Article

Fault Detection, Diagnosis, and Prognosis of a Process Operating under Time-Varying Conditions

Elena Quatrini [1], Francesco Costantino [1], Xiaochuan Li [2] and David Mba [3,*]

1   Department of Mechanical and Aerospace Engineering, Sapienza University of Rome, Via Eudossiana, 18, 00184 Rome, Italy; elena.quatrini@uniroma1.it (E.Q.); francesco.costantino@uniroma1.it (F.C.)
2   School of Electrical Engineering and Automation, Hefei University of Technology, Hefei 230002, China; xiaochuan.li0309@gmail.com
3   Institute of Creative Computing, University of the Arts London, London WC1V 7EY, UK
*   Correspondence: d.mba@arts.ac.uk

Abstract: In the industrial panorama, many processes operate under time-varying conditions. Adapting high-performance diagnostic techniques under these relatively more complex situations is urgently needed to mitigate the risk of false alarms. Attention is being paid to fault anticipation, requiring an in-depth study of prediction techniques. Predicting remaining life before the occurrence of faults allows for a comprehensive maintenance management protocol and facilitates the wear management of the machine, avoiding faults that could permanently compromise the integrity of such machinery. This study focuses on canonical variate analysis for fault detection in processes operating under time-varying conditions and on its contribution to the diagnostic and prognostic analysis, the latter of which was performed with machine learning techniques. The approach was validated on actual datasets from a granulator operating in the pharmaceutical sector.

Keywords: residual useful life prediction; contribution plot; performance estimation; prognosis; diagnosis; fault detection

1. Introduction

The increasing complexity of production systems triggers the need for condition monitoring (CM) and condition-based maintenance (CBM) techniques that can be adapted to suit specific requirements. This paper discusses a project aimed at optimizing the application of CBM in industrial settings that operate under time-varying conditions, i.e., processes whose standards change over time. Such processes are widespread in the current industrial scenario, e.g., inconsistent characteristics of the processed product, unfixed quantities of processed products throughout different production cycles, multiphase processes where different operations are carried out by a single machine without a clear shift between phases. Existing CM and CBM techniques are often limited to fixed operating conditions, and their implementation in real settings is very challenging. Consequently, changes that are inherent in the production process can often be mistaken for fault situations, thus causing false alarms. Today, companies and researchers are ever more frequently focusing on machine prognosis [1] in order to obtain a sufficiently large forecasting period before a fault onset, so as to allow the execution of maintenance tasks aimed at preventing its occurrence. This research investigates the application of multivariate statistical techniques, focusing on canonical variate analysis (CVA) alongside with machine learning algorithms for fault detection, diagnosis, and prognosis in time-varying conditions processes.

Ultimately, this research aims to define a maintenance model that incorporates the main maintenance steps and areas, i.e., fault detection, diagnosis, and prognosis. The integrated and synergistic management of all these areas not only enhances performance but also ensures a simpler and more streamlined management. Furthermore, a highly debated aspect in the maintenance field is that of the decision-making process. The importance

of decision making and its correlation with fault detection and prognosis is noticeably clear, and, for this reason, it should be always taken into account [2,3]. However, while it certainly plays an increasingly important role in achieving effective maintenance systems, the definition of a structured decision-making process is particularly challenging. In this regard, the authors believe that the contribution plots that increase the quality of the prognosis analysis can be of invaluable support to the decision-making process for maintenance activities. The role of contribution plots is particularly relevant, as it allows obtaining effective and highly robust models. The decision-making process has significant safety implications as well, especially in processes that are highly hazardous both for the plant and the operator [1,4].

The research questions to be investigated, as the focus of this paper, are:

RQ1: is it possible to define a comprehensive methodology for the joint management of fault detection, diagnosis, and prognosis, setting the focus of the analysis on the information that can be extracted from the application of CVA?

RQ2: is it possible to implement an effective selection of the process variables to be considered for the prognosis analysis, by analyzing the relationship between these variables and the CVA health indexes?

A fault detection analysis on data gathered during different production intervals is not effective, as varying conditions could change the variables that provide prognosis information. Nevertheless, it is possible to carry out the selection of variables for prognosis analyzing only one production interval. This is because the correlation logic between process variables and degradation of the machinery remains unchanged between various production intervals. Obviously, for subsequent prognosis models, it remains necessary, currently, to train a model for each subinterval. So, this paper focuses on the study of the prognosis model, and it is a continuation of a previous study regarding the applicability of CVA for fault detection in processes operating under time-varying conditions, with the focus on the process analyzed in this research [5]. The remainder of the paper is structured as follows. Section 2 presents the related literature, Section 3 discusses the methodology, and Section 4 outlines the case study, with particular attention to the prognosis model, and the obtained results. The discussion of the obtained results, the conclusions, and the suggestions for future research are summarized in Section 5.

2. Literature Review

The complexity of modern engineering systems and manufacturing processes is rapidly increasing, and their reliability management becomes more and more challenging in modern dynamic operational settings. In this context, the research area of CBM data management is still immature, with many open challenges at different levels, due to the fact that several CBM approaches are data driven [6] (Wiggelinkhuizen et al. 2008). Fault detection and diagnosis have always been the focus of analysis to maintain adequate production standards, with multivariate statistical process control proving to be an excellent tool to support these analyses [7] (Zhang et al. 2021). In recent years, the processes operating under time-varying conditions, which are at the center of this research proposal, became of great interest in the field of fault detection and maintenance [8–11]. In these processes, the assumption of the linearity and static nature of the underlying processes represents a risk in the application of multivariate statistical techniques. For this reason, there are currently some limitations for CVA in such processes, with the consequent need, for example, to implement changes to the traditional CVA. If properly managed or modified, CVA proves to be high performing in this context [12,13].

It should be noted that CVA has achieved excellent results in many maintenance applications, both in its original or modified forms [14–16]. The fundamental characteristic of CVA is that it calculates linear combinations of the past and future values of the system to maximize the correlation between them [17]. It is crucial to go beyond the assumption that process variables are time independent. The dependence of process variables on time is particularly relevant in processes that show a significant correlation between past and

future instances, such as the ones used in the chemical sector. The idea of the variables' time-independence is at the heart of many widely applied multivariate statistics techniques for fault detection, such as principal component analysis. Consequently, unlike other multivariate statistical techniques, CVA allows consideration of the time dependence of variables during process monitoring, thus making this analysis extremely suitable for fault detection in many contexts where the assumption of independence of processes from time is not valid [18].

