MDPI

Article

Machine Learning Approach for LPRE Bearings Remaining Useful Life Estimation Based on Hidden Markov Models and Fatigue Modelling

Federica Galli [1,2] , **Philippe Weber** [3], **Ghaleb Hoblos** [1,*] , **Vincent Sircoulomb** [1] , **Giuseppe Fiore** [2] and **Charlotte Rostain** [2]

1   UNIROUEN/ESIGELEC/IRSEEM, 76000 Rouen, France; federica.galli@groupe-esigelec.org (F.G.); vincent.sircoulomb@esigelec.fr (V.S.)
2   Centre National d'Etudes Spatiales (CNES), 75012 Paris, France; giuseppe.fiore@cnes.fr (G.F.); charlotte.rostain@cnes.fr (C.R.)
3   Centre de Recherche en Automatique de Nancy (CRAN), 54506 Nancy, France; philippe.weber@univ-lorraine.fr
*   Correspondence: ghaleb.hoblos@esigelec.fr

Abstract: Ball bearings are one of the most critical components of rotating machines. They ensure shaft support and friction reduction, thus their malfunctioning directly affects the machine's performance. As a consequence, it is necessary to monitor the health conditions of such a component to avoid major degradations which could permanently damage the entire machine. In this context, HMS (Health Monitoring Systems) and PHM (Prognosis and Health Monitoring) methodologies propose a wide range of algorithms for bearing diagnosis and prognosis. The present article proposes an end-to-end PHM approach for ball bearing RUL (Remaining Useful Life) estimation. The proposed methodology is composed of three main steps: HI (Health Indicator) construction, bearing diagnosis and RUL estimation. The HI is obtained by processing non-stationary vibration data with the MODWPT (Maximum Overlap Discrete Wavelet Packet Transform). After that, a degradation profile is defined and coupled with crack initiation and crack propagation fatigue models. Lastly, a MB-HMM (Hidden Markov Model) is trained to capture the bearing degradation dynamics. This latter model is used to estimate the current degradation state as well as the RUL. The obtained results show good RUL prediction capabilities. In particular, the fatigue models allowed a reduction of the ML (Machine Learning) model size, improving the algorithms training phase.

Keywords: RUL estimation; HMS; prognosis; HMM; rotating machines; bearings; reusable rocket engines

1. Introduction

Bearing Prognosis and Health Management (PHM) has a critical role in the field of rotating machines. First of all, a bearing is a key mechanical component of rotating machines. It provides shaft support, friction reduction, vibration denoising and other important features [1]. Its correct functioning is essential for the good functioning of the machine. As a consequence, its State of Health (SoH) needs to be estimated and predicted during the machine's operation to avoid failure and major damages.

According to the definition given by Hu et al. [2], PHM is a new piece of technology that can perform three main tasks: Fault/Anomaly Detection, Fault Diagnosis and SoH Prediction (prognosis) to support the maintenance decision process. Depending on when the PHM is integrated in the system or component life cycle, it can be used in different scenarios. The most common one consists in easing the maintenance planning. Thanks to the PHM, maintenance activities can be optimised according to the machine SoH, reducing costs and maximising the operating time. In addition, the knowledge acquired during

diagnosis and prognosis analysis can be used to guide refurbishing procedures during which a new set of components has to be chosen for the next mission [2]. In the most advanced scenarios, when the PHM is included from the initial product design phase, it can be applied for system control or operating point correction. Different PHM methodologies have already been proposed in the literature concerning the diagnosis and prognosis of bearings and gears in rotating machines. Their fields of application are wide and they include wind turbines, aeronautical systems and space propulsion. In order to perform the three tasks previously listed, PHM algorithms are characterised by a four-step general structure [3] including Data Collection, Health Indicator (HI) Construction, Health Stage (HS) Division and Remaining Useful Life (RUL) Estimation.

In the Data Collection step, the available monitoring data are gathered together with all the knowledge about the system. In the case of bearing PHM, the monitoring data can be collected with multiple sensors such as accelerometers, acoustic emission sensors, temperature sensors and encoders [4]. Vibration signals are generally preferred since accelerometers are reliable sensors and they can be easily installed on the bearing housing.

Depending on the type, quality and quantity of the available monitoring data, the HI is constructed and computed. As described by Lei et al. [3], HIs obtained from vibration signals can be categorised as Physical HIs and Virtual HIs. Physical HIs include all the classical statistical features like the Root Mean Square (RMS), the Kurtosis, the Crest Factor and the Impulse Factor [5]. Depending on the situation, the acquired vibration signal is treated with a sliding window or divided into segments. The HI is computed at different time instants and its evolution can be used to detect and diagnose faults or to predict failures. A deep understanding of the failure mechanism is required to establish a strong link between the variation of the HI and the failure mechanism (wear, crack propagation, spalling, etc.). To this matter, Jain et al. [6] carried out a detailed analysis of the bearing defect size variation effect on the most frequently used statistical indicators. Moreover, to cope with the masking of the degradation information, advanced signal processing techniques can be implemented to enhance the vibration content and to improve the HI computation like Empirical Mode Decomposition (EMD) [7] and Discrete Wavelet Packet Transform (DWPT) [8]. The second family of HIs, Virtual HIs, is preferred when the degradation phenomenon is too complex to understand, affecting the correct interpretation of the HI evolution over time. In this case, the HI is built using Artificial Intelligence (AI) techniques including Machine Learning (ML) [9], Deep Learning (DL) and Transfer Learning (TL) [2], etc. As an example, Mochammad et al. [10] proposed an unsupervised health indicator obtained through frequency analysis and Principal Component Analysis (PCA) feature fusion. While these techniques can extract the degradation information without the a priori degradation mechanism knowledge, their accuracy strongly depends on the quality and quantity of the available data. As explained by the authors in previous work [11], the dataset used for the algorithm training should contain clear information about the degradation dynamics during the entire life of the bearing.

Once the HI is defined, bearing diagnosis and prognosis can be performed. Based on the current HI value and its history up to the present time, the bearing SoH can be assessed to understand to what extent the component is damaged. Statistical tests as well as AI techniques can be used to perform these tasks. For example, Fan et al. [12] used a new type of CUSUM statistical test to perform Bearing Fault Detection. Chen et al. [13] proposed a deep transfer Convolutional Neural Networks (CNN) to perform bearing diagnosis. Konig et al. [14] presented a precise fault classification algorithm based on CNN and ML.

Lastly, the bearing RUL can be computed. The HI trend is captured and projected on a determined time horizon or until failure is reached. Thus, a predictive algorithm has to be used, which can be one of of three different types: model-based, data-driven or hybrid. Model-based prognosis approaches are characterised by mathematical models that describe the bearing degradation from a theoretical point of view. Since bearings are mostly subject to fatigue damages, fatigue models, like the Paris law, have been extensively used in the past to predict the size of bearing surface defects. For example, Behzad et al. [15]

combined the Paris Law with reliability methods. This kind of models requires a small amount of monitoring data and has been proven to be effective as far as crack initiation and propagation are correctly modelled. At the same time, their application is limited since they describe the degradation mechanism according to the degree of its understanding. Complex and highly non-linear study cases may not be correctly represented, leading to an estimation error during the RUL computation.

Data-driven prognosis approaches consist of the biggest group of predictive algorithms for RUL estimation. They include AI techniques (ML, DL, TL) as well as fuzzy logic techniques and statistical modelling techniques. They directly extract the degradation dynamics from the data without any previous knowledge, but their precision is directly affected by the quantity and the quality of the learning dataset. Haidong et al. [16] proposed an early fault prognosis algorithm based on complex wavelet transform and deep gated recurrent unit networks. Gao et al. [17] used PWT to construct an efficient HI and estimated the bearing RUL with an encoder–decoder Long-Short Term Memory (LSTM) neural network. Liu et al. [18] used a transformer model-based predictor to perform bearing prognosis under different degradation processes. Some interesting work was also conducted in the field of statistical modelling where Hidden Markov Model (HMM) approaches have been proposed for bearing diagnosis and prognosis. HMM models are classified as unsupervised ML models. In this field, Medjaher et al. [19] used GMM-HMM for bearing RUL estimation. Soave et al. [20] proposed a generalised GMM-HMM model for rotating machinery prognosis. Zhao et al. [21] developed a multi-feature HMM model for bearing diagnosis and prognosis.

Lastly, hybrid prognosis techniques try to increase the RUL estimation precision and reliability by combining the two previous types of algorithms. For example, Qian et al. [22] proposed a hybrid approach merging PTW and a modified Paris law to take into account fast and slow degradation dynamics.

Going further, PHM methodologies can be classified as real-time or offline methodologies. Real-time PHM methods collect the data and process them immediately before or while the next sample is available. This is the case for production lines and machine monitoring: the bearings are supervised in real time when the machine is functioning. In a different way, offline approaches collect the data during a specific time interval (missions) and treat them at the end of it. The PHM analysis is carried out in between missions and may serve as aid for maintenance planning and preparation for the next mission.

Problem Statement and Contributions

The space transportation sector has been revolutionised during the last decade by the arrival of reusable launchers and with them reusable Liquid Propellant Rocket Engines (LPRE). LPRE are very complex systems and their reusability introduces some new challenges. Engine re-ignition, trust modulation and maintenance planning are just a few of them. As far as maintenance planning is concerned, the Space Shuttle program revealed how extremely expensive and time consuming this kind of procedures can be. For this reason, PHM is a promising discipline since it allows us to optimise the maintenance actions and their planning: action will be taken only when required, towards the system end-of-life before any irreversible failure takes place. Some studies have been carried out in the field of rocket engine PHM or HMS. Park et al. [23] proposed a DNN for fault detection and diagnosis during the startup transient stage of liquid-propellant rocket engine tests. Chelouati et al. [24] proposed an extended Kalman filter to predict the SoH of a LPRE in case of crack propagation in the combustion chamber. Wang et al. [25] proposed a detailed survey of diagnosis techniques applied to rocket engines or rocket engines subsystems. Most of the research performed in this specific field concerns on-line or real-time approaches in the domain of Health Aware Control. In this case the RUL serves as key information to adjust the engine operating point. Not a lot of works have been carried out concerning LPRE turbopump bearings, except for the contribution of Pan et al. [26] where

a complex data-driven approach with meta pruning framework with attention augmented convolutions is proposed for LPRE bearing RUL estimation.

In this context, we propose an offline PHM methodology to estimate the RUL of a reusable LPRE turbopump bearing. The offline approach is motivated by the fact that the estimated RUL is intended for maintenance guidance. As a consequence no real time estimation is necessary since the computation is performed at the end of each mission. The bearing acceleration signals are decomposed using the Maximum Overlap Discrete Wavelets Packet Transform (MODWPT) and the nodal energy HI is computed. Then, diagnosis and prognosis of the bearing are performed through the use of a physics-defined Hidden Markov Model (HMM). In particular, the bearing RUL is obtained through the computation of the Markov Chain (MC) absorption time. The present work is characterised by the following contributions:

- A new methodology is proposed which is tailored and adapted to LPRE RUL estimation for preventive maintenance planning. Indeed, previous works approached LPRE from a global point of view, taking into account the entire system. When the final objective is Health Aware Control, a global approach is suitable, but in the case of maintenance planning a local approach focusing on a specific components is better, since specific degradation phenomena have to be taken into account which may not be visible if the system is monitored globally. The proposed method takes into account also other challenging aspects such as the limited amount of data and the difficult operating conditions. The application chosen by the authors seeks a yet unexplored research area .
- A degradation profile is extracted from existing data and validated with respect to fatigue models which was not presented in the literature before according to the authors' knowledge.
- The proposed methodology is based on a Multi-Level MB physics-driven HMM model. Even thought HMM are not a new piece of technology, this new version of them coupled with a physical degradation model of the bearing fatigue degradation proposes a new and efficient alternative to classical methods.

The rest of the article is organised as follows: Section 2 describes the PHM methodology step by step. Section 3 presents the simulation set-up and the obtained results when the proposed approach is applied to synthetic data. Section 4 reports the results obtained with real run-to-failure experimental data. Finally, Section 5 draws the conclusions.

2. PHM Approach

In this section, the developed approach is described in detail from the initial step of data collection until RUL estimation. The proposed data-driven approach is deployed in two main steps: unsupervised training and algorithm exploitation. In the first step, each collected acceleration signals is processed to construct its respective HI. The obtained HIs are used to train a family of multiple HMM whose structure was defined according to bearing qualitative fatigue models. During the second step, new signals are treated to compute the necessary HI and used to compute the bearing RUL. The HMM model that best represents the current HI is chosen from the trained models and the RUL is obtained through the computation of the MC absorption time. Figure 1 shows the main steps from data collection to RUL estimation. The proposed methodology is built on an off-line approach which means that it is not applied during the functioning of the bearing but right after. The unsupervised learning step is performed before the first bearing utilisation, while the algorithm exploitation is performed at each interruption of the bearing functioning.

2.1. Data Collection

The vibration produced by the bearing motion is captured by an accelerometer placed on the bearing housing. Samples of duration d_s are registered during a sampling interval p_s and with a sampling frequency f_s. Figure 2 gives an example of the data acquisition

process where $p_s = 10$ s, $d_s = 0.1$ ms and $F_s = 25.6$ kHZ. Finally, all the samples are chained to form one long signal.
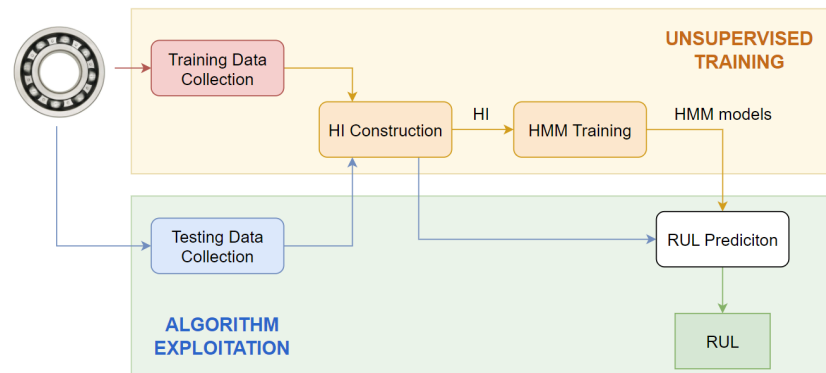

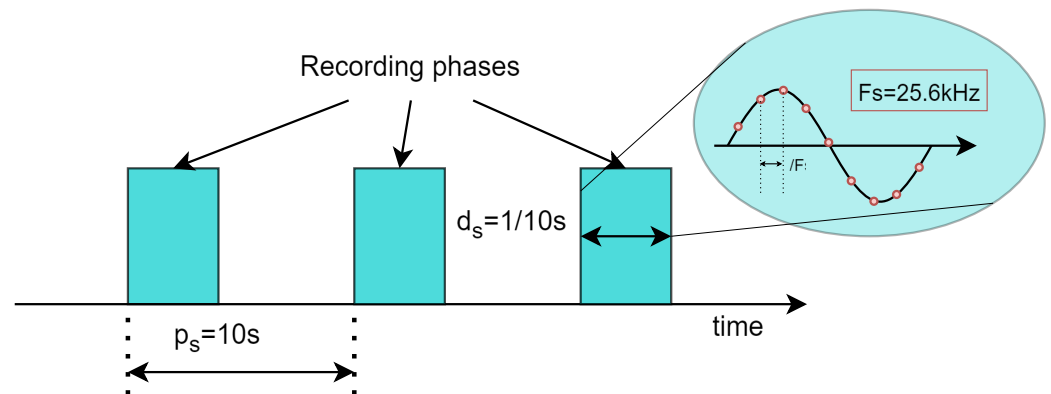**Figure 1.** General methodology scheme.

