
Article

Clustering application for condition-based maintenance in time-varying processes: a review

Elena Quatrini¹, Silvia Colabianchi¹, Francesco Costantino^{1,*}, Massimo Tronci¹

¹ Department of Mechanical and Aerospace Engineering, Sapienza University of Rome, Via Eudossiana, 18, 00184 Rome, Italy; elena.quatrini@uniroma1.it, silvia.colabianchi@uniroma1.it, francesco.costantino@uniroma1.it, massimo.tronci@uniroma1.it

* Correspondence: francesco.costantino@uniroma1.it; Tel.: +39-0644585260

Abstract: In the field of industrial process monitoring, more and more interest is being shown in specific process categories. These include time-varying processes, that is, those processes whereby the response one receives as output from the system depends on when the input signal is sent into it. There are many reasons for this process variability and such contexts are not always analyzed with this operational characteristic at their core. At the same time, interest in certain categories of techniques is also becoming more prominent, to meet certain application needs. Among these, clustering and unsupervised techniques in general are gaining ground. This is largely due to the difficulty of finding fault data with which to train, for example, supervised models. On the other hand, the clustering technique, on which this contribution focuses, also makes it possible to compensate for the lack of complete knowledge of the structure of the process itself. With these two considerations in mind, this contribution proposes a literature review on the topic of clustering applied in time-varying contexts, in the maintenance field. The aim is to present an overview of the main fields of study, the role of clustering in this context and the main clustering techniques used.

Keywords: Latent Dirichlet Allocation, Natural Language Processing, Condition based maintenance

1. Introduction

In the perspective of Industry 4.0 and the growing digitalization of industrial plants, a proper maintenance management process based on data analysis has now become a focal point not only for cost control but also to maintain a challenging level of competition, thanks to increasing productivity. This necessity goes hand in hand with the needs of a constantly evolving production production scenario that has led to more than significant innovations in the field of sensor technology. Current industrial settings are becoming increasingly complex and difficult to manage from a maintenance point of view. The complexity of current production processes, due to the high level of availability requested for the equipment and to the high number of components and interacting elements in the industrial processes, is starting to show the shortcomings of some condition monitoring and Condition-based maintenance (CBM) techniques currently widely and successfully applied in linear and static processes. Unfortunately, many industrial processes operate under time-varying conditions. Reasons that can lead to this type of situation in real industrial processes are, for example, changes in the load input into the process, physiological aging of the equipment or drift caused by disturbance [1], different operational steps within the process not easily detectable [2]. Condition-based maintenance relies on the interpretation of measured values as signals of anomaly situations that suggest maintenance interventions [3]; in time-varying process a value could be expected or anomalous

depending on the specific phase of the production, thus how to implement a proper condition-based maintenance in time-varying processes is currently an open challenge for research. It is worth noting that processes that operate under time-varying conditions are extremely common in the existing industrial scenario. Examples are the pharmaceutical sector, with processes such as the granulation, or processes involving rotating machines in their many applications. The level of complexity in these engineering systems and manufacturing processes is higher than in stationary situation. Managing the reliability of these systems becomes challenging in modern dynamic operational settings. In particular, the difficulties of monitoring time-varying processes are obvious, because the different variability characteristics present within the same process can lead to monitoring with incorrect results [4], with the consequent failure to identify faults or, more importantly, the problem of false alarms [5]. The risk, in fact, in this type of monitoring is that the model tends to develop confusion in discerning between intrinsic changes in the nature of the process, dictated precisely by its time-variant nature, and real failure situations. Data management for CBM appears still as an immature research area, with many open challenges at different levels, since several CBM approaches are data-driven [6]. Entering into the details of the time-varying conditions, it is possible to assert a significant increase of interest in recent years in the field of fault detection and maintenance in general [7–10]. As an important remark for the presented research, the application of clustering techniques in the maintenance research gained considerable. The most common motivation for the application of clustering is that supervised learning techniques are not available; in fact, there are many contexts in which access to failure data is limited [11], sometimes because of the difficulty of monitoring, but in general because failure situations arise as exceptional situations, thus implying that monitoring models can not feed supervised machine learning models. In this situations the unsupervised techniques lead to more reliable results compared to supervised learning methods [12]. Clustering is the basic task for unsupervised learning and it helps to overcome data shortage problem [13–15].

Since the evidence just introduced, the research topic focused on the coexistence of these two themes, namely clustering applied specifically in time-variant contexts. The aim of this review is, therefore, to understand how clustering has, so far, been used to support maintenance in time-variant contexts. Specifically, the purpose is to define the functions that clustering has in this context, the main areas of application and the clustering techniques that have been most investigated and applied. To do this, an approach using the natural language process (NLP) was proposed to support the literature analysis. As will be explained in more detail later in the paper, the contributions selected for analysis were grouped together using an NLP technique, highlighting the main themes, and also making it possible to understand which contributions were most representative of a thematic area. NLP techniques allow to automatically process, in this case, written information of a natural language by means of an electronic calculator. Through this phase of pre-analysis of the contributions it is possible, first, to make an already categorized and conscious reading of the contributions, with a significant indication of the most representative contributions for each identified topic. Similarly, the identification of any overlapping thematic areas facilitates a "multiple" and parallel reading of the contributions, making it possible to identify any links between thematic areas. It also speeds up the analysis of the literature because only contributions with high values of belonging to a topic can be carefully considered. Finally, it makes the analysis more objective, because manual keyword labelling and subsequent categorization on this basis could be marred by extreme subjectivity. The approach selected to implement NLP analysis is Latent Dirichlet Allocation (LDA) [16,17]. The latter creates a mixture of topics, to which each document is subsequently assigned according to probabilistic logic models, allowing a contribution to be allocated to more than one topic, with different weights, thus providing more realistic and elucidative results. The literature proposes some applications of the text mining concept to support literature analysis [18,19]. As far as the authors are aware, however,

there are no contributions on the application of NLP, specifically LDA, to support literature reviews on the topic specifically under consideration. Consequently, the fruitful applicability of the LDA and the output that can be extracted from it in the context of the literature review on this topic are also assessed in this contribution.

The sections of this paper are organized as follows. Section 2 will present the methodology followed, both from the point of view of structuring the literature review and from the point of view of applying the LDA algorithm. Section 3 will present the evidence extracted from the literature, detailing the different topics identified. Finally, the discussion and conclusions section is presented, highlighting the extracted contents, their possible application in the maintenance field, and possible future developments of the research.

2. Methodology

The literature review considered in this study was carried out on the Scopus database, the largest abstracts and citations database of peer-reviewed literature [20]. The selected key-search was set targeting a high level of inclusiveness, with regards to the topic. The key-search is ("cluster*" and ("condition based maintenance" OR "fault*" OR "useful life" OR "diagnos*" OR "prognos*" OR "failure") and ("time var*" OR "non stationary" OR "multi phase" OR "multiphase" OR "process phase")) in the title, abstract or keywords fields of the database. The final query has been run-up in September 2021, to present an analysis as complete as possible, although the topic was monitored in the last year of research. The choice of keywords was dictated by some specific considerations. Firstly, the term "time var*" was included because the authors wanted to consider contributions that explicitly considered this type of feature in their analyses. In fact, sometimes some time-variant processes are not treated as such, and these contributions were not of interest for this analysis. It has also been noted that the term "non-stationary" is often used to describe the characteristics of signals from time-varying contexts [21]. For this reason, it was included as a word in the search. The terms, on the other hand, "multi phase", "multiphase" and "process phase" were inserted to investigate the context of multi-phase processes, because it is the type of time-varying process that is at the center of a broader research path. The term "cluster*" was inserted to identify those contributions that explicitly referred to clustering as a tool used in the contribution, just as the terms "condition based maintenance" OR "fault*" OR "useful life" OR "diagnos*" OR "prognos*" OR "failure" were inserted to circumscribe the field to the field of maintenance.

Subsequently, only contributions in English were selected and following a reading of the abstracts, only contributions with a specific relevance and belonging to the topic were selected. The final analysis was therefore carried out on 72 contributions.

