Data Selection Criteria for the Application of Predictive Maintenance to Centrifugal Pumps

Núbia N. A. Silveira[1], Richard Loendersloot[1], Annemieke A. Meghoe[1], and Tiedo Tinga[1,2]

[1] *University of Twente, Drienerlolaan 5, Enschede, 7522 NB, Netherlands*
n.n.alvesdasilveira@utwente.nl
r.loendersloot@utwente.nl
a.a.meghoe@utwente.nl

[2] *Netherlands Defence Academy, PO Box 10000, Den Helder, 1780 CA, Netherlands*
t.tinga@utwente.nl

Abstract

The maintenance of vehicles and components is present in most people's daily lives, ranging from changing a private vehicle's oil to the failure prediction of an aircraft component during flight. Usually, the manufacturer's maintenance recommendation is a good solution when the cost is not too high, and the real application is used as indicated by the manufacturer. However, this recommendation can turn unfeasible when there is a significant variation in operational conditions or high maintenance costs. In these cases, the manufacturer's suggestion is typically conservative, leading to unnecessarily high costs. Therefore, the challenge is to find the best approach for optimizing a component's maintenance, given the system in which it is integrated and the associated operational and environmental conditions. Nevertheless, the available information on the loads on the component also plays a role in that choice. This paper proposes to combine case-specific information with generic degradation prediction models to obtain an acceptable but also affordable approach. The objective is to develop data selection criteria to indicate the parameters that have a high impact on the failure prediction, in this case, of a generic impeller pump. Subsequently, the approach delivers to the user an indication of the component remaining useful life using different operational scenarios.

1. Introduction

Predictive maintenance (PdM) is a topic in development, and there are many gaps to be understood and developed. PdM uses Prognostic and Health Management (PHM) technologies to avoid premature failures. Besides that, with PdM, it is possible to enable remote diagnosis, reduce secondary damage to components and maintenance costs, and optimize a future design. Thereby, it is possible to quantify the deteriorated condition, performance, and the remaining useful life (RUL) of the system (Lei & Sandborn, 2016).

Some authors use data-driven methods to generate condition monitoring or prognostics results (Parrondo, Velarde, & Santolaria, 1998; J. Wang, Zhang, Zheng, & Wang, 2019). However, the application turns to be impractical in a real case due to the complexity of the setup, storage of the data, or uncertainties in the results. An example of a physics-based approach is present by Parrondo et al. (1998), who analysed the system pressure and accelerometer data from a centrifugal pump. Their work shows which alternative best represents the cavitation behavior. They conclude that the signal frequency produced by sensors from the pressure in the pump inlet and outlet can provide enough data to evaluate the condition of the impeller. However, this approach cannot provide a deterioration prediction, besides that, it demands high storage space and a laboratory setup.

The study of maintenance strategies is portrayed in the literature to increase efficiency with consolidated approaches to proactive, preventive maintenance, and health monitoring (Ivankevich, Piterska, Shakhov, Shakhov, & Yarovenko, 2019). This strategy avoids a low quality of services performed, which causes low efficiency and high cost of components. However, this strategy lacks a methodology, and the companies do not have a science-based method to predict failure. Health or condition monitoring is an important element in predictive maintenance. In this approach, the deterioration of components is monitored. The behavior and the evolution of the failure of a component is described according to a specific condition. However, extrapolation is not straightforward, as indicated in the literature (Abdel Fatah, Hassan, Lotfy, & Dimitri, 2019). There is a need for new approaches and models that predict component failure where the condition and application can change.

The present work proposes a methodology to evaluate the importance of the parameters used in Reliability or Failure Rate Models, also called Handbook Models in this work. Such models use a historical failure database, information on the operating conditions, and some physical properties to predict the next component failure. The combination of data and models is implemented in the proposed tool which interacts with a user by asking for information or data. The type of data used to run the tool can come from different sources; besides sensor data, Enterprise Resource Planning (ERP) data, environmental conditions, technical manual information, and material properties. Although ERP systems are not designed with a maintenance focus, it is now possible to find service packages that help with maintenance management and produce useful information regarding system maintenance history (Kohli, 2017; Bumblauskas, Gemmill, Igou, & Anzengruber, 2017).

This work aims to develop data selection criteria that will be used as one of the elements in the tool to be developed. The tool can decide which model and parameters will be used to predict the component failure. This decision-making is based on sensitivity analyses and a proposed set of data selection criteria. In section 2 a handbook model is presented that is the base model to demonstrate this approach. Section 3 presents the data selection methodology proposed in this work. The proposed criteria are demonstrated in Section 4. Finally, the paper is closed with conclusions (section 5).