From the available literature, it was noted that topics concerned with residual useful life and time to failure prediction are largely covered in research areas that involve aspects of the degradation processes affecting a system. Similarly, the issue of predicting the remaining useful life of equipment was found to be associated with research on prioritizing maintenance activities [19]. Since different degradation processes require different approaches [20–23], a comprehensive review of these processes is fundamental to this research, and future research should aim at developing more accurate models. In this context, the use of machine learning is still an open topic, especially as far as residual useful life and time to failure prediction are concerned. Examples are represented by the application of support vector machine [24,25], random forests [26] or neural networks [27,28], and incremental learning [29]. In the literature, the use of CVA is described in combination with other techniques to predict the trend of degradation and the behavior of a system after the onset of a failure [30,31]. However, in practice, prognosis analysis is more often applied after a failure to monitor its progress. Considering the prediction of residual useful life in support of the decision-making process, it can be said that the understanding of the complex interactions between operating conditions and component capability is crucial to estimate the degradation of a component and predict its remaining useful life, thus allowing decision making that directly impact the results [32], even financially [33]. Predicting the health status trend of a piece of machinery allows expansion of the topic and context of decision making. It presents and reveals the possibility of adapting decision making, evaluating it in different time horizons [34].

In general, according to the existing literature, it is possible to state that extremely time-variable processes, regardless of the selected technique, are difficult to understand, and the prognostic analysis implemented is clearly affected [35]. This issue is still considered an open point in the literature, since many contributions and approaches focus only on the analysis of static operating conditions [36]. So, considering the management of machinery degradation, a lot of research is emerging and developing, all supported by an evident increase in the quality of sensors and monitoring techniques of machinery conditions. In general, however, some limitations in this regard are still evident and to be managed. One of these is the impossibility, in some contexts, to make run-to-failure machines operate, resulting in the lack of a sufficiently large historical failure dataset for subsequent analysis. This relates, then, to the need to focus prognosis analyses on specific variables that have greater predictive power with respect to degradation trends, so that accurate prognosis analyses can be achieved, reducing the inclusion of unnecessary information in the model [37].

In conclusion, the analyzed literature reveals the points of interest and innovation of the proposed contribution. First of all, the topic of the prognosis and the residual useful life of the machinery is still the starting point and the focus of extensive and heated debates, with a growing need for new contributions to this matter. Secondly, the type of process analyzed, i.e., an operating process under varying conditions over time, has many interesting facets but also numerous difficulties in the management of maintenance and data analysis, with greater repercussions on the topic of prognosis. This type of process, given the countless implicit difficulties, is currently less debated and has obtained fewer results than more linear and static processes. Ultimately, a final interesting point concerns the considered technique, CVA, and the approach proposed. On one hand, the use of CVA in this type of process is innovative because, as previously mentioned, multivariate statistics techniques still have limitations in such contexts. On the other hand, the proposed

approach is important for its dual role in fault detection and in the prognosis in conjunction with contribution plots and machine learning. Considering the results obtained in the previous contribution, the approach is innovative and even leaner than those that suggest intrinsic changes to CVA.

The main contributions of this study to the research community are summarized as follows:

(1)　This study applies state-of-the-art process control and machine learning techniques to a granulator operating in the pharmaceutical sector, with the developed models being verified with real-world data, thus offering a practical solution to the monitoring of the granulation process.

(2)　Unlike existing studies where fault prognosis was carried out relatively independent from fault detection, this study evaluates the influence of fault detection on prognosis by adding a connecting step between the two. The results show that the CVA-based contribution analysis can facilitate fault prognosis.

3. Methodology

This section briefly introduces the proposed methodology. Let $Y_N = [y_1, y_2, \ldots, y_N]$ be the training set with N number of observations obtained from the machine; by applying $p_k = [y_{t-p+1}, y_{t-p+2}, \ldots, y_t]^T$ and $f_k = [y_t, y_{t+1}, \ldots, y_{t+f-1}]^T$ at each time instance k of Y_N, one can form the past and future matrices:

$$Y_p = \left[p_p, p_{p+1}, \ldots, p_{N-f+1} \right]$$

$$Y_f = \left[f_p, f_{p+1}, \ldots, f_{N-f+1} \right]$$

Then, the sample covariance and cross-covariance of the past and future matrices can be estimated as

$$\Sigma_{pp} = \frac{1}{N-p-f+1} Y_p Y_p^T$$

$$\Sigma_{ff} = \frac{1}{N-p-f+1} Y_f Y_f^T$$

$$\Sigma_{ff} = \frac{1}{N-p-f+1} Y_f Y_f^T$$

The goal of CVA is to find a set of linear combinations between the past and future matrices such that the correlations between the two groups of linear combinations are maximized. To do so, one can perform a singular value decomposition as follows:

$$H = \Sigma_{ff}^{-1/2} \Sigma_{fp} \Sigma_{pp}^{-1/2} = U\Sigma V^T$$

Then, two monitoring indices T^2 and Q can be derived as

$$T^2 = z_t z_t^T, \; z_t = V^T \Sigma_{pp}^{-1/2} p_t$$

$$Q = z_t z_t^T, \; z_t = V^T \Sigma_{pp}^{-1/2} p_t \tag{1}$$

When new data are available, one can calculate the monitoring indices T^2 and Q, and compare their values to the corresponding failure threshold for fault detection.

The details of the application of the methodology are presented in Section 4, with a step by step description of the approach and related insights presented in Figure 1. For more details about CVA, see [38].
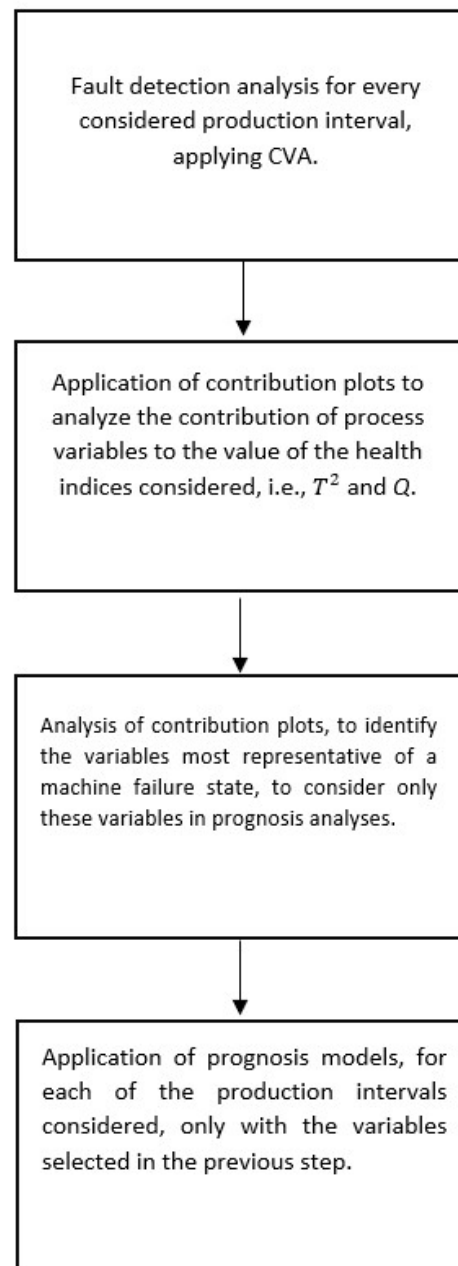
Figure 1. Methodology for the proposed model.