Figure 1 presents a clear step-by-step representation of the literature analysis developed in the paper.

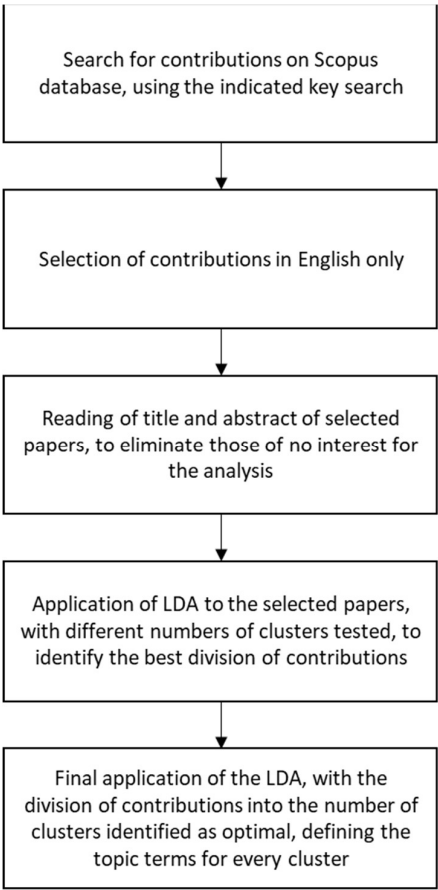


Figure 1. Literature reserach methodology

Before proceeding to the actual application of the LDA and the subsequent literature reviews, some graphical analyses were implemented to extract initial information. To do this, the authors made use of VOSviewer, a software tool for constructing and visualizing bibliometric networks. In detail, three analyses were implemented. First, a so-called network visualization was carried out, based on the keywords in the contributions (Figure 2).

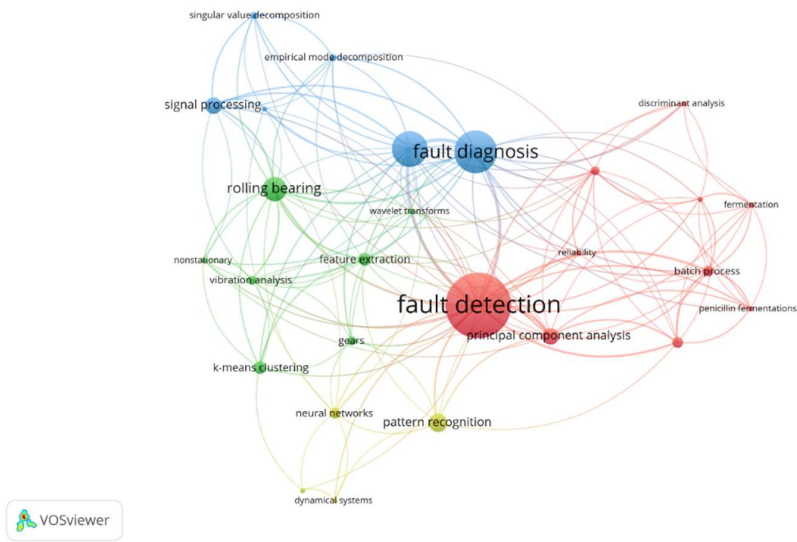


Figure 2. Network visualization

In this type of visualization, items are represented by their label and by a circle. The size of the label and the circle represent the weight of the item, meaning that the higher the weight of an item, the larger the label and the circle of the item, while the lines connecting the circles represent a link. In this specific case, the weight that was considered was that of the occurrence of the keyword within the contributions, considering only those that appeared in at least three contributions. This threshold is needed to make the analysis readable by eliminating keywords used only in specific contributions. This is an outline analysis, while the details of the contributions will be shown with the next analysis. Finally, the color of an item is determined by the cluster to which the item belongs. This analysis makes it possible to identify, quickly, which are the main concepts analysed in the thematic area, as well as the relationships between the words identified. And it is precisely the relationship between the keywords considered that shows, in this step, four groups of keywords, defined according to their co-occurrence in the documents. If you go into the details of the keywords, you will see that the most important concepts are the following (by decreasing occurrence):

- Fault detection
- Fault diagnosis
- Rolling bearing
- Signal processing
- Empirical mode decomposition
- Vibration analysis
- Support vector machines
- Singular value decomposition
- Neural networks
- Pattern recognition
- K-means clustering
- Gears
- Principal component analysis
- Reliability
- Batch process
- Penicillin fermentation
- Feature extraction

- Wavelet transforms
- Reliability
- Fermentation
- Discriminant analysis

Another evidence that can be extracted from this graphical analysis is the connections between the keywords. If one focuses on rolling bearings, one immediately notices their close correlation with the concept of vibration, feature extraction and the k-means algorithm. Another connection that jumps out at you is that between the concept of batch processes and fermentation processes, which is also related to the technique of principal component analysis and the concept of discriminant analysis. Similarly, the concept of signal processing forms a cluster with singular value decomposition and empirical mode decomposition. Finally, as one might expect, the themes of fault diagnosis and fault detection are extremely cross-cutting. This is clearly not surprising, as they are in fact two of the main objectives of condition monitoring. the concept of feature extraction, although important, is less transversal and more circumscribed than the two mentioned above.

A second analysis related to the first one, but with a different point of view, was the one called density visualization Figure 3.

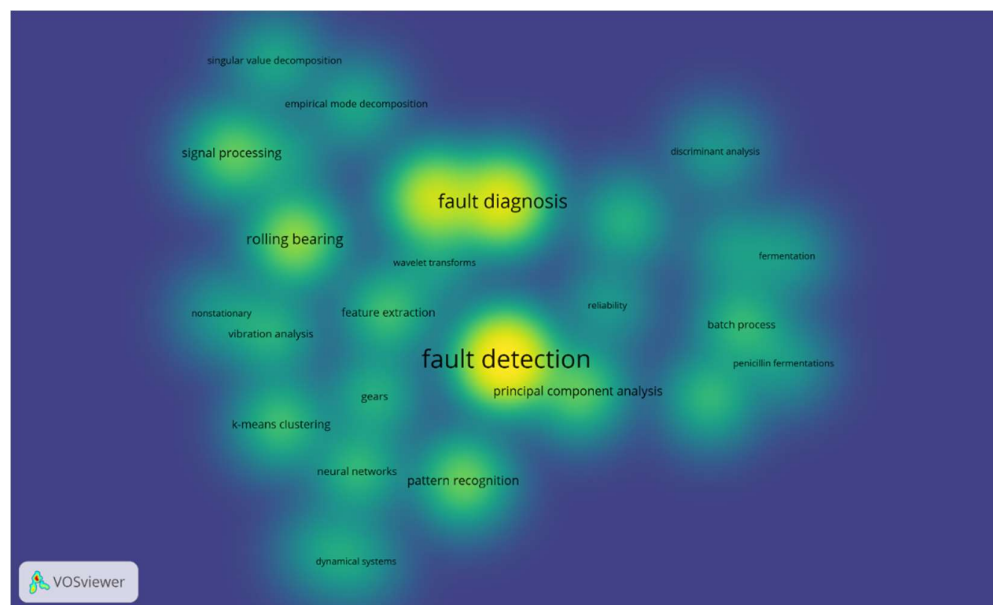


Figure 3. Density visualization

This analysis takes up the keywords considered above, but in this visualization, items are represented by a color that represents their density. By default, colors range from blue to green to yellow. The larger the number of items in the neighborhood of a point and the higher the weights of the neighboring items, the closer the color of the point is to yellow. The other way around, the smaller the number of items in the neighborhood of a point and the lower the weights of the neighboring items, the closer the color of the point is to blue. Here it can be seen that, apart from fault detection, fault diagnosis and rolling bearings, all other terms tend to have the same density of occurrence.

Finally, to deepen the themes more, the analysis of the weight of words was also extended to titles and abstracts. In this way, it was possible to extract a greater number of

concepts, application areas, and techniques with relevance to the research area investigated.