Figure 2. Data collection example [27].

2.2. Health Indicator Construction

Figure 3 shows all the steps concerning the HI construction. At first, each bearing acceleration signal is decomposed using the MODWPT (step 1). Then, the obtained sub-signals are used to compute the nodal energy HI, also known as relative energy (step 2). The obtained HIs were analysed and they led to the identification of a common recurrent degradation profile clearly showing the degradation evolution with respect time (step 3). At this point the HI are prepared for prognosis. A discretization technique is applied to enhance the degradation information (step 4). Eventually the obtained HIs are broken down into clusters according their degradation rate and sensitivity to the degradation (step 5). Having completed step 5 the HIs are ready for HMM training and RUL prediction.

2.2.1. Signal Processing (MODWPT)

In this second step, the bearing vibration signal is decomposed using the MODWPT signal processing technique. This specific type of wavelet transform was chosen since it allows us to decompose the raw signal on all its frequency spectrum without losing the time-dependent information contained in the data. Moreover, this technique allows us to avoid down-sampling thanks to the wavelets filter adaptation at each decomposition level and to the padding of the filtered signals [28]. To ensure the conservation of the signal energy the *db5* mother wavelet was chosen, which is an orthogonal function. The decomposition frequency bands can be computed using Equation (1), where f is the index of the frequency band fb and L is the decomposition level [29].

$$fb_f = \left[(f-1)\frac{F_s}{2^{L+1}}; (f)\frac{F_s}{2^{L+1}}\right] \quad \text{with} \quad f = 1, ..., 2^L \tag{1}$$
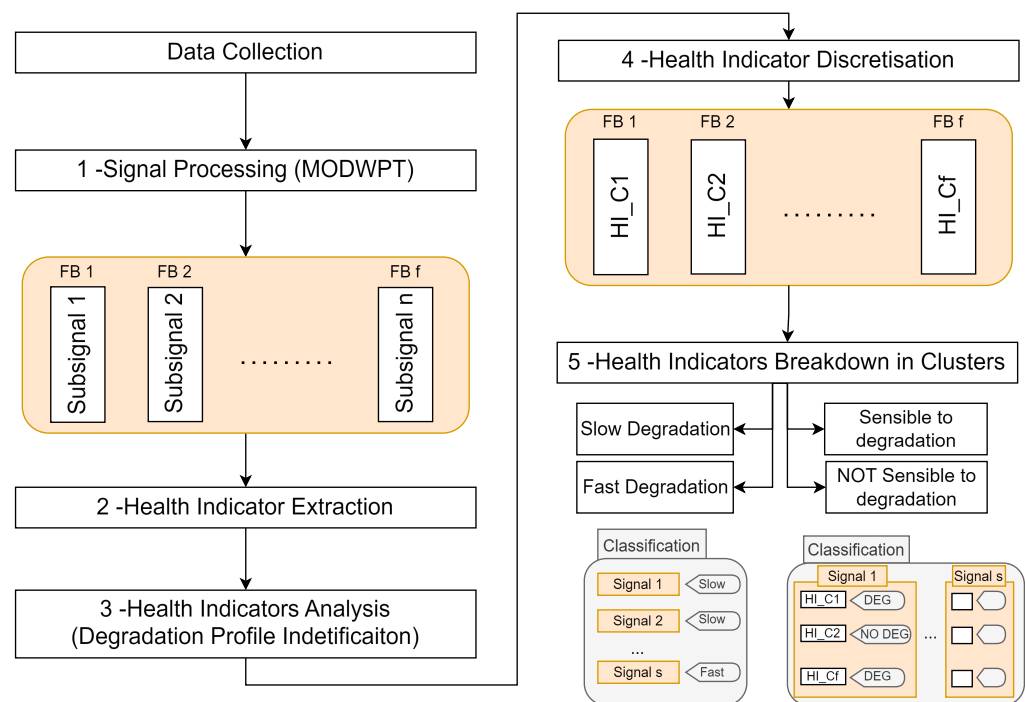
Figure 3. HI construction steps.

2.2.2. Health Indicator Extraction

The sub-signals obtained during the previous steps are now used to compute the HI using a sliding window of length w. After a careful analysis of the available statistical indicators, it was decided to proceed with the computation of the nodal energy, or relative energy, EN_{rel} as shown by Equation (2) [11]:

$$EN_{rel}(t,f) = \frac{En_{fb_f}}{En_{tot}} = \frac{\sum_{j=t}^{t+w} |x_{s,f}(j)|^2}{\sum_{j=t}^{t+w} |x_s(j)|^2} \tag{2}$$

where EN_{fb_f} and EN_{tot} are the energy of the sub-signal $x_{s,f}$ and the raw signal x_s, respectively; f is the frequency band and t is time.

2.2.3. Health Indicator Analysis

The identified degradation profile is composed of three stages, as depicted in Figure 4:

1. Nominal stage: in this early life stage, the bearing is functioning without defect. The HI is constant with small oscillations.
2. Degradation stage: here the degradation begins. The defect is initiated and slowly propagates. The HI shows an increasing or decreasing trend with variable levels of oscillations. Towards the end of this stage, the profile tends to a constant trend. Indeed, in some situations, due to friction smoothing or other factors, the degradation is momentarily slowed down. Anyway, the degradation is not stopped and it is just a matter of time before failure is reached.
3. Failure: eventually the bearing reaches its end-of-life. The HI starts to vary with a random behaviour, without a clear logic.

The change points between one stage and another are marked with blue dots in Figure 4 and are called t_1 and t_2. The first parameter, t_1, represents the duration of the nominal phase. At $time = t_1$ the degradation starts. The second parameter, t_2, marks the end of the degradation phase and the bearing end of life.
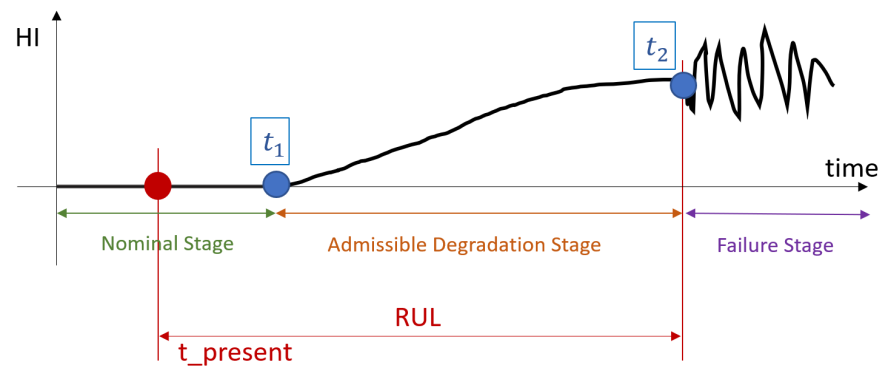
Figure 4. Bearing experimental degradation profile [11].

Depending on how the degradation affects the bearing vibration signature this profile can be observed at lower or higher frequencies at a specific frequency band [11]. Anyway, sometimes multiple frequency bands may be sensitive to the degradation. In these cases the wavelet decomposition level should be adapted to avoid cutting the spectrum near the solicited frequencies, and the obtained HIs should be analysed again carefully.

2.2.4. Health Indicator Discretisation

To be a suitable HI for HMM-based bearing diagnosis and prognosis, a discretization step is necessary. Indeed, classical HMM algorithms are suitable only for discrete observation sequences. The chosen discretization technique consists defining two discrete observation sequences for each HI and combining them.
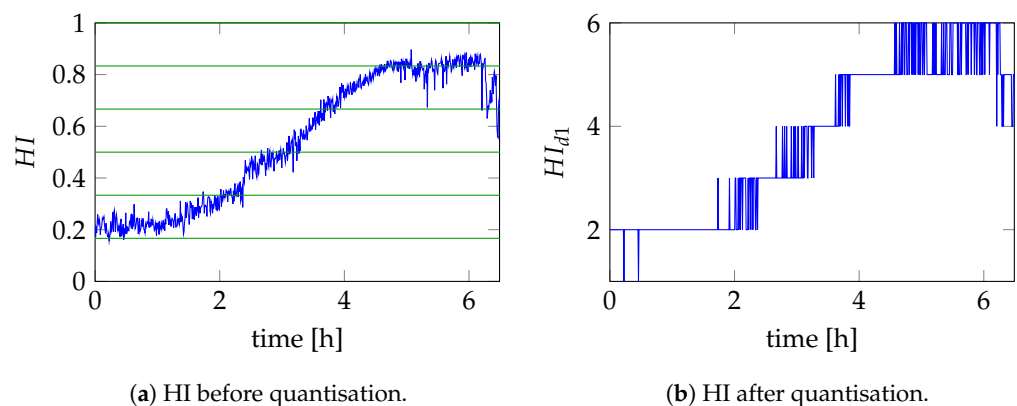
A   Discrete Sequence 1, H_{d1}: the first discrete sequence is obtained by quantisation of the HI. Since the nodal energy is a normalised quantity, its value is included in the interval $I = [0, 1]$. Such an interval is divided into sub-intervals of equal length as shown hereafter:

$$I = I_1 \cup I_2 \cup ... \cup I_\gamma \tag{3}$$

where γ is the number of sub-intervals I. Sub-intervals from I_1 to $I_{\gamma-1}$ are right open intervals while I_γ is closed on both sides. Each element of the HI is replaced by the index of the sub-interval including it:

$$HI_{d1}(t) = \iota \tag{4}$$

where $\iota = \mathrm{argmin}_\iota |HI - I_\iota|$. In other words, ι is the index of the interval that minimises the absolute difference between the health indicator HI and the boundaries of the intervals. This means that ι is the index of the interval to which HI belongs. Figure 5a,b illustrates a practical example.

(**a**) HI before quantisation.

(**b**) HI after quantisation.

Figure 5. HI discretization—first sequence.

B Discrete Sequence 2, H_{d2}: this discrete sequence is defined based on the DCS failure detection:

$$HI_{d2}(t) = \begin{cases} 1 & \textit{Failure not detected} \\ 2 & \textit{Failure detected} \end{cases} \tag{5}$$

For each initial raw signal, a family of 2^L HIs is obtained. Depending on the degradation phenomena that affected the bearing, the degradation impacts the vibration signal on the low, medium or high frequencies. This behaviour can be clearly observed thanks to the MODWPT decomposition. As stated above, the vibration signal becomes random when failure occurs. With the chosen HI, it results in an abrupt variance increase or by strong oscillations of the mean. The Dynamic Cumulative Sum (DCS) statistical test was chosen to perform failure detection for its good capability of detecting these types of signal variations both in the variance and the mean. The DCS test computes the likelihood ratio between signal segments as shown in Equation (6) [30]:

$$DCS_{S_a,S_b}(t) = \sum_{k=1}^{w} 0.5 \cdot \left(log \left(\frac{\sigma_a^2}{\sigma_b^2} \right) + (S_b(k) - \mu_b)^2 \cdot \left(\frac{1}{\sigma_a^2} - \frac{1}{\sigma_b^2} \right) \right) \tag{6}$$

In this equation, w is the window length, S_a is the window after time t, S_b is the window before time t, σ is the standard deviation and μ is the mean. When the segments S_a, S_b present similar distributions, the ratio is close to one while in the opposite case its value increases. Let us consider two sliding contiguous windows; it was demonstrated in [30] that the DCS likelihood ratio will reach its maximum when the change point, cpt, is reached. As a consequence, a decision function, g, can be established to perform the change point detection as shown in Equation (7) [30].

$$g(t) = max(DCS) - DCS(t) \tag{7}$$

Figure 6 illustrates an example of the DCS test applied to a signal that changes in mean and variance. As it can be noticed, the DCS will detect multiple change points, thus an additional analysis has to be carried out to select the failure change points. To do this two actions are taken. At first, the optimal threshold is defined using the experimental ROC curve [31]. The threshold is optimal since it minimises the false positive rate and maximises the true positive rate. Subsequently, the detected change point are analysed by the authors and classified manually. This procedure is applied to the data used during the training phase. During the testing, or in real life conditions, failure detection is not performed. Indeed, the bearing is not supposed to reach failure since as soon as its remaining useful life is shorter that the expected mission duration it will be replaced. If any failure happened during the mission, the consequences would be catastrophic and no RUL estimation procedure would be necessary.
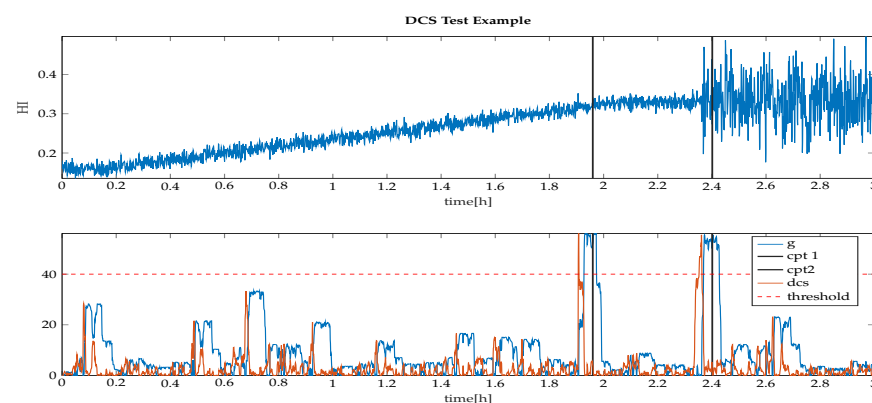

Figure 6. DCS statistical test example, cpt_2 is the failure change point.

C Discrete HIs combination, HI_C: at last, HI_{d1} and HI_{d2} are combined. Table 1 gives an example of the combination process. Given γ and knowing that the second discrete sequence can only have two values, the y vector containing all the possible combinations can be defined, which will contain $M = \gamma \cdot 2$ elements. For example, if $\gamma = 4$:

$$y = [11, 12, 13, 14, 21, 22, 23, 24] \tag{8}$$

Table 1. Sequence combination example.

HI_{d1}	1	1	2	3	2
HI_{d2}	1	1	1	2	2
HI_C	"11"	"11"	"12"	"23"	"22"

Figure 7a reports an example of the two discrete sequences, HI_{d1} and HI_{d2}. HI_{d1} conserved the information about the HI evolution trend, in this case monotonically increasing, while HI_{d2} contained specific information about its failure. If taken singularly HI_{d1} or HI_{d2} are not very strong HIs. Indeed, as observed during the HI analysis, the degradation provokes a change in mean and variance in the relative energy, which is partially lost during the HI quantisation. The increasing trend is still noticeable in HI_{d1} but the failure time is more difficult to spot. HI_{d2} contains only information about the failure time and no increasing trend is observable. Consequently, the discrete sequence HI_{d2} can only reveal if failure occurred or not and it cannot provide any information about how the degradation evolved. HI_{d1} or HI_{d2} contain complementary information. As a consequence, their combination provides a strong HI containing both information about the degradation process and the failure time. Figure 7b shows an example of HI_C.
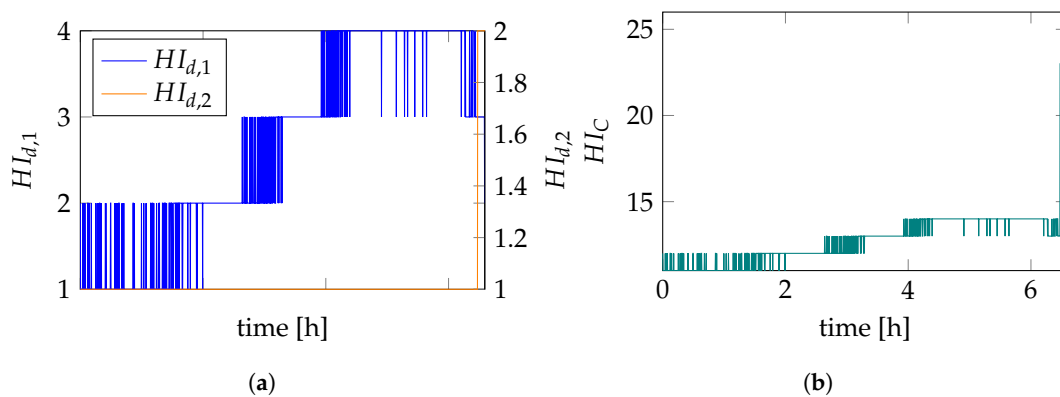

Figure 7. HI discretisation—combination. (**a**) Discrete sequences HI_{d1} (blue) and HI_{d2} (orange). (**b**) Combined discrete sequence HI_C.