Focusing attention for a moment on the first step of the methodology, it is not part of this article. This step has been detailed in another article [5]. As anticipated, the role of this step is twofold: on the one hand, implementing a fault detection analysis provides the ability to identify anomalies that cannot be identified with prognosis analysis. The purpose is then to support prognosis analysis with real-time data analysis, with the aim of making joint maintenance management more robust. On the other hand, the health indices calculated in this step are the input for the subsequent variable selection analyses for prognosis analyses.

4. Case Study

The study presented in this paper can be defined as the continuation of research proposed in a previous publication [5]. In the first stages of the prior research, a methodology for managing the non-linearities of a process was proposed to allow its monitoring by using a well-known multivariate statistics technique, i.e., CVA. As discussed in the introductory

part of the paper, countless monitoring challenges arise in the processes that operate under time-varying conditions, and, for this reason, in-depth research is necessary. Considering the level of detail with which the fault detection application is described in the previous contribution, the authors refer to that paper for further description of the implementation [5]. As previously mentioned, fault detection has a double role in this research:

- Input for the prognostic and decision-making step.
- Continuous process control to compensate for errors and inaccuracies in the prognostic model.

The production process considered in this contribution is a dataset from a real-world industrial plant. This ensures that the case study is relevant for research. As asserted before, the distinguishing features of the considered production process are the management and the structure of the production process itself. The process does not operate continuously, but it only runs in specific time windows, defined by managerial choices or necessary maintenance interventions. The production intervals taken under analysis are:

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

Another feature that exacerbates the process monitoring complexity is its division into seven sub-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.

There are 13 monitored parameters, and every data point $d = \{x_1, \dots, x_{13}\}$ represents a set of 13 measures, 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.

The results obtained in the first phase of the project paved the way for the inclusion of an additional section, with the aim to create a complete methodology that adds more value to the implemented maintenance process. Currently, the previous research focuses only on fault detection, i.e., real-time monitoring of the process to identify trends that are not consistent with what is expected, and that are consequently identified as faults. However, as mentioned at the beginning of the paper, in this study greater attention was paid to the decision-making process. To achieve this goal, the methodology was divided into two subphases:

- A phase of process structure analysis in a state of failure to circumscribe its most indicative and representative variables;
- Prognostic analysis of the failure based on information obtained in the previous phase to predict the machinery's useful life by monitoring only a subset of variables. This allows for streamlining of the analysis process and to eliminate any noise in the model caused by variables that are not representative of the failure.

Contribution plots were used to analyze the role of the variables related to failure. The role of a contribution plot is to show the contribution of each process variable to the calculated statistic, in this case, T^2 and Q. In this way, when the statistic is out of control, it is possible to identify which variable caused the anomaly. This analysis allows circumscribing the variables that are relevant in the process monitoring, which are the focal point of the following step of the prognosis. This allows not only streamlining of the analysis by speeding up its execution but also elimination of unnecessary variables that can create noise in the machine learning model leading to the achievable results. This step can be attributed a double role. The prognostic analysis of the process allows timely prediction of the time to failure, thus ensuring the implementation of maintenance interventions with the aim of preventing the occurrence of a fault or limiting its extent. Additionally, this analysis is always accompanied by a continuous and real-time monitoring of the process to implement a possible precautionary fault detection, with the consequent compensation for any possible errors in prognostic prediction. The analysis of the variables representative of the fault ensures faster circumscription of the origin of the fault, and a consequent implementation of rapid maintenance interventions, thus limiting the duration of unwanted machine downtimes. In conclusion, it is possible to state that this phase clearly plays a key role in the maintenance decision-making process. In a state of emergency, it steers maintenance inspections, it predicts the machinery's failure time, and it identifies a subgroup of variables that are representative of the failure.

For the case study under analysis, the logic analysis used in this phase differed from the one used in the fault detection phase. Despite the difficulties encountered in the latter due to the non-linearities of the process, the fault detection phase did not influence the application of the contribution plots. While the health thresholds and the relationship between health status and failure change in different production intervals, the logic inherent in the process remains unchanged. For this reason, the application of the contribution plots, aimed at skimming the variables to be input to the prognosis model, was performed only on a randomly selected production sub-interval, i.e., interval E. In addition, both the contribution plots of T^2 and Q were considered. However, to extrapolate the representability of the variables, it was deemed sufficient to consider essentially only the contribution of T^2, which represents the state space. As regards the evaluation of the contribution plot of Q, it is dependent on the considerations extracted from that of T^2. The results obtained from the application of the contribution plots are shown in Figures 2 and 3, which display only data related to a fault state.

As can be seen in Figures 2 and 3, a contour plot was applied, i.e., a plot containing the isolines of a Z matrix, with Z containing height values on the *x-y* plane. The term isolines is widely used to describe images representing different levels of a certain value. Isolines are "contour lines", i.e., lines joining all the points presenting the same value of the chosen variable. In the case under analysis, the contour lines represent the value of the influence that a variable had on the value of the index considered in the specific graph. In this case, the *x-y* plane is structured as follows:

- *y*-axis → the process variables;
- *x*-axis → dataset instances.