2. Methods

The methodology is developed for a generic centrifugal pump through Handbook Models. Further, it is analyzed how operational factors like water flow and temperature can influence cavitation during pumping. The generic models available for centrifugal pumps focus on a single failure mode, i.e., cavitation or erosion. In practice, it will be necessary to consider all possible failure modes in a centrifugal pump, but to demonstrate the approach, in this paper only the cavitation model will be considered. The case study analyzes a centrifugal pump from a seawater cooling system in different scenarios with varying operational and environmental conditions.

Firstly, a handbook model is selected based on failure rate history and physical parameters. Secondly, the necessary operational parameters and their range of values are determined. In this step, a sensitivity analysis identifies which parameters have the most significant impact on the model, indicating the necessity for more precise data. This step also ensures that the user does not collect unnecessary data for a parameter with a minor impact. Simultaneously, the user is encouraged to collect more accurate data, like time series data or verified data, for the parameters with a significant impact.

2.1. Failure Rate Model

A handbook or Failure Rate Model (FRM) is usually referred to as a collection of previous works. The NSWC-11 (Handbook of Reliability Prediction Procedures for Mechanical Equipment) presents a methodology to evaluate mechanical designs based on information from databases, aging characteristics, regression techniques, and field failure data (NSWC, 2011). The failure rate λ_{FD} (in failures per million hours) for a fluid driver, i.e. the impeller, in a centrifugal pump is given as:

$$\lambda_{FD} = \lambda_{FD,B} \cdot C_{PF} \cdot C_{PS} \cdot C_C \cdot C_{SF} \tag{1}$$

where $\lambda_{FD,B}$ is the base failure rate for the pump fluid driver; C_{PF} is the percent flow multiplying factor, C_{PS} is the operating speed multiplying factor; C_C is the contaminant multiplying factor; and C_{SF} is the service factor. The base failure rate is based on historical data for the considered pump. As it acts as a scaling factor, and its value can vary several orders of magnitude (depending on the application), it has a significant impact on the failure rate λ_{FD}. In NSWC-11 a historical failure rate is proposed that varies between 0.12 and 0.2 failures per million hours. However, this variation is based on the driver mode and impeller model type, while application, type of fluid, and environmental factors are not considered in this Handbook Model.

Note that the correction factor for the flow C_{PF} is determined by the ratio between the actual flow Q and the specified maximum flow Q_r. The value of the correction factor depends, for an ordinary volute casing, on the flow ratios according to:

$$\begin{aligned}
&\text{For } 0.1 \leq \frac{Q}{Q_r} \leq 1.0: \\
&C_{PF} = 9.94 - 0.90\left(\frac{Q}{Q_r}\right) - 10\left(\frac{Q}{Q_r}\right)^2 + 1.77\left(\frac{Q}{Q_r}\right)^3 \\
&\text{For } 1.0 \leq \frac{Q}{Q_r} \leq 1.1: \\
&C_{PF} = 1.0 \\
&\text{For } 1.1 \leq \frac{Q}{Q_r} \leq 1.7: \\
&C_{PF} = -30.6 + 36\left(\frac{Q}{Q_r}\right) - 4.5\left(\frac{Q}{Q_r}\right)^2 - 2.2\left(\frac{Q}{Q_r}\right)^3
\end{aligned} \tag{2}$$

Similarly, C_{PS} is a based on the ratio between operating Speed (V_O) and the maximum allowable design speed (V_D). According to the NSWC-11, this factor affects the cavitation damage since the increase of operating speed also increases the pump energy level. The correction factor is defined as:

$$C_{PS} = 5 \left(\frac{V_O}{V_D} \right)^{1.3} \tag{3}$$

The correction factor for contamination C_C was developed based on research from the NSWC and is given as:

$$C_C = 0.6 + 0.05 \cdot F_{AC} \tag{4}$$

where F_{AC} is the filtration level typically given in micrometers. However, the application under investigation in this work highly differs from the application from the NSWC, implying that the validity of Eq. (4) is questionable.

3. Data Selection Criteria

This paper's main goal is to establish data selection criteria for a predictive maintenance tool applied to a generic component. Figure 1 presents the flowchart for the tool under development. The user should fill the fields in green, and the software then takes the actions in blue, i.e. selects the most suitable RUL model (from a database of models), identifies the dominant parameters, advices on data collection and ultimately performs the RUL calculation. In this paper, the focus will be on the step in the red dashed box: determining the high impact parameters. Initially, the user should select the component, insert the component information, and specify available data. The latter field "specify available data" signalizes to the software which parameters are available with specific information, but in this step, the user does not yet need to provide data for the tool.

Data information combined with the models available in the software, such as the Handbook model, will enable the system to choose the most suitable physics based model for the case. However, it is not mandatory that all the available data will be provided to the chosen model. Such a method may analyze a second model, Physical Model, with considerable efficiency and parameters with high influence available. Therefore, the second model can present more significant results than the model chosen previously.

