cause of discontinuity is the process structure itself. The analyzed process consists of 7 phases, each having specific process characteristics and consequently specific anomaly situations:

- $p_1$ = Initializing;
- $p_2$ = Conditioning;
- $p_3$ = 1st Spray;
- $p_4$ = Heating;
- $p_5$ = 2nd Spray;
- $p_6$ = Drying;
- $p_7$ = Unloading.

These non-linearities of process delivery, as well as the delivery of any maintenance interventions or changes (for example, in the sensors on the production line), might lead to a lack of continuity in the process trend, with consequent different process characteristics in the different periods analyzed.

These above-mentioned features render this case study (and more generally similar production processes) highly complex, from the standpoint of monitoring their health status. This process operates clearly under time-varying conditions. All of the above makes it challenging to analyze the condition of the machinery, because it is difficult to contextualize the raw data received in input.

The final dataset was composed of 46,770 instances referring to a phase of activity of the machine.

The monitored parameters were 15, and every data point $d = \{x_1, \ldots, x_{15}\}$ was a set of 15 measures at the same time, with a time rate of 1 min:

- $x_1$ = Spray Percentage
- $x_2$ = Air IN Flow
- $x_3$ = Spray Flow
- $x_4$ = Air Pressure Spray
- $x_5$ = Microclimate Pressure
- $x_6$ = Cleaning Pressure
- $x_7$ = Air IN Temperature
- $x_8$ = Washing Air Temperature
- $x_9$ = Air OUT Temperature
- $x_{10}$ = Product Temperature
- $x_{11}$ = Cooling Temperature
- $x_{12}$ = Absolute Air IN Humidity
- $x_{13}$ = Relative Air IN Humidity
- $x_{14}$ = Product Humidity

- $x_{15} =$ Mill Speed

As far as the pre-processing of the data were concerned, the data were actually processed in a raw manner, i.e., as it was collected by the machinery. In order to evaluate the CVA's ability to manage process discontinuities, in terms of the various phases of the process, as well as the various verifiable failures, it was preferred to not label data from either point of view. The data entered and processed by the model were, therefore, pure process data, without any kind of manipulation or analysis. However, in order to better understand the progress of the process, the faults and the connections between the various phases, it was necessary to involve some process experts in the initial phase of the analysis. In fact, the comprehension of the process knowledge is absolutely essential; this involves communicating and collaborating with who manages and carries out the process, referred to as the process experts. The experts of the process:

1. They explain how healthy and unhealthy values and trends depend on what is going on in the process, thus the model requires a logical decomposition of the production cycle. This means that the experts of the process explain to the data-scientist, the process phases and their characteristics. This allows us to have a greater and more conscious understanding of the data gained.
2. They help with the identification of normal and faulty conditions of the machinery.

It is important to circumscribe the applicability, and especially the replicability, of what is presented in the following section of the study. The parameters shown are not replicable and reusable in an absolute way, in the same production process, in a different context. These parameters, in fact, depend on the type of production process in general, but also on the extremely specific characteristics of each process itself. The processed product might in fact be different, as well as the characteristics with which it is processed, e.g., the maintenance work performed on it or the sensor technology applied to it. These are just some of the characteristics that make what is proposed in the case study a verification of the applicability of CVA in the type of process under consideration, and a proposal of approach for the management of its non-linearities, but not a proposal of specific parameter settings. As can be seen in the next section, in fact, different periods of the same production process that were analyzed had different optimal values of the same parameters. As far as the faults considered were concerned, in this study, only fault detection and not fault identification was taken into consideration. We, therefore, decided to not analyze the specific faults identified by the model, and instead look at the accuracy of identifying a non-health state of the machinery. This meant that no specific analysis was carried out at present on the type of faults identified and the distribution of errors between them.

## 3. Results

Figure 2 represents the normal trend of the 15 parameters presented in Section 2.2 "Case study", during a healthy production cycle. It was important to specify that the values in Figure 2 were normalized, converting them into a range of values between 0 and 1. The purpose of this figure was simply to give an idea of the behavior of the monitored parameters, during a healthy operating phase of the machinery. The authors believe that Figure 2 clearly shows the discontinuities and non-linearities present in the process, even when there are no faults present. It is necessary to make two points regarding Figure 2, explicit—the nomenclature was consistent with what was previously presented in the text; $x_{14}$ was not presented, since it was always 0.
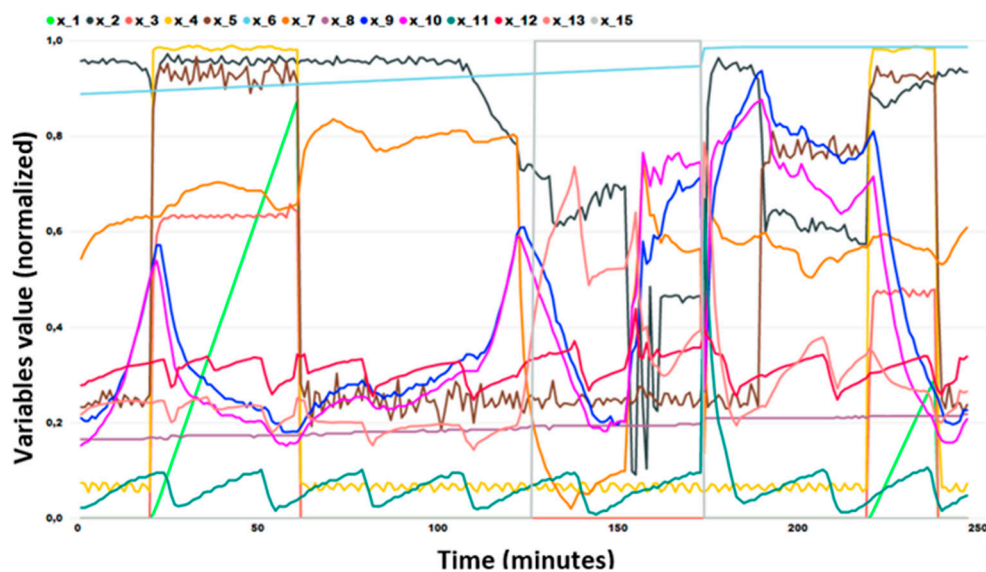
**Figure 2.** Run chart of the 15 parameters during a complete production cycle.