Furthermore, Z represents the value of the contribution of a process variable to the value of the statistic, and like-colored lines represent an area with the same value of this contribution. It should be noted that, to improve the readability of the obtained results, the logarithm of the calculated contributions was graphed, thus helping to respond to skewness towards larger values. The colored bars on the right side of the graphs represent

the scale of the magnitude of the influence of a variable on the T^2 value and on the Q value, presented in Figures 2 and 3, respectively.
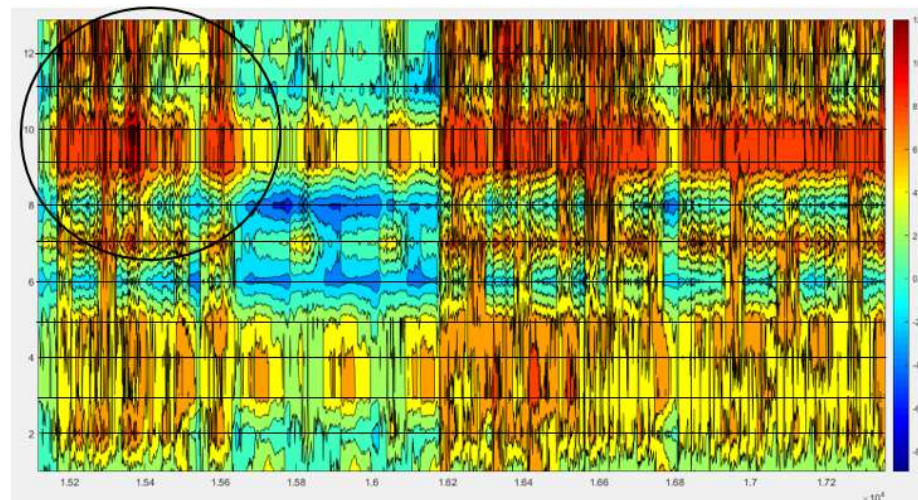

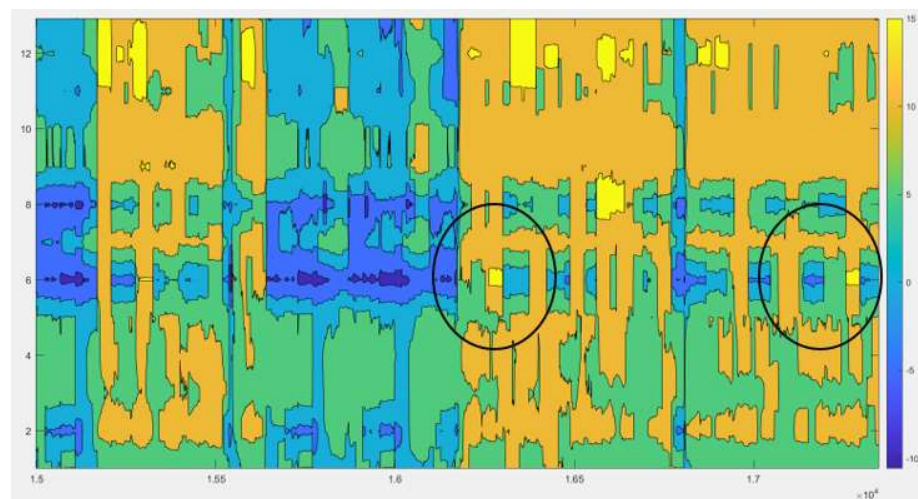**Figure 2.** Contribution plot of T^2 in faulty states.

Figure 3. Contribution plot of Q in faulty states.

After completion of the contribution plots analysis, six variables were selected for the prognostics model: x_2, x_6, x_8, x_{10}, x_{12}, and x_{13}. The selection of the relevant variables started from the analysis of the contribution plot results, and it was also based on several considerations made about the process and about the relationships between the variables. Clearly, both the x-axis and the y-axis move on discrete values. The y-axis can only take on integer values representing the variables, and the x-axis represents a time sequence at discrete intervals. What makes the view of the graph as if it were completely active is that it represents a 2D view of a matrix of three values, thus actually representing 3D space. By considering a 2D view of the isolines and choosing a "full" representation of the graph, i.e., one in which the inside of the areas delimited by the isolines are colored, it is possible to obtain the above-described view of a completely full graph. To facilitate the reading of the graphs, the authors have inserted horizontal lines in Figure 2 which represent the active zones of the graph for each of the considered variables. For convenience, in this paper such lines have only been included in Figure 2, but the considerations made so far apply to the contribution plot of Q (Figure 3) as well. Going into the details of the contribution plot analysis, it is necessary to state that the authors started by analyzing only the contribution plot for T^2.

During this analysis, the authors identified the areas where the lines added to the graph, i.e., the ones representing active areas of the graph for the considered process variables, intercepted one or more dark red areas. The colored bar shows that this is the color representing the greatest influence of a variable on the value of T^2, i.e., the first color from the top of the colored bar. However, in case of the absence of dark red-colored areas, the second level of intensity would have been used. The output of this first round of analysis selected five variables, specifically x_2, x_8, x_{10}, x_{12}, and x_{13}. The circles added within the graph in correspondence with the dark-red zones represent areas of high influence of variables on the T^2 metric. It is interesting to note that two variables comprising several areas with considerable intensity of influence on the T^2 value, but not the greatest intensity defined, are closely correlated with two of the selected variables, precisely x_9 and x_7. The process variable x_9 is linearly correlated with x_{10}, and x_7 is linearly correlated with x_{13}. Following this analysis of T^2's contribution plot, considerations were made about Q's role.

Considering the insights extracted from the previous work [5], we proceeded in the following way:

(1) The variables showing the strongest relationship with Q were identified.
(2) We analyzed which of these variables were activated in a disjointed manner with respect to variables already identified with the T^2 analysis.
(3) As a result of this analysis, variable x_6 was further selected.

As can be noticed from the circled areas in Figure 2, in no less than two areas variable x_6 proves to be the most associated variable with the value of Q, while none of the variables selected with the analysis of T^2 were activated in the same way. The considerations made for the analysis of T^2 apply for the analysis of the contribution plot of Q as well. The only difference lays in the chosen color bar. For the analysis of the contribution plot of Q, the color representing a greater influence of the variable on the metrics is yellow. For the sake of completeness, it should be noted that the authors implemented a test for the prediction of the remaining useful life of machinery without the inclusion of the last variable considered, i.e., x_6, and the obtained the results were not acceptable. According to the information obtained using contribution plots, i.e., the six variables selected, five prognosis models were developed, one for each sub-interval considered. While for the extraction phase of the process trend's most representative variables the sub-intervals were considered to have common characteristics, for the prognosis phase, the subgroups were again considered as separate, in light of the experience acquired through fault detection. To implement this phase, regression machine learning algorithms were applied, specifically approaching the learning ensemble. As far as statistics and machine learning are concerned, ensemble methods use multiple learning algorithms to achieve better predictive performance instead of using constituent learning algorithms alone. The methods selected for this phase consisted of decision trees. In retrospect, after the analysis and after tests with other algorithms, it is possible to say that this category was the one that guaranteed the best results.

