machines

IMPACT FACTOR 3.0

CITESCORE 6.1

Article

Quantifying the Reliability Impact of ISO 10816-7 Vibration Severity Zones in Centrifugal Pumps Using Proportional Hazards Models

Marc Vila Forteza, Diego Galar, Uday Kumar and Kai Goebel

Special Issue
Mechanical Systems Condition Monitoring and Fault Diagnosis

Edited by
Dr. Victor H Benitez

MDPI

https://doi.org/10.3390/machines14101203

machines

MDPI

Article

Quantifying the Reliability Impact of ISO 10816-7 Vibration Severity Zones in Centrifugal Pumps Using Proportional Hazards Models

Marc Vila Forteza [1,2,*], Diego Galar [1], Uday Kumar [1] and Kai Goebel [3]

[1] Civil, Environmental and Natural Resources Engineering, Luleå University of Technology, 97187 Luleå, Sweden; diego.galar@ltu.se (D.G.); uday.kumar@ltu.se (U.K.)
[2] Reliability Department, Petronor Oil Refinery, 48550 Muskiz, Spain
[3] Fragum Global, LLC, Mountain View, CA 90012, USA; kai.goebel@fragumglobal.com
* Correspondence: marc.vila.forteza@associated.ltu.se

Abstract

Excessive vibrations in rotating machinery accelerate wear and component failures, causing undesired downtime in industrial plants. To assess the severity of these vibration levels, ISO 10816-7 classifies pump condition into four zones. Zone A represents good condition, Zone B is acceptable for long-term operation, Zone C is generally suitable only for limited operation, and Zone D corresponds to vibration levels that may cause machine damage. However, these qualitative classifications do not quantify how vibration severity relates to equipment reliability. This paper addresses that gap for rotodynamic pumps by integrating standardized vibration severity levels within a reliability framework based on Proportional Hazards Models (PHMs). The models were fitted with observational field data from an oil refinery using the Mean Time Between Failures (MTBF) as the reliability outcome. From the resulting survival curves, reliability is summarized by the median predicted life (B50), defined as the time at which the predicted survival probability reaches 0.5. This measure can be estimated within the observed follow-up without extrapolating the censored tail of the survival curve. The model achieved a concordance of 0.870 and was then used to estimate B50 across a range of vibration levels, holding the remaining operating conditions constant. The analysis indicated an association between the vibration levels and the predicted life, consistent with the severity considerations set by the ISO standards. The results showed a 62.5% decrease in mean predicted B50 from Zone A to Zone D. Within the 0–6 mm/s RMS range, which contained 95% of the observed vibration data, prediction uncertainty remained comparatively low, whereas it increased at higher vibration levels as data density decreased. Across speed and power groups, at a given vibration level, higher rotating speeds and higher installed power are associated with higher predicted lives. American Petroleum Institute (API) 610 between-bearings pumps showed higher predicted reliability than horizontal overhung and vertical overhung OH6 configurations. By providing a quantitative mapping from ISO severity levels to predicted equipment life and its practical implications for reliability-oriented maintenance planning, the proposed approach adds numerical resolution to current standards and supports more informed, data-driven maintenance and reliability decisions for centrifugal pumps.

Keywords: centrifugal pumps; proportional hazards models; ISO 10816-7

Check for updates

Academic Editor: Victor H Benitez

Received: 18 September 2026
Revised: 6 October 2026
Accepted: 8 October 2026
Published: 16 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Condition monitoring (CM) is essential for maintaining the reliability and performance of centrifugal pumps in the oil and gas industry. Parameters such as vibration, temperature, oil condition, flow, power, and pressure can be monitored continuously or periodically to identify potential failures and support maintenance decisions. Among these techniques, vibration analysis (VA) is one of the most widely used approaches for rotating machinery. Its application to machinery condition assessment dates back to early vibration-severity criteria such as those proposed by Rathbone [1,2] and has since evolved towards modern signal-processing and diagnostic methods.

Vibration-based CM typically involves vibration measurement and preprocessing, signal processing, feature extraction and selection, and diagnosis [3]. These steps enable vibration measurements to be transformed into condition information that can support fault identification and maintenance decisions.

Recent prognostic approaches for centrifugal pumps increasingly combine vibration and condition-monitoring data with machine learning [4–9], deep learning [10–13], physics-based or hybrid [14,15], and statistical or probabilistic methods [16–22] to estimate Remaining Useful Life (RUL), defined as the useful life remaining from the present moment. These approaches commonly target individual components such as bearings, impellers, or shafts and often rely on detailed or continuous condition-monitoring data and identifiable degradation trajectories. Although effective in well-instrumented environments, such conditions are not always met in industrial practice, where instrumentation is often partial and data are collected periodically rather than continuously.

In contrast, system-level prognostics aim to characterize the health or useful life of the asset as a whole, accounting for maintenance actions, system interactions, and operational variability [23,24]. This perspective is particularly relevant for complex repairable systems where component-level degradation trajectories are difficult to isolate and maintenance decisions are made at fleet or plant level, rather than by the prediction of isolated component failures.

For repairable industrial assets, Mean Time Between Failures (MTBF) is a widely used system-level reliability indicator, unlike Mean Time To Failure (MTTF), which applies to non-repairable items. System-level pump reliability has been addressed using regression trees [25], ensemble learning [26], neural-network and SVM methods [27], Non-Homogeneous Poisson Processes [28], Proportional Hazards Models (PHMs) [29], and multivariate statistical techniques [30].

Cox PHMs are suitable for industrial applications involving complex systems because they accommodate censored observations and allow operating, condition, design, and maintenance covariates to be related to failure risk without requiring specification of the baseline hazard distribution [31]. Their ability to combine historical reliability information with condition and operational data makes them well suited to system-level reliability assessment in industrial fleets.

The Cox PHM framework used in the present study was previously developed and validated on the same pump dataset [29]. In a complementary study [32], the contribution of individual reliability predictors was subsequently assessed using several statistical and machine-learning techniques, including the partial likelihood-ratio χ^2 statistic computed from the standard Cox PHM. This analysis provided a stable and interpretable measure of variable relevance, with results consistent with the rankings obtained from the other approaches considered. Together, these previous studies provide the methodological and empirical basis for adopting the Cox PHM in this research. The present study builds on this validated framework, using reliability predictors identified from the literature [33] and

domain expertise but addresses a different objective: translating standardized vibration severity levels into quantitative reliability estimates.

When model-based prognostic tools are unavailable, standardized vibration thresholds provide a practical means of assessing machinery condition [34]. The ISO 20816 series [35] defines vibration measurement and evaluation principles for rotating machinery, while ISO 2954 [36] specifies instrumentation requirements. For centrifugal pumps, standards such as ANSI/HI 9.6.4 [37], API 610 [38], and ISO 10816-7 [39] provide vibration criteria based primarily on overall RMS vibration velocity measured on non-rotating parts.

ISO 10816-7 classifies vibration severity into four zones: Zone A represents good machine condition, Zone B is acceptable for long-term operation, Zone C generally permits only limited operation, and Zone D corresponds to vibration levels that may cause machine damage. However, pump reliability also depends on operating, design, and maintenance factors, including flow, pressure, rotating speed, bearing and sealing configuration, lubrication, and maintenance history. Overall vibration should therefore be interpreted as an integrative condition indicator rather than as an isolated cause of failure.

Although ISO 10816-7 provides practical vibration severity thresholds, its zones are based on accumulated empirical experience and are primarily intended for machinery condition assessment rather than lifetime prediction. Despite extensive research on vibration-based condition monitoring and prognostics, a direct quantitative interpretation of the ISO 10816-7 severity zones in terms of pump reliability remains unexplored, particularly at fleet level using industrial field data. Translating these severity levels into reliability estimates can provide additional value to vibration measurements by relating condition severity to expected equipment life, thereby supporting the prioritization and planning of maintenance interventions. The present study addresses this gap by translating the ISO 10816-7 severity classification into reliability estimates expressed in units of time for centrifugal pumps. Rather than introducing a new PHM formulation or vibration diagnostic technique, the study uses a Cox PHM and real operating data to determine how predicted pump lifetime varies across the standardized vibration severity range.

The relationship between vibration level and pump reliability was assessed by systematically varying overall vibration values across the range defined by ISO 10816-7. This variation should be interpreted as a structured marginal prediction analysis within an empirically calibrated hazard model. Rather than augmenting the dataset or generating artificial failure events, the approach evaluates how model-based reliability estimates change across alternative vibration scenarios while retaining the remaining pump-specific covariates at their observed values. The resulting reliability patterns were also examined across different rotating-speed, installed-power, and pump-type groups, providing additional context for interpreting vibration severity in relation to relevant equipment characteristics.

The remainder of the paper is organized as follows. First, the main vibration-related failure mechanisms of centrifugal pumps and the fundamentals of MTBF and PHMs are introduced. The methodology for model fitting, validation, and vibration-reliability mapping is then presented, followed by the case study and results obtained from the industrial dataset. Finally, the findings, limitations, practical implications, and future research directions are discussed.

2. Failures of Centrifugal Pumps and Vibration

2.1. Failures of Centrifugal Pumps in the Oil and Gas Industry

Centrifugal pumps play a fundamental role in reliability and safety in the oil and gas industry, as they handle hazardous fluids that can cause dangerous incidents. Pump failures and their causes have been extensively documented [31,40–46], and many mechanical, hydraulic, operating, and installation-related problems can manifest through increased

vibration. Vibration measurements can therefore provide valuable information on pump condition, although subsequent vibration analysis is generally required to identify the underlying failure mechanism and assess its severity.

Bloch et al. [47] identified a broad set of operating, maintenance, and design factors influencing centrifugal pump reliability. These factors include operating flow relative to the best efficiency point (BEP), Q–H curve shape, suction design, misalignment, rotor unbalance, foundation and grouting quality. Many of these conditions are also associated with increased vibration, supporting the use of overall vibration as an integrative condition indicator. Other factors, including lubrication and sealing condition, may likewise contribute to both vibration and reliability degradation.

2.2. Effect of Speed, Power, and Pump Type on Vibration Levels

Vibration in centrifugal pumps can arise from installation and maintenance deficiencies, incorrect application, hydraulic phenomena, and design or manufacturing conditions [48], as summarized in Table 1. Mechanical sources include rotor unbalance, misalignment, looseness, bearing defects, shaft deformation, and structural resonance. Hydraulic sources include off-design operation, recirculation, cavitation, turbulence, pressure pulsations, and interactions between the pump and piping system. These mechanisms generate dynamic radial, axial, inertial, frictional, and hydraulic forces that are transmitted through the rotor, bearings, casing, and support structure and ultimately appear in the measured vibration response.

Table 1. Causes of vibrations in centrifugal pumps [48].

Category	Issues
Installation maintenance	Unbalance, shaft-to-shaft misalignment, seal rubs, case distortion caused by piping loads, piping dynamic response (supports and restraints), support structural response (foundation), anchor bolts/grout, improper assembly
Application	Operating from a design point, improper speed/flow, inadequate net positive suction head, entrained air
Hydraulic	Interaction of pump (head-flow curve) with piping resonances, hydraulic instabilities, acoustic resonances (pressure pulsations), water hammer, flow distribution problems, recirculation, cavitation, flow induced excitation (turbulence), high flow velocity
Design manufacturing	Lateral critical speeds, torsional critical speeds, improper bearings or seals, rotor instability, shaft misalignment in journals, impeller resonances, bearing housing/pedestal resonances

Importantly, these mechanisms do not necessarily produce the same vibration signature: different faults may be associated with different frequency components, harmonics, broadband responses, or resonance phenomena. Consequently, overall RMS vibration provides an integrated measure of vibration severity but does not identify the specific physical mechanism that produces the vibration. In practice, the measured vibration response results from the interaction of multiple mechanisms. Rotating speed, for example, affects excitation frequencies and centrifugal forces, while flow conditions can modify turbulence, recirculation, and cavitation. These phenomena are well known and continue to be supported by recent experimental and modelling studies. Off-design operation and associated flow instabilities have been shown to significantly modify the vibration response of centrifugal pumps, including overall RMS vibration velocity [49], while cavitation and changes in hydraulic operating conditions are also reflected in measured vibration signals [50]. Mechanical sources such as shaft misalignment and impeller unbalance affect both vibration amplitude and frequency content [51], while pressure pulsations associated with blade-passing phenomena can provide an additional source of excitation and

potentially contribute to structural or acoustic resonance [52]. Structural resonances may amplify these excitations when forcing frequencies are close to the natural frequencies of the installation, and the resulting response can be further influenced by bearing condition, alignment, lubrication, foundation stiffness, and other design and maintenance factors. These interactions explain why similar overall vibration levels may originate from different physical mechanisms and may not imply identical degradation processes.