A first test implemented was to apply CVA to the entire production process, without distinction of the production phases and without considering the production interruptions present in the dataset. Despite attempts at optimization, in this context, CVA was not able to distinguish correctly neither fault situations nor normal situations, as shown in Figure 3. Both $T^2$ and $Q$ revealed an excessive number of false alarms, but above all a complete inability to recognize fault situations. In fact, from the instances after 40,000, roughly, the state of the machinery was always faulty, but the model was not able to correctly identify its state. Going into detail about Figure 3's axes:

- the *x*-axis represents the time sequence of the measurements, i.e., the number of instances contained in the dataset.
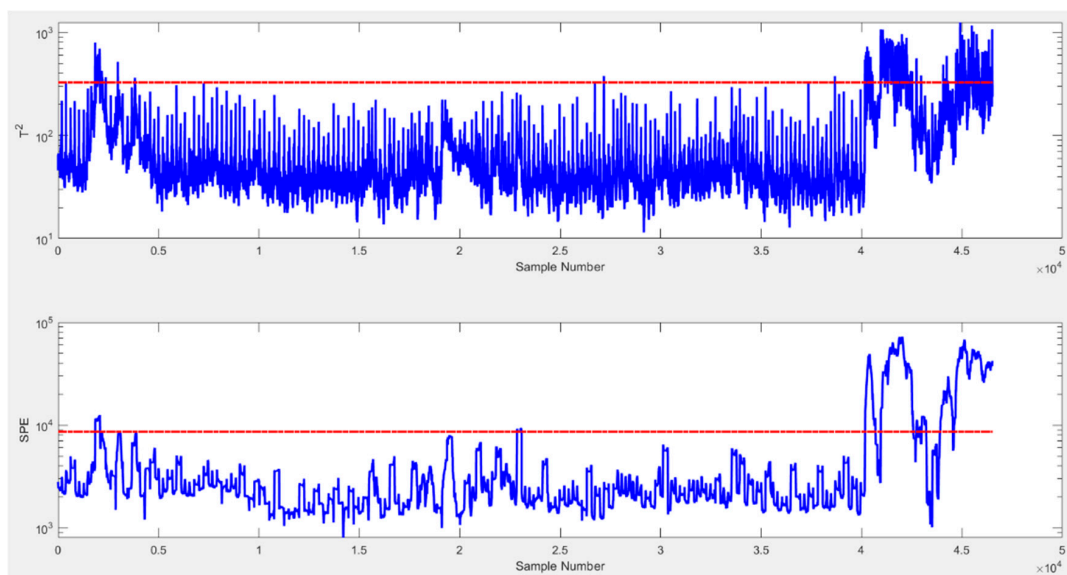- the *y*-axis represents the value of the health indicator.



**Figure 3.** Canonical variate analysis (CVA) applied to the complete dataset.

The results obtained, therefore, highlight the need to analyze the various production continuity periods separately. In doing so, the results were not affected by any maintenance work or sensor changes

that might have been carried out during production interruption periods. As a result, five different CVA models were implemented, i.e., one for the period from 2 to 12 January, one from 11 to 21 February, one from 4 to 15 May, one from 1 to 22 June, and the last one from 29 July to 17 September.

For the definition of the value of $p$, i.e., past samples, and $f$, i.e., future samples, the authors tried to apply the function of autocorrelation. The nature and composition of the production process, however, caused the results obtained from its application to not allow immediate identification of the value sought. The analyzed process was correlated with itself in time, preventing the identification of a specific value for the time intervals of interest. An example is shown in Figure 4. In Figure 4, the composition of the axes was as follows:

- the $x$-axis represents time lags, i.e., how much an instance was correlated in time with itself.
- the $y$-axis represents the degree of correlation of the instances in time.
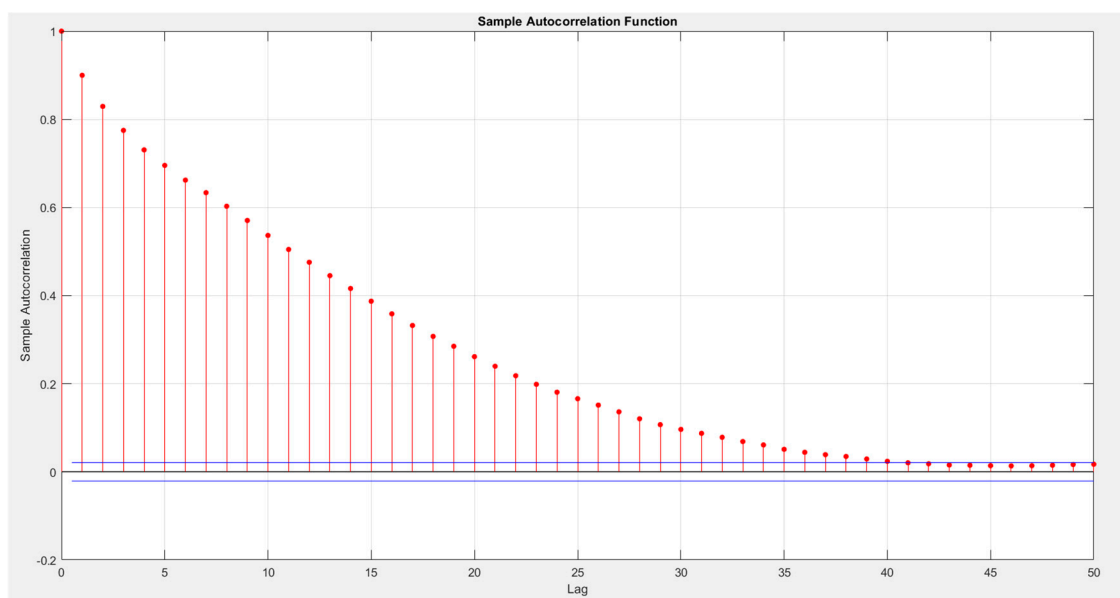


**Figure 4.** Example of the results of the autocorrelation function.

This trend, therefore, required the authors to determine the time lag of interest for each block analyzed iteratively, i.e., through tests. The starting point of the tests was a detailed analysis of the results achieved with the application of the autocorrelation function. This meant that the authors tested the achievable results with different $p$ and $f$ values for each period considered, to find out which was the best setting, i.e., the one with the smallest error rate.

The correct setting for $p$ and $f$ after a series of tests was determined to be that shown in Table 1.

**Table 1.** $p$ and $f$ values.

| Period | $p$ and $f$ Value |
|---|---|
| From 2 to 12 January | 150 |
| From 11 to 21 February | 40 |
| From 4 to 15 May | 40 |
| From 1 to 22 June | 245 |
| From 29 July to 17 September | 40 |

Different methodologies were proposed for the computation of the retained states $q$. The most popular techniques were those based on considering the dominant singular values in matrix $\Sigma$ (8). Determining $q$ based on individual values, in this case, was not a viable option. What was realistically

searched for in the graph in Figure 5 was a value of $q$, that is, the retained states, beyond which one could say with certainty that the value of the singular values, following the $q$-singular value were zero, allowing to neglect their respective state variables from the model, since the values quickly decreased with a very evident slope.
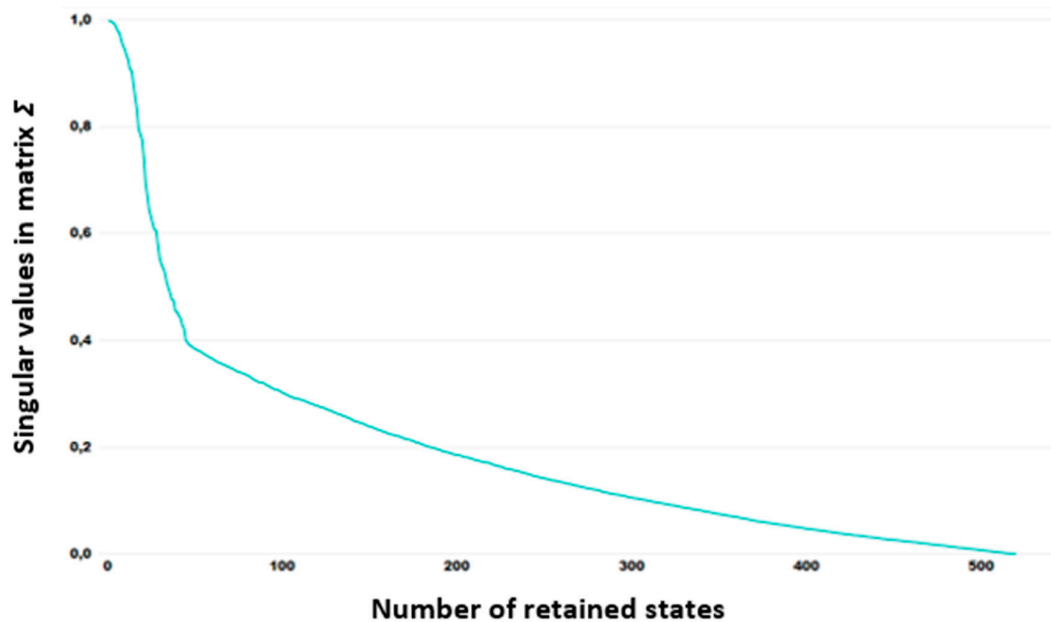


**Figure 5.** Number of retained states using dominant singular values.

As shown in Figure 5, there was no evident step of decrease in the value of the dominant singular values. Furthermore, since both performance indicators were included in this case study, the value of $q$ could not influence the fault detection process. This could be said because what was not detected by $T^2$, was detected by $Q$.

The authors then decided to determine $q$ since the false alarm rates were obtained, looking for the value that would ensure the lowest number of false alarms. After several tests, it was considered appropriate to set the value of $q$ to 50, for all considered periods. Finally, the confidence boundary adopted for the definition of the normal healthy conditions was 0.99.