Following numerous tests and numerous optimizations of the algorithms, the superiority of the results achievable with the bagged trees was proven. Bagged trees have also been proven to be suitable for prognosis in the presence of datasets that are not considered to be large in size, making them appropriate for this study. In Bagging (Bootstrap Aggregation) several models of the same type are trained on different datasets deriving from an initial dataset obtained by random sampling with replacement, reducing the variance of a decision tree. Firstly, several subsets of data are created from a randomly chosen training sample with replacement. Then, each collection of subset data is used to train its decision trees. Finally, the obtained result is an ensemble of different models, and the average of all the predictions obtained from different trees is used. This process is more robust than a process that uses single decision trees since individual decision trees tend to overfit. Bootstrap-aggregated (bagged) decision trees combine the results of many decision trees, thus reducing the effects of overfitting and improving generalization. Every tree in the ensemble was grown on an independently drawn bootstrap replica of input data, and

the observations not included in this replica are "out of the bag". The number of features was selected randomly for each decision split, so it is possible to assert that the algorithm selected was the random forest algorithm [39]. The evaluation metrics considered were as follows:

- R^2: it is the coefficient of determination and represents the ratio of the variance in the dependent variable that is predictable from the independent variables.
- **MEAN SQUARED ERROR (MSE):** it is a frequently used measure of the differences between values predicted by a model and the values observed $\rightarrow \frac{\sum_{i=1}^{n}|y_i - y_i|}{n}$

The first one, i.e., R^2, provided an insight into the quality of the predisposed model, as it allowed understanding to what extent it represents reality. The second one, i.e., MSE, allowed giving greater weight to large mistakes rather than to smaller ones. This metric was extremely important in the analysis under consideration, since the need to penalize large errors was particularly relevant.

In conclusion, it should be noted that there was no parallelism between the results of the two phases, i.e., the subgroup with the best results for fault detection was not the best for the prognosis.

Prognosis Models

For the part concerning the prognosis, a model was developed for each of the considered sub-intervals, as previously conducted in the fault detection models [5]. Furthermore, algorithm optimization was implemented to define the algorithm parameters. This means that various combinations of parameters were tested in an automatic way to identify the one that would allow obtaining the best achievable results. The metric taken into consideration to select the best configuration was the MSE. In particular, the best setting was the one that guaranteed the lowest value. As well as in the previous phase, the prognosis models were developed using MATLAB. As previously mentioned, five bagged tree models with different combinations of parameter settings were trained and tested, and three parameters were optimized:

- Number of learners: Number of trees.
- Minimum leaf size: Minimum number of observations per tree leaf.
- Number of predictors to sample: Number of variables to be selected at random for each decision split.

The number of iterations performed for the optimization of each model was 50, since several experiments showed a flattening of the improvements. Table 1 presents the best settings for all the considered models:

Table 1. Best settings for the algorithm in every considered interval.

Interval	Number of Learners	Minimum Leaf Size	Number of Predictors to Sample	MSE
A	500	1	3	258.46
B	500	1	4	1504.6
C	500	1	5	463.42
D	500	1	5	433.46
E	85	1	6	843.12

The metrics selected for each interval were then used to predict the residual useful life of the machinery in the five sub-intervals under consideration. The results obtained are shown in Figures 4–6.
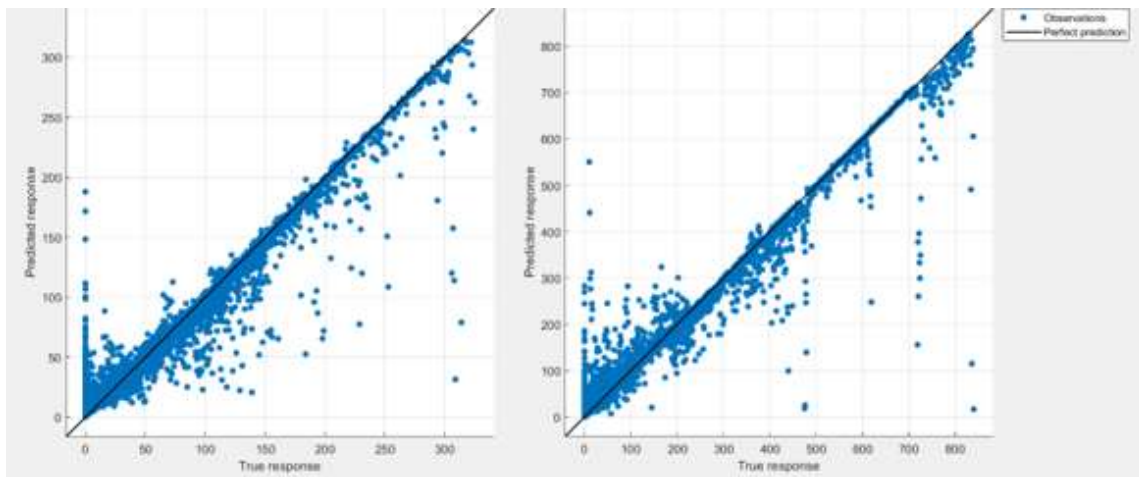
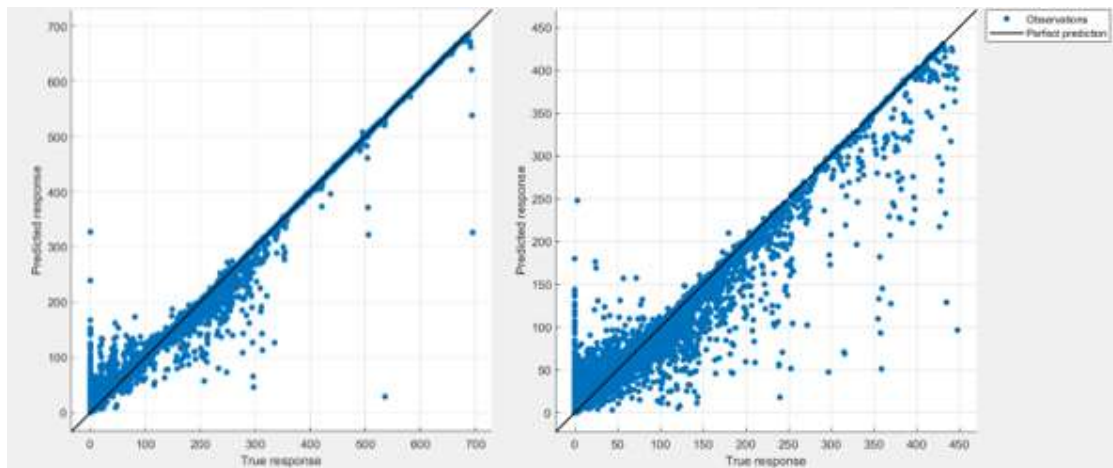**Figure 4.** Prediction trend for interval A (**left**) and B (**right**).