Identifying the specific source of vibration would require further diagnostic analysis, including frequency-domain information and other vibration analysis techniques, which are outside the scope of this study. The objective here is instead to evaluate overall vibration severity and its relationship with reliability. Overall RMS vibration is therefore used because it is the severity descriptor adopted by ISO 10816-7 and enables a consistent fleet-level assessment across the available industrial dataset. Table 2 summarizes the principal vibration mechanisms considered in this context.

Table 2. Principal causes and physical mechanisms of vibration in centrifugal pumps.

Cause	Physical Mechanism of Vibrations
Rotating speed	Higher speeds increase centrifugal forces and vibration.
Impeller unbalance	Causes radial vibration as a result of centrifugal forces.
Misalignment	Increases axial and radial vibration due to an irregular torque transmission.
Bearing wear	Leads to vibration due to periodic impacts, loss of shaft stability or surface damage.
Hydraulic forces (flow rate)	Off-design flow induces turbulence increasing hydraulic vibration.
Cavitation	Formation and collapse of vapor bubbles, generating shock waves, flow asymmetry, and unbalanced hydraulic forces.
Shaft alignment	Eccentric rotation and asymmetric loading of internal components.
Structural resonance	Amplifies vibration when natural frequency matches excitation.
Baseplate and fixation	Loose or incorrect fixation amplifies vibration due to unwanted movement between components.
Operating condition changes	Transients increase inertial effects.
Installed power	High-power motors may induce higher torques and high-energy vibrations.
Power Supply Fluctuations	Can cause torque variations and unsteady operation.

3. MTBF and Reliability Modelling Using Proportional Hazards Models

3.1. MTBF and Hazard Rate

To monitor uptime, downtime, and the effectiveness of reliability and maintenance strategies, industries rely on various performance metrics. Among the most commonly used are MTBF, Mean Time to Repair (MTTR), and MTTF. MTBF, defined in EN 15341:2019 [53], is a KPI frequently adopted in the oil and gas sector. It quantifies the average operational time an asset is expected to run between two consecutive failures and applies to assets that can be repaired, hence the term "between failures". Depending on the expected lifetime of the asset, it can be expressed in different time units, such as hours, days, or months, and can be computed for individual assets or populations using (1) and (2) respectively.

$$\mathrm{MTBF}_{\mathrm{asset}} = \frac{Operating\ time\ recorded}{Number\ of\ Failures} \tag{1}$$

$$\mathrm{MTBF}_{\mathrm{population}} = \frac{Operating\ time * Number\ of\ assets}{Number\ of\ Failures} \tag{2}$$

The reliability $R(t)$ of a component, equipment, or system is the likelihood that it will function correctly under certain conditions for a specified duration (t). It is closely related to MTBF and follows an exponential model for assets with a constant hazard rate (λ). In the reliability distribution, represented by (3), the MTBF is the reciprocal of λ as shown in (4). These equations can be applied when: (1) the probability of failure per time unit stays constant, independently of the asset's age or usage; and (2) assuming the time between failures follows an exponential distribution, indicating each failure occurs independently of the previous one, λ remains constant.

$$R(t) = e^{-\lambda t} \tag{3}$$

$$\text{MTBF}_{\text{asset}} = \frac{1}{\lambda} \tag{4}$$

However, in many practical situations, the hazard rate λ changes over time and becomes time-dependent ($\lambda(t)$). In such cases, the MTBF cannot be directly calculated as the inverse of the hazard rate, and the reliability function $R(t)$ must be integrated using $\lambda(t)$ as shown in (5). The Weibull distribution is commonly used to model failure times due to its flexibility in representing various hazard rate behaviors. By adjusting the shape parameter β and the scale parameter η, it can capture different failure patterns. The time-dependent nature of the hazard rate is governed by β, while η determines the expected timing of failures. Their combined influence on the hazard rate is defined by (6).

$$R(t) = e^{-\int_0^t \lambda(u)du} \tag{5}$$

$$\lambda(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \tag{6}$$

3.2. PHMs

PHMs are widely used in survival analysis because they provide a flexible framework for analyzing time-to-event data and assessing the effects of predictors on the hazard. They have been applied across fields such as medical research, social sciences, and engineering.

PHMs assume the hazard function for an individual can be expressed as the product of a baseline hazard function and a relative risk function dependent on certain covariates. Mathematically, the hazard function at time t for an individual with covariates $\boldsymbol{X}$ is given by (7).

$$h(t|\mathrm{X}) = h_0(t)e^{(\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p)} \tag{7}$$

In this equation, $h_0(t)$ is the baseline hazard function, which is non-parametric and captures the underlying risk over time, while the exponential term captures the relative risk associated with the p covariates $\mathbf{X}$ [54].

The proportional hazards assumption implies that hazard ratios between individuals remain approximately constant over time, such that covariates act multiplicatively on the baseline hazard function without modifying its shape. The validity of this assumption can be explored using scaled Schoenfeld residuals. The Cox PHM also assumes a common baseline hazard function across individuals, differing only by the relative risk component. Other conditions are that the survival times of individuals are assumed to be independent of each other, and the model accommodates non-informative right-censoring. The last condition means some subjects do not experience the event by the end of the study, as the censoring is independent of the event risk.

Regression coefficients in PHMs are usually estimated by maximizing the partial likelihood function. The partial likelihood is formulated by evaluating the probability

of a specific event occurring at a given moment, based on the group of individuals still at risk at that time. The case with no ties is standard [47], but the presence of tied event times is common in practice [55]. In that case, exact partial likelihoods are available but computationally demanding [56,57]; therefore simpler approximations are widely used. Breslow's and Efron's approximations are the most common, with Efron generally more accurate when ties are frequent [58,59]. The Efron approximation was used for fitting the Cox PHM in the present study.

Building on this, the Cox PHM and the MTBF are connected by the hazard rate. Within the PHM framework, it is represented as $h(t \mid \mathbf{X})$ in (8), where $h_0(t)$ is the baseline hazard function (the hazard when all covariates are zero), $h(t \mid \mathbf{X})$ is the hazard rate at time t for a specific set of covariates $\mathbf{X}$, and β represents the coefficients that quantify the effect of each covariate. In the PHM framework, the hazard therefore varies according to the covariate values.

$$h(t|\mathrm{X}) = h_0(t)\mathrm{e}^{\beta \mathrm{X}} \tag{8}$$

3.3. Relationship Between Reliability and Hazard Rate

An illustrative relationship between the reliability function $R(t)$ and hazard rate $h(t)$ is shown in Figure 1. The hazard rate represents the instantaneous risk of failure at time t, while the reliability function represents the probability of surviving beyond that time. These quantities are related through the cumulative hazard, such that the survival probability decreases as failure risk accumulates over time. In the subsequent analysis, reliability is summarized using the median predicted life B50, corresponding to $S(t) = 0.5$, as detailed in Section 4.2.

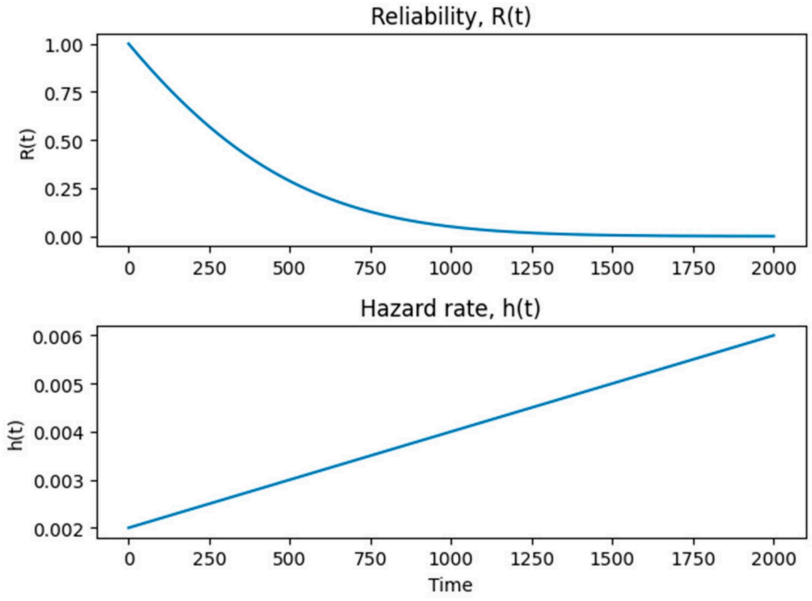

Figure 1. Relationship between hazard rate and reliability function. Values are representative for conceptual visualization only.

In many industrial assets, failure risk depends on operating conditions, degradation mechanisms, maintenance history, and other covariates. Different survival modelling approaches can incorporate these factors. Parametric models, such as Weibull, log-normal, and generalized gamma formulations, require specification of the underlying survival-time distribution, while machine-learning alternatives such as Random Survival Forests (RSF) can capture complex nonlinear relationships and interactions. In contrast, the Cox PHM leaves the baseline hazard function unspecified while accommodating censored observations and quantifying the association between covariates and failure risk [60,61]. This

makes the Cox PHM particularly suitable when the underlying failure-time distribution is unknown or difficult to justify, as is often the case in industrial applications involving heterogeneous operating conditions.

4. Methodology

This section describes how a previously validated reliability framework was applied to quantify the relationship between vibration and the predicted reliability of centrifugal pumps. The contribution of this work does not lie in the model-fitting procedure, which follows the framework validated in [29,32], but in its practical application to map ISO 10816-7 vibration severity zones onto reliability predictions. The methodology is organized into two stages. In the first, the Cox PHM is fitted and validated using the real pump fleet and the variables described in Section 5, including the fitting of candidate models, data-reduction strategies, and selection of the final model based on predictive performance and control of overfitting. In the second, the fitted model is evaluated across the ISO vibration severity range using a prediction grid generated by varying vibration while retaining the remaining pump-specific covariates at their observed values. Importantly, this prediction grid is used only to query the already fitted model and does not enter model fitting or validation.

4.1. Reliability Model Fitting and Validation

The procedure used to fit and select the Cox PHM follows the sequence of steps shown in Figure 2. The objective is to fit a model with a balanced predictive capacity while minimizing the likelihood of overfitting. Overfitting arises when a model captures not only the underlying structure of the data but also its random noise. As a result, the model may perform well on the training data but fail to perform on unseen data because the learned patterns are not generalizable.

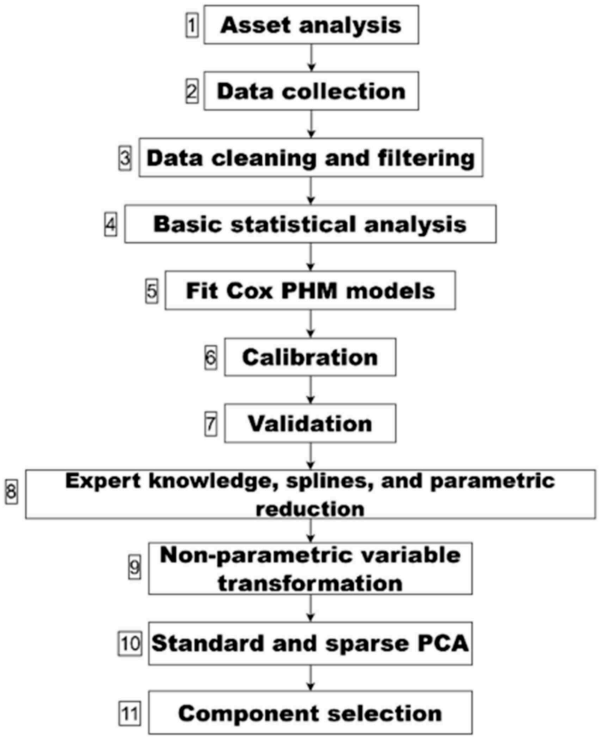

Figure 2. Steps of the method used for model fitting and data reduction [29].

The first step is to select and analyze the characteristics of the assets to be modelled, in this case, a fleet of centrifugal pumps. This approach is suitable for modelling with multiple potential predictors, for which data reduction may be required to limit the risk of

overfitting. The second step is to collect data for the analyzed assets. Historical data from the maintenance and reliability departments should be collected together with the design parameters available in the API datasheets of the assets.

The response variable of the model is the MTBF of each pump. For pumps with at least one recorded failure, it is computed as the operating time over the observed timeframe divided by the number of failures registered in that period. Pumps that did not fail during the observed period are included as right-censored observations at their total operating time. This right-censoring, which the Cox PHM handles natively, avoids discarding the reliability information of those pumps that did not fail.

In the third step, the data should be reviewed and the measurement units harmonized to be consistent for all observations in the dataset. Items with incomplete observations or those whose working conditions are infrequent or too extreme may be removed. The objective of this step is to remove observations that are not representative of the general operating conditions of the fleet. The fourth step is to perform a basic statistical analysis to detect possible correlations between the predictors and redundant variables. Correlations between the predictors and the output variable are evaluated, and basic statistical parameters are computed to assess the quality of the data and their suitability for the regression model.