A sensitivity analysis with the range of the data and the selected model is done first. This step indicates the parameters that have a high influence on the model output. After that, the tool analyzes if there is sufficient data for the parameters identified with high influence. The user needs to specify the parameters to allow the system to identify whether sufficient data is provided in the first step. If not all data is, or can be, specified by the user, reference values will be used. The tool calculates the Remaining Useful Life, but also communicates to the user which data is necessary to improve the RUL estimation, in case incomplete data was provided.

The feature importance technique commonly used in data mining does not apply to this case. For cases with extensive features, some data are excluded due to the irrelevance or redundant contribution (Razmjoo, Xanthopoulos, & Zheng, 2017). Note that in case of large datasets, many (advanced) techniques are available to select the most important features (e.g. Principal Component Analysis). However, in this case it is assumed that such a large dataset is not available (yet), and the (physical) reliability model is used as starting point.

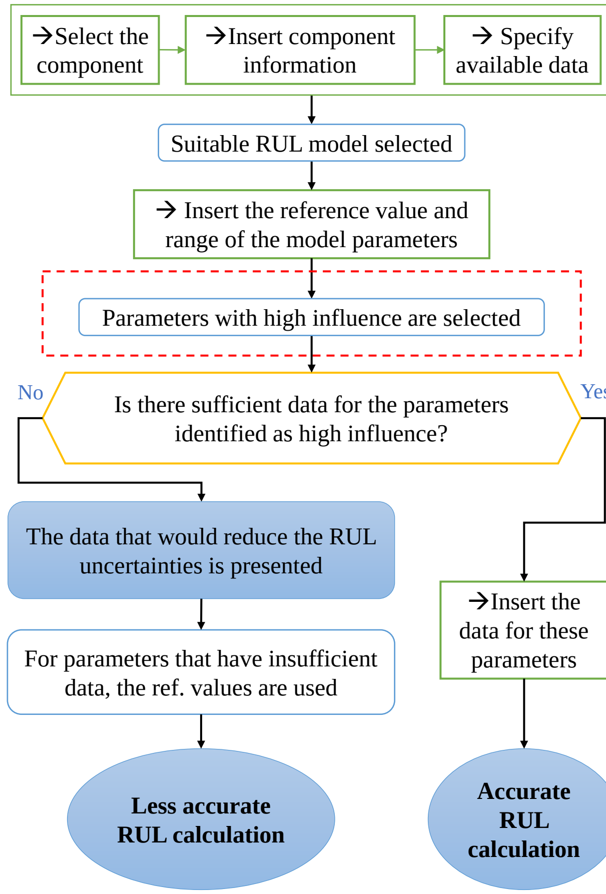

Figure 1. Tool flowchart. The actions made by the user are in green/rectangle, and the decision made by the software in blue/rectangle with rounded corners. The solid colors present what will be displayed for the user.

It is proposed here to use a sensitivity analysis to select the most relevant parameter. There are many types of sensitivity analysis methodologies, and each methodology is specifically suitable for a certain application. There are many factors considered to select the suitable method; however, there is no indication for a unique methodology for each case. Tian (2013) characterizes the Morris method with high computational effort, many inputs, and qualitative parameters analysis. This work aims to create a generic model to apply to a generic hydraulic system. The Morris method meets this specification since it covers the high amount of parameters and high computational effort of an entire system.

The Morris method provides not only the influence of each parameter, but also the correlation between the parameters (Meghoe, Loendersloot, & Tinga, 2020; Forrester, Sobester, & Keane, 2008). Each input (k) varies along a pre-determined number of levels (p) in the input factor space. The domain of experimentation (Ω) is a regular k-dimensional p-level grid. For a given vector $\boldsymbol{X} = [x_1, ..., x_k]$ the elementary effect Ee_i of the i^{th} parameter x_i to the response value y is presented as (Meghoe et al., 2020; C. Wang, Peng, & Xia, 2020):

$$Ee_i(\mathbf{X}) = \frac{y(x_1, ..., x_i + \Delta, x_{i+1}, ..., x_k) - y(x_1, ..., x_k)}{\Delta} \tag{5}$$

where $\boldsymbol{X} \in \Omega$, when $x_i \leq 1 - \Delta$ (Morris, 1991).

According to Campolongo, Cariboni, and Saltelli (2007), an even number of levels p and Δ defined as:

$$\Delta = \frac{p}{2(p-1)} \tag{6}$$

is a convenient choice because it ensures symmetric treatment of inputs, with which it is likely to have a more even distribution of values in $\boldsymbol{X}$. Thereby, the probability of a severely imbalanced sample decreases in every column. Since this work intends to apply this methodology in different models with different amounts of inputs, this is the recommendable approach.