2.2.5. Health Indicators Breakdown in Clusters

The obtained HI were carefully analysed and broken down into categories in order to organise the information contained in them and ease the training of the predictive model. Two different classifications were performed according to the following criteria:

- Degradation Speed: the degradation process may take place at different speeds. In this work a faster degradation dynamic damaging the bearing in approximately 2.5 h and a slower dynamic damaging the bearing in approximately 5 h are considered. Signals are labelled as "*SLOW*" or "*FAST*". These values were chosen based on the degradation behaviour observed during two different test campaigns [27,32].
- Sensitivity of the HI to the presence of degradation: as explained in Section 2.2.1, the HI are computed after having performed the MODWPT. The application of this signal processing technique allows us to isolate the effect of the degradation in a specific frequency band enhancing the variation of the health indicator. As a consequence,

some frequency bands are more sensitive than others to perform RUL estimation. The application of this classification criterion is based on the assumption that given stable working conditions (speed and radial load), the health indicator affected by the presence of a degradation will show significant variations while the others will remain stable. HIs are labelled as "*SENSIBLE*" and "*NOT-SENSIBLE*".

All the available data are classified according to the described criteria and two sets of clusters are obtained. Each one has a specific purpose. The clusters obtained with the first criteria are used to select a specific structure of the predictive model (HMM) that will know a priori that two types of degradation dynamics exist, a faster and a slower one. More details can be found by the author in the following Section 2.3.1. The clusters obtained from the second criterion are useful to select the right HI for RUL prediction as explained in Section 2.3.3.

2.3. RUL Prediction

2.3.1. Multi-Level Multi-Branch Hidden Markov Models

HMMs are an extension of Markov Chains. A Markov chain is a probabilistic model containing information about sequences of random variables [33] based on the following Markov assumption [34]: "*When predicting the future (time t + 1), the past (time t − 1) does not count, only the present (time t) counts*". Given a list of states $S = [S_1, ..., S_N]$ and a sequence of observed states, $Q = [q_1, ..., q_T]$, the previous assumption can be formalised as reported in Equation (9):

$$P(q_i = a|q_1, ..., q_T) = P(q_i = a|q_{i-1}) \tag{9}$$

MC are used to describe observable discrete stochastic processes. When the process is not observable, the MC model is extended to a HMM [34,35]. Assuming the chosen observation variable varies with the state and that it depends only on the state that generated it:

$$P(o_i|q_1, ..., q_i, ..., q_T, o_1, ..., o_i, ..., o_T) = P(o_i|q_i) \tag{10}$$

Given a list of states, $S = [S_1, ..., S_N]$ and a list of observations $O = [o_1, ..., o_M]$, the classical HMM model, $\lambda = (A, B, \pi)$, is defined by the following elements [34]:

1. Transition matrix (A): N × N matrix, where N is the number of hidden states. The element $a_{i,j}$ is the probability of going from state *i* to state *j*:

$$a_{i,j} = P(q_t = S_j|q_{t-1} = S_i), 1 < i, j < N \tag{11}$$

$$\sum_{j=1}^{N} a_{ij} = 1 \tag{12}$$

2. Emission Matrix (B): N × M matrix, where M is the number of possible observations. The element $b_j(k)$ is the probability of measuring observation k, $1 < k < M$, with hidden state *j*:

$$b_j(o_k) = P(O(t) = o_k|q_t = S_j), 1 \leq k \leq M \tag{13}$$

$$\sum_{k=1}^{M} b_j(o_k) = 1 \tag{14}$$

3. Initial distribution vector (π) : 1 × N vector. The element π_i is the probability of being in state *i* at the initial time:

$$\pi_i = P(q_1 = S_i), 1 < i < N \tag{15}$$

Depending on how the HMM is applied, three different problems can be defined [34]:

1. Likelihood Problem: given the observations sequence $O = [o_1, ..., o_M]$ and the model λ, compute $P(O|\lambda)$, the probability of observing sequence $O = [o_1, ..., o_M]$ given λ.

2. Decoding Problem: given the observations sequence $O = [o_1, ..., o_M]$ and the model λ, find the most probable sequence of hidden states, $Q = [q_1, ..., q_T]$.

3. Learning Problem: given the observations sequence $O = [o_1, ..., o_M]$ and a sequence of hidden states $S = [S_1, ..., S_N]$, find the model λ.

2.3.2. Model Definition

Many variations of the classical HMM model previously described have been proposed over the years. Among them, it is possible to find I/O HMM models, Semi HMM, GMM HMM, Hierarchical HMM, etc. The HMM chosen for the present work is a Multi-Level Multi-Branch Physics-defined HMM:

- Physics-defined: The classical formulation from [34] is kept, but the structure defining the A matrix is defined according to a fatigue model describing the bearing degradation process. The qualitative degradation model proposed by [36] was chosen for this purpose. As shown in Figure 8, the model is composed of five degradation stages: running-in, steady state, defect initiation, defect propagation and damage growth. Assuming that the running-in stage is much shorter than the others, the selected model can be reduced to four degradation stages. On this basis, a left-to-right HMM model can be defined with the structure presented in Figure 9. In this scheme, no backward transition is allowed because the degradation process is believed to be irreversible. The last state, S_4, is called the absorbent state since once it is reached it is not possible to leave it.
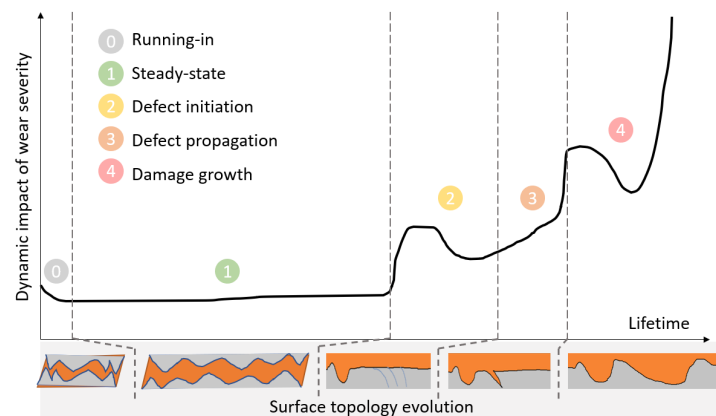

Figure 8. Dynamic behaviour and surface topology due to wear evolution [36].

- Multi-Branch: Classical HMM models extract the average dynamics according to the defined states and the available observation. When the observations are generated by different dynamics, a simple HMM is not enough to represent the hidden stochastic process. Indeed, its parameters would be trained to be somewhere in the middle or closer to the most frequent degradation dynamic. Following the steps of Le et al. [35], a multi-branch HMM is proposed to catch the multiple dynamics that may provoke the bearing degradation which is composed of b left-to right HMM models. Figure 9 shows the general structure of the proposed model.

- Multi-Level: Depending on the working conditions but also on how the degradation process solicited the bearing harmonics, it was decided to introduce a MB-HMM model for each frequency band obtained during the wavelets decomposition (step 2 in Figure 1). In this way, the proposed predictive model is provided with enough flexibility to isolate and predict the bearing degradation. Figure 10 shows a small scheme of the general structure of the proposed predictive model.
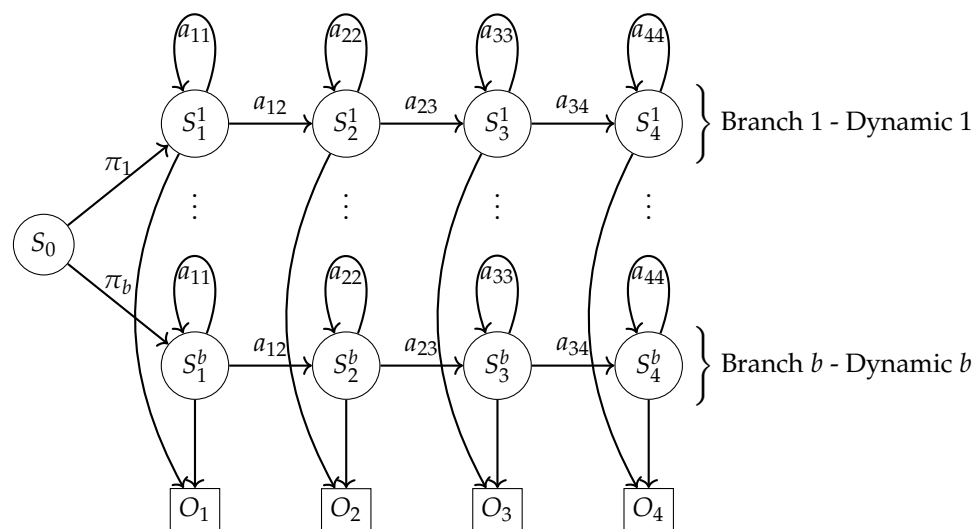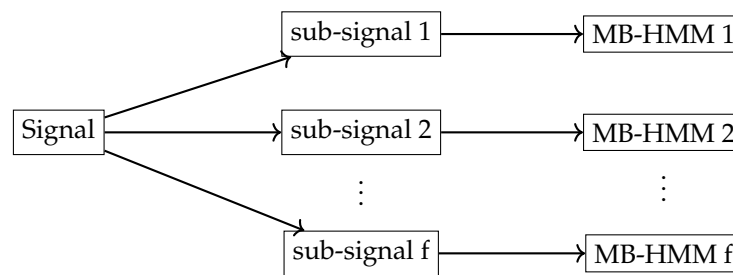
Figure 9. Proposed MB-HMM model.

Figure 10. Multi-Level MB-HMM Structure.

Given the scheme in Figure 9 and the HI discretization proposed in Section 2.2.4, it is possible to write the general formulation, applicable to each branch b in the MB-HMM model, of matrix A_b and B_b as reported by Equation (16):

$$
A_b = \begin{bmatrix} a_{1,1} & a_{1,2} & 0 & 0 \\ 0 & a_{2,2} & a_{2,3} & 0 \\ 0 & 0 & a_{3,3} & a_{3,4} \\ 0 & 0 & 0 & 1 \end{bmatrix}; B_b = \begin{bmatrix} b_1(y(1)) & b_1(y(2)) & \dots & b_1(y(M)) \\ b_2(y(1)) & b_2(y(2)) & \dots & b_2(y(M)) \\ b_3(y(1)) & b_3(y(2)) & \dots & b_3(y(M)) \\ b_4(y(1)) & b_4(y(2)) & \dots & b_4(y(M)) \end{bmatrix}; \tag{16}
$$

The model λ_b of each branch in the Multi-Level MB-HMM is trained using the Baum–Welch Algorithm. The training dataset for each branch is chosen according to the classification performed on the available data.

2.3.3. Health Indicator Classification

Once the Multi-Level MB-HMM is ready, the signals in the testing dataset are treated by applying the same steps from Figure 3. A series containing f HIs is obtained for each of them. The HIs showing very low values of relative energies are discarded by using a threshold. Indeed, if the relative energy is too low, no clear variation can be observed, meaning that no degradation can be observed in that specific frequency band. The next step consists of understanding which health indicator among the remaining ones actually reveals the presence of fatigue degradation. To do this, two additional HMM models were trained. Both models are classical discrete left-to-right HMM models, λ_h and λ_d. The matrices A_h, A_d, B_h and B_d present the same structure as in Equation (16). The model λ_d is trained with HI from frequency bands where the degradation is clearly visible, while model λ_h is trained with HIs, whose profile did not present any significant energy variation.

Each new HI is classified in the following way:

1. The likelihood probabilities $P(O|\lambda_h)$ and $P(O|\lambda_d)$ are computed using the decoding algorithm [34].
2. The HI class is defined according to the model presenting the highest likelihood probability.

2.3.4. RUL Computation

Having selected the best HI, the RUL estimation can be preformed by using the corresponding trained Multi-Level MB-HMM, $(\lambda_1,...,\lambda_b)$. The posterior probability of each branch is computed using the Bayes theorem [35]:

$$P(\lambda_b|O) = \frac{P(O|\lambda_b)P(\lambda_b)}{\sum_{b=1}^{n_b} P(O|\lambda_b)P(\lambda_b)} \tag{17}$$

where n_b is the number of branches, $P(O|\lambda_b)$ is the likelihood obtained by solving the likelihood problem described in Section 2.3.1 and $P(\lambda_b)$ is the a priori probability computed using Equation (18).

$$P(\lambda_b) = \frac{number\ of\ sequences\ used\ for\ training\ \lambda_b}{total\ number\ of\ available\ sequences} \tag{18}$$

Once the branch with the highest probability is found, the Remaining Useful Life (RUL) is obtained based on the formulation proposed by [37] for Mean Time To Failure (MTTF) computation. In [37], the MTTF is computed using the transition matrix A which contains transition rates between the different states of the defined MC. Following the approach proposed in some previous work [38], the MTTF formulation from [37] is re-proposed for bearing RUL estimation. In this case, the matrix A contains the transition probabilities between multiple degradation states (transient states) converging to one failure state (absorption state). As a consequence, the proposed approach consists of estimating the RUL by computing the time to absorption of the HMM model defined in Figure 9. The detailed mathematical formulation together with its demonstration can be found by the reader in Appendix A. Equation (19) reports directly the matrix formulation used for MTTF computation [37].

$$MTTF = \frac{1}{\Delta} \begin{vmatrix} 0 & P_{01} & ... & P_{0l} \\ 1 & a_{11} & ... & a_{1l} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_{l1} & ... & a_{ll} \end{vmatrix} \tag{19}$$

As it can be observed, the MTTF depends on the initial condition probabilities. If the initial state is known, the P_{0i} probability becomes 1, otherwise an initial probability distribution $[P_{01}, ..., P_{0l}]$ has to be defined. In the present work, Equation (19) was used in an iterative way to estimate the absorption time at each time t. Given a discrete observation sequence (HI_C), the trained HMM model is used to obtain the posterior state probabilities ($P(S_i)$) by applying the Viterbi algorithm [34]. The vector $P(S_i)$ is used as the initial probability vector to compute the absorption time T_a given the present degradation state. Following a similar approach proposed by Shahin et al. [38], the new modified matrix equation is presented in Equation (20):

$$T_a(t) = \frac{\begin{vmatrix} 0 & P(S_j)(t) \\ 1 & P(S_j|S_{j-1}) \end{vmatrix}}{|P(S_j|S_{j-1})|} \tag{20}$$

Since Equation (20) computes the average time the system spends in each hidden state, T_f is constant if the degradation state does not change. To take into account the fact

that the bearing already spent time in the current degradation state, the estimation from Equation (20) is corrected with a time counter, d_t:

$$RUL_t = T_a(t) - d_t(t);\tag{21}$$

3. Application to Simulated Data

The performance of the algorithm was first tested on a synthetic dataset. In this section, the simulation set-up as well as the obtained results are described.