Considering the 5 periods examined, the ratio between healthy and unhealthy data for each of them was as follows (Table 2):

**Table 2.** Ratio between healthy and unhealthy data for every period.

| Period | Healthy Data (%) | Unhealthy Data (%) |
|---|---|---|
| From 2 to 12 January | 80% | 20% |
| From 11 to 21 February | 92% | 8% |
| From 4 to 15 May | 93% | 7% |
| From 1 to 22 June | 77% | 23% |
| From 29 July to 17 September | 90% | 10% |

It is necessary to underline that, in Figures 6–10, the test datasets contained both healthy and faulty data. To improve the comprehension of the results, the authors believed it was necessary to better explain what is shown in Figures 6–10. The red dotted line represents the threshold of the health status of the machinery, i.e., the value of the health indicator beyond which the machinery was in a state of failure. The data presented were initially representative of a total health status, specifically up to the claims presented for each period in Table 3:
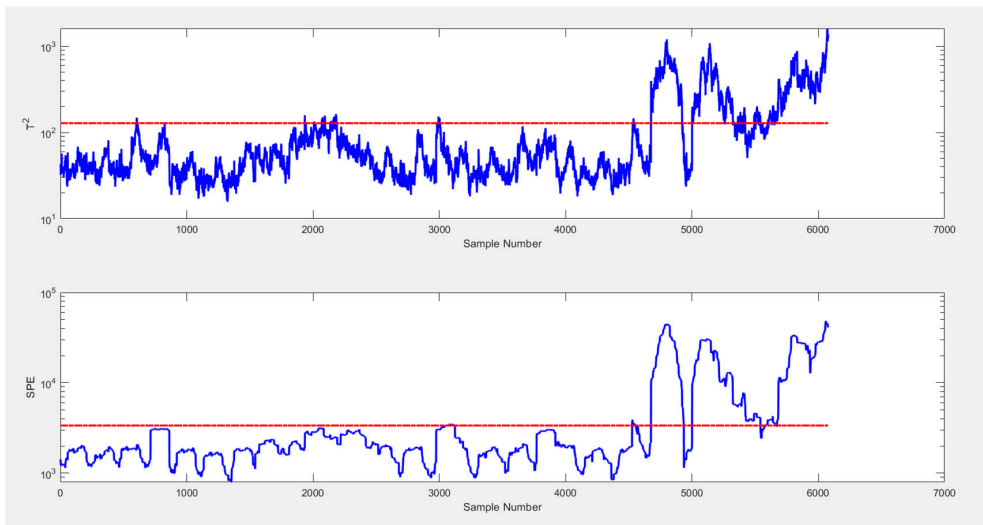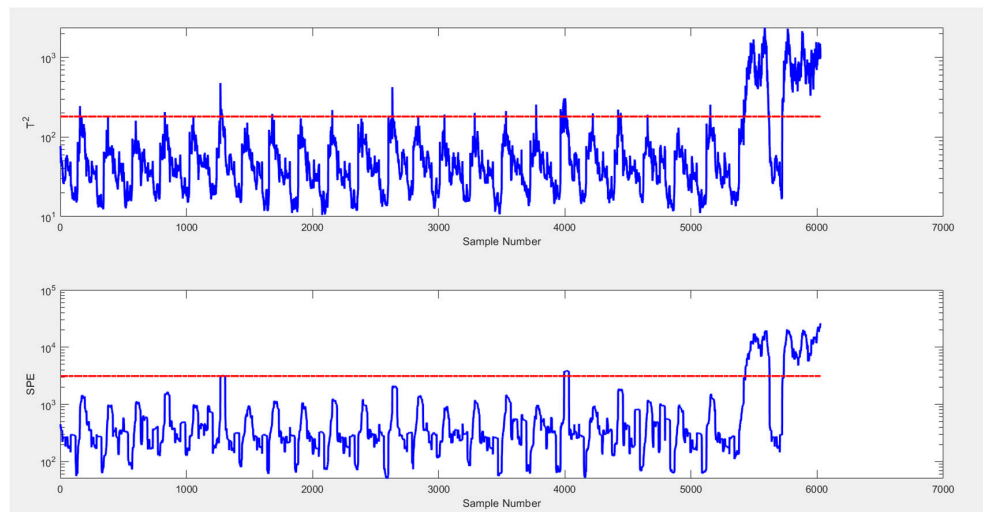
**Figure 6.** $T^2$ and $Q$ for period $A$.
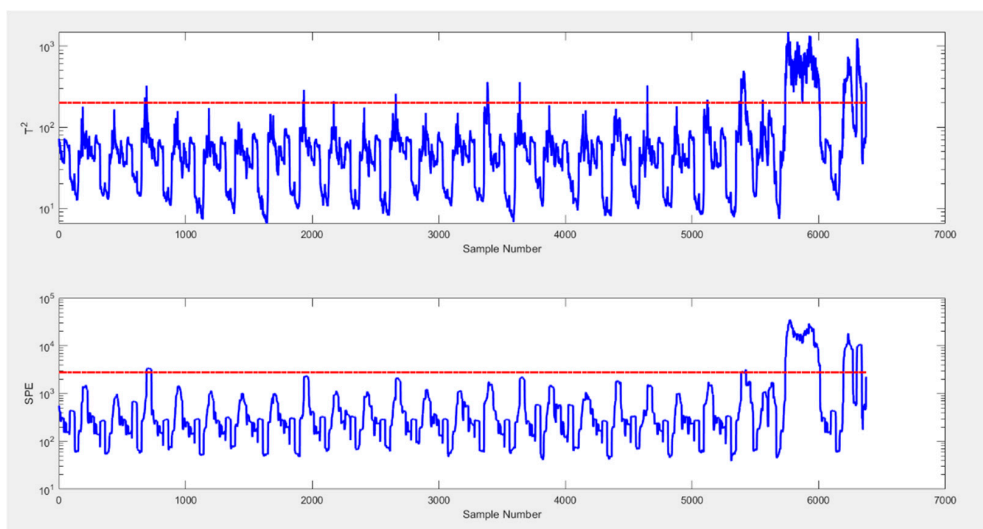