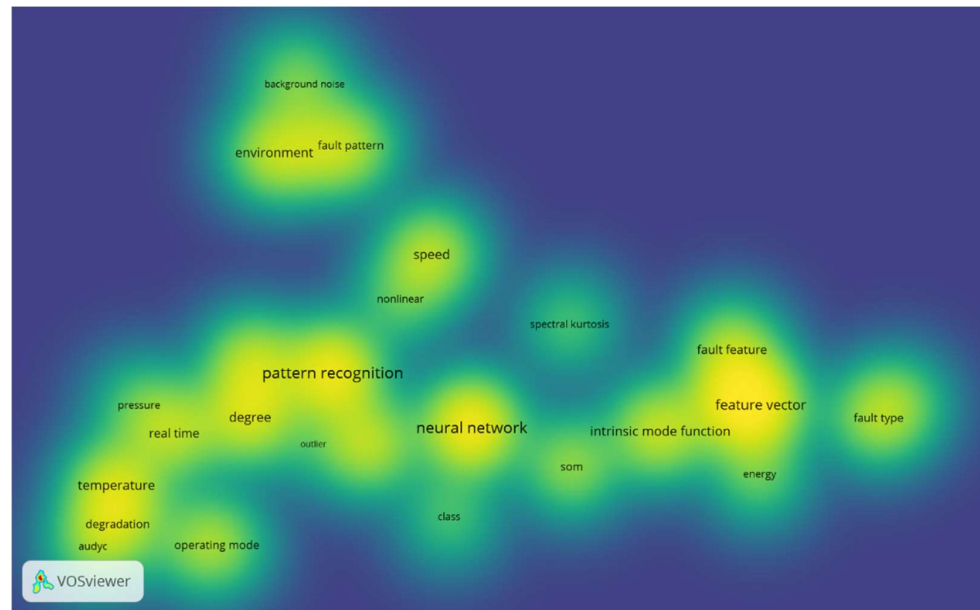


Figure 4. Density visualization for title and abstract

In the third analysis of Figure 4, it was decided not to consider the words that were already highlighted in the previous analysis. In this way, it was possible to identify several new concepts, more specific than those identified with the keywords. Precisely because of the specificity of these terms, it was decided not to consider the relationships between the words, as done in Figure 3, since it was felt that this could lead to unhelpful and misleading considerations. Indeed, considering the thematic breadth, nuances and connections would be lost. As can be seen, in this analysis, specific process parameters can be identified, such as speed, pressure, temperature, or energy. Beyond that, the problem of background noise becomes apparent, related to the concept of the working environment and the non-linear processes. If attention is focused, however, on the concept of fault detection and fault diagnosis, some more peculiar concepts emerge, such as pattern recognition, degradation, class, with the meaning of fault class, degree, that is the level of gravity of a fault, fault feature, feature vector, fault pattern, fault type or operating mode. As far as techniques and methods are concerned, the presence of the AUDyC algorithm, the spectral kurtosis, the self-organizing maps, and the intrinsic mode function is revealed. Finally, the concept of real-time application and outlier is identified by the analysis. This analysis is a first investigation of the thematic area, based, in this case, on the occurrence of terms. Thus, through the LDA it will be possible to analyse the contributions in more depth, also showing the thematic connections. The results obtained from this first investigation, however, already allow us to have a wide-ranging vision of themes that are very present and treated in the contributions considered.

Finally, it was decided to implement a geographical analysis of contributions. This is because both from the point of view of the journals of publication, or conferences, and from the point of view of the authors, the field of research is still very diverse and an analysis in this sense would have contributed little to the value of the analysis. In fact, publication still takes place in very scattered and varied journals, just as no specific cluster is currently identified for authors. For this reason, it was decided to shift the focus to the

identified with the term corpus, with M documents, in which a specific document d has N_d words, one can state that for each document in D the LDA algorithm models D according to the following generative process [16]:

1. Select a multinomial distribution φ_t for topic t ($t \in \{1, \dots, T\}$) from a Dirichlet distribution with parameter β .
2. Select a multinomial distribution θ_d for document d ($d \in \{1, \dots, M\}$) from a Dirichlet distribution with parameter α .
3. For a word w_n ($n \in \{1, \dots, N_d\}$) in document d ,
 - Select a topic z_n from θ_d .
 - Select a word w_n from φ_{z_n} .

The assumption behind the LDA technique is that a document is written by authors as a set of different topics and each document has its own peculiar distribution of topics [17]. There are two parameters that characterize the LDA:

- α , that represents, considering each of the topics identified in the document, the probability that the document under consideration belongs to each of the topics.
- β , that represents, considering each of the tokens, what is the probability that the topic under consideration contains each of the tokens.

In the process presented above, the only variables actually observed are the words in the documents. φ and θ are, in fact, latent variables and α and β are hyperparameters [22]. The next step is to calculate and maximise the probability of the observed data D , which following what has been proposed in the previous contribution [22], follows (1):

$$p(D | \alpha, \beta) = \prod_{d=1}^M \int p(\theta_d | \alpha) \left(\prod_{n=1}^{N_d} p(z_n | \theta_d) p(w_{d_n} | z_{d_n}, \varphi) \right) P(\phi | \beta) d\theta_d d\phi \quad (1)$$

An issue to be considered, if attention is drawn to (1), is that of exact inference in LDA, which turns out to be intractable due to the coupling between θ and φ [22]. Some approximate algorithms are used in this respect, such as as variational inference [16] or Markov chain Monte Carlo [23].

The optimum number of topics is chosen using a perplexity score, a statistical measure of how well a statistical model describes a dataset, with lower perplexity denoting a better probabilistic model [22], and a coherence score, that represents the semantic similarity between high scoring words in the topic value [24].

To define perplexity, it evaluates how well the dataset is described by the statistical model. Perplexity is defined as (2) [16]:

$$\text{Perplexity} = \exp \left\{ - \frac{\sum_{d=1}^M \log p(w_d)}{\sum_{d=1}^M N_d} \right\} \quad (2)$$

For the coherence score, higher values are sought, considering that the value moves between 0 and 1. In the specific case of this research, the documents analyzed for each paper, as also previously stated, are its abstract, its title and its keywords, all together.

For what concern the coherence score (3), reference is made to a specific contribution [25]

$$\varphi_{S_i}(\vec{u}, \vec{w}) = \frac{\sum_{i=1}^{|W|} u_i \cdot w_i}{\|\vec{u}\|_2 \cdot \|\vec{w}\|_2} \quad (3)$$

Where

$$\vec{u}(W') = \{\sum_{w_i \in W'} NPMI(w_i, w_j)^{\gamma}\}_{j=1, \dots, |W'|} \quad (4)$$

$$NPMI(w_i, w_j)^{\gamma} = \left(\frac{\log \frac{P(w_i, w_j) + \epsilon}{P(w_i) \cdot P(w_j)}}{-\log(P(w_i, w_j) + \epsilon)} \right) \quad (5)$$

The final coherence score is the result of the arithmetic mean of all φ_{S_i} .

3.1 Tokenization

To proceed with the analysis of a written text, it is paramount to identify the units that compose it, specifically the linguistic ones. The linguistic units into which the text being analyzed is broken down are called tokens, hence the term tokenization. There is, obviously, no single way to define what a token is within the considered text. This choice depends on the objectives of the research and the structure of the text. In the specific case of this contribution, it was decided to consider spaces as token delimiters. As a cleaning of the text, all punctuation has been removed. It will then be these units that will go on to define, after being processed, the thematic areas of interest in the analyzed domain.

3.2 Sparse terms removal and stop-words removal

After the identification of the tokens that make up the analyzed text, a further phase must be implemented to eliminate superfluous words. The term superfluous, which is then captured in the terms "stop words", refers to two concepts of words. On the one hand, there are words which in themselves do not convey any meaning of their own, such as adverbs or articles, and on the other, words which have no meaning in the analysis being implemented. In the details of this contribution, for example, all words that do not refer to algorithms, fields of application or functionalities of the techniques identified were eliminated. In this type of analysis, therefore, it is relevant to identify which objectives are being pursued, to be able to guide the various steps that make up the analysis. On the contrary, sparse terms are instead those ones that occur in less than 1% of the documents [17]. Sparse terms are instead terms that occur in less than 1% of the documents. In general, to avoid problems when analyzing the text, it is a good rule of thumb to eliminate words present in the search key used. In the research presented in this contribution, the stop words removal has been performed twice: immediately after the tokenization phase and after the stemming & lemmatization phase, which will be presented in the following paragraphs.