In the fifth step, different Cox models are fitted: (1) assuming linear covariate effects and (2) representing these effects using restricted cubic splines (RCS) with a maximum of five knots per covariate, placed at Harrell's recommended quantile positions [60] p. 28. The number of knots assigned to each variable is later adjusted according to its relative importance, measured by its penalized χ^2 likelihood-ratio in the full model. Therefore, the linearity of the covariate effects is not assumed but tested.

Model calibration was addressed in step six by assessing the agreement between predicted and observed survival probabilities. No separate hold-out set was used. Instead, following Harrell's validation strategy, Efron's bootstrap with 1000 resamples of the 666 observations was applied for internal validation [60] p. 114. Bootstrap validation was preferred over a single train–test split because it allows the complete dataset to be used for model development while estimating the optimism associated with model performance. In each bootstrap iteration, the model fitted to the resampled dataset is evaluated both in the bootstrap sample and in the original dataset. The average difference in performance is then used to obtain optimism-corrected validation indices, providing an internal assessment of model performance and overfitting while making full use of the available data.

In step seven, the optimism-corrected validation indices are used to assess model performance and the potential for overfitting. If the model is likely to be overfitted, step eight is initiated, and the formal data reduction process starts, using a data reduction approach consisting of: (1) removal of variables according to expert knowledge, (2) reduction of the number of degrees of freedom (d.f.) of the previous fitted RCS, and (3) parameterization of those variables whose relationship with the log-hazard ratio is similar to known functions. After this reduction, steps five to seven are performed again, and if the models do not present a risk of overfitting, they are selected. However, if the models continue to show potential overfitting, step nine begins. At this stage, non-parametric transformations of the variables are performed to further reduce the dimension of the models. After this variable transformation, step ten is initiated. Two PCA techniques are applied on the transformed variables: standard PCA and sparse PCA. The aim of the tenth step is to further reduce the number of d.f. to mitigate overfitting while maintaining acceptable performance indices of the models. Finally, in step 11, the principal components group whose model shows the best balance between predictive capacity and minimal risk of overfitting is selected. Once the final sparse PCA Cox PHM is selected, the proportional hazards assumption is

assessed using scaled Schoenfeld residuals, both globally and for each principal component included in the final model.

All computations were carried out in R (RStudio 2024.12.1): the survival, rms and Hmisc packages were used for model fitting and validation; glmnet for the penalized comparisons; pcaPP for the robust and sparse PCA; and splines and acepack for the spline and ACE/AVAS transformations. The complete list of R packages is provided in [29].

4.2. Mapping Vibration Severity to Predicted Reliability

Once the model has been fitted and validated as described in the previous section, it is used to quantify how predicted reliability responds to vibration. Figure 3 summarizes the workflow of this mapping procedure, whose steps are described below.

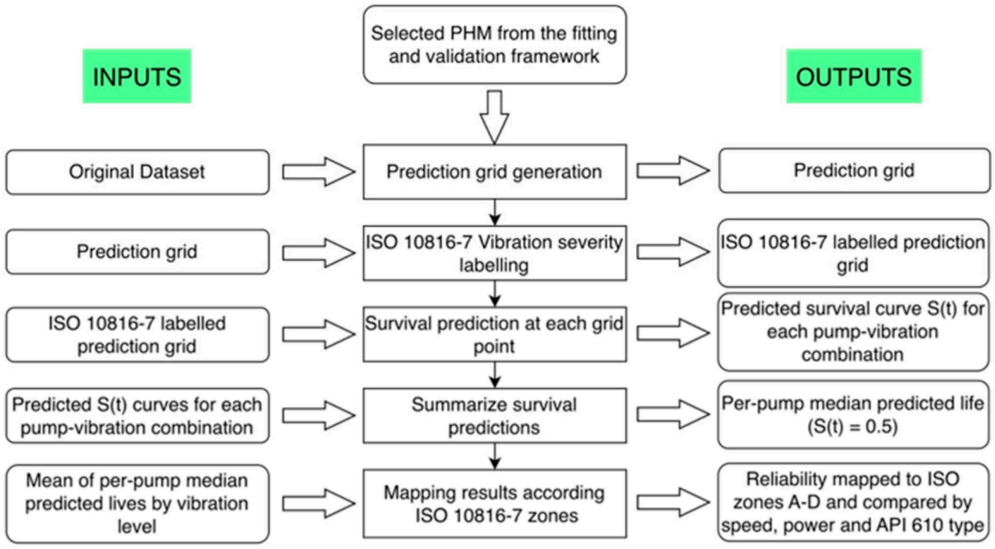

Figure 3. Flowchart for mapping vibration effects on reliability of centrifugal pumps.

In the first step, a set of prediction observations (prediction grid) is generated by evaluating the model at sixteen discrete overall vibration levels selected to cover the four ISO 10816-7 severity zones across their full range (0.5, 1.5, 2, 2.75, 3.25, 3.75, 4.25, 5.25, 6, 6.75, 7.25, 8, 8.5, 9.5, 10.5 and 12 mm/s RMS) while holding all remaining covariates of each pump at their observed values. For each vibration level in the prediction grid, the corresponding spline-transformed predictor vector was constructed, standardized using the centering and scaling parameters obtained during model fitting, and projected onto the sparse PCA loadings. Thus, the PC scores were recalculated for each grid observation and used as inputs to the fitted Cox PHM.

The overall RMS vibration level does not act as an isolated cause but as an integrative condition indicator: mechanical conditions such as unbalance, misalignment, looseness, or bearing wear can manifest themselves through changes in vibration. Therefore, the vibration descriptor may reflect, in aggregate, the contribution of multiple underlying mechanical conditions. This interpretation is also consistent with the use of overall vibration as the severity descriptor in ISO 10816-7, whose severity zones are defined according to the measured vibration level. This simplification is further discussed in Section 7. Accordingly, varying overall RMS vibration in the prediction grid is intended to evaluate the model response across the same severity descriptor used by ISO 10816-7, rather than to simulate the individual physical mechanisms responsible for the measured vibration.

The empirical structure of the original dataset also supports this interpretation. Redundancy analysis showed that overall RMS vibration was only partially explained by the remaining measured covariates (R^2 = 0.421). Installed power showed the strongest associa-

tion with vibration (R^2 = 0.324), while the remaining covariates increased the explained variability only slightly (R^2 = 0.097). This indicates that the measured covariates do not fully account for the observed variability in vibration and is consistent with vibration also reflecting the mechanical conditions referenced above, which are not explicitly represented in the dataset.

Each observation in the prediction grid is assigned to its ISO 10816-7 severity zone (A, B, C, or D) [39], which is derived from the overall RMS vibration level and the installed power following the standard, rather than introduced as an independent variable. To examine differences across rotating-speed ranges, observations are also grouped into three speed ranges (low, medium, and high) and by power rating (below and above 200 kW). These speed and power groups are defined only for the presentation and comparison of results, not for model fitting, enabling the subsequent analysis of how power and vibration levels relate to reliability.

Next, for each observation in the prediction grid, the survival function is computed using the fitted model from Subsection A. The baseline survival function is estimated consistently with the Efron method used to fit the Cox model, and the individual survival curves are obtained by combining this baseline estimate with the corresponding predictors. This procedure constitutes a controlled prediction analysis that evaluates the modelled association between vibration and predicted reliability, while the remaining measured covariates are retained at their observed values, and does not attempt to reproduce the full physical state of a specific pump operating at a given vibration level.

Reliability is then summarized by the median predicted life (B50), defined as the time at which the predicted survival probability reaches 0.5. This quantity is distinct from the mathematical mean of the predicted survival-time distribution and was selected because it can be estimated within the observed follow-up, whereas estimating the unrestricted mean would require assumptions about the survival curve beyond the observation period. For each vibration level, the B50 was obtained for each pump from its individual predicted survival curve, and the mean of the resulting per-pump B50 values was reported as the representative reliability estimate at that vibration level. Accordingly, MTBF refers throughout this study to the observed reliability outcome used to fit the Cox PHM, whereas B50 refers to the model-based lifetime summary derived from the predicted survival curves.

In addition, for each grid point, prediction uncertainty is assessed through the 95% confidence interval (CI) of the mean reliability estimate and the individual 95% confidence intervals derived from the fitted survival curves. Together with the data density of the original dataset across the vibration range, these quantities indicate where predictions are well supported by observed data and where they should be interpreted with caution.

The predictions obtained from the prediction grid also enable comparisons across relevant equipment characteristics. Specifically, predicted B50 is compared between pumps rated below and above 200 kW; across low-, medium-, and high-speed ranges; and among API 610 [37] pump types. Finally, the predicted values are grouped according to the A, B, C, and D severity zones established by ISO 10816-7.

The application described in this section was implemented in R (RStudio 2024.12.1). The predicted survival function at each grid point was obtained from the fitted Cox PHM using the survest function of the rms package (with Hmisc and the underlying survival package). The tidyverse (dplyr) tools were used to construct the prediction grid, assign the ISO 10816-7 severity labels and group the results by operating speed, installed power and pump type, while ggplot2 was used for the survival and reliability plots. The software environment used for model fitting is the one described in Section 4.1, which is used in this step for prediction and post-processing.

5. Case Study

5.1. Dataset and Variables Used for Model Fitting

The approach described in the previous section was used on field data consisting of 675 pumps at an oil refinery. The facility operates a fleet of 3200 rotating machines. Approximately 1200 are centrifugal process pumps, representing 38% of the entire rotating machinery assets. Each pump was constructed in accordance with API 610. The dataset covers the main API 610 pump configurations represented at the refinery, including horizontal overhung pumps (OH1, OH2, OH3, and OH5), vertical OH6 pumps, and between-bearings pumps (BB1, BB2, BB3, and BB5). These configurations represent different mechanical arrangements and operating characteristics, providing a heterogeneous industrial population for the reliability analysis.

The pumps are placed in different areas of the facility and operate under a wide range of conditions, enhancing the representativeness of the dataset. Table 3 presents the distribution of pumps across production areas, both for the whole refinery and for the subset corresponding specifically to the dataset. The percentages indicate the proportion of pumps located in each area relative to the total number of pumps in the refinery and the dataset, respectively. This breakdown helps contextualize the operational environment and highlights the concentration of assets in key areas, such as atmospheric distillation and conversion. The selected dataset was reviewed, and nine entries whose working conditions were infrequent or too extreme were removed. Consequently, a subset of 666 pumps was used to fit the model.

Table 3. Dataset composition by production area.

	Dataset		Refinery	
	% Pumps	**Number of Pumps**	**% Pumps**	**Number of Pumps**
Atmospheric distillation area	44.8%	303	41.3%	503
Conversion area	35.8%	241	26.3%	320
Fuel reduction area	12.2%	82	10.8%	132
Tanks and dock	7.3%	49	21.6%	263

Vibration measurements in the original dataset were acquired in accordance with ISO 20816-3 [35], using portable data collectors equipped with piezoelectric accelerometers with a measurement bandwidth of 0.2–10 kHz. Vibration measurements were performed on the bearing housing at the coupling side, both for overhung pumps and between-bearings configurations. The overall vibration severity, expressed as RMS velocity in mm/s, was computed directly by the portable collector following the same ISO standard requirements.

Twenty-nine potential predictors were selected (see Table 4). All of them are commonly referenced in the selection and evaluation of industrial centrifugal pumps, available in API 610 datasheets and maintenance records, and known to influence the MTBF of centrifugal pumps [25,39,59]. For analytical clarity, they were grouped into six categories: operating conditions (vibration values measured with non-intrusive piezoelectric accelerometers), hydraulics, mechanics, sealing, age, and maintenance history, as shown in Table 4. This categorization facilitated the structured description of the predictors included in the analysis. Rotating speed and installed power were included as continuous covariates in the fitted model. For the presentation of results, assets were subsequently grouped into speed and power ranges.

Table 4. Potential predictors considered for the PHM.

Category	Variables
Operating conditions	Fluid type, bottom pump, ISO 10816-7 severity zone, flow ratio, NPSH margin, relative density, dynamic viscosity, vapor pressure, discharge pressure, fluid temperature, vibration level
Hydraulics	Double suction, tip speed, diameter ratio, efficiency, suction specific speed, specific speed, suction specific speed ratio
Mechanical	Power, rpm, bearing type, lube type
Sealing	Seal arrangement, seal type, API 682 plan, number of seals
Age	API 610 edition
Maintenance	Lube work orders, ordinary work orders

5.2. Prediction Grid

Next, the required prediction grid was created to support the subsequent analysis at different vibration levels using the model fitted previously with the real dataset. For each of the 666 pumps, sixteen prediction-grid observations were generated, corresponding to overall vibration levels ranging from 0.5 to 12 mm/s RMS, while the remaining pump-specific covariates retained their observed values. This resulted in 666 × 16 = 10,656 prediction-grid observations. Rotating speed was subsequently categorized into three ranges: low (≤1500 rpm), medium (>1500–3000 rpm), and high (>3000 rpm), based on the fact that 1500 rpm and 3000 rpm are the most common synchronous speeds in regions where the electrical frequency is 50 Hz, corresponding to motors with four and two poles, respectively. Installed power was similarly categorized as below and above 200 kW. Predicted B50 values were then obtained for the 10,656 grid observations using the previously fitted model. The structure of both the original dataset and the prediction grid, classified by pump type, is shown in Table 5.