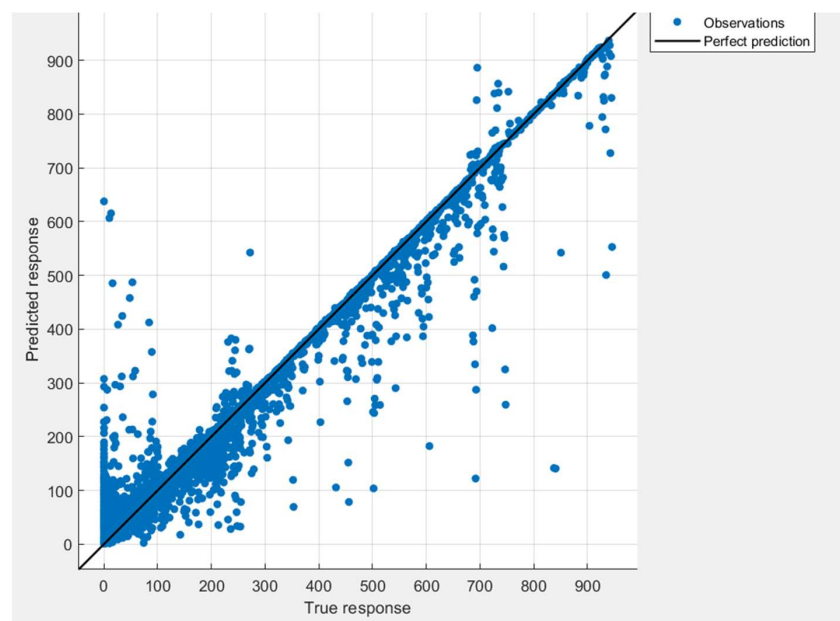**Figure 5.** Prediction trend for interval C (**left**) and D (**right**).

Figure 6. Prediction trend for interval E.

Figures 4–6 represent the trends of the prediction of the RUL for all the analyzed intervals. This highlights the relationship between the predicted value of the RUL on the *y*-axis and the actual value of the RUL, on the *x*-axis. To analyze the obtained results, in addition to the previously mentioned MSE value (see Table 1), the R^2 metric was considered. As previously described, this metric represents the extent to which the model can accurately represent actual operational data. The closer the value is to 1, the higher the model accuracy; on the contrary, values close to 0 represent an inaccurate model. The R^2 values of all the intervals were as follows:

- Interval A: 0.94;
- Interval B: 0.96;
- Interval C: 0.98;
- Interval D: 0.93;
- Interval E: 0.97.

It was noted that the algorithm was able to satisfactorily approximate the actual operations considered, demonstrating good predictions of the RUL. Although in some of the analyzed intervals the results were slightly lower than relative to other cases, the overall results were considered to produce substantially favorable outcomes. The lowest forecast value obtained was 0.93. However, the analysis of forecasting trends allowed extrapolation, strengthening the analysis of the results obtained. Evidence from Figures 4–6 reveals that the number of errors detected was low relative to the total number of cases examined and predicted. This observation offered important considerations on the combined use of fault detection, diagnosis, and prognosis, and it also allowed evaluations to manage forecasting errors. By looking at the presented data, it was observed that the errors reduced as the prediction time window increased.

For this purpose, due to the need to break the dataset into sub-intervals both in the fault detection phase and for the prognosis phase, it was not possible to define a univocal threshold for all the sub-intervals. For example, an optimal threshold for interval C could be 300, while the same threshold would be too large for interval A, where a 150 threshold would be sufficient. The definition of such thresholds within the forecasting models allowed a significant increase in the achievable results, which were already excellent. Furthermore, another consideration that added further value to the achieved results was the fact that most errors were caused by an underestimation of the residual useful life of the machinery. This approach could be defined as "precautionary", as it tends to direct towards early rather than late maintenance. The cases of underestimation of the forecast are extremely small, almost negligible. Finally, considering all the above analysis, it is possible to state that CVA successfully contributes to improving the performance of the residual useful life prediction model. Without the prior analysis of the dataset with CVA and the consequent application of the contribution plots, the achievable results would have been different. An example of the achievable results for the prognosis without CVA in interval C is shown in Figure 7.

After the optimization process, and without the prior application of CVA, the results achieved for interval C were the following:

- MSE: 470.19;
- R^2: 0.93.

As can be seen, there was an increase in the MSE, but the most striking result was the worsening of the R^2 index from 0.98 to 0.93. This last consideration was also evident in a graphic analysis, comparing Figures 5 and 7, where it is clear that the model was less able to predict and reproduce the trend of the residual useful life of the machinery. Even though the results achievable without the application of CVA were excellent, with the application of CVA it was possible to significantly improve the performance of the residual useful life prediction model of the machinery.
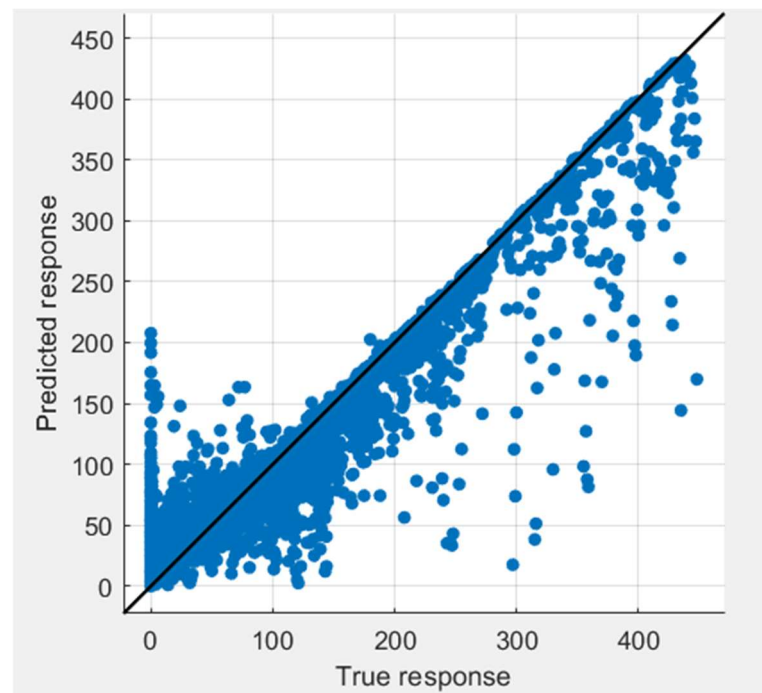
Figure 7. Prediction trend for interval C, without CVA optimization.

5. Discussion

Before discussing the results obtained in this research, it is useful to describe once again CVA's dual role in maintenance:

- The input of a prognosis model, which is used to predict the residual useful life of a machine, thus allowing action before the onset of a fault;
- Real-time fault detection, which is used to monitor a machine continuously, thus allowing identification of any unforeseen faults or, in general, to make up for any inconsistencies and errors in the prognosis model.