3.1. Simulation Set-Up

The model used for synthetic data generation consists of a 5 Degrees of Freedom (DoF) lumped parameters model capable of reproducing the bearing non-stationary vibration behaviour. As previously described by the authors [39,40], the chosen model is made of a set of motion equations to represent the inner ring, outer ring and bearing housing displacements and a set of equations coming from the Hertzian contact theory used to compute the contact force of the touching surfaces. This model includes a non-stationary degradation profile which allows us to take into account the evolution of surface defects that affect the bearing functioning. Figure 11 shows a scheme of the proposed model, while Figure 12 shows the structure of the simulation loop. In the presented scheme the orange blocks represent the model input parameters, the blue blocks represent the output and the red blocks represent the coupled variables computed during the simulation.
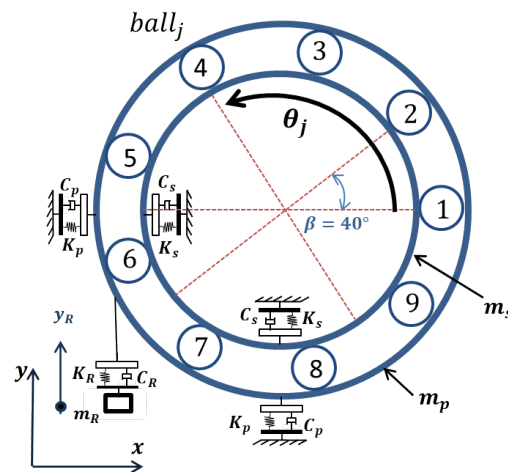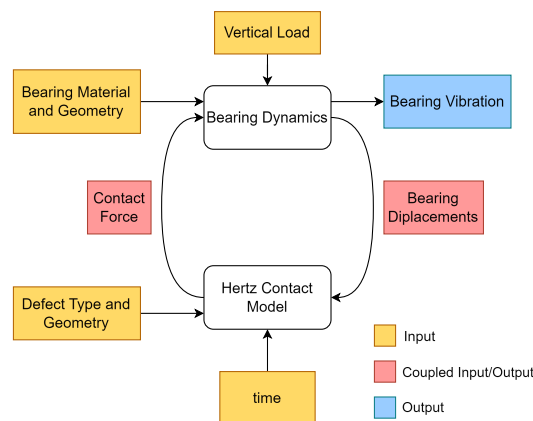

Figure 11. 5 DoF bearing model [39,41].

Figure 12. Simulation block diagram [39].

The proposed model can simulate a great variety of simulation scenarios characterised by different defect shapes and degradation evolution profiles. In this particular case, a

rectangular defect was considered on the Outer Ring whose width evolves over time with a linear profile. This very simple scenario was chosen because the objective of the synthetic data generation is not simply to produce bearing vibration signals similar to real ones but it is also to generate a meaningful dataset that can test the performance of the proposed algorithm. By choosing this simple scenario, the authors are able to generate vibration signals whose content is known in detail, reducing the dataset uncertainty level. This choice will ease the interpretation of the obtained results and will allow the identification of possible algorithm problems. A complete list of all the utilised bearing parameters is provided in Tables 2 and 3.

Table 2. Bearing parameters [39,42].

Sym	Quantity	Value	Sym	Quantity	Value
m_s	Shaft Mass	3.2638 kg	K_b	Load Deflection Factor	1.89×10^{10} N/m$^{1.5}$
m_p	Bearing Box Mass	6.638 kg	ϕ_{slip}	Ball Slip Angle	0.01 rad
m_R	Resonator Mass	1 kg	n_b	Ball Number	9
R_s	Shaft Damping	1.3768×10^3 Ns/m	D_p	Pitch Diameter	3.932×10^{-2} m
R_p	Bearing Box Damping	2.2107×10^3 Ns/m	D_b	Ball Diameter	7.94×10^{-3} m
R_R	Resonator Damping	9.4248×10^3 Ns/m	c	Bearing Clearance	0 m
K_s	Shaft Stiffness	7.42×10^7 N/m	α	Contact Angle	0°
K_p	Bearing Box Stiffness	1.51×10^7 N/m	F_y	Vertical Force	5500 N
K_R	Resonator Stiffness	8.8826×10^9 N/m	ω_s	Shaft Speed	600 rpm

Table 3. Defect parameters.

Sym	Quantity	Value	Sym	Quantity	Value
B	Defect Width	20×10^{-4} m	λ	Ratio of ϕ_1 to $\Delta\phi_d$	0.2
H	Defect Depth	6×10^{-4} m	k	Order of the Defective Ball	9
B_f	Defect width at failure	0.0165 m	t_{deg}	Failure time	[1.3–3.04] h
t_1	Degradation beginning instant	[1.3–1.5] h			

The rotational speed and the vertical load reported in Table 2 were selected during the model validation process [39]. The vertical load is in the order of magnitude of previous reusable rocket engines. For example, the SSME liquid oxygen turbopump bearings were submitted to this kind of radial loads [43]. The typical rotational speed of reusable rocket engines are much higher, ranging from 20,000 rpm to 90,000 rpm. The selected rotational speed for data generation, 600 rpm, is quite far from the classical nominal LPRE one, but given that the model has not been validated yet on higher speed simulation conditions, it was decided to first test the proposed algorithm on a lower rotational speed for which the correct vibration generation has been ensured before. Later on, with a transfer learning approach, the proposed methodology and model will be adapted to real rocket engine vibration signals. Table 3 includes all the parameters used to model the bearing surface defect. The defect is represented by a rectangular shape of width B and depth H. At the beginning of the simulation, no defect is present and a healthy signal is obtained. At time t_1 the degradation appears. The value of H is fixed, while the value of B increases over time as shown by the plots in Figure 13. The time t_{deg} is the moment when B becomes wide enough so that two bearing balls fall inside the defect simultaneously. This produces an interference in the contact forces that cancel each other out or are amplified in an unpredictable way. The authors estimated that this is the time at which failure occurs and used this value as reference to asses the precision of the RUL estimation.
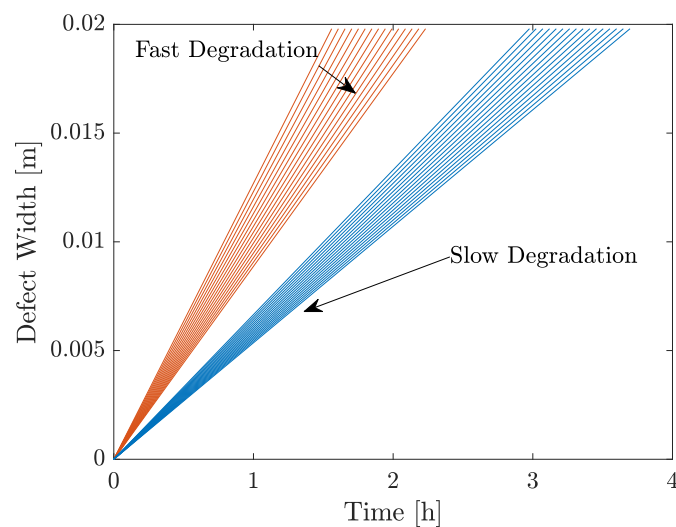
Figure 13. Degradation profile—evolution of the bearing defect width over time.

3.2. Computation Set-Up

The Multi-Level MB-HMM scheme from Figure 9 needs an a priori definition of the number of branches. In this case a MB-HMM model composed of two branches was chosen as presented in Figure 14. The available data were split in two groups of 30 signals and each branch was trained separately using the Baum–Welch algorithm. The transition matrices, A_1 and A_2, are 4×4 matrices since four hidden states were selected. The HI discretization was performed using six intervals. Thus the emission matrices, B_1 and B_2, are 4×12 matrices. Equation (22) shows the initialisation matrices A_0 and B_0.

$$A_0 = \begin{bmatrix} 0.7 & 0.3 & 0 & 0 \\ 0 & 0.7 & 0.3 & 0 \\ 0 & 0 & 0.7 & 0.3 \\ 0 & 0 & 0 & 1 \end{bmatrix}; B_0 = \begin{bmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{bmatrix}; \tag{22}$$

The initial transition matrix A_0 has the same structure as previously presented in Equation (16). The values on the diagonal were set to 0.7 and the other to 0.3 following the initial assumption that it is more probable to stay in the current state than transitioning. This hypothesis is in line with the proposed left-to-right HMM structure.

A bootstrap approach was selected to conduct the algorithm training and testing. One signal is selected from the training datasets and put aside. This signal is the testing signal. The remaining ones are processed and used to train the Multi-Level MB-HMM model. When the latter is ready, the testing signal is divided in segments whose length is equal to the length of a mission. A loop was implemented to retrieve a segment, chain it to the previous segments and compute the RUL. This approach reproduces the real-life scenario in which the reusable LPRE performs multiple missions and at the end of each mission the bearing SoH is verified.

The observed computation time to perform one loop is only a few minutes. Since the proposed method is applied off-line the observed computation time is not an issue. Of course, the RUL estimation process should be as quick as possible in order to avoid delays in the maintenance planning. Given an usual turnaround times of days or weeks in the case of reusable launchers, a computational time of some minutes is acceptable. The most time-consuming step is the wavelet decomposition, which is logical since the vibration signals are quite long. The rest of the steps concerning the HI construction and the RUL computation are fast. In particular, the authors have observed very quick RUL estimation times, $1 \text{ s} \leq t_{c,av} \leq 6 \text{ s}$. This time includes the MB-HMM training, the HI decoding and the RUL estimation. From a theoretical point of view, the computational complexity of the

training, decoding and RUL computation is directly linked to the size of matrix A and matrix B. The Baum–Welch training algorithm as an asymptotic complexity of *iter* $\times$ O(N^2t) and the decoding algorithm has an asymptotic complexity of O(N^2t), where *iter* is the number of iterations, N is the number of states and t is the time steps. The computation of the determinant has an asymptotic computation complexity of O(n^3) where *n* is the matrix size. Even though the presented asymptotic complexity is not very fast, the fact that the proposed HMM model has a minimised size compensates for that and in the end the computation time has an order of magnitude of a few seconds. Indeed, thanks to the reduced size of the *A* and *B* matrices the computation speed is increased for both algorithm training and testing. This aspect highlights the computation efficiency of the proposed algorithm, which is one of its main strengths. The interesting aspect of the proposed approach is that a complex degradation phenomenon could be approached with a small model. Indeed, given $N = 4$ and $M = 12$, a total of 48 parameters was estimated during the learning step, which is small compared to AI approaches characterised by hundreds of parameters.
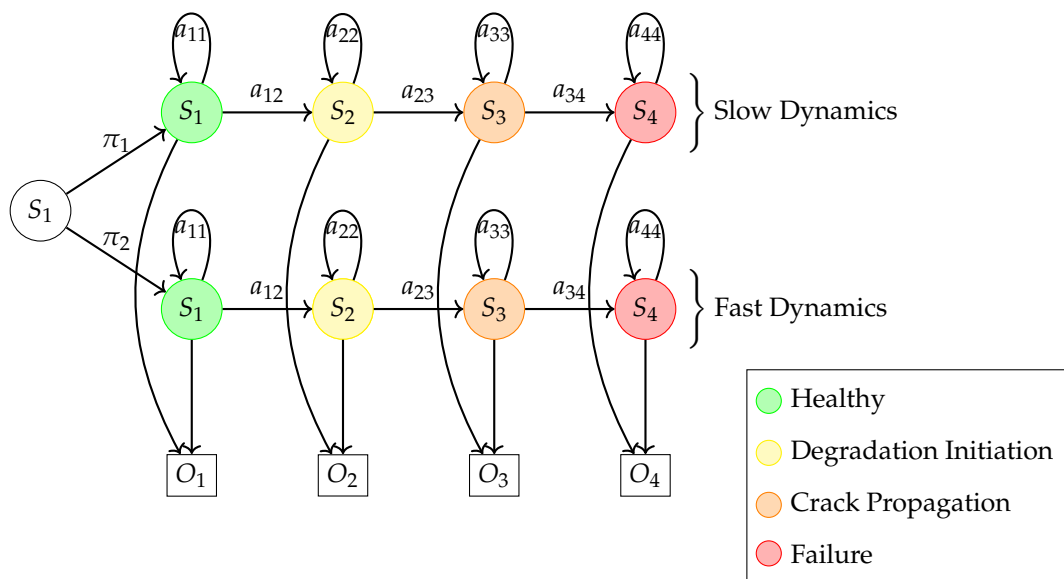

Figure 14. MB-HMM model.

3.3. Results and Discussion

Following the simulation set-up presented in the previous section, 116 bearing vibration signals were generated. Figure 15 shows, as an example, the vibration signal number 10. By looking at the figure, it is possible to clearly distinguish the vibration produced during the nominal functioning of the bearing and the vibration produced during its degraded operation. Starting from $t = 1.3$ h , the signal magnitude increases in a gradual way until $t = 2.78$ h when its behaviour becomes random. When the magnitude becomes unpredictable, failure is reached.
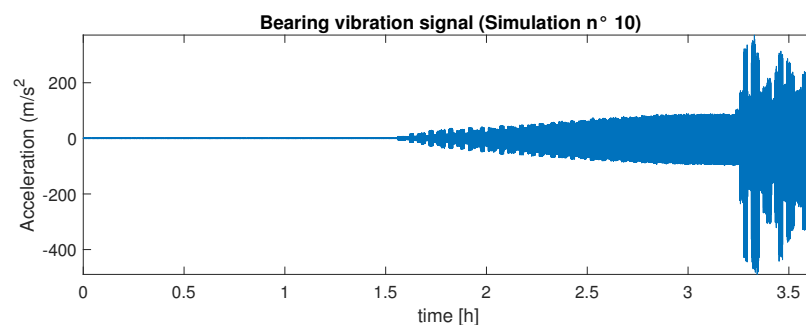

Figure 15. Example of a generated vibration signal, $t_1 = 1.3$ h, $t_{deg} = 1.48$ h.