It is essential to define criteria to select or exclude some parameters using multiple sensitivity analyses, since the aim is to build a generic model for an arbitrary centrifugal pump. According to Campolongo et al. (2007), 20 is a sufficient number of random analysis. Convergence analyses of three different models and some literature findings were used to determine the p-level grid (Ruano, Ribes, Ferrer, & Sin, 2011; Campolongo et al., 2007; Meghoe et al., 2020). Values between four and eight discrete levels along each dimension are pointed out to be sufficient according to the literature. The convergence analyses also indicated values of $p = 8$ as sufficient. Therefore, the authors choose to use eight levels.

The results of a Morris analysis are visualized by plotting the standard deviation versus the mean of the samples in a screening plan, as for example shown in Figure 2. The importance of each of the variables is defined by the combination of both the standard deviations and the mean. There are no well-defined rules to select which parameters should be included or excluded. An automation of the selection of parameters is desired, to avoid human intervention at this point in the process. It is proposed here to draw a quarter ellipse, crossing the horizontal and vertical axis at 50% of the mean value and 50% of the standard deviation as the limit. The general validity of this method is tested on the findings of a number of sources from the literature. Touhami, Lardy, Barra, and Bellocchi (2013) analyzed 28 parameters with the purpose of identifying the most influential parameters for the grassland system in specific contexts. The authors divided the plot into four quadrants; however, they do not define by which criteria the quadrants are limited. On the other hand, the representative variables are at the quadrant opposite from the origin, and this matches with the ellipse approach.

Morris (1991)analyzed a problem with 20 parameters. An artificial computational model was used to demonstrate the method. In this analysis, the ten last inputs are clustered with means and standard deviation close to 0. The first ten inputs are considered significant. However, from these first inputs, seven have a higher impact. The degree of importance of three of the first set of ten parameters is not defined; the selection criteria proposed in the present a solution for this question.

Forrester et al. (2008) analyzed ten parameters to identify the inputs with little and high impact on the objective function. They evaluated the impact of the inputs of a function to optimize the weight of the wing of a light aircraft. The authors pointed out that the values far from the origin have the most significant impact. The analyzes match the approach propose in this work.

C. Wang et al. (2020) analyzed 24 parameters to identify the essential parameters for a marine passive residual heat removal system. The parameters indicated as being of greater importance appear out of the ellipse when the technique is applied. However, there is some inconsistency in the ellipse approach versus the assessment of the authors: firstly, and sometimes they take parameters with minor relevance, sometimes a parameter does not appear very important in the same analysis with a different normalization.

Finally, Meghoe et al. (2020) analyzed 8 parameters to identify which parameter affect the wear number of a rail. The data selection criterion proposed here confirms the parameters indicated as those with a strong influence and correlation. However, one of the analyses considered one parameter as relevant that presents a relatively minor impact compared with the highest impact parameter. The ellipse approach still largely matches the findings of Meghoe et al. (2020).

With the ellipse selection method proposed in this work, the software points out which parameters have a high impact. A value of 50% of half of the maximum mean value and 50% of the maximum standard deviation for the axis of the ellipse proved to give satisfactory results when applied to a number of examples from literature. The results of the analysis are collected in Table 1.

4. Demonstration, Results and Discussion

In this section, the described methodology of data selection is applied to an impeller of a pump system. The demonstration

Table 1. Overview of the correspondence of the ellipse based method with findings in the open literature.

Paper reference	Num. of evaluated parameters	Purpose	Outcome
Touhami et al. (2020)	28	Identification of the most influential parameters for the grassland system in specific contexts	Good correlation
Morris (1991)	20	Demosntration of the Morris Method	Good correlation
Forrester et al. (2008)	10	Identification of inputs with little and high impact on the objective function	Good correlation
C. Wang et al. (2020)	24	Identification of important parameters for a marine passive residual heat removal system	Good correlation - small inconsistency
Meghoe et al. (2020)	8	Identification of which parameter affect the wear number	Good correlation - small inconsistency

is based on the data from a NISM 125-250/01 pump installed in a seawater cooling system (Janse, 2009). The particle distribution and sediment concentration were based on previous studies (Sheldon, Prakash, & Sutcliffe Jr, 1972; Lauwaert et al., 2016). The installation details, like pipe length and liquid level height; and time-series data, like temperature and flow, were assumed for a realistic situation.

4.1. Sensitivity Analysis

The sensitivity analysis (SA) is an essential function in this work; it selects the parameters with a large influence on the results and therewith prevents that time and cost are wasted on parameters that have a low impact. The Morris method calculates the means and standard deviation of the elementary effect (Ee) distribution for each parameter. The mean of the Ee distribution represents how influential a parameter is. When a parameter returns a high mean value, that parameter has a significant influence. The standard deviation indicates how a parameter deviates from the distribution. So, if a parameter highly deviates, it has a high correlation with one or more other parameters. Considering that, the results of a Morris method are usually plotted to visualize the relation between the parameter's influence and correlation in a so-called screening plan. Table 2 presents the parameter ranges for the handbook model as estimated for the demonstration case of this work. Minimum (min) and maximum (max) values represent the range of probable values during operation. The reference value (Ref) is either the manufacturer recommended value for operation or the most probable value in the case of environmental and operational factors.