3.3 Stemming & lemmatization

For the definition of stemming and lemmatization, reference is made to [26]. Stemming is the procedure of reducing a word to its germ, i.e., to the terminological root of that term. Lemmatization is directly linked to the stemming phase. Specifically, it uses part of speech tagging for each identified word, (e.g., verb, noun, etc). The phases are directly linked because the lemmatization phase allows the definition of the tags to be attributed to the tokens identified previously. Subsequently, based on the lemmatization, specific stemming rules are applied to the tokens.

3.4 N-grams tagging

The term N-grams refers to a group of N words that often appear together in the analyzed document. In this contribution are considered only sequence of two words, i.e., b-grams, and sequence of three words, i.e., tri-grams.

3.5 Text features extraction

This phase permits the definition of the importance of the remaining terms, following the implementation of the steps presented above. This extraction is based on two concepts:

- How many times a word appeared in each document.
- How much a word appears in documents, giving little weight to those that appear in many and more weight to those that appear in a few.

4. Findings from the literature review

After discussing the application steps of the LDA, this section of the paper will go into detail about the evidence found in the literature review. Beyond this, the results obtained by applying LDA will be evaluated, defining whether the categorization of contributions obtained corresponds to the real content of the papers.

The analysis carried out with the LDA led to the definition of four clusters within the analyzed thematic area. This value was defined because of the highest coherence score obtained. As explained above, each contribution was assessed, resulting in a value being assigned to each of the four identified topics. Each paper was attributed to the topic for which it had a higher percentage of membership. The graphical display of the topics, following the proposal in [27], and this is representative of how, according to the application of the LDA, the thematic area investigated is clearly clustered. In that graph, each bubble represents a topic, considering that the bigger is the bubble, the more prevalent that topic is. This means that in **Figure 6**. Topics distributionFigure 6 the size of the bubbles represent, with respect to the 30 most frequent terms derived from the linguistic analysis, how relevant, in complexity, the specific topics are in the definition of these terms. But this is not the reason for the inclusion of this figure. In fact, a model is considered effective when the bubbles are fairly big, non-overlapping and scattered throughout the chart instead of being clustered in one quadrant. To be more specific, the identified topics are plotted in a two-dimensional space, where the centers are determined by calculating the distances between the topics. The inter-topic distances are then projected using multidimensional scaling onto a two-dimensional space [28]. The authors decided to utilize the default settings [27], that are Jensen-Shannon divergence for computing inter-topic distances and to Principal Components for scaling the set of inter-topic distances. The clustering of the identified topics is clear and well-defined, but this does not mean that the identified thematic sub-areas do not have points in common. This is not the case, and it would be strange to think otherwise, considering the many aspects that make up such a rich in content and complex research area. From this initial observation it can be deduced that, according to the logic of the LDA, the terminology makes it easy to identify specific thematic areas, which certainly have their own peculiarities, but also points of contact with other topics. Figure 6 shows the clustering performed in the thematic area.

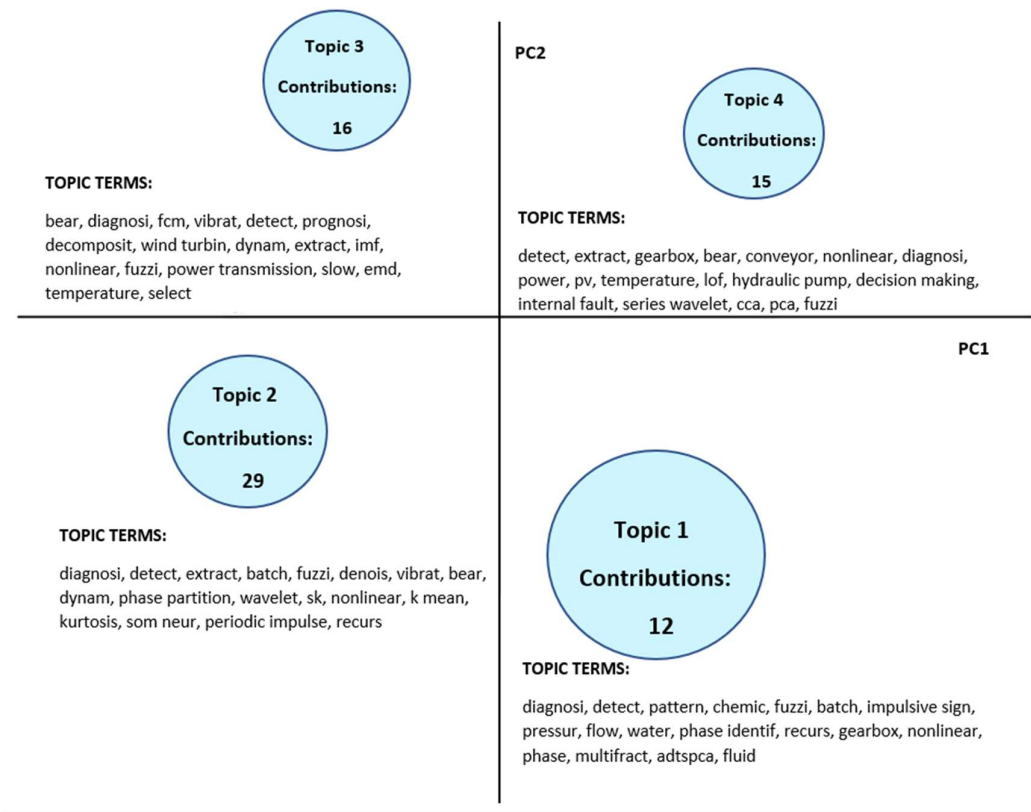


Figure 6. Topics distribution

At first glance, it can already be seen that, apart from topic 2, which has the most contributions, the other three topics tend to be of the same size. All acronyms will be detailed in the appendix, so as not to create difficulties in reading the paper.

The next step was to check that these keywords did indeed describe areas of interest, also considering the coherence between them and the papers labelled with them. A first comparison shows that some keywords are present in several clusters, indicating thematic transversality. An obvious example is that of the keywords “diagnosis”, “detect”, “pattern” and “extract”. These keywords are objective keywords, in the sense of the objective being pursued in the research in which they are encountered. This plurality should not come as a surprise, as it is quite evident that these activities occur in various fields of application and with different techniques. In general, considering also that each paper is allocated to each of the tasks with a specific percentage of membership, it is important to consider that the allocations are indicative of the topic that is mostly contained in that contribution. However, some contributions present transversal themes and have been allocated to a specific contribution either because of the connection with the other themes of the selected topic or because that particular thematic area is predominant in it. In general, therefore, this clustering cannot be considered a watertight compartmentalization since the various thematic areas talk to each other. Another point of attention is the application context. As can be seen, for example, the term “bear”, which is nothing more than the representation of the root of the word “bearing”, is present in three clusters, just as the term gearbox is present in two. This should not be taken as an error in categorizing the contributions, but as evidence of the fact that transversal and broad application changes such as that of bearings or gearboxes translates into various thematic areas, in relation to the algorithms used or the context in which the specific component is inserted.

For the same reason, after a careful reading of the papers, some contributions were moved from one topic to another, based on the authors' considerations. Considering the proportion between the number of contributions and papers moved by the authors, this is not considered a limitation of the model. Some words may be misinterpreted by the model.

In general, therefore, the predominant role of the LDA in this type of research, from the experience gained for the analysis presented in this contribution, is to highlight points of attention not only within individual clusters, but of the thematic field in general. This makes it possible to understand trends within the research, but also how the same theme may have been approached with different modalities.