Table 5. Composition of the original dataset and prediction grid by API 610 pump type.

Observations	API 610 Pump Type								
	BB1	**BB2**	**BB3**	**BB5**	**OH1**	**OH2**	**OH3**	**OH5**	**OH6**
Original	34	87	37	5	130	262	8	10	89
Prediction grid	544	1392	592	80	2080	4192	128	160	1424
Total original	666 observations								
Total pred. grid	10,656 observations								

5.3. Fitted Model and Validation

The methodology described in the previous section was applied to fit and compare candidate Cox PHMs. After data preparation and preliminary statistical analysis, two initial models were fitted using all available predictors: one assuming linear effects of the continuous covariates and another representing these effects using RCS with up to five knots.

Model calibration was assessed at a 1500-day prediction horizon using Efron's bootstrap with 1000 resamples of the 666 observations. The 1500-day horizon was selected because it is close to the central tendency of the observed lifetime distribution and lies within a well-populated region of the follow-up period. The bootstrap procedure was also used to obtain optimism-corrected validation indices. Next, the validation process was applied to the same resampled datasets. While the model using the RCS approximation

showed better performance in metrics such as Somers' D_{xy} and explained variance R^2, it also indicated a higher risk of overfitting, as evidenced by a lower calibration slope than the linear model. Both models had slopes below 0.9, highlighting the need for shrinkage techniques or data reduction.

Initial data reduction was guided by domain expertise and redundancy analysis, reducing the d.f. from 60 to 35 for the linear model and from 103 to 75 for the spline model. Non-linear effects were then assessed using Chi-square statistics; this revealed evidence of non-linearity, supporting the use of RCS representations. Further dimension reduction involved simplifying continuous variables by reducing knot numbers and reparametrizing covariates based on the shape of relative hazard curves. Likelihood ratio χ^2 statistics were used to evaluate the impact of each reduction, with most covariates showing improved or unaffected explanatory power. Despite these reductions, validation metrics remained suboptimal, leading to the application of PCA and additional variable transformations.

Standard and sparse [62] PCA was applied to four sets of covariates: raw variables, splined (RCS) variables, a set transformed using the Alternating Conditional Expectation (ACE) method [63], and a new set computed using the Additive Variance Stabilizing (AVAS) transformation [64]. The goal was to reduce dimensionality while preserving predictive performance, using AIC values to evaluate model fit as the number of components was progressively increased. Spline-based PCA models achieved the lowest AICs but showed no significant improvement beyond 40 components. AVAS-transformed models, while less complex, offered a favorable balance of model simplicity and performance, achieving competitive AICs with fewer components. Sparse PCA yielded AIC results similar to those for standard PCA but with lower d.f., allowing better interpretability of the principal components since many loadings were exactly zero, making it possible to identify which spline-transformed variables contributed most to each component.

Although the ACE- and AVAS-transformed models were also evaluated, as described in detail in [29], their additional transformation complexity reduced the interpretability of the vibration effects without providing a clear performance advantage over the spline-based sparse PCA approach. The final sparse PCA Cox PHM retained 22 principal components. The number and combination of components were selected by balancing model fit and dimensionality reduction within the survival modelling procedure, using AIC to assess the trade-off between goodness of fit and model complexity. This selection was intended to retain components providing relevant information for the survival model while limiting the number of model parameters and the associated risk of overfitting. Therefore, the latter was selected for the present vibration-reliability analysis, offering a favorable balance of complexity and predictive performance (see Table 6).

Figure 4 presents a heatmap displaying the principal components included in the final model, together with the loadings (weights) of each variable, either their spline basis components or their parametrized functions. These loadings quantify the contribution of each standardized predictor or spline basis function to the corresponding principal component. Prior to PCA, all predictors were standardized to zero mean and unit variance to ensure comparability across variables and spline components. Larger absolute loading values therefore indicate greater contribution to the corresponding principal component.

Table 7 provides the quantitative specification of the final sparse PCA Cox PHM, reporting the explained variance of each of the 22 selected principal components together with their corresponding hazard ratios, 95% CIs, and p-values. Because the principal-component scores have different variances, hazard ratios are expressed for a one-standard-deviation increase in each component score, providing a common scale for comparison across components. These hazard ratios refer to the complete PC scores and should therefore

not be interpreted as the effects of individual original predictors. The explained-variance values represent the proportion of the total variance associated with each component.

Table 6. Main indices of the model used for prediction. All validation indices are optimism-corrected (bootstrap 1000 resamples).

Description	PHM Sparse PCA Raw Splined Variables		
Model tests	**LR χ^2**	**d.f.**	
	690.35	22	
Discrimination indices	**R^2**	**D_{xy}**	**$R^2_{df,425}$**
	0.645	−0.740	0.737
Predictive discrimination	**Concordance**	***AIC***	***AIC* χ^2 scale**
	0.870	5312.91	646.35
Validation	**LR R^2**	**D_{xy}**	**Calibration slope**
	0.617	−0.725	0.923
Calibration	**Mean |err|**	**0.9 quantile |err|**	
	0.0077	0.0182	
Proportional hazards diagnostic	**Schoenfeld global χ^2**	113.88	
	***p*-value**	<0.001	
	Significant components	C7, C13, C14, C29, C32, C39, C40	
	Vibration-related PC	C1: $p = 0.181$; C36: $p = 0.117$; C44: $p = 0.176$	

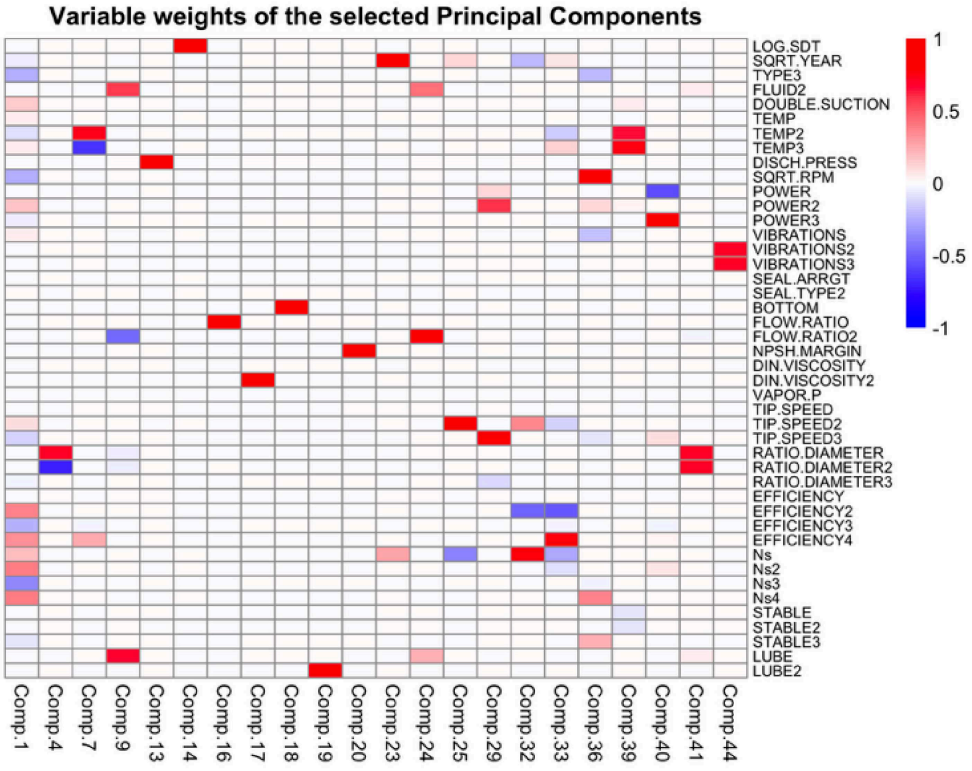

Figure 4. Weights of each splined and parameterized variable for each component of the model.

Table 7. Selected principal components, explained variance, and Cox PHM estimates.

PC	Explained Variance (%)	HR per 1-SD Increase	95% CI	*p*-Value
C1 [1]	12.708	0.615	0.415–0.911	0.015
C4	3.732	0.049	0.013–0.188	<0.001
C7	3.190	0.508	0.214–1.203	0.124
C9	2.286	0.459	0.245–0.859	0.015
C13	2.273	1.025	1.017–1.033	<0.001
C14	2.273	6.372	5.306–7.653	<0.001
C16	2.273	0.677	0.175–2.622	0.572
C17	2.273	15.903	2.031–124.529	0.008
C18	2.273	1.295	0.936–1.792	0.119
C19	2.273	0.368	0.094–1.437	0.150
C20	2.273	1.003	1.000–1.007	0.045
C23	2.227	5.146	2.775–9.543	<0.001
C24	2.179	2.545	1.115–5.811	0.027
C25	2.174	0.269	0.116–0.625	0.002
C29	2.041	0.336	0.170–0.664	0.002
C32	1.457	0.598	0.372–0.963	0.034
C33	1.418	2.056	1.255–3.368	0.004
C36 [1]	1.086	0.969	0.930–1.010	0.139
C39	0.840	0.322	0.170–0.611	<0.001
C40	0.827	2.422	1.258–4.665	0.008
C41	0.814	5.998	2.327–15.461	<0.001
C44 [1]	0.517	7.341	3.395–15.872	<0.001

[1] Components containing non-zero loadings for the vibration spline-basis terms.

To further interpret the role of vibrations, the spline-basis loadings were examined across the selected components. Non-zero vibration loadings were identified only in C1, C36, and C44 as indicated in Table 7. The corresponding loadings for VIBRATIONS, VIBRATIONS2, and VIBRATIONS3 were (0.0575, 0, 0), (−0.1903, 0, 0), and (0, 0.7071, 0.7071), respectively. VIBRATIONS represents the linear spline-basis term, whereas VIBRATIONS2 and VIBRATIONS3 represent the two nonlinear basis terms. Importantly, the explained variance of a principal component reflects its contribution to variability in the predictor space and should not be interpreted as a measure of its prognostic relevance. The final components were selected according to their contribution to the Cox survival model rather than according to a cumulative explained-variance threshold.

Finally, the top plot of the summary panel in Figure 5 shows the evolution of the AIC as a function of the number of principal components included in the model, while the bottom plot displays the proportion of variance explained by the selected components. This visualization highlights the trade-off between model fit and complexity.

The assessment of the proportional hazards assumption using scaled Schoenfeld residuals indicates evidence of departure from strict proportionality for the final model (global test: χ^2 = 113.88, 22 d.f., $p < 0.001$). This departure is not uniform across the principal components, with statistically significant time-dependent effects identified for

C7, C13, C14, C29, C32, C39, and C40. Importantly for the vibration-reliability analysis, the principal components containing the vibration spline terms (C1, C36, and C44; Table 7) showed no statistically significant evidence of non-proportionality (p = 0.181, 0.117, and 0.176, respectively).

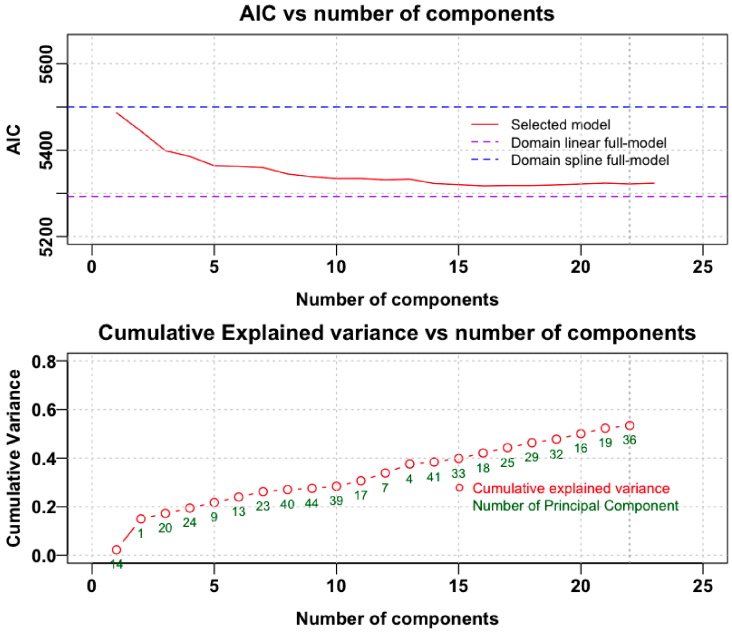

Figure 5. AIC (**top**) and cumulative proportion of total explained variance (**bottom**) as a function of the number of principal components included in the model.