The data from Table 2 are used as input in the Morris method described before. The method calculates the elementary effect for each parameter using the handbook model. The means and standard deviations are calculated from the results of the elementary effect and presented in a screening plan, Figure 2, as an output. Additionally, the ellipse as proposed before, connecting the half values of the maximum mean and standard deviations, is plotted in red. This indicates that filter size (F_{AC}), operational flow (Q), historical failure ($\lambda_{FD,B}$), and

Table 2. Parameter values handbook model for sensitivity analysis.

Parameter	Min	Max	Ref	Unit
$\lambda_{FD,B}$	2.425	7.275	4.85	failures/million hrs
Q	1	400	260	m^3/h
V_O	1700	1800	1750	rpm
F_{AC}	1	128	32	µm
C_{SF}	1	2	1.5	-

service factor (C_{SF}) have a high correlation and influence. Therefore, it is necessary to focus more on these parameters, as a slight variation in these parameters can substantially impact the result. For operational flow (Q), it is necessary to present time series data since it is a directly measured parameter. The user chooses the service factor (C_{SF}) according to the severity of the operation; 1.25 for uniform load, 1.5 for moderate shock, 1.75 for heavy shock, and 2 for extreme shock (NSWC, 2011). The C_{SF} is 1.5 for this work because it is considered an operation with moderate shock and actual operational hours. $\lambda_{FD,B}$ is the base failure rate of the pump fluid driver. For $\lambda_{FD,B}$ the historical failure rate of such a component can be used. In case that it is not available, it is suggested to use the data provided by the manufacturer (SINTEF, 2002).

4.2. Remaining Useful Life (RUL)

The models introduced in section 2.1 will now be applied to calculate the RUL of the pump in various scenarios. In the simulations, the decrease of the pump condition will be modelled, where a condition equal to 1 represents the initial state of the component, as new from the manufacturer, and 0 represents the component wholly degraded. For the model, condition 0 corresponds to the occurrence of the first failure, for which the time can be obtained from inverting the failure rate. As the parameters can change over time, the "condition" decreases with steps corresponding to the actual values of the parameters. Time steps of 3 seconds are used in this work.

Figure 3 presents the results of four analyses with the failure

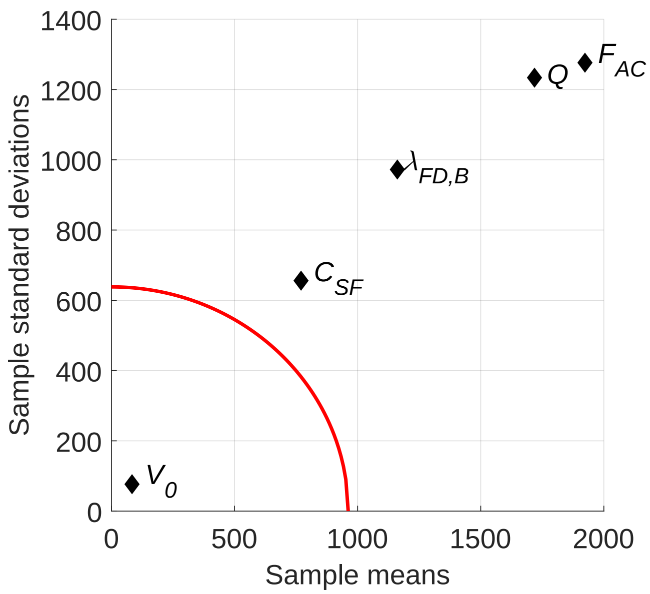

Figure 2. Sensitivity analysis result: screening plan of impeller failure rate model (eq. 1).

rate model. The black line represents the model without adjustment; in this analysis, the model was used with the parameter values as recommended in the handbook. The user provided the filter size, 3000 µm, to avoid animals, vegetation, and trash from the sea entering the pump. However, this F_{AC} is not representative for the particles that can be expected in the flow since the seawater particle diameter varies from 1 to 128 µm (Sheldon et al., 1972). Besides that, the base failure rate used is 0.2 failure/million hours. This $\lambda_{FD,B}$ is also unrealistic since the generic handbook model does not cover the specific application and environmental factors considered in this case. Applying the model without adjusting the parameters, for the type of component and application, can lead to an unexpected failure.