In the following paragraphs, the different topics will be analyzed without mentioning all the selected contributions as some of them deal with related issues and would be redundant, making the reading of the paper more difficult. Moreover, considering that the application of the LDA allows a degree of belonging of the contributions to the specific topic, it is possible to describe each topic only through the contributions with the highest levels of belonging.

4.1 Topic 1

The predominant role in this topic is played by chemical plants, which are the focus of research in various ways. This is not surprising, as the literature shows that chemical processes tend to be the ones with the most time-varying characteristics. Interestingly, in the chemical context, clustering helped several applications, such as the identification of fault subclasses in contexts where faults exhibit time-varying characteristics [29]. Similarly, in penicillin production processes, clustering has been used to identify in the batch process the specific production stage [30]. This contribution also ties in with the keywords in topic 2. In that topic phase partitioning is applied with clustering. In this contribution [30], it is highlighted that the analyzed process belongs to the world of chemical processes, while in topic 2 the focus is on phase recognition, and the production process assumes a pure case study. Moreover, this contribution [30] is the only one applying the AdTsPCA. As further analysis of the chemical context, more specifically petrochemicals, clustering allows the identification of the operating modes present in the process, based on a process variable highlighting the change, to develop specific models for the operating modes identified [31]. Entering into the field of water distribution and treatment, clustering again showed interesting performance. In general, water distribution systems are assimilated to chemical plants because of the frequent use of chemical substances to ensure the water potability. These systems are fragile due to their high number of components. There are several flow regimes, and clustering allows their modeling and recognition, through self-organizing maps and the analysis of pressure [32]. Similarly, wastewater treatment processes have also seen the application of clustering applied to predict the trend of the data pattern, specifically the fuzzy c-means [33]. Considering instead the bearing analyses attributed to this cluster, they are mainly based on the analysis of impulsive vibration signals. In this case, the role of clustering, that is a generalization of k-means, is to support the separation of the impulsive components of the vibration signals [34]. Finally, to conclude the application areas presented in this topic of contributions, the area of gearboxes is treated by considering vibrations as a tool for fault diagnosis [35]. Focusing on the techniques in this topic, it is to mention the proposal of the AdTsPCA technique [30]. This technique is highlighted as innovative compared to other techniques, even when considering the other identified topics, while other techniques are not highlighted as they do not predominate in the sub-thematic area. The last two concepts that are marked are the recursiveness of some of the proposed algorithms and methods [29,31,33] and the fact that

they are multifractal [35]. The issue of non-linearity is actually transversal to the various topics and not specific to this one. It is precisely this transversality of the concept of non-linearity of the process that will be of interest in defining future research developments

4.2 Topic 2

This topic presents the highest number of contributions. This makes it more complex to analyze. The first theme to be discussed in this topic is that of phase partitioning, which is closely linked to the concept of the batch and the concept of the production phase. In fact, batch processes often have the characteristic of being multi-phase and clustering algorithms focus on this. An interesting algorithm is the density peaks clustering [36] that manages not only the individual phases but also the transition between them. A problem that arises when one wants to do phase partitioning is that some techniques, such as fuzzy c-means clustering, require an a priori definition of the number of phases to be identified. This is not always possible. To overcome this limitation, an adaptive fuzzy c-means is proposed [37]. Another problem is the variable batch duration time, a feature that makes a phase analysis based on the duration tricky, tending to be impossible. Sequence-constrained fuzzy c-means handles this situation [2]. In this contribution, in addition to uneven batch durations, attention is also paid to the identification of steady and transitional phases within production batches. This latter feature is also the focus of analysis in other contributions, which propose, for example, kernel entropy component analysis as an algorithm for phase partitioning [38], based on the divergence of kernel entropy. Specific subjects in this topic are bearings, data denoising, and vibrations. Bearings, as mentioned above, have been a much-investigated subject, with focuses on specific characteristics. Bearing-related contributions use vibration as a central parameter and the problem of noise and its consequent data denoising as a research focus. In these contributions, with an unusual application, a k-means clustering helps to evaluate the efficacy of the proposed model (e.g. a denoising model)[39]. Another contribution, again with the help of k-means, suggests the application of K-SVD-based sparse representation method clustering to implement dictionary learning and extract information from vibration impulses [40]. Finally, on the subject of feature extraction, another contribution proposes the application of clustering based on spectrum overlap coefficient and kurtosis index to implement the extraction of periodic pulses from bearing vibration signals [41]. This last contribution highlights another of the keywords presented for this topic, namely the topic of periodic impulses. The importance of this concept is due to the fact that the occurrence of this type of impulse in the vibration signals presumably indicates the presence of a fault [41]. Other concepts that appears in this topic are the previously mentioned data denoising, signal decomposition, self-organizing maps, and wavelet transforms. Specifically, there are some contributions with the joint application of signal decomposition approaches and self-organizing maps for fault diagnosis, with the ensemble empirical mode decomposition [42] as well as wavelet transform [43,44]. Other contributions in this topic highlighting the presence of noise are [45], in which spectrum entropy clustering is applied to extract features to be used for fault diagnosis analysis in a gear system. Another solution, on the other hand, proposes the application of a recursive clustering algorithm for diagnostic analysis [46]. The importance of clustering for noise management lies in the fact that it allows analyses to be implemented without having prior knowledge of the process. Another application in this respect that is suggested is the algorithm of mean shift clustering [47]. Two other keywords that carry a common concept are kurtosis and spectral kurtosis (SK). One of the previously mentioned contributions already highlighted the concept of kurtosis, specifically by using a kurtosis index for the process of extracting periodic vibration impulses [41]. In another contribution, kurtosis is used to create two indicators of the health status of a grinding chatter, defining a kind of kurtosis possibility density function benchmark against which to compare the kurtosis of the test data. In this context,

clustering, specifically k-means, has the role of assessing the goodness of the results [48]. Spectral kurtosis was also used for diagnostic purposes in a gearbox, defining a diagnostic feature. Here too, clustering takes on a controlling role, being used to define the probability of a correct diagnosis [49]. Among the keywords, k-means also stands out, as among the clustering algorithms present in this one, it is the one with the highest number of hits. In this topic, as in topic 1, weight is again given to the fact that some of the proposed solutions involve a recursive approach [46].

4.3 Topic 3

In this topic we found several contributions about power generation and transmission, with exclusion of photovoltaic application (see Topic 4). Studying faults in power transmission lines, clustering permits discriminant models using k-means, to identify faults and correctly managing inspections and safety [50]. The latter is a non-negligible issue in this context. The topic includes contributions explicitly linked to prognosis analyses. The role of clustering is exploratory, to extract information that set the probability of failure occurrence over specific time frame [51,52]. Interestingly, these two contributions share two important points: the first is the clustering algorithm used, i.e., AUto-adaptive and Dynamical Clustering, and the second is the concept of slow fault evolution. Slow evolution of a failure tends to be more problematic to analyze, making it necessary to structure the entire analysis on this specific characteristic. Researches into fault management in wind turbines is becoming increasingly important, and here too the concept of clustering and variability over time is being investigated. Wind turbines present high level of complexity due to the high variability of their operating conditions. Based on this, clustering, e.g. k-means, is proposed for identifying the operating conditions of wind turbines [53]. The topic of bearings is also addressed, with more contributions. In fact, bearings are one of the most investigated area in the condition maintenance literature. In this topic, compared to e.g., topic 1, it is addressed without focusing on impulsive signals that can be extracted from vibration measurements. In this sub-area, while still using vibration analysis, clustering takes on a more diagnostic role, used as a tool to identify the nature of the fault [54]. Similar analyses, using the same algorithm but with different parameters, specifically the energy ratios of the intrinsic mode functions of the bearings, have also been carried out in other research projects [55]. The same parameter is found, with the same diagnostic objectives, in other contributions to bearing analysis [56]. Picking up on the theme of slowly evolving faults, a contribution on temperature controllers shows the application of the AUto-adaptive and Dynamical Clustering technique to estimate the probability of occurrence of a fault. This gives insight into how this algorithm is currently performing [57]. Analyzing the contributions from the point of view of the techniques, the authors notice in the keywords that the technique of fuzzy c-means is highlighted, mainly applied for the bearings monitoring [54,55], but even in different contexts [58]. The same contributions, and some more [56,59] show a fruitful application of the decomposition of monitored signals into intrinsic mode functions, in order to use some extracted information as a discriminative feature within the analysis. Another important technique in this topic is the empirical mode decomposition, for the decomposition of monitored signals [54–56,59,60]. The temperature theme, on the other hand, highlighted among the keywords, relates to the topic of temperature controller monitoring [51,57].