6. Results

6.1. Effect of Vibrations on Predicted Life (B50)

Predicted B50 values were obtained for the 10,656 observations of the prediction grid referenced in Section 5.2. using the model explained in the previous section. The objective was to evaluate how predicted pump reliability varied across vibration levels and to assess whether the resulting trends were consistent with the severity considerations of ISO 10816-7.

Figure 6 represents the predicted B50 values produced by the model versus the overall vibration level in mm/s RMS. For each vibration value, the mean of the per-pump median predicted lives was used to construct the B50 versus vibration curve. The 95% CI around the mean reliability estimate reflects the uncertainty of the group-level estimate, while the mean limits of the individual 95% CIs provide an indication of the average prediction uncertainty at the individual-pump level. This distinction allows both aggregate and individual-level prediction uncertainty to be visualized on the same curve. The figure also includes the data density from the real dataset and the vibration range for 95% of the observations to help identify the region in which the predictions are best supported by the observed data.

From an engineering perspective, Figure 6 indicates a non-linear relationship between vibration severity and predicted B50. At low vibration levels, from approximately 0.5 to 3.5 mm/s RMS, the predicted B50 decreases progressively as vibration increases. From 3.5 to 6 mm/s RMS, the curve becomes considerably flatter, indicating relatively small changes in predicted B50 over this range.

Figure 7 shows the predicted B50 values for each ISO 10816-7 severity zone (according to Table 8), together with the 95% CI of the group-level estimate and the corresponding individual-level prediction uncertainty. In line with this standard, all equipment in the dataset was classified as Category I, which typically includes pumps requiring a high level of reliability, availability, or safety, such as those used in critical applications, including

oil and gas, specialized chemical processes, nuclear facilities, and power plants. Finally, Table 9 presents a summary of the predicted B50 values corresponding to the ISO vibration severity levels. For each vibration category, the table includes the relative deviation (%) of the corresponding 95% CIs, both for the group-level prediction and for the individual predictions within the grid. Table 9 also reports the results of a sensitivity analysis in which the same 22 principal components were retained, while the effects of the 7 components showing evidence of non-proportionality (C7, C13, C14, C29, C32, C39, and C40) were allowed to vary with log time, such that $\beta_j(t) = \beta_j + \gamma_j \log(t)$. The relative deviation of these sensitivity estimates from the corresponding primary model B50 values is reported in brackets. Consistent with Figure 6, predicted B50 decreases from Zone A to Zone B, shows only a slight additional decline from Zone B to Zone C, and then drops sharply from Zone C into Zone D. However, the greater uncertainty associated with Zone D should be considered when interpreting this decline.

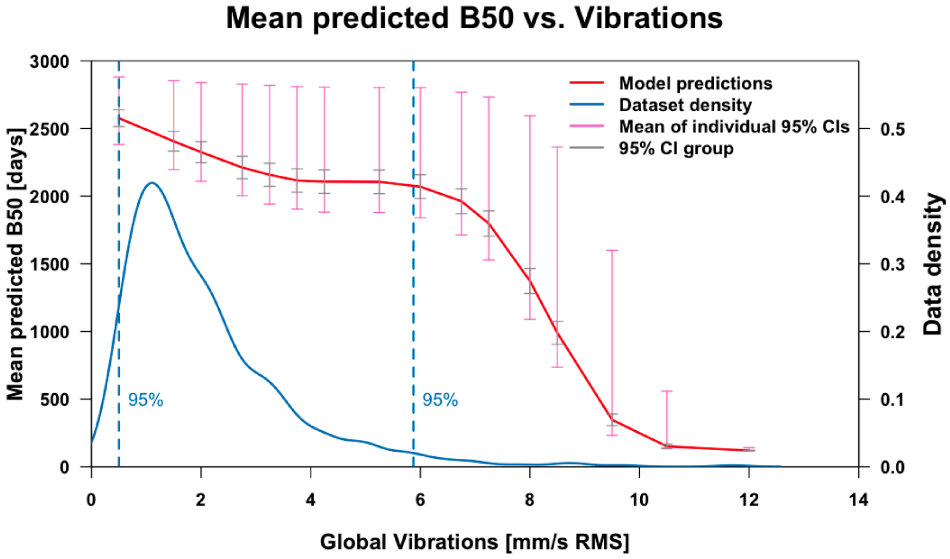

Figure 6. Mean predicted B50 vs. vibration levels of the prediction grid, including confidence intervals, and data density of the dataset.

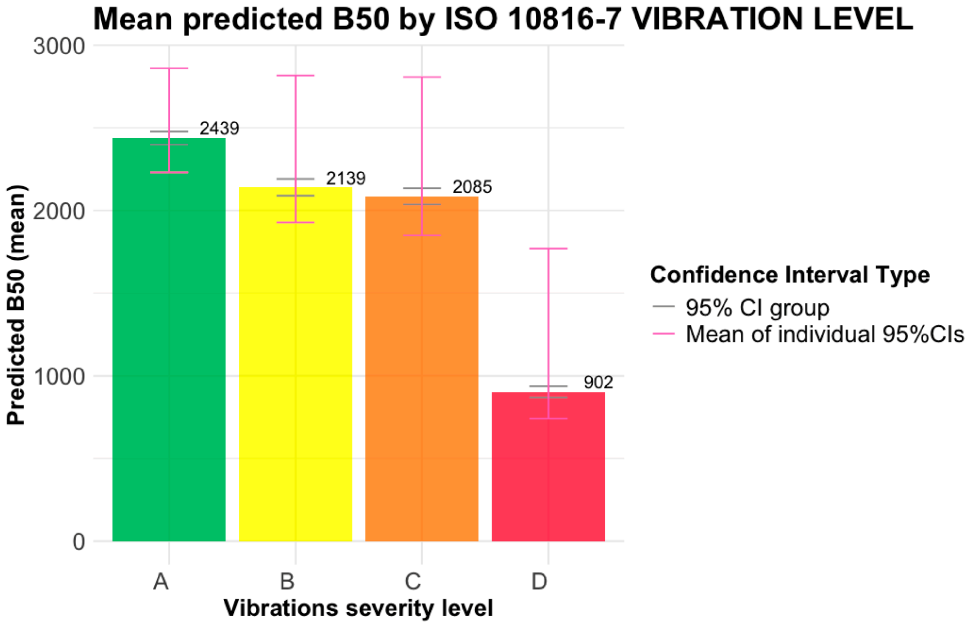

Figure 7. Mean predicted B50 vs. vibration severity levels of the prediction grid.

Table 8. Vibration severity zones ISO-10816-7 (Category I).

Zone	Power ≤ 200 kW (mm/s RMS)	Power > 200 kW (mm/s RMS)
A	≤2.5	≤3.5
B	>2.5 to ≤4.0	>3.5 to ≤5.0
C	>4.0 to ≤6.5	>5.0 to ≤7.6
D	>6.5	>7.6

Table 9. Mean predicted B50 and 95% confidence intervals by ISO severity zone, including relative deviations (%) and B50 estimates from the time-dependent coefficient sensitivity analysis.

Severity Zone	Mean B50 (Days)	Group Mean: Lower Dev. (%)	Group Mean: Upper Dev. (%)	Individual: Lower Dev. (%)	Individual: Upper Dev. (%)	Sensitivity B50 (Days) [Δ%]
A	2439	−1.61	1.61	−8.55	17.29	2379 [−2.45%]
B	2139	−2.37	2.37	−9.91	31.68	2105 [−1.58%]
C	2085	−2.35	2.35	−11.30	34.61	2048 [−1.78%]
D	902	−3.73	3.73	−17.96	96.08	984 [+9.09%]

6.2. Results Grouped by Operating Speed Ranges

To analyze the differences in predicted B50 across rotating-speed ranges, the prediction grid observations were grouped into three speed categories: low (0–1500 rpm), medium (>1500–3000 rpm), and high (above 3000 rpm). Predicted B50 values across vibration levels were then compared among these groups. As shown in Figure 8, all three groups follow a similar vibration–B50 pattern to that observed in Figure 6, with slightly higher predicted B50 values observed in the higher-speed groups at comparable vibration levels.

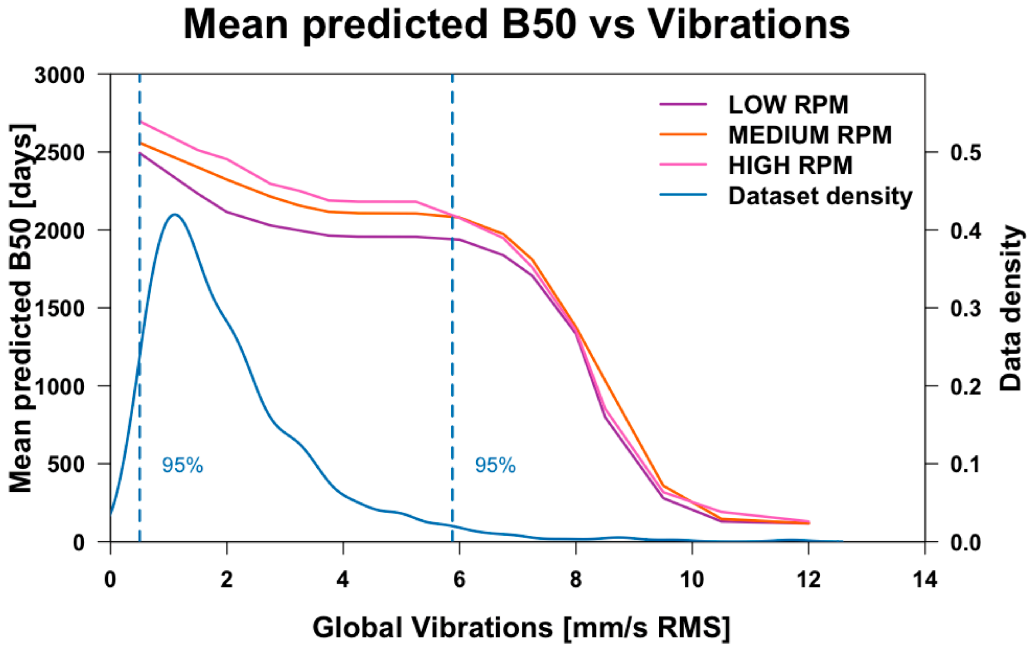

Figure 8. Mean predicted B50 at different vibration levels and at different speeds.

6.3. B50 Results Grouped by Installed Power According to ISO 10816-7

Similarly to the approach taken with rotating speed, the B50 results were grouped by installed power. In accordance with ISO 10816-7, the dataset was divided into two categories: ≤200 kW and >200 kW. Severity zones were then assigned by considering that

all equipment was classified as Category I. The results, referenced to overall RMS vibration (Figure 9) and ISO severity zones (Figure 10), follow the same pattern observed in Figure 6. For a given vibration level, differences in predicted B50 between the two power groups are more pronounced than those observed among the rotating-speed groups.

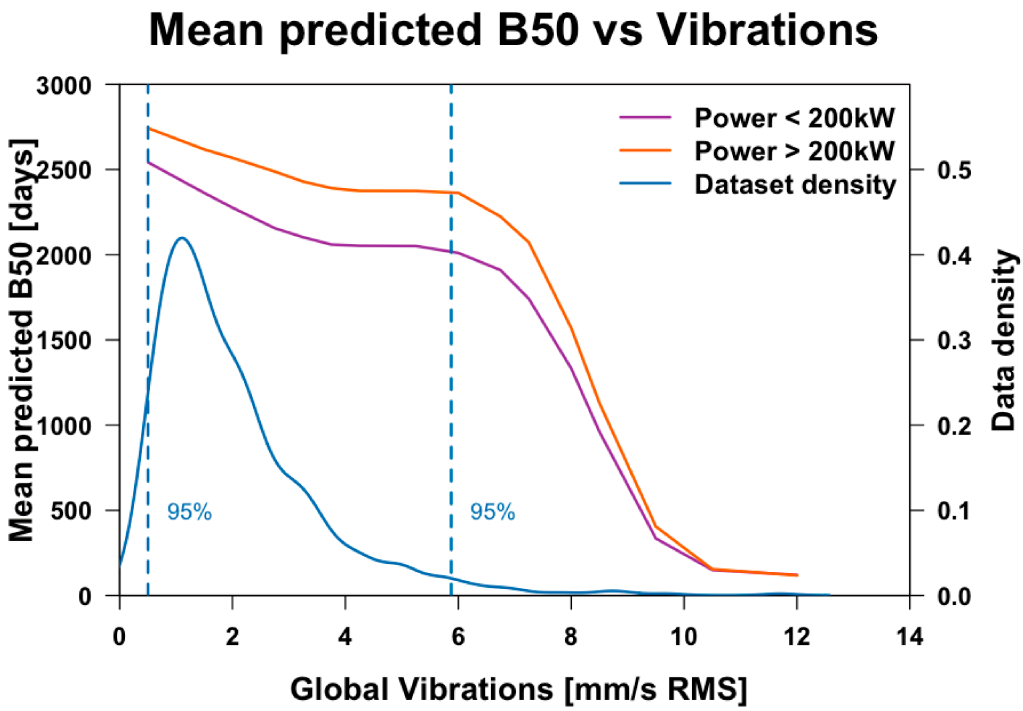

Figure 9. Mean predicted B50 of the prediction grid for installed power ≤200 kW and >200 kW.