The green line represents the model with $\lambda_{FD,B}$ = 4.85 failures/million hours, where the base failure rate has been adjusted to apply for the specific case in this work. The parameters Q and C_{SF} are realistic if applied as recommended by the model. However, as can be observed in Figure 2, the filtration level (F_{AC}), also has a high correlation and influence, which still results in an unrealistic model. Applying the partially adjusted model is still very conservative, and may lead to premature replacement of components.

The blue line (“Model entirely adjusted” in Figure 3), considers the realistic reference data for all parameters, **Ref** value, as listed in Table 2. The red line (“Realistic Model in Figure 3), presents the model for actual measured data of Q. As presented in Figure 4, there is only data for the first 1500 operating hours, and there is only a slight variation of this parameter. This also yields a slight variation on the estimated degradation in Figure 3, where beyond 1500 hours the dashed red line represents the extrapolation of the initial curve. The reference value is applied for F_{AC} since there is no time-series data for this parameter. The authors intend to use the probability distribution of this parameter in future work. The reference value is also used for $\lambda_{FD,B}$ since there is no historical data; the authors used the most likely data from SINTEF.

The present pump works on average 1946 hours per year. The model entirely adjusted indicated that the impeller would work 6000 hours before the first failure. It indicates that the impeller lifetime is around three years. According to Rivas (2008), proper coverage of the NPSH rate can guarantee a lifetime of 40000 calendar time hours, i.e., around 4.5 years. The handbook model presents a conservative approach to determine the failure, besides evaluating the actual failure rate. However, the predicted lifetime looks reasonable. Validation of the model with accurate data will take place in future works.

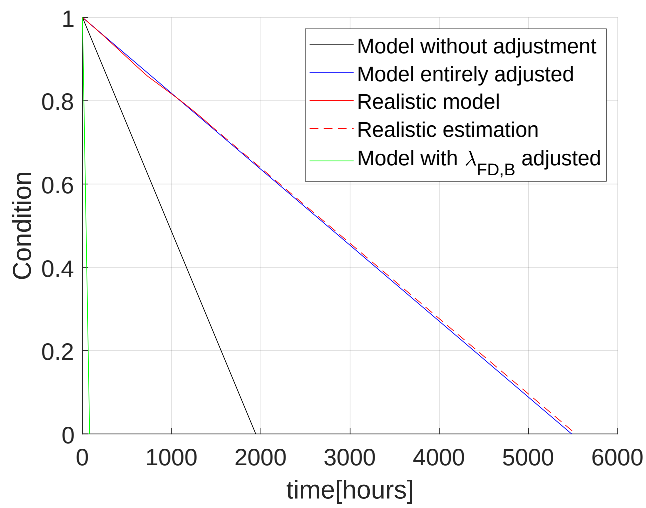

Figure 3. RUL Failure rate model for with different usage of data.

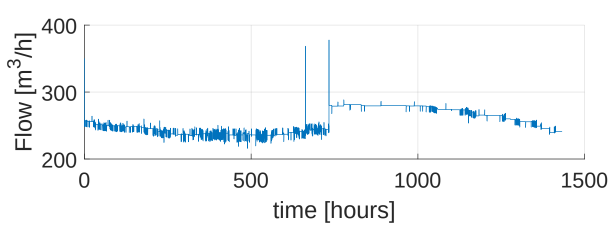

Figure 4. Historical variation of the flow, as obtained from measurements.

Figure 5 presents the RUL for the scenarios using the F_{ACmin} (1 µm) and F_{ACmax} (128 µm) . The reference values were used for the other parameters. In these analyses, a significant difference is observed in the RUL prediction for the minimum and maximum values. This confirms that F_{AC} indeed is a high impact parameter, as was derived from its position in Figure 2.

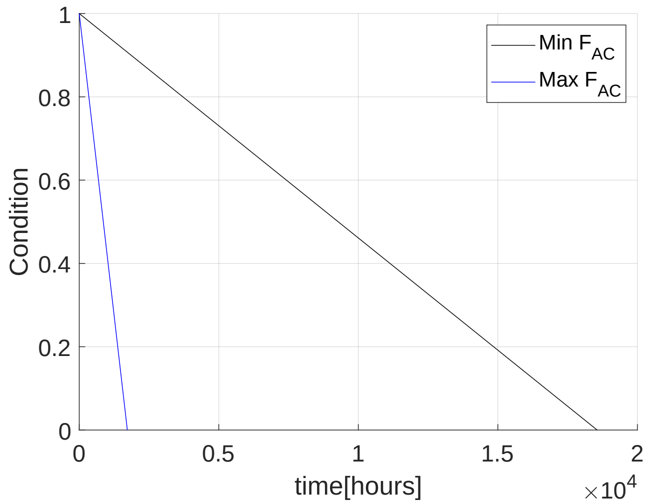

Figure 5. RUL Failure rate model F_{AC} minimum and F_{AC} maximum.