The described signal was processed with the MODWPT and the obtained sub-signals were used to compute the family of health indicators whose plots are reported in Figure 16a–h. The analysis of these figures reveals that the higher energy density can be found at the lower frequencies, between 0 and 3.2 kHz. The first and second frequency bands are the most affected by the degradation. The health indicators in Figure 16a,b show clear signs of variation all along the bearing life. The other health indicators all show an abrupt decrease in energy at t_1 and remain constant before and after. As a consequence, they are not good candidates for RUL estimation and prediction.
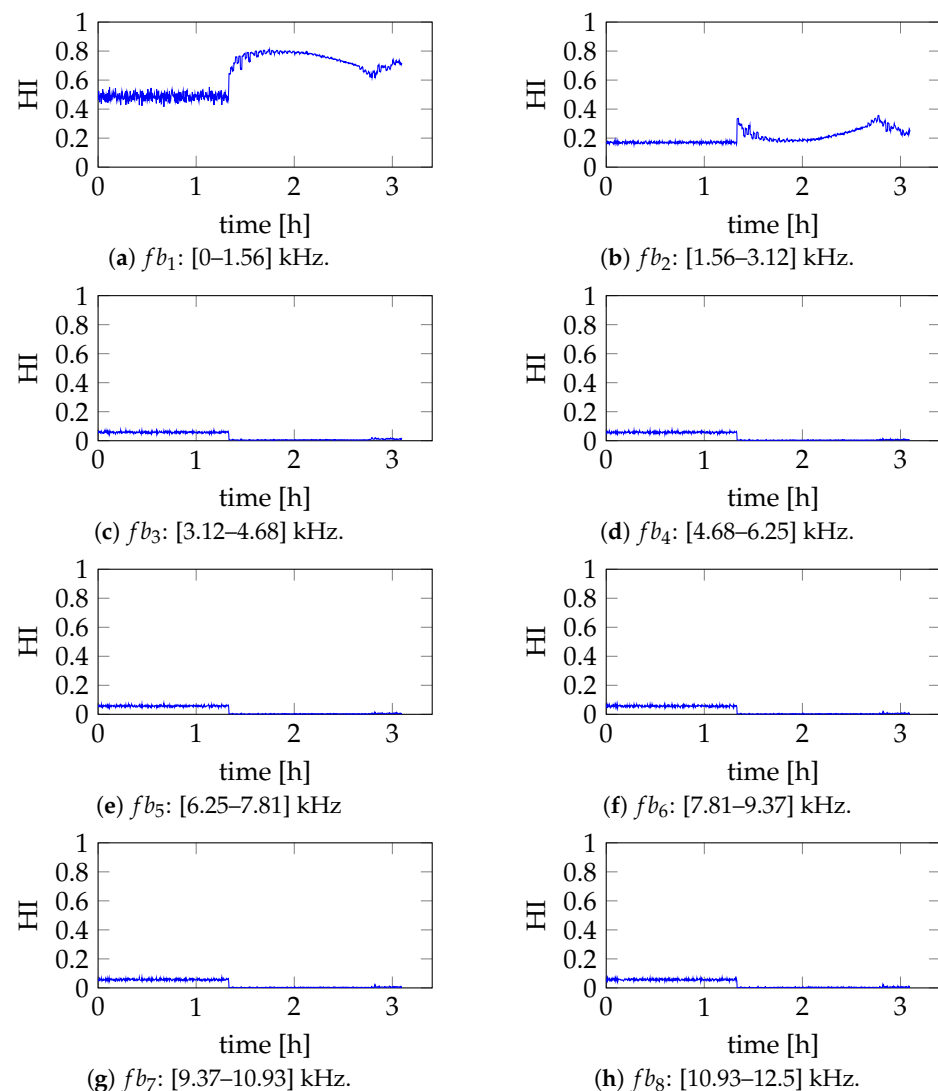

Figure 16. Relative energy health indicators computed for Signal 10 (Figure 15).

All the obtained HI underwent the failure detection step. Figures 17 and 18 show the result of the DCS statistical test for the HIs from Figure 16a and Figure 16b, respectively. The results concerning the other frequency bands are not reported here since no failure could be detected. In both Figures 17 and 18 the upper plot shows the concerned health indicator together with the detected failure. The lower plot reports the DCS statistical indicator (orange) and the decision function (blue). As can be observed, the DCS presents picks of different amplitude over time and the highest do not always corresponds to the time instant when failure occurs. Anyway, the failure is properly detected, showing that the combination of DCS statistical test described in Section 2.2.4 is capable of correctly performing the detection step even in such difficult conditions.
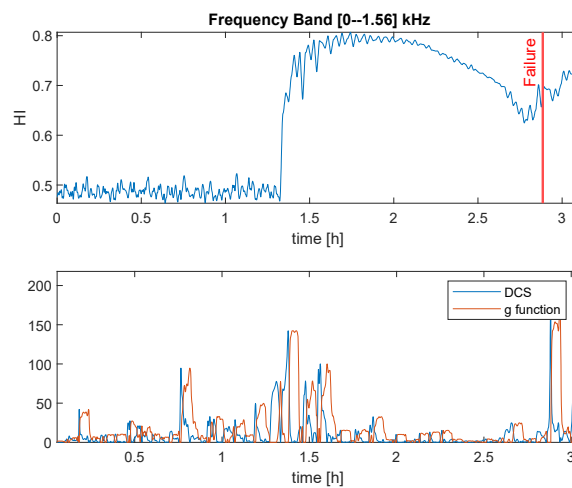
Figure 17. DCS statistical test applied to the HI from the frequency band [0–1.56] kHz.
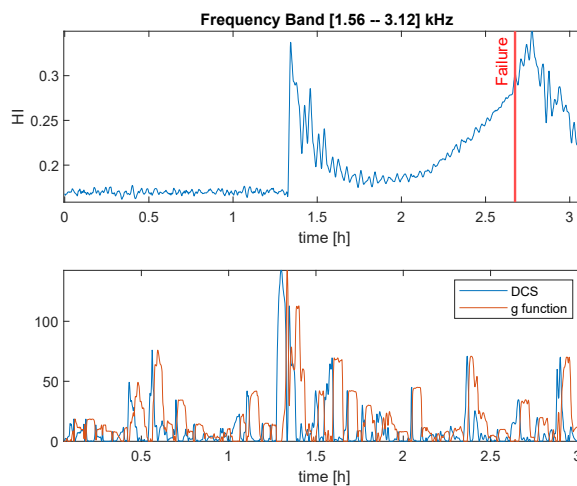

Figure 18. DCS statistical test applied to the HI from the frequency band [1.56–3.12] kHz.

Figure 19a,b shows the results of the HI discretization. As before, all the available health indicator were treated but only the ones concerning the lower frequency are reported since the other did not present interesting features. In both Figure 19a,b, the upper plot shows the two discrete sequences, $HI_{d,1}$ and $HI_{d,2}$, and the lower plot shows the final discrete sequence, HI_C. The $HI_{d,1}$ indicator (blue) is the result of a quantisation of the initial health indicator considering six intervals of equal distance. As can be noticed, the relative energy increase is still noticeable but the failure instance is less visible. Increasing the number of the quantisation intervals could help, but, at the same time, it would increase the size of the HMM B matrix. The introduction of the $HI_{d,2}$ secondary discrete HI allows us to construct a better final discrete sequence where the regions before and after failure are visible. Two examples of the obtained HI_C discrete sequences are presented in green in Figure 19a,b to prove that the proposed discretization technique enhances the degradation profile contained in the initial HI.
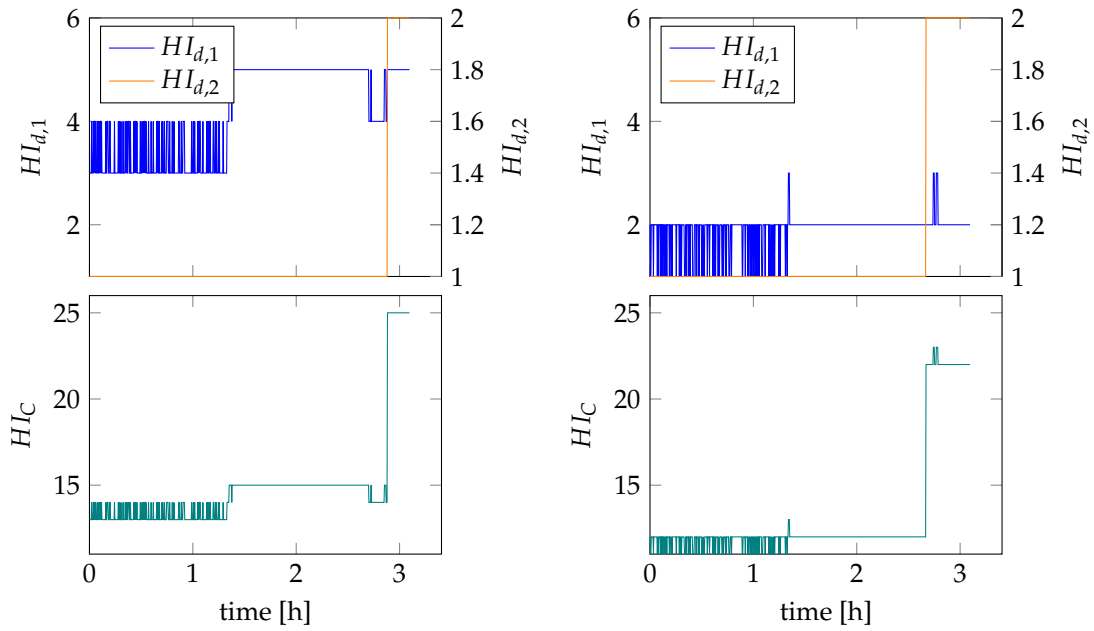
(**a**) Frequency band (0–1.56 kHz).

(**b**) Frequency band (1.56–3.12 kHz).

Figure 19. HI discretization, Signal 10.

All the available synthetic signals were treated like Signal 10 and the obtained discrete sequences were used to perform the MB-HMM training. Figure 20 shows some of the used training sequences. In the figure, the different degradation speeds can be easily observed. The models, λ_1 and λ_2, were obtained by applying the Baum–Welch training algorithm and they are reported here. Equations (23) and (24) concern the fast dynamic branch:

$$A_{1,fb1} = \begin{bmatrix} 0.9975 & 0.0025 & 0 & 0 \\ 0 & 0.7905 & 0.2095 & 0 \\ 0 & 0 & 0.9977 & 0.0023 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{23}$$

$$B_{1,fb1} = \begin{bmatrix} 0 & 0 & 0.70596 & 0.2940 & 3.130 \times 10^{-35} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4.2728 \times 10^{-5} & 0.9998 & 0.0002 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1.5316 \times 10^{-49} & 0.0519 & 0.9481 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.2414 & 0.7586 & 0 \end{bmatrix} \tag{24}$$

Equations (25) and (26) concern the slow dynamic branch:

$$A_{2,fb1} = \begin{bmatrix} 0.9978 & 0.0022 & 0 & 0 \\ 0 & 0.9216 & 0.07839 & 0 \\ 0 & 0 & 0.9987 & 0.00126 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{25}$$

$$B_{2,fb1} = \begin{bmatrix} 0 & 0 & 0.75091 & 0.2491 & 7.3028 \times 10^{-36} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4.2725 \times 10^{-10} & 0.9986 & 0.0114 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2.6067 \times 10^{-79} & 0.0576 & 0.9423 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.2063 & 0.7937 & 0 \end{bmatrix} \tag{26}$$
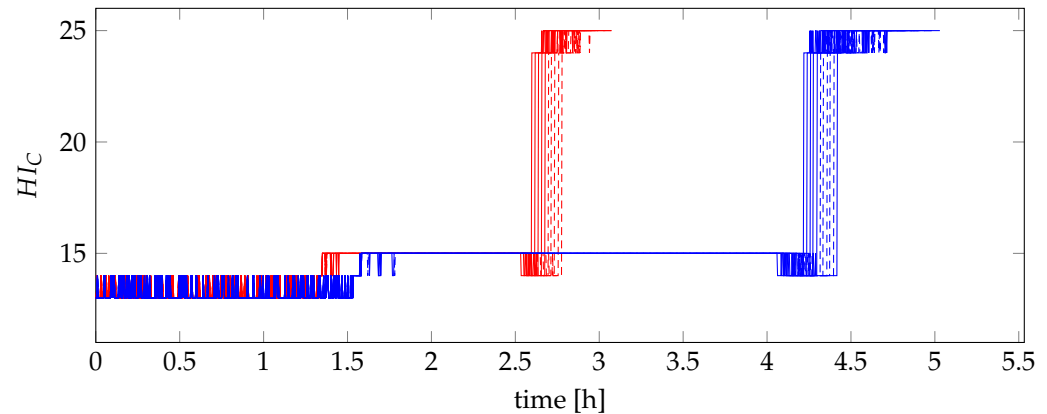
Figure 20. Training sequences: fast dynamic (red) and slow dynamic (blue).

Figure 21 and Table 4 illustrate the final results for Signal 10. Figure 21 shows the estimated RUL (red), the real RUL (black), the error margins at +30% (cyan), −30% (brown) and the mission duration limit (green). The estimated RUL is represented with round markers while the other variables are reported with solid lines. This different graphical representation is to underline the fact that the RUL is estimated at the end of each mission and not with a real-time approach. By looking at the figure, it is possible to notice that the RUL is predicted in an accurate way all along the bearing life. Table 4 reports the data concerning the selection of the level (frequency band presenting the best degradation profile) and the selection of the HMM model (fast or slow). Both the correct frequency band, fb_1, and the correct branch in the MB-HMM model, "Fast", are selected. Moreover, it is possible to notice how the posterior probability of the selected HMM model increases mission after mission. This result shows that the proposed algorithm lowers the uncertainty in the model selection step as time goes on. Indeed, the sequences retrieved after each mission are chained, thus a longer sequence containing all the bearing history is provided to the algorithm which can perform a better informed choice. The mean average percentage error was computed to give an indication of the RUL estimation accuracy. In this case, it is equal to 2.5%. This result highlights the good prediction capabilities of the proposed MB-HMM model.
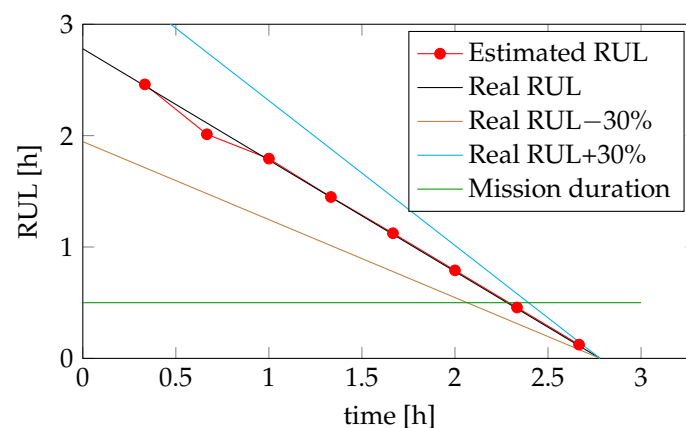
Figure 21. RUL estimation for Signal number 10.

The proposed algorithm was tested as well with other synthetic signals. In particular, Figure 22 shows the low frequency HI_Cs for signals 20, 30, 90 and 110. Signal 10 is reported as well as a reference case. Signals 20 and 30 are both characterised by a slow degradation dynamics. In particular, signal 31 presents the slowest fast degradation among the simulated ones. Signals 90 and 110 are characterised by slow degradation dynamics.

Table 4. Signal 10: posterior probability of the selected HMM model for frequency bands fb_1 to fb_8.

Mission	fb1	Dynamic	fb2 to fb8
1	0.6281	Fast	0
2	0.7892	Fast	0
3	0.7885	Fast	0
4	0.8471	Fast	0
5	0.9140	Fast	0
6	0.9470	Fast	0
7	0.9678	Fast	0
8	0.9806	Fast	0

Figure 22. Tested sequences.

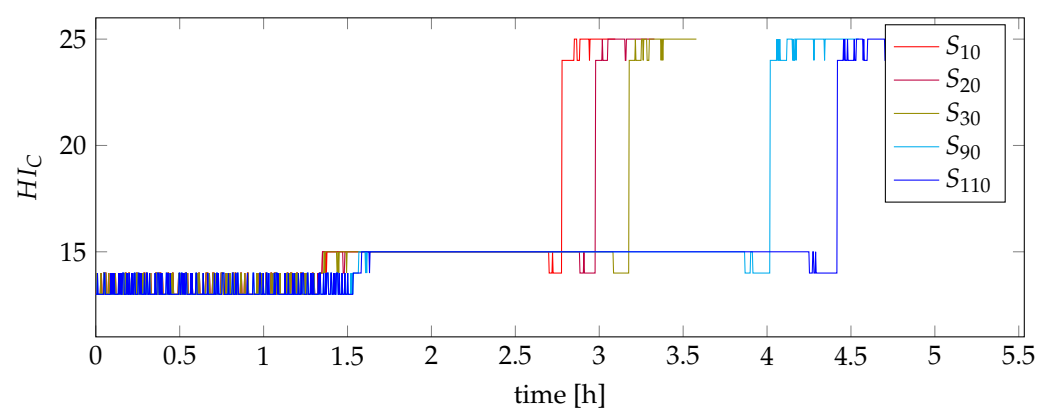Figure 23 and Table 5 report the obtained results concerning signal 20. In this case, the RUL estimation error is higher than what is observed in Figure 21. The RUL is underestimated from mission 1 to mission 9 and respects the -30% error limit until mission 7. The computed percentage error is 21.67%.