**Figure 7.** $T^2$ and $Q$ for period $B$.



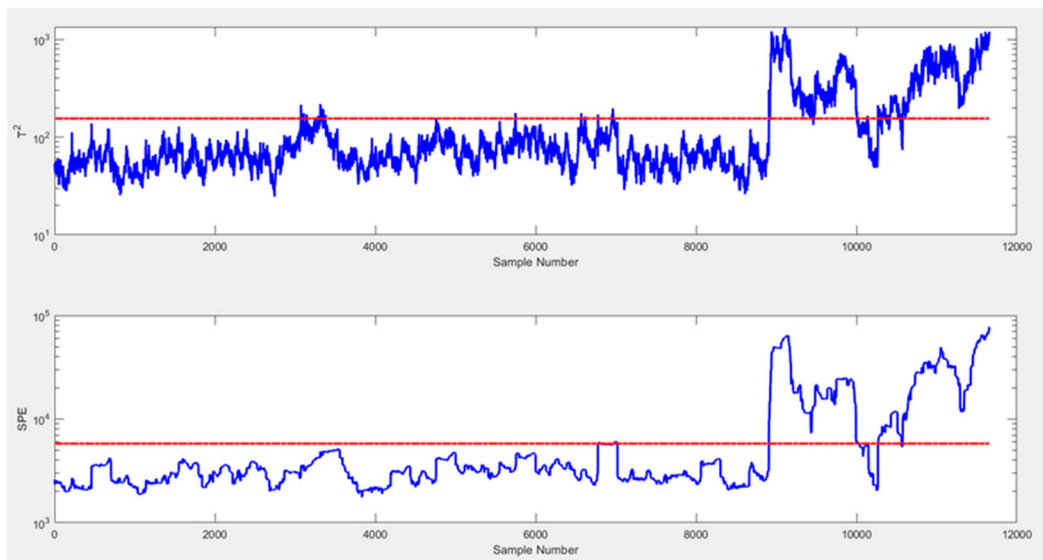**Figure 8.** $T^2$ and $Q$ for period $C$.

**Figure 9.** $T^2$ and $Q$ for period $D$.



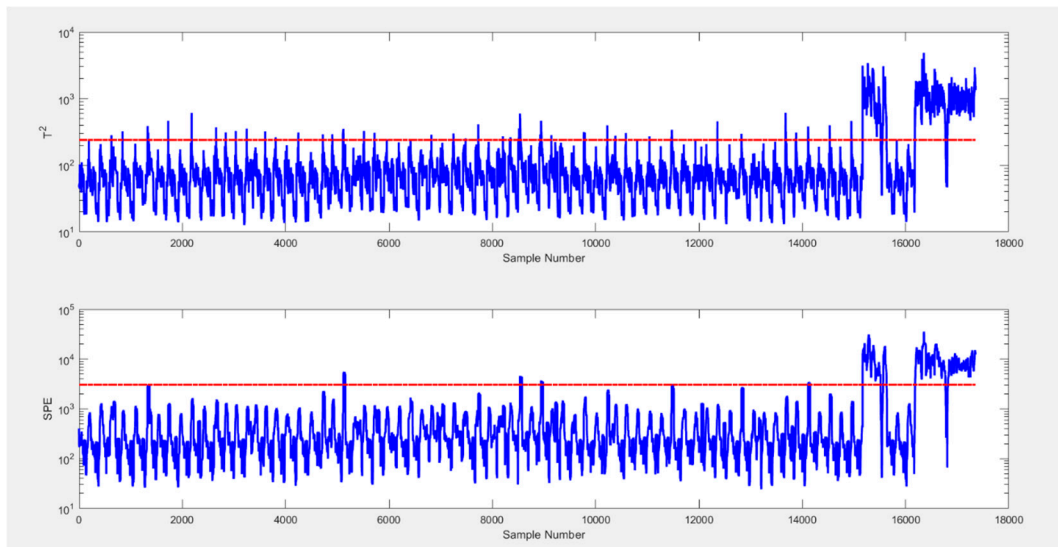**Figure 10.** $T^2$ and $Q$ for period $E$.

**Table 3.** Explanation of healthy data for the period from A to E.

| Period | Healthy Instances |
| --- | --- |
| From 2 to 12 January | 4824 |
| From 11 to 21 February | 5456 |
| From 4 to 15 May | 5774 |
| From 1 to 22 June | 9151 |
| From 29 July to 17 September | 15,207 |

Subsequently, the data present were representative of the alternating states of failure and states of health, which is commented on later.