4.4 Topic 4

The last identified topic is now analyzed. The first highlighted content is photovoltaic (PV) systems. Considering the energy context, power generation appears as descriptive of the topic, as in Topic 3. The currents generated and present in PVs have non-stationary characteristics. In this case, the LOF (local outlier factor) technique is used not only to

identify the fault, but also to assess its degree of severity [61]. It is interesting to note that, from the contributions analyzed, this technique appears to be applied only in the context of photovoltaic installations, despite being an established technique in the study of anomaly identification. In this context, clustering is proposed to apply a wavelet cluster-based bandpass filtering in the data pre-processing phase to obtain the fault sensitive envelope signal, for hydraulic pumps [62]. This topic picks up on a theme already presented in topic 1, namely the gearboxes. As with the bearings, there are also differences that have led to their separate analysis. In this topic, different techniques are used for the detection and diagnosis of gearbox faults, such as series wavelet analysis and fuzzy clustering method [63]. Analyses always start with vibration analysis, but then evolve with different techniques. This deviates from the characteristics outlined in the contribution assigned to topic 1. In addition, bearings and gears are explicitly considered separately in this contribution, which was not made explicit in the previous contributions allocated in other topics. Another interesting evidence is that in this topic the analyses of gearboxes are closely linked to those of conveyor belts. Clustering is here applied starting not only from the most common vibration analyses [64], but also from temperature analyses [65]. Taking up one of the terms highlighted as representative of this topic, namely the concept of decision making, this is explicitly stressed in the context of conveyor belts [64,65]. Another recurring theme in topic 4 is bearings. In this case, the commonality in this subgroup of articles is that local mean decomposition is applied to the analysis of bearing health [66–68]. The cluster takes on the role of a tool for fault diagnosis [66] or a technique to support fractional lower order p-function optimization [67,68]. The topic 4 applications of clustering in fault diagnosis focus on the analysis of inrush current and internal faults [69]. The latter relates to the concept of "internal faults", which is considered representative of this topic. Techniques that characterized topic 4 are curvilinear component analysis (CCA), applied in the field of bearings [70] and PCA, which is used for dimensionality reduction and elimination of collinearity, where clustering is instead used to identify known or new events [71].

5. Discussion and conclusions

Following the analysis of the four topics, conclusions can be drawn not only on the current state of clustering research, but also on the current role of LDA in this type of analysis. Starting from the analysis of the literature context, it is interesting and important to note that, although research is evolving, there are still many scenarios and contexts to be investigated. In detail, although some sectors such as chemicals and petrochemicals present some analyses, there is more of a tendency to analyze specific components rather than industrial processes in their entirety. The main exception to this is batch processes, where clustering tends to identify the sub-phases or operations modes of the process. In general, however, it is possible to say that clustering has mainly the purpose of helping to extract information for further analysis or has the role of a diagnostic tool. The application of the LDA shows that time-variant contexts often present evidence of non-linear trends. This is evident studying the keywords extracted by the LDA algorithm, where the word "non-linear" is present in all four topics. The same applies to the term "fuzzy", which is present in all four topics; fuzzy clustering encompasses numerous techniques that are highly applied in various contexts. As also mentioned during the contribution, bearings, consistent with maintenance analyses in general, are the most common mechanical element studied in time-varying situations. Another piece of evidence is that k-mean and fuzzy c-means clustering are the most widely used clustering algorithms, due to their widespread use and ease of application compared to other techniques. However, their limitations are also emphasized, and new solutions are proposed.

Analyzing what has been obtained from the literature and relating it to some of the research topics investigated by the contributors, some further evidences stand out. First,

researchers still do not give particular attention to the multi-components systems, which are extremely present in real industrial contexts, and mechanical system are commonly considered as single element. In general, except for PCA, multivariate statistical techniques related to clustering are scarcely present. This ties in with the evidence that the contributions present little insights to deal the process non-linearities.

The results of the application of LDA could be considered certainly positive. As mentioned in the introduction phase, some contributions were shifted between the various clusters, but this did not compromise the successful implementation of the LDA. LDA, regardless of the topic allocation, made it possible to have a view of some thematic areas of interest and of peculiar techniques, before reading the contributions,. Based on this, the reading of the papers was certainly more conscious and rational. The reason is twofold. Firstly, the reading of the papers focuses to the concepts expressed by the keywords of the specific topic, allowing to identify, for example, correlations between techniques and application contexts that otherwise could have been missed. Similarly, having already extracted the keywords of all the topics, it was easy to identify incorrect allocations, as well as transversal themes between the various topics. In addition, considering a cross-cutting theme such as bearings, it was possible to quickly identify small thematic sub-areas within the larger theme. In addition, the availability of a measure for every content of the contribution to the topic let the researcher to start with those with higher values of belonging. In this way, when the contributions with a lower level of belonging were read, it was possible to identify immediately whether it was necessary to modify their allocation and whether they took up a theme already presented in other read contributions. In general, therefore, the application of LDA in this case should be seen as a valid pre-processing phase for the subsequent literature reviews. The issue of topic clustering is interesting and relevant, but, according to the authors, the most important outcome is the definition of keywords, as a guide for the reading of the contributions. For this reason, extreme emphasis must be given to the stop words that are inserted, which will outline how the keywords will then be defined. Therefore, the aim of the review is a priority before starting LDA, to insert in the stop words the terms that do not fit into the search. This process is clearly iterative, because only with subsequent applications of the LDA will there be a global and complete overview of the words to be eliminated.

As regards the analysis of future research developments in clustering applied to time-varying process monitoring, some interesting insights emerged. Certainly, it is worthwhile to further emphasize the concept of non-linearity of processes by analyzing various specific monitoring techniques for handling this feature, such as kernel principal component analysis. Phase partitioning and identification of the operating modes of a process emerge as promising research areas for clustering. The idea, which is relevant to test, is to use clustering to divide the whole production process in steps or activities, overcoming the most common batch view. What differentiates this approach from the identification of operational modes is twofold: firstly, the use of all the variable measured in the process and not just a few of these; secondly, a conceptual difference lead from the recognition of process phases by operational modes to a recognition of the trends sufficiently homogeneous to be analyzed together with a dedicated model, suggesting a process phase. Different operational modes could be analyzed jointly, just as the same operational mode in different configurations. Finally, considering the actual limitations in the application of techniques such as canonical variate analysis and multivariate statistics in general for time-varying processes, a possible research development could be to use clustering as a pre-processing stage for these techniques. The two last proposed research scenarios can converge into one, considering the recognition of the process phases with clustering just as a pre-processing step for multivariate statistics.

Appendix

This section details all the acronyms in the paper, in alphabetical order.