This result can be justified by considering the HMM model from a practical point of view. In particular, the focus should be put on the transition matrix $A_{1,fb1}$. By definition each element is the probability of passing from one state to another and its inverse is the mean duration of the hidden state itself. The HMM model captures the average behaviour based on the training data set which is provided. As a consequence, when the λ_1 model is used for RUL prediction, it will work perfectly for signals that are very close to the average behaviour and it will be less accurate for signals that are far from the detected dynamic. Based on these remarks, the results presented in Figure 23 can be explained by the fact that the signal is characterised by a slower dynamics than the one detected in the training dataset. As a result, the HMM underestimates the RUL. However, given a mission duration of 0.5 h, the fact that towards the end the RUL estimation is less precise does not bother the decision-making process. Indeed, the bearing would have been already discarded or sent to maintenance.

Figure 24 and Table 6 report the obtained results concerning signal 30. This signal was generated considering the slowest of the fast degradation dynamics. As a consequence, even if the correct health indicator and HMM model are selected, the RUL estimation respects the margins at $\pm30\%$ until mission 5, so only for half of the bearing life. The obtained percentage error is equal to 38.72%. As before, this result is motivated by the fact that HMM models capture the average degradation dynamics and they cannot properly decode and predict sequences that are too far from it. A possible solution to this issue could be the definition of additional branches in the MB-HMM model. Consequently, intermediate dynamics could be taken into account, improving the RUL estimation.
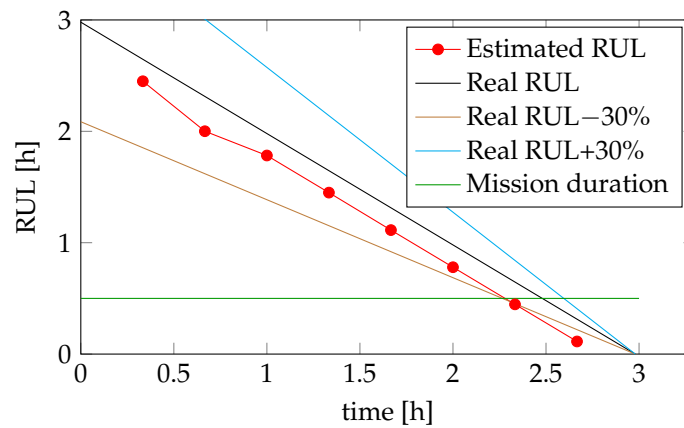
Figure 23. RUL Estimation for signal number 20.

Table 5. Signal 20: posterior probability of the selected HMM model for frequency bands fb_1 to fb_8.

Mission	fb1	Dynamic	fb2 to fb8
1	0.6280	Fast	0
2	0.7892	Fast	0
3	0.7885	Fast	0
4	0.8471	Fast	0
5	0.9289	Fast	0
6	0.9580	Fast	0
7	0.9755	Fast	0
8	0.9858	Fast	0
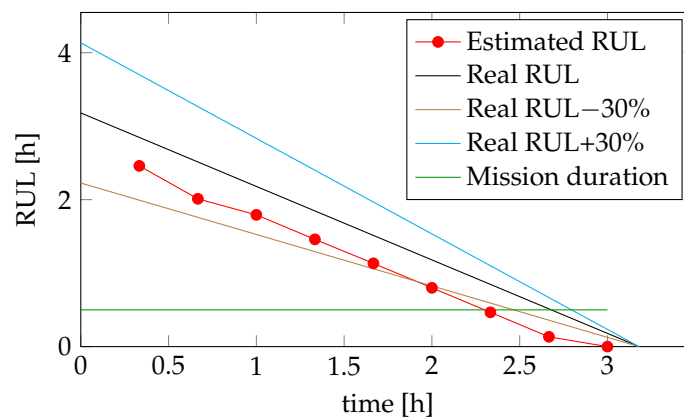
Figure 24. RUL estimation for signal number 30.

Figure 25 and Table 7 report the obtained results concerning signal 110. As for signal 10, both the frequency band and the HMM branch are correctly selected. Except for a couple of hesitations, the proposed methodology is able to clearly recognise that the current degradation dynamic is the slow one.

Similarly to observations for signals 30 and 20, when the degradation rate becomes faster, the algorithm accuracy decreases. This is the case for signal 90, which is far from the training sequences. Figure 26 and Table 8 show that the frequency band selection is correct but that towards the bearing end-of-life, the model selection becomes critical. As shown in this table, the fast dynamic is chosen instead of the slow one, provoking a strong underestimation of the RUL. Again, it occurs when the RUL is close to the mission duration limit. Since this signal falls in the region between the two training datasets it is

not surprising that the algorithm has difficulties in selecting the correct degradation branch model in the MB-HMM. In general, a conservative behaviour is observed: the model has the tendency to classify a slow dynamic as "Fast" while the opposite rarely happens. This aspect is very promising and proves that the present methodology can be considered for later application in real rocket engine data. The percentage error is directly affected by the algorithm hesitation and the computed values are 110% for signal 90 and 27.31% for signal 110. This aspect will be addressed first in future works since the method precision and performance directly depends on it. The obtained percentage errors prove it clearly.

Table 6. Signal 30: posterior probability of the selected HMM model.

Mission	fb1	Dynamic	fb2 to fb8
1	0.6399	Fast	0
2	0.7975	Fast	0
3	0.7970	Fast	0
4	0.8229	Fast	0
5	0.8273	Fast	0
6	0.8873	Fast	0
7	0.9282	Fast	0
8	0.9550	Fast	0
9	0.9721	Fast	0

Table 7. Signal 110: posterior probability of the selected HMM model for frequency bands fb_1 to fb_8.

Mission	fb1	Dynamic	fb2 to fb8
1	0.5415	Fast	0
2	0.6903	Slow	0
3	0.9023	Slow	0
4	0.9605	Slow	0
5	0.9710	Slow	0
6	0.9530	Slow	0
7	0.9247	Slow	0
8	0.8813	Slow	0
9	0.8180	Slow	0
10	0.7311	Slow	0
11	0.6220	Slow	0
12	0.5011	Fast	0
13	0.8740	Slow	0
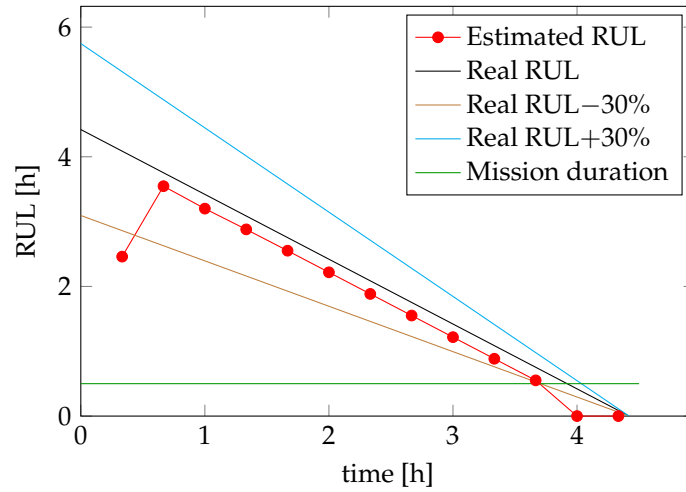
Figure 25. RUL estimation for signal number 110.
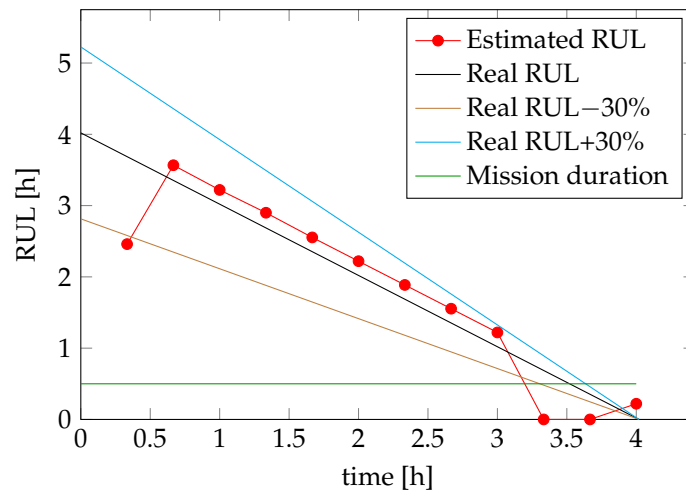
Figure 26. RUL estimation for signal number 90.

Table 8. Signal 90: posterior probability of the selected HMM model for frequency bands fb_1 to fb_8.

Mission	fb1	Dynamic	fb2 to fb8
1	0.5424	Fast	0
2	0.6900	Slow	0
3	0.9026	Slow	0
4	0.9607	Slow	0
5	0.9196	Slow	0
6	0.8750	Slow	0
7	0.8108	Slow	0
8	0.7240	Slow	0
9	0.6163	Slow	0
10	0.5042	Fast	0
11	0.6242	Fast	0
12	0.9494	Slow	0

4. Application to Experimental Data

The methodology previously applied to the synthetic dataset is here re-proposed and applied to some real experimental data coming from the XJTU-SY bearing dataset.

4.1. XJTU-SY Test Bench Set-Up

As described by Wang et al. [32], this dataset was provided by the Institute of Design Science and Basic Component at Xi'an Jiaotong University (XJTU), Xi'an, China and the Changxing Sumyoung Technology Co., Ltd. (SY), Changxing, China. A total of 15 LDK UER204 bearings were tested under three different operating conditions reported in Table 9.

Table 9. Testing conditions—XJTU-SY dataset.

	TC1	TC2	TC3
RadialLoad (kN)	12	11	10
Rotating Speed (rpm)	2100	2250	2400

The bearings were charged radially and the vibration was recorded using two accelerometers of type PCB 352C33 placed on the vertical and horizontal directions. The sampling frequency was set to $F_s = 25.6$ kHz, the sample duration to $d_s = 1.28$ s and the sampling interval to $p_s = 1$ min. At the end of each experiment, an inspection was conducted which revealed the faulty bearing component (Inner ring, Outer ring, Cage and Ball). The computation setup defined for this study case is similar to the one defined for synthetic signals. A MB-HMM model composed of two branches was chosen as previously presented in Figure 14. The available data were split in two groups of eight signals (fast dynamics) and five signals (slow dynamics). Each branch was trained separately using the Baum–Welch algorithm. The transition matrices, A_1 and A_2, are 4×4 matrices since four hidden states were selected. The HI discretization was performed using four intervals. Thus the emission matrices, B_1 and B_2, are 4×8 matrices.

4.2. Results and Discussion

The methodology presented in Section 2 was applied to all the signals in the dataset. The results obtained for Bearing 1–3 are presented here. Figure 27a–h shows the HIs obtained after the application of the wavelets decomposition. By looking at the plots, it can be observed that the lower frequency bands are again the most sensitive to the presence of degradation. In particular, the first frequency band, [0–1.6 kHz] presents a very clear degradation profile.

Going further, Figure 28a,b shows the discrete sequences used for the MB-HMM training of the first frequency band. The figure contains the fast sequences (red) and the slow sequences (blue). The available training signals presented much more variability with respect to the previously available synthetic data. Now, the obtained discrete sequences present very different profiles. As a result, it is more difficult for the HMM model to learn the parameters of the A and B matrices. To achieve the same accuracy as before, a larger dataset should be used. Anyway, even if the dataset size has a very important role, providing high quality sequences can compensate in the case of small datasets like the chosen one. The obtained models λ_1 and λ_2 are reported hereafter.

Equations (27) and (28) concern the fast dynamic branch, $b = 1$:

$$A_{1,fb1} = \begin{bmatrix} 0.9875 & 0.01253 & 0 & 0 \\ 0 & 0.9908 & 0.00919 & 0 \\ 0 & 0 & 0.9906 & 0.0094 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{27}$$

$$B_{1,fb1} = \begin{bmatrix} 0.9518 & 0.04823 & 8.7285 \times 10^{-9} & 9.5273 \times 10^{-148} & 0 & 0 & 0 & 0 \\ 0.0023 & 0.6988 & 0.2989 & 1.1724 \times 10^{-15} & 0 & 0 & 0 & 0 \\ 0.00117 & 0.0061 & 0.7761 & 0.2166 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.00238 & 0.0795 & 0.7073 & 0.2108 \end{bmatrix} \tag{28}$$

Equations (29) and (30) concern the slow dynamic branch, $b = 2$:

$$A_{2,fb1} = \begin{bmatrix} 0.9994 & 0.0006 & 0 & 0 \\ 0 & 0.9991 & 0.0009 & 0 \\ 0 & 0 & 0.9985 & 0.0015 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{29}$$

$$B_{2,fb1} = \begin{bmatrix} 0.4149 & 0.5843 & 0.00078 & 0 & 0 & 0 & 0 & 0 \\ 0.8594 & 0.1406 & 4.5603 \times 10^{-10} & 0 & 0 & 0 & 0 & 0 \\ 1.03901 \times 10^{-44} & 0.2094 & 0.7662 & 0.024403 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.4078 & 0.5668 & 0.02539 \end{bmatrix} \tag{30}$$

(**a**) fb_1: [0–1.6] kHz

(**b**) fb_2: [1.6–3.2] kHz

(**c**) fb_3: [3.2–4.8] kHz

(**d**) fb_4: [4.8–6.4] kHz

(**e**) fb_5: [6.4–8] kHz

(**f**) fb_6: [8–9.6] kHz

(**g**) fb_7: [9.6–11.2] kHz

(**h**) fb_8: [11.2–12.8] kHz
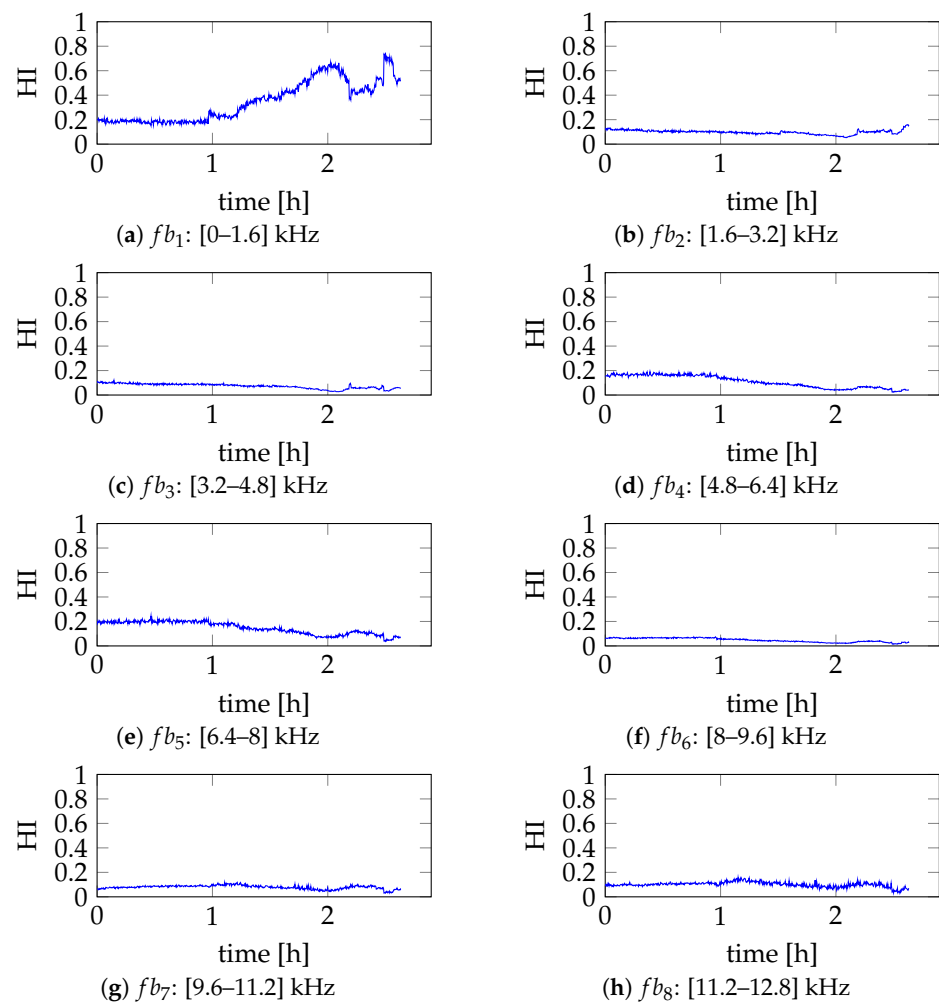
Figure 27. Relative energy health indicators computed for Bearing 1–3.