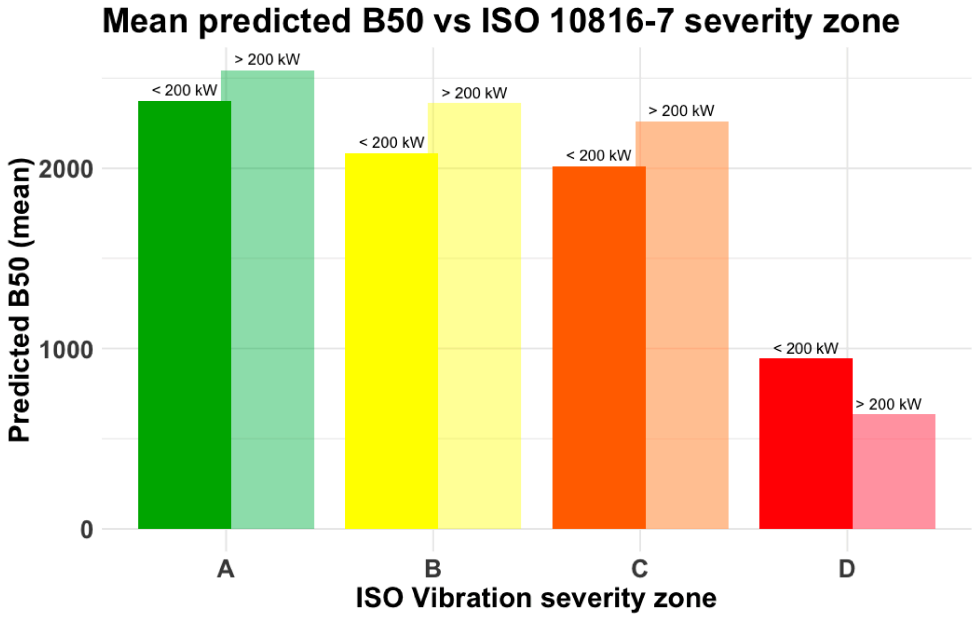

Figure 10. Mean predicted B50 of the prediction grid corresponding to each ISO 10816-7 severity zone grouped by installed power.

6.4. B50 Results Grouped by Pump Type Based on API 610

This section reports the B50 results grouped by pump type according to API 610 [37]. The original dataset contained nine pump types classified under API 610, as detailed in Table 10. During model fitting, pump type was combined with lubrication method (forced lubrication, oil ring, or oil mist [65]) to reduce the number of d.f., resulting in the eight

categories shown in Table 10. For the subsequent comparison of predicted reliability, the prediction-grid observations were grouped into three broader API 610 configurations: overhung, between-bearings, and vertical pumps. This broader grouping was used to provide sufficient observations within each category and facilitate comparison of the predicted B50 patterns.

Table 10. Pump types grouped by lubrication methods.

Type	BB1	BB2	BB3	BB5	OH1	OH2	OH3	OH5	OH6
Grouped 1	BB12		BB35		OH12		OH35		OH6
Lube method [1]	OM	OR	OM	OR	OM	OR	OM		FL
Grouped 2	BB12 OM	BB12 OR	BB35 OM	BB35 OR	OH12OM	OH12 OR	OH35 OM		OH6 FL

[1] OM = oil mist, OR = oil ring, FL = forced lubrication.

The results depicted in Figure 11 show similar patterns across the three pump configurations, with all groups exhibiting the general decrease with increasing vibration observed in Figure 6. The exception is the vertical OH6 group, which shows a more pronounced B50 reduction between 2 and 4 mm/s RMS than the other types at comparable vibration levels.

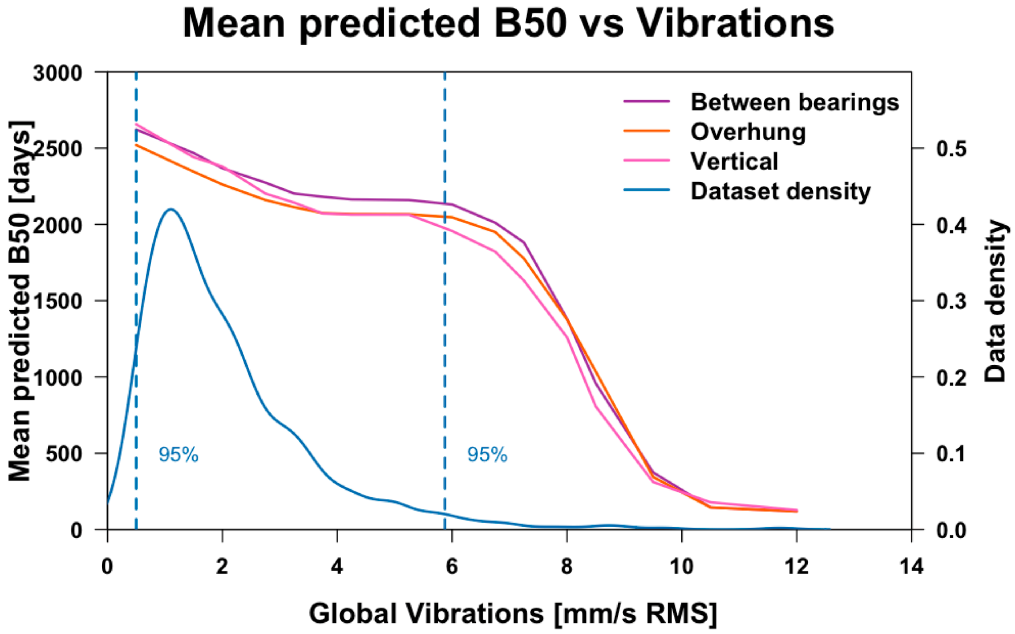

Figure 11. Mean predicted B50 of the prediction grid grouped by pump type.

7. Discussion

The present study investigated the relationship between overall vibration values, measured in mm/s RMS according to ISO 20816-3 guidelines [35], and the median predicted life B50 of a sample of centrifugal pumps. Vibration levels were also categorized into the four severity zones (A, B, C, D) defined by ISO 10816-7 [39]. To examine whether the predicted reliability patterns differed across relevant equipment characteristics, the results were further analyzed by rotating-speed range, installed power, and API 610 pump type.

A Cox PHM was fitted using MTBF as the reliability outcome on a real 8-year dataset of 666 pumps. The fitted model was subsequently used to obtain B50 estimates from the predicted survival curves across the vibration grid. It must be noted that higher B50 predictions could potentially be obtained if the model were trained on a larger time-based dataset. The risk of overfitting associated with the large number of covariates was addressed through RCS transformations and sparse robust PCA. The effectiveness of this approach has been demonstrated elsewhere [29].

The model used for this research showed an optimism-corrected concordance of 0.870, a mean absolute calibration error of 0.0077, and a calibration slope of 0.923. Although the global proportional-hazards test indicated departures from strict proportionality, internal validation showed adequate predictive discrimination and calibration at the prespecified 1500-day horizon. Accordingly, the model is used here as an internally validated prediction model rather than as a basis for time-invariant causal interpretation of individual coefficients. No statistically significant evidence of non-proportionality was detected for the principal components containing the vibration terms, although these components also contain contributions from other predictors. Within this framework, the fitted model was evaluated for every pump across the full range of vibration levels, with all other variables held at their observed values, to obtain the corresponding B50 predictions.

Although vibration is varied independently in the prediction grid, overall RMS vibration should be interpreted as an integrative condition indicator rather than as an isolated physical cause. In actual operation, several mechanical and operational conditions can manifest themselves through changes in overall vibration. The analysis therefore isolates the modelled vibration descriptor, while the measured vibration level may reflect the combined influence of multiple underlying degradation and operating mechanisms. This is consistent with the use of overall RMS vibration as the condition descriptor in the ISO 10816-7 severity classification.

Ninety-five percent of the 666 observations (obtained after data filtering) in the real dataset registered vibration values ranging from 0 to 6 mm/s RMS. Therefore, the results obtained within this range are supported by substantially greater data density and show comparatively low prediction uncertainty. Within this range, the relative deviations of the 95% CIs for the mean B50 predictions were below $\pm 3.68\%$. Similarly, the relative deviations of the individual 95% CI ranged from -9.73% to 23.28%. Comparable patterns were observed in the grouped B50 analysis by ISO vibration severity zones (A, B, C, D), where relative deviations of the group-mean predicted B50 values fell below $\pm 4\%$. Regarding the 95% CI, the dispersion increased with vibration severity from -8.55% and 17.29% in zone A to -17.96% and 96.08% in zone D, highlighting the growing uncertainty associated with higher vibration levels. Taken together, these results indicate comparatively low prediction uncertainty within the vibration range containing most of the observed data, while uncertainty increases substantially at higher severities. The larger intervals at high vibration levels are consistent with the comparatively low density of observations above 6 mm/s RMS, reinforcing the need to interpret predictions in this region cautiously.

The sensitivity analysis further supports the stability of the predicted vibration–reliability pattern. Allowing the effects of the seven components showing evidence of non-proportionality to vary with time resulted in only small changes in mean predicted B50 for Zones A, B, and C (-2.45%, -1.58%, and -1.78%, respectively). A larger difference was observed for Zone D (+9.09%), although this zone remained associated with substantially lower predicted B50 than the other severity zones. Therefore, the same overall reliability ordering (A > B > C > D) was retained under the time-dependent specification. The greater sensitivity of the Zone D estimate is consistent with the lower data density and wider prediction uncertainty already observed at the highest vibration levels.

The results obtained across different speed ranges showed only slightly higher predicted B50 values for equipment operating at higher rotating speeds at comparable vibration levels. Given the limited magnitude of these differences, this pattern should not be interpreted as evidence that increasing rotating speed improves pump reliability. A possible explanation may lie in the different process services associated with the speed groups. High-speed centrifugal pumps are commonly used in light hydrocarbon services. For a given differential pressure, lower-density fluids require a greater developed head [41],

which may favor higher rotational speeds according to the pump affinity laws [66], depending on the pump geometry and system requirements. Conversely, lower-speed pumps are more frequently associated with heavier and more viscous products. These service differences may also affect other reliability-relevant components, particularly mechanical seals. Cleaner services may provide more favorable sealing conditions, whereas abrasive or contaminated fluids may require additional flushing and conditioning arrangements to maintain an appropriate environment at the seal faces [67]. These service-related factors may also contribute to the small differences observed among the speed groups rather than indicating a direct beneficial effect of rotating speed on reliability.

A more pronounced separation was observed between the installed power groups, with pumps above 200 kW showing higher predicted B50 at comparable vibration levels. Rather than a direct effect of installed power, this difference may reflect the mechanical design and auxiliary systems typically associated with larger and more demanding pump services. Higher-power and higher-speed pumps are more likely to require hydrodynamic bearings since API 610 considers rotational speed, bearing life, and pump energy density when defining bearing selection requirements [38]. Larger and more demanding pump services may also incorporate pressurized lubrication systems with main and standby oil pumps, providing redundancy in the lubricant supply to the bearings, as considered in API 614 [68]. These characteristics may therefore contribute to the observed reliability pattern between the two groups. The difference between power groups is also consistent with ISO 10816-7, which establishes higher overall vibration levels for pumps above 200 kW [39]. However, the individual contribution of these characteristics cannot be isolated from the present analysis.

The results obtained for different pump types are also consistent with this interpretation. Between-bearings configurations provide greater rotor support than overhung arrangements, where the impeller is located beyond the bearing support. Rotor stiffness is particularly relevant in overhung pumps, and API 610 explicitly considers shaft flexibility as a design parameter for these configurations [38]. Greater rotor stiffness can reduce shaft deflection and sensitivity to mechanical conditions such as unbalance or misalignment, providing a plausible physical basis for the different vibration and reliability patterns observed between pump configurations. Differences in bearing arrangement, lubrication systems, installed power, and typical service conditions may also contribute. Overall, these engineering considerations provide physically plausible explanations for the grouped results but should not be interpreted as causal effects of rotating speed, installed power, or pump type since these characteristics and their associated operating and design conditions were not independently varied in the present analysis.