- **AdTsPCA** = Adjointed Time Series Principal Component Analysis
- **CBM** = Condition Based Maintenance
- **CCA** = Curvilinear Component Analysis
- **CM** = Condition Monitoring
- **EMD** = Empirical Mode Decomposition
- **FCM** = Fuzzy C-Means
- **IMF** = Intrinsic Mode Function
- **LDA** = Latent Dirichlet Allocation
- **LOF** = Local Outlier Factor
- **NLP** = Natural Language Processing
- **PCA** = Principal Component Analysis
- **PV** = Photovoltaic
- **SK** = Spectral Kurtosis
- **SVD** = Singular Value Decomposition

References

1. Shang, C.; Yang, F.; Huang, B.; Huang, D. Recursive slow feature analysis for adaptive monitoring of industrial processes. *IEEE Trans. Ind. Electron.* **2018**, *65*, 8895–8905.
2. Luo, L. Monitoring Uneven Multistage/Multiphase Batch Processes using Trajectory-Based Fuzzy Phase Partition and Hybrid MPCA Models. *Can. J. Chem. Eng.* **2019**, *97*, 178–187.
3. Quatrini, E.; Costantino, F.; Di Gravio, G.; Patriarca, R. Condition-based maintenance-An extensive literature review. *Machines* **2020**, *8*.
4. Gao, Y.; Wang, X.; Wang, Z.; Zhao, L. Fault detection in time-varying chemical process through incremental principal component analysis. *Chemom. Intell. Lab. Syst.* **2016**, *158*, 102–116.
5. 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.
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. Trans. ASME* **2008**, *130*, 0310041–0310049.
7. Wang, T.; Zhang, L.; Qiao, H.; Wang, P. Fault diagnosis of rotating machinery under timevarying speed based on order tracking and deep learning. *J. Vibroengineering* **2020**, *22*, 366–382.
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 2018; 2018; 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.* **2018**, *66*, 509–518.
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.* **2017**, *14*, 894–904.
11. Morteza, S.; Mohammad, P.-M.; Ali, R.; Ahmad, G. Design for reliability of complex system: Case study of horizontal drilling equipment with limited failure data. *J. Qual. Reliab. Eng.* **2014**, *2014*.
 12. Amruthnath, N.; Gupta, T. Fault Diagnosis Using Clustering. What Statistical Test To Use for Hypothesis Testing? *arXiv* **2019**.
 13. Mahmoodian, A.; Durali, M.; Saadat, M.; Abbasian, T. A life clustering framework for prognostics of gas turbine engines under limited data situations. *Int. J. Eng. Trans. C Asp.* **2021**, *34*, 728–736.
 14. Wei, J.; Huang, H.; Yao, L.; Hu, Y.; Fan, Q.; Huang, D. New imbalanced fault diagnosis framework based on Cluster-MWMOTE and MFO-optimized LS-SVM using limited and complex bearing data. *Eng. Appl. Artif. Intell.* **2020**, *96*.
 15. 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. J.* **2019**, *82*.
 16. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent Dirichlet allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.
 17. Chiarello, F.; Melluso, N.; Bonaccorsi, A.; Fantoni, G. A text mining based map of engineering design: Topics and their trajectories over time. In Proceedings of the Proceedings of the International Conference on Engineering Design, ICED; 2019; Vol. 2019-Augus, pp. 2765–2774.
 18. Felizardo, K.R.; Salleh, N.; Martins, R.M.; Mendes, E.; Macdonell, S.G.; Maldonado, J.C. Using visual text mining to support the study selection activity in systematic literature reviews. In Proceedings of the International Symposium on Empirical Software Engineering and Measurement; 2011; pp. 77–86.
 19. Nahar, K.M.O.; Alhindawi, N.; Al-Hazaimah, O.M.; Alkhatib, R.M.; Al-Akhras, A.M. NLP and IR based solution for confirming classification of research papers. *J. Theor. Appl. Inf. Technol.* **2018**, *96*, 5269–5279.
 20. Elsevier Scopus, an Eye on Global Research. **2018**.
 21. Feng, Z.; Liang, M.; Chu, F. Recent advances in time-frequency analysis methods for machinery fault diagnosis: A review with application examples. *Mech. Syst. Signal Process.* **2013**, *38*, 165–205.
 22. Zhao, W.; Chen, J.J.; Perkins, R.; Liu, Z.; Ge, W.; Ding, Y.; Zou, W. A heuristic approach to determine an appropriate number of topics in topic modeling. *BMC Bioinformatics* **2015**, *16*.
 23. Shivashankar, S.; Srivathsan, S.; Ravindran, B.; Tendulkar, A.V. Multi-view methods for protein structure comparison using latent dirichlet allocation. *Bioinformatics* **2011**, *27*.
 24. Lau, J.H.; Baldwin, T.; Newman, D. On collocations and topic models. *ACM Trans. Speech Lang. Process.* **2013**, *10*.
 25. Syed, S.; Spruit, M. Full-Text or abstract? Examining topic coherence scores using latent dirichlet allocation. In Proceedings of the Proceedings - 2017 International Conference on Data Science and Advanced Analytics, DSAA 2017; 2017; Vol. 2018-Janua, pp. 165–174.
 26. Leopold, E.; Kindermann, J. Text categorization with support vector machines. How to represent texts in input space? *Mach.*

Learn. **2002**, *46*, 423–444.

27. Sievert, C.; Shirley, K. LDAvis: A method for visualizing and interpreting topics. *Proc. Work. Interact. Lang. Learn. Vis. Interfaces* **2015**, 63–70.
28. Chuang, J.; Ramage, D.; Manning, C.D.; Heer, J. Interpretation and trust: Designing model-driven visualizations for text analysis. In Proceedings of the Conference on Human Factors in Computing Systems - Proceedings; 2012; pp. 443–452.
29. Yu, W.; Zhao, C. A multi-model exponential discriminant analysis algorithm for online probabilistic diagnosis of time-varying faults. In Proceedings of the 2017 IEEE 56th Annual Conference on Decision and Control, CDC 2017; 2018; Vol. 2018-Janua, pp. 5751–5756.
30. Tanatavikorn, H.; Yamashita, Y. Batch process monitoring based on fuzzy segmentation of multivariate time-series. *J. Chem. Eng. Japan* **2017**, *50*, 53–63.
31. Yong, G.; Xin, W.; Zhenlei, W. Fault detection for a class of industrial processes based on recursive multiple models. *Neurocomputing* **2015**, *169*, 430–438.
32. Bai, B.; Zhang, S.; Zhao, L.; Zhang, X.; Guo, L. Online recognition of the multiphase flow regime. *Sci. China, Ser. E Technol. Sci.* **2008**, *51*, 1186–1194.
33. Dovžan, D.; Škrjanc, I. Recursive fuzzy c-means clustering for recursive fuzzy identification of time-varying processes. *ISA Trans.* **2011**, *50*, 159–169.
34. Sun, C.; Yin, H.; Li, Y.; Chai, Y. A Novel Rolling Bearing Vibration Impulsive Signals Detection Approach Based on Dictionary Learning. *IEEE/CAA J. Autom. Sin.* **2021**, *8*, 1188–1198.
35. Zhang, X.; Zhao, J.; Li, H.; Ni, X.; Sun, F. Gearbox Fault Diagnosis Based on Multifractal Detrended Fluctuation Analysis and Improved K Means Clustering. In Proceedings of the Proceedings - 2018 Prognostics and System Health Management Conference, PHM-Chongqing 2018; 2019; pp. 527–531.
36. Yan, H.; Yang, W.; Zhang, H.; Tao, B.; Zheng, Y. Density peaks clustering based sub-phase partition and monitoring for batch process. In Proceedings of the Proceedings of 2017 IEEE 6th Data Driven Control and Learning Systems Conference, DDCLS 2017; 2017; pp. 297–301.
37. Gao, X.; Cui, N.; Qi, Y.; Wang, P. Fault detection of multi-phase batch process based on adaptive FCM. In Proceedings of the Proceedings of the 33rd Chinese Control Conference, CCC 2014; 2014; pp. 3088–3093.
38. Peng, C.; Junfei, Q.; Xiangyu, Z.; Ruiwei, L. Phase Partition and Fault Diagnosis of Batch Process Based on KECA Angular Similarity. *IEEE Access* **2019**, *7*, 125676–125687.
39. He, Q.; Wang, X.; Zhou, Q. Vibration sensor data denoising using a time-frequency manifold for machinery fault diagnosis. *Sensors (Switzerland)* **2013**, *14*, 382–402.
40. Yu, Q.; Li, J.; Li, Z.; Zhang, J. A clustering K-SVD-based sparse representation method for rolling bearing fault diagnosis. *Insight Non-Destructive Test. Cond. Monit.* **2021**, *63*, 160–167.