Figure 6 presents the RUL for the scenarios using V_{Omin} (1700 rpm) and V_{Omax} (1800 rpm). Again, the reference values were used for the other parameters. In these analyses, a smaller difference in the RUL prediction for the minimum and maximum of this values range is observed. As expected from Figure 2, V_O has a much lower impact than F_{AC}. However, this parameter can still reduce uncertainties although the effect is rather small. This reduction makes some authors consider the available data from parameters indicated with a small impact in a sensitivity analysis (Meghoe et al., 2020).

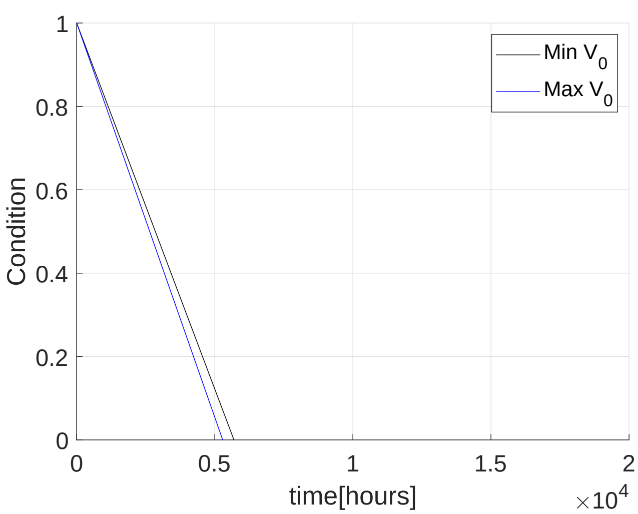

Figure 6. RUL Failure rate model V_O minimum and V_O maximum.

Finally, Figure 7 presents the result of the model applied to three different applications. The scenario of the pink line presents a prediction of RUL in an area where the particle diameter is 64 µm. The green line represents a scenario where the particle diameter is 8 µm. Finally, the dashed line presents the same component applied in two different environments. In the first 3000 hours, the particle diameter is 8 µm and after that, operation is continued in an area with a particle diameter of 64 µm. This analysis demonstrates that the model can be applied in different environments with a data update. It is not necessary to restrict to only one scenario. Once the data is updated, the RUL is also updated.

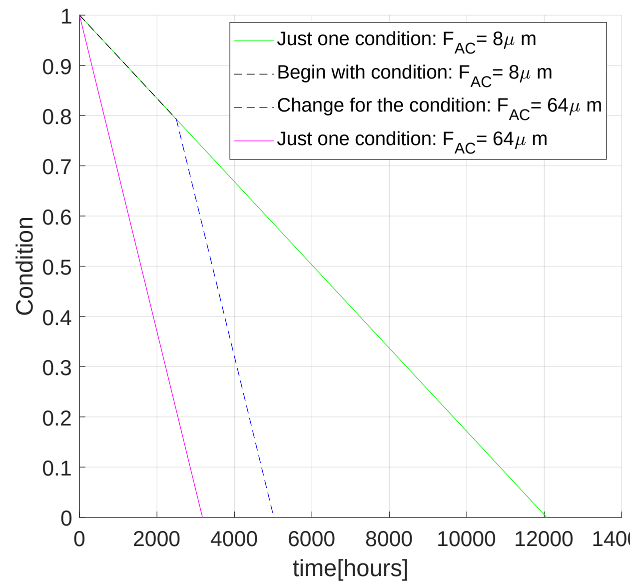

Figure 7. RUL Failure rate model F_{AC} In two different applications.

5. Conclusion

The data selection criterion proposed in this work to reduce prediction uncertainties proves to be effective. The prediction is based on failure rate, but rather than taking a single value for it, it is derived from a model containing a number of parameters. A sensitivity analysis is then used to identify which of these parameters play a significant role in the estimation of the actual failure rate and hence the prediction of the RUL. The method proposed to identify the significant parameters in a Morris sensitivity analysis, using an ellipse, with the ratio at half the value of the largest standard deviation and mean, proved to be effective. When compared to other studies, the method proved to select the same parameters, apart from a few specific cases, which seemed a deliberate choice of the authors of those studies. This work also demonstrates that this approach works for different applications, and the method can calculate the new prediction considering the previous stage. Therefore, it is concluded that the criterion proposed in this work meets the requirement of data selection for failure prediction.

ACKNOWLEDGMENT

The authors would like to thank the MaDeSi project for the financial support and acknowledge Chris Rijsdijk from the Netherlands Defence Academy (NLDA) for the fruitful discussions on pump failures and data acquisition. We also thank the PrimaVera project team for the insightful discussions.

REFERENCES

Abdel Fatah, A. A., Hassan, M. A., Lotfy, M., & Dimitri, A. S. (2019). Health monitoring of centrifugal pumps using digital models. *Journal of Dynamic Systems, Measurement, and Control*, *141*(9).