At the level of ISO severity zones, the B50–vibration relationship suggests a progressively steeper slope as severity increases. Specifically, a moderate reduction is observed from Zone A to Zone B (≈12.5% decrease in B50), followed by a marginal change from Zone B to Zone C, whereas the transition towards Zone D is characterized by a pronounced drop. The limited change between Zones B and C occurs within a range containing a comparatively large proportion of the field observations, indicating that sparse data alone are unlikely to explain the flatter model response. However, the contribution of model smoothing or the spline–PCA representation cannot be completely excluded. At moderate vibration levels, this flatter response may reflect operating conditions in which other reliability-related factors represented in the model, such as sealing, lubrication, or process conditions, may also contribute less to the overall failure risk, resulting in a comparatively lower combined risk of failure.

8. Limitations and Threats to Validity

This study has limitations that should be considered when interpreting the results, presented below in order of importance.

- Data density. Above 6 mm/s RMS, the number of available observations decreases substantially, increasing prediction uncertainty and requiring greater caution when interpreting results in this range, particularly those associated with the highest vibration severity levels. Increasing the number of observations at higher vibration levels would improve empirical support and reduce prediction uncertainty in this region.
- Scope. This study considers Category I pumps as defined in ISO 10816-7, i.e., pumps requiring a high level of reliability, availability or safety (e.g., hazardous/toxic/flammable services; critical applications in oil and gas, special chemical, nuclear or power). The investigated population therefore represents demanding and highly reliable industrial pump applications and should not be considered representative of all populations of centrifugal pumps. All the analyzed assets were designed according to API 610, providing a comparatively standardized design context. Extending the analysis to other pump categories and industrial services would allow the applicability of the observed vibration–reliability patterns to broader pump populations to be assessed.
- Marginal effect analysis of overall vibration. The prediction grid varies vibration while retaining all other modelled covariates at their observed values. In actual operation, vibration rarely changes independently of other mechanical, operational, and maintenance-related conditions, and unobserved or unmodelled factors may also influence reliability. Consequently, the analysis isolates the modelled relationship between vibration and reliability rather than reproducing how these conditions evolve simultaneously in an individual pump. Explicit modelling of the interactions between vibration and the covariates would allow these dependencies to be investigated in future applications.
- External validity. The dataset was obtained from a single refinery with site-specific operating practices and maintenance strategies; therefore, the generalizability of the absolute reliability estimates to other industrial populations may be limited. External validation using data from other refineries and pump populations would allow the generalizability of the observed reliability patterns to be assessed.
- Vibration analysis domain. Vibration was represented by overall RMS values; no spectral features (e.g., band-limited RMS, kurtosis, or fault-specific indicators) were incorporated. Overall RMS is appropriate for the present objective because it is the severity descriptor used by ISO 10816-7, but it does not identify the physical mechanism responsible for the measured vibration. Future studies incorporating spectral or fault-specific vibration features could complement the severity-based analysis and investigate the reliability implications of individual vibration mechanisms.
- Proportional hazards assumption. The global Schoenfeld residual test indicates departures from strict proportionality in a subset of principal components of the final Cox PHM. However, no significant evidence of non-proportionality was observed for the components containing the vibration terms. To assess the potential impact of these departures, an additional sensitivity analysis was performed using a time-dependent coefficient Cox model, in which the effects of the seven components showing evidence of non-proportionality were allowed to vary with log time. Although the results indicate that the principal vibration–reliability findings are relatively insensitive to the observed departures from proportionality within the present dataset, further evaluation of time-dependent specifications using independent data would be valuable to determine whether this stability extends beyond the investigated pump population.

- Failure definition and censoring. The reliability outcome used to fit the Cox PHM was derived from repairable-system events observed over an 8-year window, with pumps without recorded failures treated as right-censored. Consequently, the finite observation window and right-censoring may affect the resulting survival predictions and absolute B50 estimates; the analysis therefore places greater emphasis on relative reliability patterns than on their absolute magnitude. Longer observation periods and a greater number of recorded failure events would provide more complete information on asset life and reduce the influence of censored observations.

9. Conclusions and Future Work

This paper provides a data-driven, population-level quantification of the reliability penalty associated with the vibration severity zones defined in ISO 10816-7 for centrifugal pumps. The study also contributes to the relatively limited application of Cox PHMs to centrifugal pump reliability, an area in which conventional vibration-based diagnostic and component-level prognostic approaches are more established. The proposed approach combines a multivariable Cox PHM, capable of accounting for censored observations and heterogeneous operating, design, and maintenance conditions, with a prediction grid that systematically evaluates reliability across specified overall RMS vibration levels while retaining the observed pump-specific covariate profiles. This provides a continuous quantitative mapping between vibration severity and predicted lifetime.

The main practical contribution is the translation of this relationship into the severity zones defined by ISO 10816-7. While the standard provides qualitative vibration-severity categories for machinery condition assessment, the proposed approach associates these categories with quantitative reliability estimates expressed in units of time using a model fitted with real industrial field data. The study therefore complements the existing severity classification with quantitative reliability information relevant to asset-management decisions.

The main results show a clear association between increasing overall vibration severity and decreasing predicted pump reliability, although the relationship is nonlinear across the investigated vibration range. The decline is more pronounced within the range of 0.5 to 4 mm/s RMS, whereas between 4 and 6 mm/s RMS the rate of decrease becomes considerably more gradual. When the results are interpreted in terms of the severity zones defined by ISO 10816-7 [39], the reduction in mean predicted B50 from Zone A to Zone B is 12.33%, that from Zone A to Zone C is 14.53%, and that from Zone A to Zone D is 62.50%. The sensitivity analysis using time-dependent coefficients retained the same overall reliability pattern across the ISO severity zones, supporting the stability of these findings to the observed departures from proportionality. However, uncertainty increases towards the upper vibration range, particularly in and around Zone D, where observations in the original dataset are comparatively sparse.

The analysis also identified differences in predicted reliability across rotating speed, installed power, and pump type groups. These results indicate that vibration severity should be interpreted together with relevant equipment and service characteristics rather than as an isolated reliability indicator. Overall, the results support the use of vibration severity not only as a condition indicator but also as a quantitative input for fleet-level reliability assessment.

From a practical perspective, these findings may support more evidence-based maintenance decisions by complementing generic manufacturers' recommendations with reliability information derived from observed operating data. Quantifying the relationship between vibration levels and reliability may support the tailoring of inspection frequencies and condition-monitoring thresholds to specific asset types, moving beyond one-size-fits-all criteria. Moreover, the identification of differences in predicted reliability patterns among

pump types may inform maintenance strategies, design specifications, procurement criteria, and pump-selection decisions. Vibration management, traditionally focused on compliance with severity thresholds, can also consider the reliability implications associated with different vibration levels. During the design and selection phases, this information may also support the definition of vibration targets aimed at maintaining equipment within ranges associated with more favorable predicted reliability under typical operating conditions. In hydrocarbon services, these implications extend beyond maintenance planning since changes in pump reliability may also affect process availability and risk exposure.

Future studies could focus on addressing the current limitations and expanding the model's scope; the following measures are recommended to achieve this:

- Fit and validate new models using the proposed methodology, drawing on datasets that include a greater number of observations at vibration levels above 6 mm/s RMS. Greater data density in this range would reduce prediction uncertainty and provide a more robust assessment of the reliability patterns associated with ISO severity zones C and D.
- Develop a robust methodology for modelling and interpreting interactions between vibration and other relevant covariates within the spline–sparse-PCA framework, accounting for interactions among spline basis functions and their subsequent PCA representation. Incorporating these interactions would help determine how the relationship between vibration and reliability changes under different conditions.
- Further develop and validate Cox PHMs incorporating time-varying effects, including their predictive discrimination and calibration, using independent datasets. The sensitivity analysis performed in the present study indicates that the vibration–reliability pattern is relatively stable when the observed departures from proportionality are explicitly modelled; however, further work could assess whether time-varying specifications provide consistent predictive advantages and generalize to other pump populations.
- Implement the model using periodically acquired vibration measurements together with continuously monitored process variables from plant information systems to enable dynamic updating of reliability predictions. From a practical deployment perspective, process variables such as pressure, temperature, and flow can be automatically retrieved from SCADA or process historians, while vibration measurements can be incorporated either periodically from condition-monitoring routes or continuously where online monitoring is available. Model predictions could be updated as new measurements become available, whereas model retraining could be performed less frequently, once sufficient new operating and failure data have accumulated. Since prediction with a fitted Cox PHM has a relatively low computational cost, this strategy is compatible with scheduled or frequent reliability assessment at fleet level, with the main implementation effort associated with data integration and synchronization.
- Evaluate which variables contribute to differing pump behaviors, focusing on different covariate groups such as operating conditions, hydraulic design, mechanical design, age, sealing, and maintenance history. Identifying these differences could provide further insights into the equipment and service characteristics associated with the different responses observed among overhung, between-bearings, and vertical pumps.
- Extend the research by incorporating frequency-domain and fault-specific vibration features. This would complement the overall RMS severity assessment by allowing separate investigation of different vibration signatures and their associated degradation mechanisms.

Author Contributions: Conceptualization, M.V.F., D.G. and U.K.; methodology, M.V.F. and D.G.; software, M.V.F.; validation, M.V.F. and D.G.; formal analysis, M.V.F.; investigation, M.V.F.; resources, M.V.F., D.G. and U.K.; data curation, M.V.F.; writing—original draft preparation, M.V.F.; writing—review and editing, K.G.; visualization, M.V.F.; supervision, D.G.; project administration, D.G. and U.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data supporting the findings of this study are not publicly available due to confidentiality and proprietary restrictions associated with the industrial dataset.

Conflicts of Interest: Author Marc Vila Forteza was employed by the company Petronor. Author Kai Goebel was employed by the company Fragum Global. The other authors declare no conflicts of interest.

References

1. Rathbone, T.C. Vibration tolerance. *Power Plant Eng.* **1939**, *43*, 721–724.
2. Available online: https://www.dmc.pt/en/historia-da-medicao-de-vibracoes-em-manutencao-preditiva/ (accessed on 15 February 2026).
3. Tiboni, M.; Remino, C.; Bussola, R.; Amici, C. A Review on Vibration-Based Condition Monitoring of Rotating Machinery. *Appl. Sci.* **2022**, *12*, 972. [CrossRef]
4. Khalil, A.F.; Rostam, S. Machine Learning-based Predictive Maintenance for Fault Detection in Rotating Machinery: A Case Study. *Eng. Technol. Appl. Sci. Res.* **2024**, *14*, 13181–13189. [CrossRef]
5. Jardine, A.; Lin, D.; Banjevic, D.A. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mech. Syst. Signal Process.* **2006**, *20*, 1483–1510. [CrossRef]
6. Yu, R.; Li, X.; Tao, M.; Ke, Z. Fault Diagnosis of Feedwater Pump in Nuclear Power Plants Using Parameter-Optimized Support Vector Machine. V001T03A013. In Proceedings of the 2016 24th International Conference on Nuclear Engineering, Charlotte, NC, USA, 26–30 June 2016. [CrossRef]
7. Adraoui, I.E.; Gziri, H.; Mousrij, A. Prognosis of a degradable hydraulic system: Application on a centrifugal pump. *Int. J. Progn. Health Manag.* **2020**, *11*, 1–11.
8. Zurita Millan, D.; Delgado-Prieto, M.; Saucedo-Dorantes, J.; Cariño-Corrales, J.; Osornio-Rios, R.; Ortega, J.; Romero-Troncoso, R. Vibration Signal Forecasting on Rotating Machinery by means of Signal Decomposition and Neurofuzzy Modelling. *Shock Vib.* **2016**, *2016*, 2683269. [CrossRef]
9. Wang, J.; Zhang, L.; Zheng, Y.; Wang, K. Adaptive prognosis of centrifugal pump under variable operating conditions. *Mech. Syst. Signal Process.* **2019**, *131*, 576–591. [CrossRef]
10. Souza, R.; Sperandio, N.; Erick, G.; Miranda, U.; Silva, W.; Lepikson, H. Deep learning for diagnosis and classification of faults in industrial rotating machinery. *Comput. Ind. Eng.* **2020**, *153*, 107060. [CrossRef]
11. Kumar, A.; Gandhi, C.; Zhou, Y.; Kumar, R.; Xiang, J. Improved deep convolution neural network (CNN) for the identification of defects in the centrifugal pump using acoustic images. *Appl. Acoust.* **2020**, *167*, 107399. [CrossRef]
12. Zhao, L.; Wang, X. A Deep Feature Optimization Fusion Method for Extracting Bearing Degradation Features. *IEEE Access* **2018**, *6*, 19640–19653. [CrossRef]
13. Zhang, Y.; Zhou, T.; Huang, X.; Cao, L.; Zhou, Q. Fault diagnosis of rotating machinery based on recurrent neural networks. *Measurement* **2020**, *171*, 108774. [CrossRef]
14. Vila Forteza, M.; Jimenez Cortadi, A.; Diez Olivan, A.; Seneviratne, D.; Galar Pascual, D. Advanced Prognostics for a Centrifugal Fan and Multistage Centrifugal Pump Using a Hybrid Model. In Proceedings of the 5th International Conference on Maintenance, Condition Monitoring and Diagnostics 2021, Oulu, Finland, 16–17 February 2021. [CrossRef]
15. Cubillo, A.; Perinpanayagam, S.; Esperon Miguez, M. A review of physics-based models in prognostics: Application to gears and bearings of rotating machinery. *Adv. Mech. Eng.* **2016**, *8*, 1687814016664660. [CrossRef]
16. Fouladirad, M.; Belhaj Salem, M.; Deloux, E. Variance Gamma process as degradation model for prognosis and imperfect maintenance of centrifugal pumps. *Reliab. Eng. Syst. Saf.* **2022**, *223*, 108417. [CrossRef]
17. Forrester, T.; Harris, M.; Senecal, J.; Sheppard, J. Continuous Time Bayesian Networks in Prognosis and Health Management of Centrifugal Pumps. In Proceedings of the Annual Conference of the PHM Society, Scottsdale, AZ, USA, 23–26 September 2019; p. 11. [CrossRef]
18. Zhang, Q.; Hua, C.; Xu, G. A mixture Weibull proportional hazard model for mechanical system failure prediction utilising lifetime and monitoring data. *Mech. Syst. Signal Process.* **2014**, *43*, 103–112. [CrossRef]