(**a**) Fast dynamics sequences

(**b**) Slow dynamics sequences
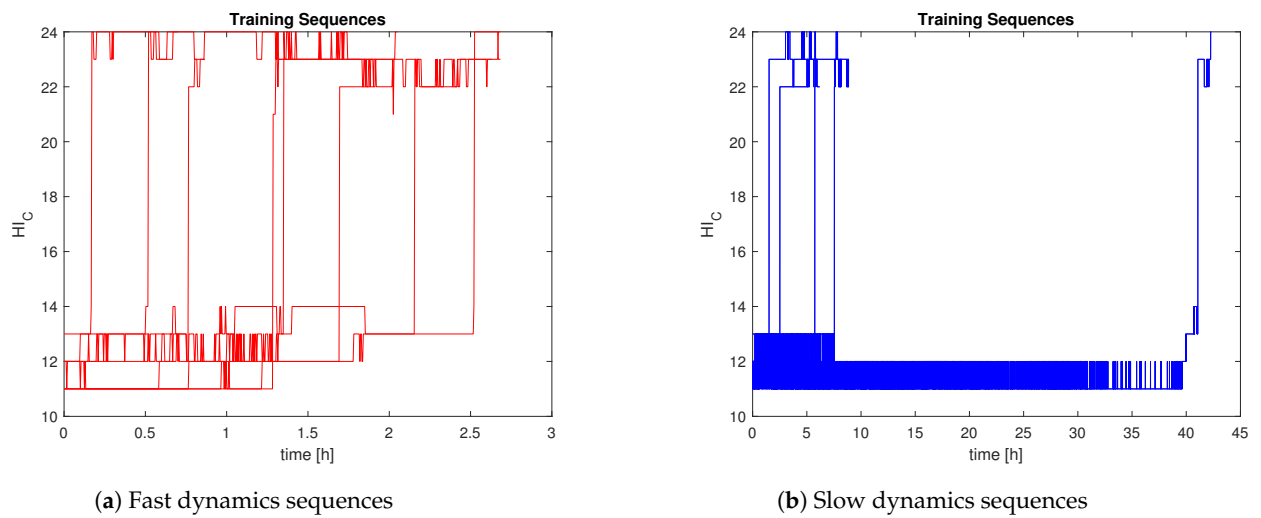
Figure 28. Training sequences.

Figure 29 and Table 10 include the RUL estimation results of Bearing 1–3. The first thing to be noticed is that during the first three missions, the algorithm does not detect the presence of degradation. As a result, the estimated RUL is a fixed value corresponding to the expected duration of three missions, in this case 1.5 h. Starting from mission 4, the degradation is properly detected and the right branch from the trained MB-HMM model is selected. The predicted RUL respects the defined ±30% error margins. The posterior probability reported in Table 10 oscillates between 0.64 and 0.67 instead of showing an increasing trend like before. The computed mean percentage error is 24.4%. This aspect can be explained considering the strong variability of the training sequences. Indeed, since the training sequences are quite different from each other, it is more difficult for the HMM decoding algorithm to select the appropriate model with a low degree of uncertainty. This aspect can be resolved by increasing the training dataset size or by selecting an appropriate family of sequences to be given as inputs. Anyway, the purpose of the present work is to develop a complete chain from data collection to RUL estimation and to show its capabilities with respect to synthetic and experimental data. The presented results confirm that the proposed methodology can be effectively applied in both cases. These results are promising and motivate the authors to continuing exploring this approach to improve its robustness.
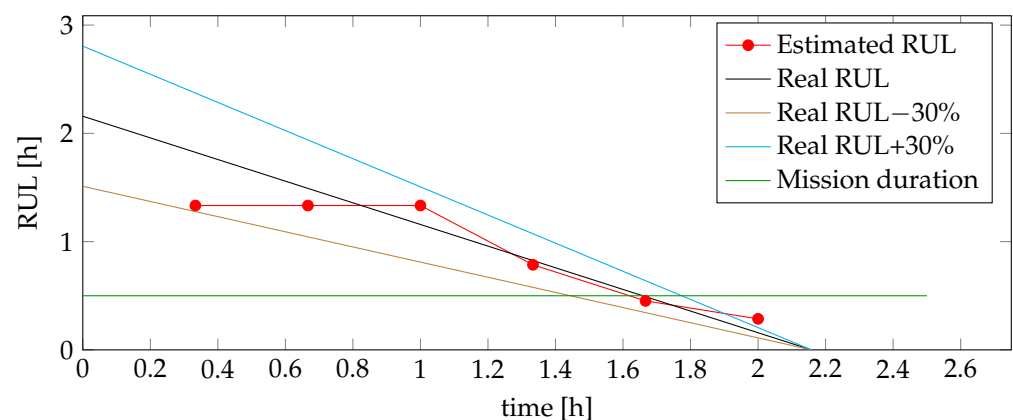

Figure 29. RUL Estimation for Bearing 1–3.

Table 10. Posterior probability of the selected HMM model.

Mission	fb1	Dynamic	fb2 to fb8
1	[]	[]	0
2	[]	[]	0
3	[]	[]	0
4	0.6453	Fast	0
5	0.6675	Fast	0
6	0.6597	Fast	0

5. Conclusions

In the context of reusable rocket engines, a complete PHM methodology is proposed to estimate the RUL of the turbopump bearing. The final objective consists in helping the decision process concerning the maintenance planning and execution between one mission and another. The proposed methodology consists of a data-driven off-line approach collecting monitoring data at the end of the mission and using it to infer about the bearing health condition. The proposed approach is composed of five main steps: data collection, HI construction, HI classification, Multi-Level MB-HMM training and RUL estimation.

The bearing is surveyed with accelerometers which record the bearing vibration during its functioning. After that, the obtained signals are processed using the MODWPT and used to extract a new and effective HI. The proposed HI is the nodal energy, or relative energy, which effectively captures different kinds of bearing degradation and allows us to isolate it in the frequency domain without losing information from the time domain. In particular, a recurrent degradation profile is identified and proposed to perform RUL prediction. Even thought the relative energy has already been the object of previous work, the authors propose a detailed interpretation of it linked to fatigue degradation phenomena.

The diagnosis and prognosis tasks are performed using HMM statistical models. Based on a detailed analysis of the available datasets, HIs are categorised according to their sensitivity to the presence of degradation and according to their degradation rate. Then, two HMM models are trained and later they are used to categorise each HI and discard the ones that are not useful. In addition to that a new type of Multi-Level Multi-Branch physics-defined HMM model is proposed to perform RUL estimation which takes into account different degradation rates that may affect the bearing. The parameters of the model have been carefully defined according to a qualitative degradation profile composed of four main stages: the Nominal stage, Crack Initiation, Crack Propagation and Failure. The introduction of fatigue models in the definition of the HMM structure allowed us to reduce the number of parameters to be estimated to an optimal one. At the same time the correct representation of the hidden degradation process is ensured.

The final RUL estimation is performed as the computation of the MC absorption time. The detailed calculation and demonstration were provided which were not available in the literature, at least according to the authors' knowledge.

The proposed methodology was applied to a synthetic dataset and to an experimental dataset. In both cases, the results proved the capability of the algorithm to accurately estimate and predict the bearing RUL. Given the strong understanding of all the blocks composing the present methodology, all the results characterised by lower accuracy could be justified. Their analysis revealed the strong features and the limits of the proposed approach.

In conclusion, an end-to-end off-line methodology was developed which can effectively estimate the Bearing RUL and is compatible with the demanding field of reusable LPREs. The obtained results are promising and further investigation will concern the algorithm robustness improvement.

Author Contributions: Conceptualisation, F.G., V.S., G.H., P.W. and G.F.; methodology, F.G., V.S., G.H., G.F. and P.W.; software, F.G.; validation, V.S., G.H., G.F., C.R. and P.W.; formal analysis, F.G.; investigation, F.G.; resources, F.G.; data curation, F.G.; writing—original draft preparation, F.G.; writing—review and editing, F.G., V.S., G.H., P.W., G.F. and C.R.; visualisation, F.G.; supervision, V.S., G.H., P.W., G.F. and C.R.; project administration, G.H., P.W. and G.F.; funding acquisition, G.H., P.W. and G.F. All authors have read and agreed to the published version of the manuscript

Data Availability Statement: Data is contained within the article.

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

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
CNN	Convolutional Neural Networks
DCS	Dynamic Cumulative Sum
DL	Deep Learning
DNN	Deep Neural Network
DoF	Degrees of Freedom
DTW	Dynamic Time Warping
DWPT	Discrete Wavelet Packet Transform
EMD	Empirical Mode Decomposition
HI	Health Indicator
HMM	Hidden Markov Model
HMS	Health Monitoring System
HS	Health Stage
LPRE	Liquid Propellant Rocket Engines
LSTM	Long-Short Term Memory
MB	Multi-Branch
MC	Markov Chain
ML	Machine Learning
MODWPT	Maximum Overlap Discrete Wavelets Packet Transform
MTTF	Mean Time To Failure
PCA	Principal Component Analysis
PHM	Prognosis and Health Management
RMS	Root Mean Square
RUL	Remaining Useful Life
SoH	State of Health
SSME	Space Shuttle Main Engine
TL	Transfer Learning

Appendix A

As explained in [37], the MTTF is obtained starting from the Laplace transform of the Markov chain probabilities, $L(s)$, computed in $s = 0$:

$$MTTF = L(0) = \sum_{i=1}^{N} L_i(O) \tag{A1}$$

In Equation (A1), i is the state index, $i = 1, ..., l$, with l the number of transient states. Starting from the system of equations that describes the variation of the probabilities over time, it can be proven that:

$$L(s) = P_0[sI - A_l]^{-1} \tag{A2}$$

where A_l is a $l \times l$ sub-matrix of the transition matrix A. A transient state is a state from which it is possible to go out. It is different from an absorbent state from which it is not possible to leave. For $s = 0$ Equation (A2) becomes:

$$MTTF = P_0[-A_l]^{-1} = -P_0[A_l]^{-1} \tag{A3}$$

Equation (A3) can be rewritten taking into account the following definition of A_l^{-1}:

$$[A_l]^{-1} = \frac{(-1)^{i+j}det(\overline{Al_{ji}})}{det(Al)} \tag{A4}$$

where $\overline{Al_{ji}}$ is the co-factors matrix.

Let us now consider a system with four states: 3 transient states and one absorbent state. A_l is a 3×3 matrix. The MTTF can be computed using Equation (A3) as follows:

$$MTTF = -\sum_{j=1}^{l}\left[P_{01}\ P_{02}\ P_{03}\right] \cdot \begin{bmatrix} c_{1j} \\ c_{2j} \\ c_{3j} \end{bmatrix} \tag{A5}$$

The vector $\begin{bmatrix} P_{01}\ P_{02}\ P_{03} \end{bmatrix}$ contains the initial probability distribution and the matrix, while $\begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}$ contains the co-factors of A_l. Equation (A5) can be rewritten using sums:

$$MTTF = -\left(P_{01}\left(\sum_{j=1}^{l}c_{1j}\right) + P_{02}\left(\sum_{j=1}^{l}c_{2j}\right) + P_{03}\left(\sum_{j=1}^{l}c_{3j}\right)\right) = -\sum_{i=1}^{l}P_{0i}\left(\sum_{j=1}^{l}c_{ij}\right) \tag{A6}$$

So for any A_l matrix :

$$MTTF = -\sum_{i=1}^{l}P_{0i}\left(\sum_{j=1}^{l}c_{ij}\right) \tag{A7}$$

By applying the definition in Equation (A4) it is obtained that:

$$MTTF = -\frac{1}{\Delta}\sum_{i=1}^{l}P_{0i}\left(\sum_{j=1}^{l}(-1)^{i+j}det(\overline{Al_{ji}})\right) = -\frac{1}{\Delta}\sum_{i=1}^{l}\left(\sum_{j=1}^{l}P_{0i}(-1)^{i+j}det(\overline{Al_{ji}})\right) \tag{A8}$$

where Δ is the A_l matrix determinant, $det(A_l)$. The obtained formulation can be further rewritten until a clean and easy to use form is obtained. Let us thus focus on the product inside the sums. Such product can be rewritten as the determinant of matrix A_l where the jth line has been replaced by the vector $[0...P_{0i}...0]$, where P_{0i} is the initial probability of state i. By developing, it is obtained:

$$MTTF = -\frac{1}{\Delta}\sum_{i=1}^{l}\sum_{j=1}^{l}\begin{vmatrix} a_{11} & ... & a_{1i} & ... & a_{1l} \\ \vdots & & \vdots & & \vdots \\ 0 & ... & P_{0i} & ... & 0 \\ \vdots & & \vdots & & \vdots \\ a_{l1} & ... & a_{li} & ... & a_{ll} \end{vmatrix} = -\frac{1}{\Delta}\sum_{j=1}^{l}\sum_{i=1}^{l}\begin{vmatrix} a_{11} & ... & a_{1i} & ... & a_{1l} \\ \vdots & & \vdots & & \vdots \\ 0 & ... & P_{0i} & ... & 0 \\ \vdots & & \vdots & & \vdots \\ a_{l1} & ... & a_{li} & ... & a_{ll} \end{vmatrix} \tag{A9}$$

$$= -\frac{1}{\Delta} \sum_{j=1}^{l} 1 \cdot \begin{vmatrix} a_{11} & ... & a_{1i} & ... & a_{1l} \\ \vdots & & \vdots & & \vdots \\ P_{01} & ... & P_{0i} & ... & P_{0l} \\ \vdots & & \vdots & & \vdots \\ a_{l1} & ... & a_{li} & ... & a_{ll} \end{vmatrix} = -\frac{1}{\Delta} \sum_{j=1}^{l} 1 \cdot (-1)^{j-1} \begin{vmatrix} P_{01} & ... & P_{0i} & ... & P_{0l} \\ a_{11} & ... & a_{1i} & ... & a_{1l} \\ \vdots & & \vdots & & \vdots \\ a_{j-1,1} & ... & a_{j-1,i} & ... & a_{j-1,l} \\ a_{j+1,1} & ... & a_{j+1,i} & ... & a_{j+1,l} \\ \vdots & & \vdots & & \vdots \\ a_{l1} & ... & a_{li} & ... & a_{ll} \end{vmatrix} = \tag{A10}$$