41. Li, J.; Yao, X.; Wang, H.; Zhang, J. Periodic impulses extraction based on improved adaptive VMD and sparse code shrinkage denoising and its application in rotating machinery fault diagnosis. *Mech. Syst. Signal Process.* **2019**, *126*, 568–589.
42. Wang, H.; Gao, J.; Jiang, Z.; Zhang, J. Rotating Machinery Fault Diagnosis Based on EEMD Time-Frequency Energy and SOM Neural Network. *Arab. J. Sci. Eng.* **2014**, *39*, 5207–5217.
43. Benítez-Pérez, H.; Benítez-Pérez, A. The use of wavelets feature extraction and self organizing maps for fault diagnosis. *Int. J. Innov. Comput. Inf. Control* **2010**, *6*, 4923–4936.
44. Cheng, G.; Chen, X.-H.; Shan, X.-L.; Liu, H.-G.; Zhou, C.-F. A new method of gear fault diagnosis in strong noise based on multi-sensor information fusion. *JVC/Journal Vib. Control* **2016**, *22*, 1504–1515.
45. Shao, R.; Li, J.; Hu, W.; Dong, F. Multi-fault clustering and diagnosis of gear system mined by spectrum entropy clustering based on higher order cumulants. *Rev. Sci. Instrum.* **2013**, *84*.
46. Inacio, M.; Lemos, A.; Caminhas, W. Fault diagnosis with evolving fuzzy classifier based on clustering algorithm and drift detection. *Math. Probl. Eng.* **2015**, 2015.
47. Fong, S.; Harmouche, J.; Narasimhan, S.; Antoni, J. Mean shift clustering-based analysis of nonstationary vibration signals for machinery diagnostics. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 4056–4066.
48. Wang, L.; Pan, J.; Shao, Y.; Zeng, Q.; Ding, X. Two new kurtosis-based similarity evaluation indicators for grinding chatter diagnosis under non-stationary working conditions. *Meas. J. Int. Meas. Confed.* **2021**, 176.
49. Gelman, L.; Kolbe, S.; Shaw, B.; Vaidhianathasamy, M. Novel adaptation of the spectral kurtosis for diagnosis of gearboxes in non-stationary conditions. In Proceedings of the 13th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, CM 2016/MFPT 2016; 2016.
50. Wu, J.; Shao, S.; Jiang, D.; Yao, H.; Ma, C.; Xu, K. Research on Adaptive Power Transmission Line Fault Inspection. In Proceedings of the IOP Conference Series: Earth and Environmental Science; 2021; Vol. 787.
51. Traore, M.; Duviella, E.; Lecoecue, S. Comparison of two prognosis methods based on neuro fuzzy inference system and clustering neural network. In Proceedings of the IFAC Proceedings Volumes (IFAC-PapersOnline); 2009; pp. 1468–1473.
52. Traore, M.; Chammas, A.; Duviella, E. Supervision and prognosis architecture based on dynamical classification method for the predictive maintenance of dynamical evolving systems. *Reliab. Eng. Syst. Saf.* **2015**, *136*, 120–131.
53. Zhang, J.; Jiang, N.; Li, H.; Li, N. Online health assessment of wind turbine based on operational condition recognition. *Trans. Inst. Meas. Control* **2019**, *41*, 2970–2981.
54. Zeng, R.; Zhang, S.; Zeng, R.; Shen, H.; Zhang, L. A method of fault detection on diesel engine based on EMD-fractal dimension and fuzzy C-mean clustering algorithm. In Proceedings of the Proceedings of the 29th Chinese Control and Decision Conference, CCDC 2017; 2017; pp. 7679–7683.
55. Wang, C.J.; Li, H.Y.; Xiang, W.; Zhao, D. *A new signal classification method based on EEMD and FCM and its application in bearing fault diagnosis*; 2014; Vol. 602–605; ISBN 9783038351948.

56. Tian, G.; Tian, H.; Liu, G.-S.; Zhao, J.-H.; Luo, L.-P. *Compound-fault diagnosis of bearing based on order tracking wavelet packet and rough sets*; 2012; Vol. 130–134; ISBN 9783037852866.
57. Traore, M.; Duviella, E.; Lecoecue, S. Dynamical clustering technique to estimate the probability of the failure occurrence of process subjected to slow degradation. In Proceedings of the ICINCO 2009 - 6th International Conference on Informatics in Control, Automation and Robotics, Proceedings; 2009; Vol. 2 RA, pp. 360–365.
58. Huang, N.; Fang, L.; Cai, G.; Xu, D.; Chen, H.; Nie, Y. Mechanical fault diagnosis of high voltage circuit breakers with unknown fault type using hybrid classifier based on LMD and time segmentation energy entropy. *Entropy* **2016**, *18*.
59. Wang, Z.; lu, C.; Wang, Z.; Liu, H.; Fan, H. Fault diagnosis and health assessment for bearings using the Mahalanobis–Taguchi system based on EMD-SVD. *Trans. Inst. Meas. Control* **2013**, *35*, 798–807.
60. Zheng, J.; Huang, S.; Pan, H.; Jiang, K. An improved empirical wavelet transform and refined composite multiscale dispersion entropy-based fault diagnosis method for rolling bearing. *IEEE Access* **2020**, *8*, 168732–168742.
61. Ding, H.; Ding, K.; Zhang, J.; Wang, Y.; Gao, L.; Li, Y.; Chen, F.; Shao, Z.; Lai, W. Local outlier factor-based fault detection and evaluation of photovoltaic system. *Sol. Energy* **2018**, *164*, 139–148.
62. Jiang, W.; Zhu, Y.; Wang, M. Fault diagnosis method for hydraulic pump based on order tracking. In Proceedings of the Proceedings of 2015 International Conference on Fluid Power and Mechatronics, FPM 2015; 2015; pp. 1276–1280.
63. Baozhen, C.; Hongxia, P. The method of scale - Energy fuzzy clustering based on series wavelet analysis in Studying fault diagnosis of gearbox. In Proceedings of the ICSPS 2010 - Proceedings of the 2010 2nd International Conference on Signal Processing Systems; 2010; Vol. 2.
64. Wodecki, J.; Stefaniak, P.; Michalak, A.; Wylomanska, A. Automatic calculation of thresholds for load dependent condition indicators by modelling of probability distribution functions - Maintenance of gearboxes used in mining conveying system. In Proceedings of the Vibroengineering Procedia; 2017; Vol. 13, pp. 67–72.
65. Wodecki, J.; Stefaniak, P.; Polak, M.; Zimroz, R. *Unsupervised anomaly detection for conveyor temperature SCADA data*; 2018; Vol. 9;.
66. Liu, Y.; Tang, B.; Duan, L.; Fei, F. Feature Extraction for Rolling Bearing Diagnosis Based on Improved Local Mean Decomposition. In Proceedings of the Proceedings - 2018 Prognostics and System Health Management Conference, PHM-Chongqing 2018; 2019; pp. 297–302.
67. Xu, Q.; Liu, K.; Xu, Z. *Research on Fractional Lower Order Feature Extraction of Bearing Vibration Signals Under Alpha Stable Noise Conditions*; 2019; Vol. 543; ISBN 9789811336621.
68. Xu, Q.; Liu, K. A New Feature Extraction Method for Bearing Faults in Impulsive Noise Using Fractional Lower-Order Statistics. *Shock Vib.* **2019**, 2019.
69. Panigrahi, B.K.; Samantaray, S.R.; Dash, P.K.; Panda, G. Discrimination between inrush current and internal faults using pattern recognition approach. In Proceedings of the 2006 International Conference on Power Electronics, Drives and Energy Systems, PEDES '06; 2006.

70. Cirrincione, G.; Hérault, J.; Randazzo, V. The on-line curvilinear component analysis (onCCA) for real-time data reduction. In Proceedings of the Proceedings of the International Joint Conference on Neural Networks; 2015; Vol. 2015-Septe.
71. Liu, J. Fault detection and classification for a process with multiple production grades. *Ind. Eng. Chem. Res.* **2008**, *47*, 8250–8262.