19. Zhang, S.; Hodkiewicz, M.; Ma, L.; Mathew, J. Machinery Condition Prognosis Using Multivariate Analysis. In *Engineering Asset Management*; Mathew, J., Kennedy, J., Ma, L., Tan, A., Anderson, D., Eds.; Springer: London, UK, 2006. [CrossRef]
20. Mosallam, A.; Medjaher, K.; Zerhouni, N. Data-driven prognostic method based on Bayesian approaches for direct remaining useful life prediction. *J. Intell. Manuf.* **2014**, *27*, 1037–1048. [CrossRef]
21. Hu, J.; Tse, P. A Relevance Vector Machine-Based Approach with Application to Oil Sand Pump Prognostics. *Sensors* **2013**, *13*, 12663–12686. [CrossRef] [PubMed]
22. Cao, S.; Hu, Z.; Luo, X.; Wang, H. Research on fault diagnosis technology of centrifugal pump blade crack based on PCA and GMM. *Measurement* **2020**, *173*, 108558. [CrossRef]
23. Tamssaouet, F.; Nguyen, K.T.P.; Medjaher, K.; Orchard, M.E. System level failure prognostics: Literature review and main challenges. *Proc. Inst. Mech. Eng. Part O J. Risk Reliab.* **2023**, *237*, 524–545. [CrossRef]
24. Kim, S.; Choi, J.H.; Kim, N.H. Challenges and opportunities of system level prognostics. *Sensors* **2021**, *21*, 7655. [CrossRef] [PubMed]
25. Bevilacqua, M.; Braglia, M.; Montanari, R. The classification and regression tree approach to pump failure rate analysis. *Reliab. Eng. Syst. Saf.* **2003**, *79*, 59–67. [CrossRef]
26. Braglia, M.; Castellano, D.; Frosolini, M.; Gabbrielli, R.; Marrazzini, L.; Padellini, L. An ensemble-learning model for failure rate prediction. *Procedia Manuf.* **2020**, *42*, 41–48. [CrossRef]
27. Orrù, P.F.; Zoccheddu, A.; Sassu, L.; Mattia, C.; Cozza, R.; Arena, S. Machine learning approach using MLP and SVM algorithms for the fault prediction of a centrifugal pump in the oil and gas industry. *Sustainability* **2020**, *12*, 4776. [CrossRef]
28. Sudadiyo, S. Nonhomogeneous Poisson Process Model for Estimating Mean Time Between Failures of the JE01-AP03 Primary Pump Implemented on the RSG-GAS Reactor. *Nucl. Technol.* **2025**, *211*, 645–660. [CrossRef]
29. Vila Forteza, M.; Galar, D.; Kumar, U.; Goebel, K. Data Reduction in Proportional Hazards Models Applied to Reliability Prediction of Centrifugal Pumps. *Machines* **2025**, *13*, 215. [CrossRef]
30. Braglia, M.; Carmignani, G.; Frosolini, M.; Zammori, F. Data classification and MTBF prediction with a multivariate analysis approach. *Reliab. Eng. Syst. Saf.* **2012**, *97*, 27–35. [CrossRef]
31. Gorjian, N.; Sun, Y.; Ma, L.; Yarlagadda, P.; Mittinty, M. Remaining useful life prediction of rotating equipment using covariate-based hazard models–Industry applications. *Aust. J. Mech. Eng.* **2017**, *15*, 36–45. [CrossRef]
32. Vila Forteza, M.; Galar, D.; Goebel, K.; Kumar, U. Towards predictive reliability: Evaluating influential variables in centrifugal pumps' MTBF for oil and gas applications. *Eng. Res. Express* **2026**, *8*, 075224. [CrossRef]
33. Bloch, H.P.; Budris, A.R. *Pump User's Handbook, Life Extension*, 4th ed.; The Fairmont Press Inc.: Atlanta, GA, USA, 2014; p. 103.
34. Delprete, C.; Gastaldi, C. On the effectiveness of standards application to threshold setting in vibration condition monitoring in industrial machinery. *Proc. Inst. Mech. Eng. Part O J. Risk Reliab.* **2025**, *239*, 207–221. [CrossRef]
35. *ISO 20816-1*; Mechanical Vibration—Measurement and Evaluation of Machine Vibration—Part 1: General Guidelines. International Organization for Standardization: Geneva, Switzerland, 2016.
36. *ISO 2954*; Mechanical Vibration of Rotating and Reciprocating Machinery—Requirements for Instruments for Measuring Vibration Severity. International Organization for Standardization: Geneva, Switzerland, 2012.
37. *ANSI/HI 9.6.4*; Rotodynamic Pumps for Vibration Measurements and Allowable Values. Hydraulic Institute: Parsippany, NJ, USA, 2016.
38. *API Standard 610*; Centrifugal Pumps for Petroleum, Petrochemical and Natural Gas Industries. American Petroleum Institute: Washington, DC, USA, 2021.
39. *ISO 10816-7*; Mechanical Vibration—Evaluation of Machine Vibration by Measurements on Non-Rotating Parts—Part 7: Rotodynamic Pumps for Industrial Applications with Nominal Power Above 1 kW. International Organization for Standardization: Geneva, Switzerland, 2009.
40. Beckerle, P.; Schaede, H.; Rinderknecht, S. Fault Diagnosis and State Detection in Centrifugal Pumps—A Review of Applications. In *Proceedings of the 9th IFToMM International Conference on Rotor Dynamics, 26–27 May 2015*; Springer: Cham, Switzerland, 2015; Volume 21. [CrossRef]
41. Bloch, H.P. *Pump Wisdom: Problem Solving for Operators and Specialists*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2011.
42. He, D.; Longoria, R.G.; Ji, J. Centrifugal Pump Mechanical Seal and Bearing Failure Modes. In Proceedings of the 12th International Probabilistic Safety Assessment and Management Conference (PSAM12), Honolulu, HI, USA, 22–27 June 2014; pp. 1–12.
43. Marscher, W.D. Avoiding Failures in Centrifugal Pumps. In *Proceedings of the 31st Turbomachinery Symposium*; Texas A&M University: Houston, TX, USA, 2002; pp. 101–110.
44. MA, Q.; Wang, K.; Liu, H.; Cheng, Z. Influence of Shaft Combined Misalignment on Vibration and Noise of Centrifugal Pump. *J. Low Freq. Noise Vib. Act. Control* **2021**, *41*, 1286–1306. [CrossRef]
45. Nelik, L.; Jackson, C. Effect of Mechanical Unbalance on Vibrations Forces and Reliability of a Single-Stage Centrifugal Pumps. 911 Metallurgist 2008. Available online: https://www.911metallurgist.com/wp-content/uploads/2016/01/Effect-of-Mechanical-Unbalance-on-Vibrations-Forces-and-Reliability-of-a-Single-Stage-Centrifugal-Pumps.pdf (accessed on 20 September 2026).

46. Nelson, E.P. Vane Pass Vibration Source Assessment and Correction: A Practical Guide for Centrifugal Pumps. In *Proceedings of the 33rd Turbomachinery Symposium*; Texas A&M University: Houston, TX, USA, 2004; pp. 123–130.
47. Bloch, H.P.; Geitner, F.K. *Improving Machinery Reliability*, 3rd ed.; Elsevier: Amsterdam, The Netherlands, 1998; Volume 1.
48. Lobanoff, V.S.; Ross, R.R. *Centrifugal Pumps: Design and Application*, 1st ed.; Gulf Publishing Co.: Houston, TX, USA, 1985.
49. Kong, W.; Li, Y.; Lei, M.; Wang, H.; Li, J. Experimental Study on the Frequency-Domain Characteristics of Flow-Induced Vibration in a High-Specific-Speed Centrifugal Pump. *Flow Meas. Instrum.* **2026**, *112*, 103492. [CrossRef]
50. Han, Y.; Zou, J.; Presas, A.; Luo, Y.; Yuan, J. Off-Design Operation and Cavitation Detection in Centrifugal Pumps Using Vibration and Motor Stator Current Analyses. *Sensors* **2024**, *24*, 3410. [CrossRef] [PubMed]
51. Yang, C.; Fan, C.; Liu, J.; Li, X.; Zhou, R. Vibration Characteristics of the Flexible Rotor System of a Centrifugal Pump with the Shaft Misalignment and Impeller Unbalance. *Proc. Inst. Mech. Eng. Part K J. Multi-Body Dyn.* **2026**, 14644193261432510. [CrossRef]
52. Karaśkiewicz, K. Vibrations in centrifugal pumps induced by pressure pulsations with blade frequency. *Instal* **2024**, *11*, 33–37. [CrossRef]
53. *EN 15341:2019*; Maintenance—Maintenance Key Performance Indicators. CEN: Brussels, Belgium, 2019.
54. Cai, J.; Zeng, D. Cox Proportional Hazard Model. In *Wiley StatsRef: Statistics Reference Online*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2014; pp. 1–10. [CrossRef]
55. Lee, E.T. *Statistical Methods for Survival Data Analysis*, 3rd ed.; John Wiley & Sons: New York, NY, USA, 2003.
56. Kalbfleisch, J.D.; Prentice, R.L. *The Statistical Analysis of Failure Time Data*, 1st ed.; Wiley: New York, NY, USA, 1980.
57. Delong, D.M.; Guirguis, G.H.; So, Y.C. Efficient Computation of Subset Selection Probabilities with Application to Cox Regression. *Biometrica* **1994**, *81*, 607–611. [CrossRef]
58. Breslow, N.E. Covariance analysis of censored survival data. *Biometrics* **1974**, *30*, 89–99. [CrossRef] [PubMed]
59. Efron, B. The Efficiency of Cox's Likelihood Function for Censored Data. *J. Am. Stat. Assoc.* **1977**, *72*, 557–565. [CrossRef]
60. Harrell, F.E., Jr. *Regression Modelling Strategies, Springer Series in Statistics*; Springer International Publishing: Cham, Switzerland, 2016.
61. Kleinbaum, D.G.; Klein, M. *Survival Analysis: A Self-Learning Text*, 3rd ed.; Springer: Berlin/Heidelberg, Germany, 2012. [CrossRef]
62. Zou, H.; Hastie, T.; Tibshirani, R. Sparse principal component analysis. *J. Comput. Graph. Stat.* **2006**, *15*, 265–286. [CrossRef]
63. Breiman, L.; Friedman, J.H. Estimating Optimal Transformations for Multiple Regression and Correlation. *J. Am. Stat. Assoc.* **1985**, *80*, 580–598. [CrossRef]
64. Tibshirani, R. Estimating transformations for regression via additivity and variance stabilization. *J. Am. Stat. Assoc.* **1988**, *83*, 394–405. [CrossRef]
65. Bloch, H.P.; Shamim, A. *Oil Mist Lubrication: Practical Applications*, 1st ed.; The Fairmont Press: Lilburn, GA, USA, 1998.
66. Karassik, I.J.; Messina, J.P.; Cooper, P.; Heald, C.C. *Pump Handbook*, 4th ed.; McGraw-Hill: New York, NY, USA, 2008.
67. *API Standard 682*; Pumps—Shaft Sealing Systems for Centrifugal and Rotary Pumps. American Petroleum Institute: Washington, DC, USA, 2014.
68. *API Standard 614*; Lubrication, Shaft-Sealing, and Oil-Control Systems and Auxiliaries. American Petroleum Institute: Washington, DC, USA, 2022.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.