The results obtained for each period considered are summarized in Table 4. The results in the table show the accuracy of the fault detection performed. The percentages show the results of the sum of instances labelled as faulty when they represent a stage of health of the machinery, and those

labelled as representative of a state of health are instead represented a failure. It could therefore be concluded that the percentage of error, for every period, was calculated as:

$$1 - \frac{correctly\ labeled\ instances}{total\ instances\ labeled}$$

**Table 4.** Percentage of error for every considered period.

| Period | Percentage of Error |
|---|---|
| From 2 to 12 January | 7.5% |
| From 11 to 21 February | 1.5% |
| From 4 to 15 May | 2% |
| From 1 to 22 June | 3.9% |
| From 29 July to 17 September | 0.7% |
| AVERAGE | 3.12% |

The total error rate for the fault detection process was the same both for $T^2$ and $Q$, $QQ$ although it was possible to say that $Q$ was slightly less affected by false alarms and $T^2$ under no circumstances could fail to identify a fault situation. Precisely, $T^2$ compensated for $Q$ by not identifying a fault in the same way that $Q$ managed to compensate for some false alarms $T^2$. These performance characteristics of the two indices permitted the authors to state that the percentage of error was essentially all due to the cases of false alarms rather than due to non-identification of dangerous situations. Regarding false alarms, no specific circumstances or characteristics were identified at this time that would identify a false alarm. This made it possible to state that critical situations of the machinery, which could not actually be labeled as a fault, tended to be managed in a precautionary manner by the model, erroneously identifying them as a fault. This made it possible to assume that the correct management of false alarms could allow precautionary maintenance, making it possible to further reduce the occurrence of faults.

## 4. Discussion and Conclusions

The study proposed the application of CVA to identify faults on a granulator used in pharmaceutical productions. The peculiarities of this production process were as follows:

- The same machine sequentially delivered 7 different production phases, without there being evidence of the transition from one phase to another.
- Each phase had specific characteristics and different failure situations.
- Production was not delivered continuously, but only in certain periods.

Furthermore, as already mentioned in the introduction, chemical processes were highly time-dependent, concerning the identification of a fault and the consequent analysis of its performance.

The need to implement a technique through which to manage these characteristics of the process simultaneously was mentioned. These characteristics and considerations led us to decide to apply CVA to identify faults, to verify its applicability and the results that can be achieved with it, in such processes.

Therefore, the innovativeness of what was proposed, in addition to the considerations extracted, could be traced back mainly to two points:

- The very structure of the process itself, which was extremely particular and discontinuous, so that no studies were carried out in this respect regarding the monitoring of this type of process with CVA.
- The approach chosen for the management of discontinuities. Specifically, it was not considered appropriate, at this stage, to implement changes on CVA, but to propose a methodological approach to manage some non-linearities in the process. This also made it possible to evaluate the discontinuities that could actually influence the achievable performance and the ones that did not.

The results obtained highlight an important potential of CVA in such application contexts, as well as interesting reflections. The first consideration concerned the management of a production process that was not continuous over time. This discontinuity inhibited the possibility of application of CVA to the entire dataset, as shown in Figure 3, with the consequent need to apply a specific CVA for each production period.

In general, it was possible to confirm that CVA could manage the numerous non-linearities present in the process, as well as the different faults that could occur, due to production stoppages or different production phases. In the same context, the applicability of CVA and its efficiency in the context of batch processes was confirmed, allowing us to compensate and manage the problems related to this category of processes.

Going into the details of the process discontinuities, the first assumption that could be made was that any maintenance work, sensor changes, or machine modifications insert discontinuity elements in the process, which cause the need to re-train the fault detection system. To be clear, there is evidently a compromise in the results achievable for process monitoring, if the various production periods are analyzed globally. It is interesting to note how it goes from a total inability to manage faults to an almost perfect identification of them. It is possible to state, on the contrary, that the 7 different phases that make up the process, as well as different faults that can occur, do not in any way compromise the performance of the model. As previously mentioned, in fact, for the application of the CVA for fault detection, these two pieces of information are neglected. This is because in many multiphase processes, it is not easy to define the process phase in running, with the consequent need for a substantial data pre-processing phase to do so. As a first consideration, there is a non-negligible lengthening of the monitoring process, with a slowdown in the timeliness of fault detection. In addition, somehow binding the fault detection to a labeling step in the pre-processing of the data can lead to errors, due to incorrect execution of the recognition of production phases. To obtain a monitoring model that does not require this labeling, according to the authors, reduces the range of possible errors of evaluation. With regards to identification or diagnosis of the phase of the fault, the authors believe that, for the same reasons as that explained for the identification of phases, it is useful to postpone this phase to the detection phase of a fault state. In fact, it is believed that a condition-monitoring model that is not specific to one type of failure, but that is generic to the monitored process, is more streamlined and, above all, less prone to errors. As is clarified later, future research related to what is presented in this paper focuses on the definition of a fault identification model. This model aims to circumscribe the root causes of a fault, addressing specific maintenance interventions.

Second, interesting considerations on the sensitivity of the two indices for false alarms and errors to detect faults were also reached. As evidenced from Figures 6–10, the $Q$ index had fewer false alarms than the $T^2$ index, thereby, managing to reduce the percentage of error due to this situation. Similarly, $T^2$ compensated for the inability to identify faults in certain situations by $Q$—the first index was, in fact, able to correctly detect all fault situations, effectively canceling the errors of $Q$. It was, therefore, clear that in this type of process, a weighted and correct use of both indices was fundamental, to obtain the best achievable error rate. It could be concluded that in such a time-varying process, monitoring could not only be based on state-space, but also needed residual-space analysis to obtain more valuable results. In conclusion, it was, therefore, necessary to analyze both indices critically, based on the considerations extracted from this case study, to draw more structured conclusions on the status of the process.