The effectiveness of CVA in fault detection in a process such as the one considered in this contribution was extensively discussed in a previous contribution [5]. The paper not only made it possible to verify the applicability of CVA in this category of processes, but it also laid the foundations for some important considerations. The most noteworthy consideration is the one referring to the structure of the process itself, which needs some special care to be successfully monitored with CVA. In line with what occurred with fault detection, temporal segmentation of the dataset was used for the prognosis as well, meaning that separate production intervals were considered as separate processes. The machine learning model for the prediction of the residual useful life of the machinery was structured with one model for each sub-interval.

As explained in the case study presented, the choice of the algorithm was the result of an iterative process. Firstly, after comparing the different results achievable with the different algorithms, decision trees were considered as the most appropriate choice. Subsequently, to define the specific parameters of the regression algorithm, an optimization process was chosen. The last step, based on the results of the optimization process, was to apply the regression algorithms again to the five sub-intervals to validate their effectiveness and extrapolate the results.

The results obtained, presented at the end of the case study, show promise of an improved performance. As expected, not all the analyzed intervals reached the same accuracy of forecast. It is also interesting to note that there was no consistency between the results achieved for fault detection and those achieved for the prognosis. To be clearer, the interval in which the best results were obtained for fault detection, for example, did not coincide with the interval in which the best results were obtained for the prognosis.

Furthermore, from the figures, it was noted that the obtained value could be further increased by making a few adjustments. Firstly, the trends showed that by increasing the forecast window, errors were reduced. Consequently, it is possible to state that with a well-managed condition-based and predictive maintenance policy that extends the life expectancy intervals, the results achievable with the presented model are undoubtedly better than those obtained overall. This result is consistent with a failure anticipation policy, rather than with a failure management policy. Forecasting the residual useful life of the machinery with little advance notice does not allow for the ability to implement precautionary maintenance interventions but only ensures better preparedness for its outbreak and its management. Conversely, the management of short time and unexpected events is left to fault detection models based on CVA, which ensures the real-time monitoring of the model. Moreover, the model tended to underestimate rather than overestimate the residual useful life of the machinery, thus allowing anticipation of maintenance operations. It is possible to avoid situations in which the overestimation of the residual useful life makes it necessary to resort to emergency maintenance operations identified through real-time monitoring. The underestimation of the residual useful life leads to a more cautious maintenance approach, thus remaining consistent with the desire for early intervention on machinery. Real-time monitoring based on CVA ensures reflectively better protection, and it allows for prompt identification of any errors in the determination of the prognosis model and unexpected failure.

As far as the contribution plots are concerned, they had a dual role. As described in the previous sections of this paper, they allowed isolation of the variables that were most representative of the fault trends, making it possible to obtain even more efficient and powerful models for the prediction of the machinery's residual useful life. Moreover, as far as emergency maintenance was concerned, i.e., the application of CVA for fault detection, the contribution plots ensured an even higher reactivity during fault identification. They allowed a quick demarcation of the variable(s) that caused the alert state. Consequently, by linking the variable(s) obtained from the contribution plots to one or more system components, it is possible to target maintenance work. In conclusion, the combination of CVA and contribution plots have shown invaluable application for prognosis analysis.

6. Conclusions

This study presented the use of canonical variate analysis applied to a granulator operating in the pharmaceutical sector for the purpose of fault detection. Moreover, to facilitate fault prognosis, CVA-based contribution plots were employed to identify key contributing variables for the detected fault, following which a machine learning model was built to predict the RUL of the machine under faulty conditions. The results showed that the CVA model successfully contributed to the improvement of the performance of the residual useful life prediction.

Author Contributions: Writing—original draft, E.Q., F.C., X.L. and D.M. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Quatrini, E.; Costantino, F.; Di Gravio, G.; Patriarca, R. Machine learning for anomaly detection and process phase classification to improve safety and maintenance activities. *J. Manuf. Syst.* **2020**, *56*, 117–132. [CrossRef]
2. Korbicz, J. Robust fault detection using analytical and soft computing methods. *Bull. Pol. Acad. Sci. Tech. Sci.* **2006**, *54*, 75–88.

3.   Shumsky, A. Redundancy relations for fault diagnosis in nonlinear uncertain systems. *Int. J. Appl. Math. Comput. Sci.* **2007**, *17*, 477–489. [CrossRef]

4.   Patriarca, R.; Falegnami, A.; De Nicola, A.; Villani, M.L.; Paltrinieri, N. Serious games for industrial safety: An approach for developing resilience early warning indicators. *Saf. Sci.* **2019**, *118*, 316–331. [CrossRef]

5.   Quatrini, E.; Li, X.; Mba, D.; Costantino, F. Fault diagnosis of a granulator operating under time-varying conditions using canonical variate analysis. *Energies* **2020**, *13*, 4427. [CrossRef]

6.   Wiggelinkhuizen, E.; Verbruggen, T.; Braam, H.; Rademakers, L.; Xiang, J.; Watson, S. Assessment of condition monitoring techniques for offshore wind farms. *J. Sol. Energy Eng.* **2008**, *130*, 0310041–0310049. [CrossRef]

7.   Zhang, C.; Yu, J.; Wang, S. Fault detection and recognition of multivariate process based on feature learning of one-dimensional convolutional neural network and stacked denoised autoencoder. *Int. J. Prod. Res.* **2021**, *59*, 2426–2449. [CrossRef]

8.   Chakour, C.; Hamza, A.; Elshenawy, L.M. Adaptive CIPCA-based fault diagnosis scheme for uncertain time-varying processes. *Neural Comput. Appl.* **2021**, *33*, 15413–15432. [CrossRef]

9.   Elshenawy, L.M.; Mahmoud, T.A. Fault diagnosis of time-varying processes using modified reconstruction-based contributions. *J. Process Control* **2018**, *70*, 12–23. [CrossRef]

10.   Liu, Y.; Zeng, J.; Xie, L.; Kruger, U.; Luo, S.; Su, H. Structured sequential Gaussian graphical models for monitoring time-varying process. *Control Eng. Pract.* **2019**, *91*, 104099. [CrossRef]

11.   Mou, W.; Jin, H.; Wang, H.; Dai, M.; Wang, J.; Zhao, C. Dissimilarity Analytics for Monitoring of Nonstationary Industrial Processes with Stationary Subspace Decomposition. In Proceedings of the 2020 Chinese Automation Congress, CAC 2020, Shanghai, China, 6–8 November 2020; pp. 651–656. [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.-L.; Liu, J.-C.; Tan, S.-B.; Wang, G.-Z. Recursive canonical variate analysis for fault detection of time-varying processes. *Dongbei Daxue Xuebao J. Northeast. Univ.* **2016**, *37*, 1673. [CrossRef]