Knowing that:

$$(-1)^{+1} \cdot (-1)^{j-1} = (-1)^{j} \tag{A11}$$

It is possible to write:

$$MTTF = \frac{1}{\Delta} \sum_{j=1}^{l} 1 \cdot (-1)^{j} \begin{vmatrix} P_{01} & ... & P_{0i} & ... & P_{0l} \\ a_{11} & ... & a_{1i} & ... & a_{1l} \\ \vdots & & \vdots & & \vdots \\ a_{j-1,1} & ... & a_{j-1,i} & ... & a_{j-1,l} \\ a_{j+1,1} & ... & a_{j+1,i} & ... & a_{j+1,l} \\ \vdots & & \vdots & & \vdots \\ a_{l1} & ... & a_{li} & ... & a_{ll} \end{vmatrix} \tag{A12}$$

Knowing that:

$$(-1)^{j} = (-1)^{j-1+1} = (-1)^{j} \cdot (-1)^{-1} \cdot (-1)^{+1} = (-1)^{j} \cdot (-1)^{+1} \cdot (-1)^{+1} = (-1)^{j+2} \tag{A13}$$

The previous formulation becomes:

$$MTTF = \frac{1}{\Delta} \sum_{j=1}^{l} 1 \cdot (-1)^{j+2} \begin{vmatrix} P_{01} & ... & P_{0i} & ... & P_{0l} \\ a_{11} & ... & a_{1i} & ... & a_{1l} \\ \vdots & & \vdots & & \vdots \\ a_{j-1,1} & ... & a_{j-1,i} & ... & a_{j-1,l} \\ a_{j+1,1} & ... & a_{j+1,i} & ... & a_{j+1,l} \\ \vdots & & \vdots & & \vdots \\ a_{l1} & ... & a_{li} & ... & a_{ll} \end{vmatrix} \tag{A14}$$

$$MTTF = \frac{1}{\Delta} \sum_{j=2}^{l+1} 1 \cdot (-1)^{j+1} \begin{vmatrix} P_{01} & ... & P_{0i} & ... & P_{0l} \\ a_{11} & ... & a_{1i} & ... & a_{1l} \\ \vdots & & \vdots & & \vdots \\ a_{j-1,1} & ... & a_{j-1,i} & ... & a_{j-1,l} \\ a_{j+1,1} & ... & a_{j+1,i} & ... & a_{j+1,l} \\ \vdots & & \vdots & & \vdots \\ a_{l1} & ... & a_{li} & ... & a_{ll} \end{vmatrix} \tag{A15}$$

$$MTTF = \frac{1}{\Delta} \sum_{j=1}^{l+1} b_{j,1} \cdot (-1)^{j+1} \begin{vmatrix} P_{01} & ... & P_{0i} & ... & P_{0l} \\ a_{11} & ... & a_{1i} & ... & a_{1l} \\ \vdots & & \vdots & & \vdots \\ a_{j-1,1} & ... & a_{j-1,i} & ... & a_{j-1,l} \\ a_{j+1,1} & ... & a_{j+1,i} & ... & a_{j+1,l} \\ \vdots & & \vdots & & \vdots \\ a_{l1} & ... & a_{li} & ... & a_{ll} \end{vmatrix} \tag{A16}$$

$$b_{j,1} = \begin{cases} 0 & if \ j = 1 \\ 1 & if \ 2 \le j \le l+1 \end{cases} \tag{A17}$$

In conclusion, the MTTF can be easily computed:

$$MTTF = \frac{1}{\Delta} \begin{vmatrix} 0 & P_{01} & ... & P_{0l} \\ 1 & a_{11} & ... & a_{1l} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_{l1} & ... & a_{ll} \end{vmatrix} \tag{A18}$$

References

1. Cheng , H.; Zhang, Y.; Lu, W.; Yang, Z. Research on ball bearing model based on local defects. *Appl. Sci.* **2019**, *1*, 1219. [CrossRef]
2. Hu, Y.; Miao, X.; Si, Y.; Pan, E.; Zio, E. Prognostics and health management: A review from the perspectives of design, development and decision. *Reliab. Eng. Syst. Saf.* **2022**, *217*, 108063. [CrossRef]
3. Lei, Y.; Li, N.; Guo, L.; Li, N.; Yan, T.; Lin, J. Machinery health prognostics: A systematic review from data acquisition to RUL prediction. *Mech. Syst. Signal Process.* **2018**, *104*, 799–834. [CrossRef]
4. Wu, G.; Yan, T.; Yang, G.; Chai, H.; Cao, C. A review on rolling bearing fault signal detection methods based on different sensors. *Sensors* **2022**, *22*, 8330. [CrossRef] [PubMed]
5. Zhou, H.; Huang, X.; Wen, G.; Lei, Z.; Dong, S.; Zhang, P.; Chen, X. Construction of health indicators for condition monitoring of rotating machinery: A review of the research. *Expert Syst. Appl.* **2022**, *203*, 117297. [CrossRef]
6. Jain, P.; Bhosle, S. Analysis of vibration signals caused by ball bearing defects using time-domain statistical indicators. *Int. J. Adv. Technol. Eng. Explor.* **2022**, *9*, 700–715. [CrossRef]
7. Sun, Y.; Li, S.; Wang, X. Bearing fault diagnosis based on EMD and improved Chebyshev distance in SDP image. *Meas. J. Int. Meas. Confed.* **2021**, *176*, 109100. [CrossRef]
8. Yang, D.M. The detection of bearing incipient fault with maximal overlap discrete wavelet packet transform and sparse code shrinkage denoising. In Proceedings of the Proceedings—2020 International Symposium on Computer, Consumer and Control, IS3C, Taichung, Taiwan, 13–16 November 2020; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2020; pp. 216–219. [CrossRef]
9. Schwendemann, S.; Amjad, Z.; Sikora, A. A survey of machine-learning techniques for condition monitoring and predictive maintenance of bearings in grinding machines. *Comput. Ind.* **2021**, *125*, 103380. [CrossRef]
10. Mochammad, S.; Noh, Y.; Kang, Y.J. Bearing Fault Degradation Modeling Based on Multitime Windows Fusion Unsupervised Health Indicator. *IEEE Sens. J.* **2023**, *23*, 19623–19634. [CrossRef]
11. Galli, F.; Fiore, G.; Sircoulomb, V.; Hoblos, G.; Weber, P. Identification of degradation profiles for turbopump bearings in reusable liquid propellant rocket engines. In Proceedings of the Aerospace Europe Conference 2023—10th EUCASS—9th CEAS, Lausanne, Switzerland, 9–13 July 2023. [CrossRef]
12. Fan, W.; Zheng, X.; Chen, C.; Li, Y.; Liu, X.; He, C. Dynamic CUSUM Chart With an Integrated Indicator for Bearing Condition Monitoring. *IEEE Sens. J.* **2023**, *23*, 15400–15412. [CrossRef]
13. Chen, Z.; Cen, J.; Xiong, J. Rolling Bearing Fault Diagnosis Using Time-Frequency Analysis and Deep Transfer Convolutional Neural Network. *IEEE Access* **2020**, *8*, 150248–150261. [CrossRef]
14. König, F.; Sous, C.; Ouald Chaib, A.; Jacobs, G. Machine learning based anomaly detection and classification of acoustic emission events for wear monitoring in sliding bearing systems. *Tribol. Int.* **2021**, *155*, 106811. [CrossRef]
15. Behzad, M.; Arghan, H.A.; Bastami, A.R.; Zuo, M.J. Prognostics of rolling element bearings with the combination of Paris law and reliability method. In Proceedings of the 2017 Prognostics and System Health Management Conference (PHM-Harbin), Harbin, China, 9–12 July 2017; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2017; ISBN 9781538603703. [CrossRef]
16. Haidong, S.; Junsheng, C.; Hongkai, J.; Yu, Y.; Zhantao, W. Enhanced deep gated recurrent unit and complex wavelet packet energy moment entropy for early fault prognosis of bearing. *Knowl. Based Syst.* **2020**, *188*, 105022. [CrossRef]
17. Gao, S.; Xiong, X.; Zhou, Y.; Zhang, J. Bearing Remaining Useful Life Prediction Based on a Scaled Health Indicator and a LSTM Model with Attention Mechanism. *Machines* **2021**, *9*, 238. [CrossRef]
18. Liu, Y.; Chen, J.; Wang, T.; Li, A.; Pan, T. A variational transformer for predicting turbopump bearing condition under diverse degradation processes. *Reliab. Eng. Syst. Saf.* **2023**, *232*, 109074. [CrossRef]
19. Medjaher, K.; Tobon-Mejia, D.; Zerhouni, N.; Medjaher, K.; Tobon-Mejia, D.A.; Zerhouni, N. Remaining useful life estimation of critical components with application to bearings. Remaining useful life estimation of critical components with application to bearings. *IEEE Trans. Reliab.* **2012**, *61*, 292–302. [CrossRef]
20. Soave, E.; D'Elia, G.; Dalpiaz, G. Prognostics of rotating machines through generalized Gaussian hidden Markov models. *Mech. Syst. Signal Process.* **2023**, *185*, 109767. [CrossRef]
21. Zhao, W.; Shi, T.; Wang, L. Fault Diagnosis and Prognosis of Bearing Based on Hidden Markov Model with Multi-Features. *Appl. Math. Nonlinear Sci.* **2020**, *5*, 71–84. [CrossRef]

22. Qian, Y.; Yan, R.; Gao, R.X. A multi-time scale approach to remaining useful life prediction in rolling bearing. *Mech. Syst. Signal Process.* **2017**, *83*, 549–567. [CrossRef]
23. Park, S.Y.; Ahn, J. Deep neural network approach for fault detection and diagnosis during startup transient of liquid-propellant rocket engine. *Acta Astronaut.* **2020**, *177*, 714–730. [CrossRef]
24. Chelouati, M.; Jha, M.S.; Galeotta, M.; Theilliol, D. Remaining Useful Life Prediction for Liquid Propulsion Rocket Engine Combustion Chamber. In Proceedings of the Conference on Control and Fault-Tolerant Systems, SysTol, Saint-Raphael, France, 29 September–1 October 2021; IEEE Computer Society: Piscataway, NJ, USA, 2021; pp. 225–230. ISSN 21621209. [CrossRef]
25. Wang, T.; Ding, L.; Yu, H. Research and Development of Fault Diagnosis Methods for Liquid Rocket Engines. *Aerospace* **2022**, *9*, 481. [CrossRef]
26. Pan, T.; Zhang, S.; Li, F.; Chen, J.; Li, A. A meta network pruning framework for remaining useful life prediction of rocket engine bearings with temporal distribution discrepancy. *Mech. Syst. Signal Process.* **2023**, *195*, 110271. [CrossRef]
27. Nectoux, P.; Gouriveau, R.; Medjaher, K.; Ramasso, E.; Chebel-Morello, B.; Zerhouni, N.; Varnier, C. PRONOSTIA: An experimental platform for bearings accelerated degradation tests. In Proceedings of the IEEE International Conference on Prognostics and Health Management, PHM'12, Denver, CO, USA, 18–22 June 2012; pp. 1–8.
28. Yang, Y.; He, Y.; Cheng, J.; Yu, D. A gear fault diagnosis using Hilbert spectrum based on MODWPT and a comparison with EMD approach. *Meas. J. Int. Meas. Confed.* **2009**, *42*, 542–551. [CrossRef]
29. Galli, F.; Sircoulomb, V.; Hoblos, G.; Weber, P.; Galeotta, M. *Remaining Useful Life Estimation Based on Wavelet Decomposition: Application to Bearings in Reusable Liquid Propellant Rocket Engines*; Springer: Cham, Switerland, 2023; pp. 107–119. [CrossRef]
30. Mustafa, O. Fault Detection by Means of DCS Algorithm Combined with Filters Bank—Application to the Tennessee Eastman Challenge Process. In Proceedings of the Fifth International Conference on Informatics in Control, Automation and Robotics, Signal Processing, Systems Modeling and Control, Funchal, Madeira, Portugal, 11–15 May 2008.
31. Fawcett, T. An introduction to ROC analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [CrossRef]
32. Wang, B.; Lei, Y.; Li, N.; Li, N. A Hybrid Prognostics Approach for Estimating Remaining Useful Life of Rolling Element Bearings. *IEEE Trans. Reliab.* **2020**, *69*, 401–412. [CrossRef]
33. Jurafsky, D.; Martin, J.H. *Speech and Language Processing–An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*; Prentice Hall: Hoboken, NJ, USA, 2000.
34. Rabiner, L.R. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proc. IEEE* **1989**, *77*, 257–286. [CrossRef]
35. Le, T.T.; Chatelain, F.; Bérenguer, C. Hidden Markov Models for diagnostics and prognostics of systems under multiple deterioration modes. In *Safety and Reliability: Methodology and Applications*; CRC Press: Boca Raton, FL, USA, 2014; pp. 1197–1204, ISBN 9781315736976.
36. El-Thalji, I.; Jantunen, E. A descriptive model of wear evolution in rolling bearings. *Eng. Fail. Anal.* **2014**, *45*, 204–224. [CrossRef]
37. Pagès, A. *System Reliability: Evaluation and Prediction in Engineering*; Springer: New York, NY, USA, 1986.
38. Shahin, K.I.; Simon, C.; Weber, P. Online remaining useful life management considering operating conditions to match the given maintenance date. In Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference, Venice, Italy, 1–5 November 2020; Research Publishing: Singapore, 2020; pp. 3525–3532, ISBN 9789811485930. [CrossRef]
39. Galli, F.; Sircoulomb, V.; Fiore, G.; Hoblos, G.; Weber, P. Dynamic modelling for non-stationary bearing vibration signals. In Proceedings of the 2023 31st Mediterranean Conference on Control and Automation, MED, Limassol, Cyprus, 26–29 June 2023; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2023; pp. 49–54. [CrossRef]
40. Savas, E.H.; Galli, F.; Hoblos, G.; Ponti, F.; Fiore, G. Modeling and Analysing Multiple Defects in Ball Bearings for Aerospace Applications (accepted). In Proceedings of the 8th International Conference on Control, Automation and Diagnosis (ICCAD24), Paris, France, 15–17 May 2024.
41. Ruan, D.; Chen, Y.; Gühmann, C.; Yan, J.; Li, Z. Dynamics Modeling of Bearing with Defect in Modelica and Application in Direct Transfer Learning from Simulation to Test Bench for Bearing Fault Diagnosis. *Electronics* **2022**, *11*, 622. [CrossRef]
42. Sawalhi, N.; Randall, R.B. Simulating gear and bearing interactions in the presence of faults. Part I. The combined gear bearing dynamic model and the simulation of localised bearing faults. *Mech. Syst. Signal Process.* **2008**, *22*, 1924–1951. [CrossRef]
43. Gibson, H.; Thom, R.; Moore, C.; Team, T.; Haluck, D.; Whitney, P. History of space shuttle main engine turbopump bearing testing at the marshall space flight center. In Proceedings of the 57th JANNAF Joint Propulsion Meeting, Colorado Springs, CO, USA, 3–7 May 2010.