Obviously, there were limitations on what was proposed in this research. First, it would be interesting to apply what was proposed in another production process, using the same characteristics as the production process examined. This would further strengthen what was proposed, highlighting further strengths, but also critical implementation issues. Furthermore, the parameters considered in this model could not be considered universal. Even if applied to the same type of production process, the proposed methodology still required calculating specific parameters of the process in question. However, this did not affect its applicability. In fact, parameter setting was not a complex procedure,

allowing the model to not only be easily reproducible in other processes, but also that the changes to the existing models did not require a lot of time and resources. Clearly, like all process monitoring procedures, their performance was closely related to the quality of the data extrapolation process, i.e., the quality of the sensors considered. This required periodic interventions to evaluate the status of the sensors, to identify any problems at an early stage, without compromising the monitoring evaluations. The ease of application of the methodology also lay in its being totally free from data pre-processing steps. This made it possible to automatically apply it, immediately during production, with the possibility of interpreting the results directly from the line operator, further reducing evaluation and intervention times.

The results obtained made it possible to investigate several future scenarios. With regards to diagnosis, the CVA and the two indices considered could be used to identify the source of the fault and circumscribe the area of intervention. This rendered the execution of maintenance work faster and more detailed. What you want to achieve, following the identification of a fault, is a model that can evaluate the link between the monitored variables and the health indices considered. This made it possible to define the root causes of a fault, thus making it possible to promptly define the component of the machinery that was currently unhealthy. Through this, quickly and with full knowledge of the facts, it was possible to manage and schedule targeted maintenance interventions.

Additionally, it was possible to use the results obtained in the fault identification phase and those extractable for diagnosis, to implement a prognostic analysis through which to predict and manage the useful life of a machine or its time to failure. The objective was therefore to implement an initial stage of fault detection, through which to obtain information on the dynamics that regulate the monitored process. This meant the extraction of the most representative variables of the state of the machinery, through which to develop prognosis models. This would make it possible to move from a maintenance logic of fault detection, and therefore, of intervention downstream of its onset, to an anticipatory maintenance approach. In fact, prognostic analyses made it possible to predict the onset of a fault, making maintenance a tool to avoid its onset and not to manage it. To do this, future research will focus on the analysis of the CVA application, together with machine-learning algorithms for predicting the useful life of the machinery.

## Abbreviations

In this section, all abbreviations used in this the paper are summarized.

| Abbreviation | Meaning |
| --- | --- |
| CVA | Canonical Variate Analysis |
| PCA | Principal Component Analysis |
| KDE | Kernel Density Estimation |
| Period A | From 2 to 12 January |
| Period B | From 11 to 21 February |
| Period C | From 4 to 15 May |
| Period D | From 1 to 22 June |
| Period E | From 29 July to 17 September |
| $p_1$ | Initializing |
| $p_2$ | Conditioning |
| $p_3$ | 1st Spray |
| $p_4$ | Heating |
| $p_5$ | 2nd Spray |
| $p_6$ | Drying |

| | |
|---|---|
| $p_7$ | Unloading |
| $x_1$ | Spray Percentage |
| $x_2$ | Air IN Flow |
| $x_3$ | Spray Flow |
| $x_4$ | Air Pressure Spray |
| $x_5$ | Microclimate Pressure |
| $x_6$ | Cleaning Pressure |
| $x_7$ | Air IN Temperature |
| $x_8$ | Washing Air Temperature |
| $x_9$ | Air OUT Temperature |
| $x_{10}$ | Product Temperature |
| $x_{11}$ | Cooling Temperature |
| $x_{12}$ | Absolute Air IN Humidity |
| $x_{13}$ | Relative Air IN Humidity |
| $x_{14}$ | Product Humidity |
| $x_{15}$ | Mill Speed |
| $p$ | Past samples |
| $f$ | Future samples |
| $q$ | Retained states |

## References

1. Jiang, B.; Huang, D.; Zhu, X.; Yang, F.; Braatz, R.D. Canonical variate analysis-based contributions for fault identification. *J. Process Control* **2015**, *26*, 17–25. [CrossRef]
2. Ko, S.J.; Lee, J.-H.; Kang, C.-Y.; Park, J.-B. Granulation development in batch-to-batch and continuous processes from a quality by design perspective. *J. Drug Deliv. Sci. Technol.* **2018**, *46*, 34–45. [CrossRef]
3. Nguyen, T.L.; Djeziri, M.; Ananou, B.; Ouladsine, M.; Pinaton, J. Fault prognosis for batch production based on percentile measure and gamma process: Application to semiconductor manufacturing. *J. Process Control* **2016**, *48*, 72–80. [CrossRef]
4. Yoo, C.; Lee, J.-M.; Vanrolleghem, P.; Lee, I.-B. On-line monitoring of batch processes using multiway independent component analysis. *Chemom. Intell. Lab. Syst.* **2004**, *71*, 151–163. [CrossRef]
5. Yao, Y.; Gao, F. Subspace identification for two-dimensional dynamic batch process statistical monitoring. *Chem. Eng. Sci.* **2008**, *63*, 3411–3418. [CrossRef]
6. Cao, Y.; Hu, Y.; Deng, X.; Tian, X. Quality-Relevant Batch Process Fault Detection Using a Multiway Multi-Subspace CVA Method. *IEEE Access* **2017**, *5*, 23256–23265. [CrossRef]
7. Wang, P.; Wang, T.; Zhang, L.; Qiao, H. Fault diagnosis of rotating machinery under time-varying speed based on order tracking and deep learning. *J. Vibroeng.* **2020**, *22*, 366–382. [CrossRef]
8. Athulya, K. Inter Turn Fault Diagnosis in Wound Rotor Induction Machine Using Wavelet Transform. In Proceedings of the 2018 International CET Conference on Control, Communication, and Computing (IC4), Kerala, India, 5–7 July 2018; Volume 8530887, pp. 22–27.
9. Luo, B.; Wang, H.; Liu, H.; Li, B.; Peng, F. Early Fault Detection of Machine Tools Based on Deep Learning and Dynamic Identification. *IEEE Trans. Ind. Electron.* **2019**, *66*, 509–518. [CrossRef]
10. Rato, T.J.; Blue, J.; Pinaton, J.; Reis, M.S. Translation-Invariant Multiscale Energy-Based PCA for Monitoring Batch Processes in Semiconductor Manufacturing. *IEEE Trans. Autom. Sci. Eng.* **2016**, *14*, 894–904. [CrossRef]
11. Kan, M.S.; Tan, A.C.; Mathew, J. A review on prognostic techniques for non-stationary and non-linear rotating systems. *Mech. Syst. Signal Process.* **2015**, *62*, 1–20. [CrossRef]
12. Li, X.; Yang, Y.; Bennett, I.; Mba, D. Condition monitoring of rotating machines under time-varying conditions based on adaptive canonical variate analysis. *Mech. Syst. Signal Process.* **2019**, *131*, 348–363. [CrossRef]
13. Shang, L.; Liu, J.; Zhang, Y. Recursive Fault Detection and Identification for Time-Varying Processes. *Ind. Eng. Chem. Res.* **2016**, *55*, 12149–12160. [CrossRef]
14. Jiang, B.; Zhu, X.; Huang, D.; Paulson, J.A.; Braatz, R.D. A combined canonical variate analysis and Fisher discriminant analysis (CVA–FDA) approach for fault diagnosis. *Comput. Chem. Eng.* **2015**, *77*, 1–9. [CrossRef]
15. Lu, Q.; Jiang, B.; Gopaluni, R.B.; Loewen, P.; Braatz, R.D. Locality preserving discriminative canonical variate analysis for fault diagnosis. *Comput. Chem. Eng.* **2018**, *117*, 309–319. [CrossRef]