14.   Lan, T.; Tong, C.; Shi, X.; Luo, L. Dynamic statistical process monitoring based on generalized canonical variate analysis. *J. Taiwan Inst. Chem. Eng.* **2020**, *112*, 78–86. [CrossRef]

15.   Shang, L.; Yan, Z.; Li, J.; Qiu, A.; Zhang, H. Canonical residual based incipient fault detection method for industrial process. In Proceedings of the 32nd Chinese Control and Decision Conference, CCDC 2020, Hefei, China, 22–24 August 2020; pp. 987–992. [CrossRef]

16.   Sun, D.; Gong, X.; Chen, Y. Integrating canonical variate analysis and kernel independent component analysis for tennessee eastman process monitoring. *J. Chem. Eng. Jpn.* **2020**, *53*, 126–133. [CrossRef]

17.   Han, X.; Jiang, J.; Xu, A.; Huang, X.; Pei, C.; Sun, Y. Fault Detection of Pneumatic Control Valves based on Canonical Variate Analysis. *IEEE Sens. J.* **2021**, *21*, 13603–13615. [CrossRef]

18.   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]

19.   Huo, J.; Zhang, J.; Chan, F.T.S. A fuzzy control system for assembly line balancing with a three-state degradation process in the era of Industry 4.0. *Int. J. Prod. Res.* **2020**, *58*, 7112–7129. [CrossRef]

20.   Chiachío, J.; Jalón, M.L.; Chiachío, M.; Kolios, A. A Markov chains prognostics framework for complex degradation processes. *Reliab. Eng. Syst. Saf.* **2020**, *195*, 106621. [CrossRef]

21.   Elsheikh, A.; Yacout, S.; Ouali, M.-S.; Shaban, Y. Failure time prediction using adaptive logical analysis of survival curves and multiple machining signals. *J. Intell. Manuf.* **2020**, *31*, 403–415. [CrossRef]

22.   Li, N.; Gebraeel, N.; Lei, Y.; Fang, X.; Cai, X.; Yan, T. Remaining useful life prediction based on a multi-sensor data fusion model. *Reliab. Eng. Syst. Saf.* **2021**, *208*, 107249. [CrossRef]

23.   Liu, H.; Song, W.; Niu, Y.; Zio, E. A generalized cauchy method for remaining useful life prediction of wind turbine gearboxes. *Mech. Syst. Signal Process.* **2021**, *153*, 107471. [CrossRef]

24.   Kozłowski, E.; Mazurkiewicz, D.; Żabiński, T.; Prucnal, S.; Sęp, J. Machining sensor data management for operation-level predictive model. *Expert Syst. Appl.* **2020**, *159*, 113600. [CrossRef]

25.   Liu, S.; Hu, Y.; Li, C.; Lu, H.; Zhang, H. Machinery condition prediction based on wavelet and support vector machine. *J. Intell. Manuf.* **2017**, *28*, 1045–1055. [CrossRef]

26.   Chen, Z.; Liang, K.; Yang, C.; Peng, T.; Chen, Z.; Yang, C. Comparison of several data-driven models for remaining useful life prediction. In Proceedings of the 2019 11th CAA Symposium on Fault Detection, Supervision, and Safety for Technical Processes, SAFEPROCESS 2019, Xiamen, China, 5–7 July 2019; pp. 110–115. [CrossRef]

27.   Ren, L.; Sun, Y.; Cui, J.; Zhang, L. Bearing remaining useful life prediction based on deep autoencoder and deep neural networks. *J. Manuf. Syst.* **2018**, *48*, 71–77. [CrossRef]

28.   Xiao, L.; Chen, X.; Zhang, X.; Liu, M. A novel approach for bearing remaining useful life estimation under neither failure nor suspension histories condition. *J. Intell. Manuf.* **2017**, *28*, 1893–1914. [CrossRef]

29.   Saucedo-Dorantes, J.J.; Delgado-Prieto, M.; Osornio-Rios, R.A.; Romero-Troncoso, R.D.J. Industrial Data-Driven Monitoring Based on Incremental Learning Applied to the Detection of Novel Faults. *IEEE Trans. Ind. Inform.* **2020**, *16*, 5985–5995. [CrossRef]

30.   Li, X.; Mba, D.; Lin, T. A Similarity-based and Model-based Fusion Prognostics Framework for Remaining Useful Life Prediction. In *2019 Prognostics and System Health Management Conference, PHM-Qingdao 2019*; IEEE: Piscataway, NJ, USA, 2019. [CrossRef]

31. Pilario, K.E.S.; Cao, Y.; Shafiee, M.; Lao, L. Reconstruction based fault prognosis in dynamic processes using canonical variate analysis. In Proceedings of the ICAC 2019—2019 25th IEEE International Conference on Automation and Computing, Lancaster, UK, 5–7 September 2019. [CrossRef]

32. Mosallam, A.; Medjaher, K.; Zerhouni, N. Data-driven prognostic method based on Bayesian approaches for direct remaining useful life prediction. *J. Intell. Manuf.* **2016**, *27*, 1037–1048. [CrossRef]

33. Dourado, A.; Viana, F.A.C. Early life failures and services of industrial asset fleets. *Reliab. Eng. Syst. Saf.* **2021**, *205*, 107225. [CrossRef]

34. Kusiak, A. Convolutional and generative adversarial neural networks in manufacturing. *Int. J. Prod. Res.* **2020**, *58*, 1594–1604. [CrossRef]

35. Wang, J.; Zhang, L.; Zheng, Y.; Wang, K. Adaptive prognosis of centrifugal pump under variable operating conditions. *Mech. Syst. Signal Process.* **2019**, *131*, 576–591. [CrossRef]

36. Mehringskotter, S.; Preusche, C. Consideration of Variable Operating States in a Data-Based Prognostic Algorithm. In Proceedings of the IEEE Aerospace Conference, Big Sky, MT, USA, 2–9 March 2019. [CrossRef]

37. Xiao, L.; Xia, T.; Pan, E.; Zhang, X. Long-term predictive opportunistic replacement optimisation for a small multi-component system using partial condition monitoring data to date. *Int. J. Prod. Res.* **2020**, *58*, 4015–4032. [CrossRef]

38. Quatrini, E.; Costantino, F.; Di Gravio, G.; Patriarca, R. Condition-based maintenance—An extensive literature review. *Machines* **2020**, *8*, 31. [CrossRef]

39. Breiman, L. Random Forest. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]