16. Shang, L.; Liu, J.; Zhang, Y.; Wang, G. Efficient recursive canonical variate analysis approach for monitoring time-varying processes. *J. Chemom.* **2016**, *31*, e2858. [CrossRef]
17. Yin, S.; Ding, S.X.; Xie, X.; Luo, H. A Review on Basic Data-Driven Approaches for Industrial Process Monitoring. *IEEE Trans. Ind. Electron.* **2014**, *61*, 6418–6428. [CrossRef]
18. Ma, L.; Dong, J.; Peng, K. A Novel Hierarchical Detection and Isolation Framework for Quality-Related Multiple Faults in Large-Scale Processes. *IEEE Trans. Ind. Electron.* **2019**, *67*, 1316–1327. [CrossRef]
19. Zhu, B.; Xu, Y.; He, Y.; Zhu, Q. Canonical Variate Analysis Based Regression for Monitoring of Process Correlation Structure. In Proceedings of the 2019 Chinese Automation Congress (CAC), Hangzhou, China, 22–24 November 2019; Volume 8997374, pp. 1328–1333.
20. Li, X.; Yang, X.; Yang, Y.; Bennett, I.; Mba, D. A novel diagnostic and prognostic framework for incipient fault detection and remaining service life prediction with application to industrial rotating machines. *Appl. Soft Comput.* **2019**, *82*, 105564. [CrossRef]
21. Russell, E.L.; Chiang, L.H.; Braatz, R.D. Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis. *Chemom. Intell. Lab. Syst.* **2000**, *51*, 81–93. [CrossRef]
22. Li, X.; Yang, X.; Yang, Y.; Bennett, I.; Collop, A.; Mba, D. Canonical variate residuals-based contribution map for slowly evolving faults. *J. Process Control* **2019**, *76*, 87–97. [CrossRef]
23. Jiang, B.; Braatz, R.D. Fault detection of process correlation structure using canonical variate analysis-based correlation features. *J. Process Control* **2017**, *58*, 131–138. [CrossRef]
24. Simoglou, A.; Martin, E.; Morris, A. Statistical performance monitoring of dynamic multivariate processes using state space modelling. *Comput. Chem. Eng.* **2002**, *26*, 909–920. [CrossRef]
25. Alizadeh, R.; Allen, J.K.; Mistree, F. Managing computational complexity using surrogate models: A critical review. *Res. Eng. Des.* **2020**, *31*, 275–298. [CrossRef]
26. Carcel, C.R.; Cao, Y.; Mba, D.; Lao, L.; Samuel, R. Statistical process monitoring of a multiphase flow facility. *Control. Eng. Pract.* **2015**, *42*, 74–88. [CrossRef]
27. Odiowei, P.-E.; Cao, Y. Nonlinear Dynamic Process Monitoring Using Canonical Variate Analysis and Kernel Density Estimations. *IEEE Trans. Ind. Inform.* **2009**, *6*, 36–45. [CrossRef]
28. Li, X.; Duan, F.; Loukopoulos, P.; Bennett, I.; Mba, D. Canonical variable analysis and long short-term memory for fault diagnosis and performance estimation of a centrifugal compressor. *Control Eng. Pract.* **2018**, *72*, 177–191. [CrossRef]
29. Samuel, R.T.; Cao, Y. Kernel Canonical Variate Analysis for Nonlinear Dynamic Process Monitoring. *IFAC-PapersOnLine* **2015**, *48*, 605–610. [CrossRef]
30. Hotelling, H. Relations between two sets of variates. *Biometrika* **1936**, *28*, 321–377. [CrossRef]