Bumblauskas, D., Gemmill, D., Igou, A., & Anzengruber, J. (2017). Smart maintenance decision support systems (smdss) based on corporate big data analytics. *Expert systems with applications*, *90*, 303–317.

Campolongo, F., Cariboni, J., & Saltelli, A. (2007). An effective screening design for sensitivity analysis of large models. *Environmental modelling & software*, *22*(10), 1509–1518.

Forrester, A., Sobester, A., & Keane, A. (2008). *Engineering design via surrogate modelling: a practical guide*. John Wiley & Sons.

Gertler, J. J. (1988). Survey of model-based failure detection and isolation in complex plants. *IEEE Control systems magazine*, *8*(6), 3–11.

Ivankevich, A., Piterska, V., Shakhov, A., Shakhov, V., & Yarovenko, V. (2019). A proactive strategy of ship maintenance operations. In *2019 ieee 14th international conference on computer sciences and information technologies (csit)* (Vol. 3, pp. 126–129).

Janse, D. (2009). *Technisch handboek centrifugal pumps bsmi 1500* (Tech. Rep.). Glacisstraat 165, 4381 SE Vlissingen, The Netherlands: Damen Schelde Naval Shipbuilding.

Kohli, M. (2017). Using machine learning algorithms on data residing in sap erp application to predict equipment failures. *International Journal of Engineering & Technology*, *7*(2.28), 312–319.

Lauwaert, B., De Witte, B., Devriese, L., Fettweis, M., Martens, C., Timmermans, S., … Vanlede, J. (2016). *Synthesis report on the effects of dredged material dumping on the marine environment (licensing period 2012-2016)* (Tech. Rep.). RBINS-ILVO-AMT-AMCS-FHR report BL/2016/09, 107p.

Lei, X., & Sandborn, P. A. (2016). Phm-based wind turbine maintenance optimization using real options. *Int J Progn Health Manag*, *7*(1), 1–14.

Meghoe, A., Loendersloot, R., & Tinga, T. (2020). Rail wear and remaining life prediction using meta-models. *International Journal of Rail Transportation*, *8*(1), 1–26.

Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, *33*(2), 161–174.

NSWC. (2011). *Handbook of reliability prediction procedures for mechanical equipment* (Tech. Rep.). Naval Surface Warface Center Carderock Division.

Parrondo, J. L., Velarde, S., & Santolaria, C. (1998). Development of a predictive maintenance system for a centrifugal pump. *Journal of Quality in Maintenance Engineering*.

Razmjoo, A., Xanthopoulos, P., & Zheng, Q. P. (2017). Online feature importance ranking based on sensitivity analysis. *Expert Systems with Applications*, *85*, 397–406.

Rivas, M. A. (2008). High-pressure solution for reverse osmosis applications. *World Pumps*, *2008*(496), 20–22.

Ruano, M., Ribes, J., Ferrer, J., & Sin, G. (2011). Application of the morris method for screening the influential parameters of fuzzy controllers applied to wastewater treatment plants. *Water Science and Technology*, *63*(10), 2199–2206.

Sheldon, R., Prakash, A., & Sutcliffe Jr, W. (1972). The size distribution of particles in the ocean 1. *Limnology and oceanography*, *17*(3), 327–340.

SINTEF, I. M. (2002). *Oreda - offshore reliability data handbook* (Tech. Rep.). SINTEF Insdustrial Management.

Tian, W. (2013). A review of sensitivity analysis methods in building energy analysis. *Renewable and Sustainable Energy Reviews*, *20*, 411–419.

Touhami, H. B., Lardy, R., Barra, V., & Bellocchi, G. (2013). Screening parameters in the pasture simulation model using the morris method. *Ecological Modelling*, *266*, 42–57.

Wang, C., Peng, M., & Xia, G. (2020). Sensitivity analysis based on morris method of passive system performance under ocean conditions. *Annals of Nuclear Energy*, *137*, 107067.

Wang, J., Zhang, L., Zheng, Y., & Wang, K. (2019). Adaptive prognosis of centrifugal pump under variable operating conditions. *Mechanical Systems and Signal Processing*, *131*, 576–591.

BIOGRAPHIES

Núbia Silveira obtained her Bachelor of Science degree in Mechatronic Engineering at the Pontificia Universidade Catolica (PUC) of Minas Gerais in 2010. Between 2009 and 2015, she worked as a Maintenance engineer responsible for corrective, preventive, and predictive maintenance of heavy-duty machines in civil construction and in the mining process. In 2019 she received her Master of Science degree in Aeronautical and Mechanical Engineering at the Aeronautical Design, Aerospace Structures and Systems department of the Aeronautical Technological Institute (ITA). Nowadays, she is working as a Ph.D candidate in the chair of Dynamics Based Maintenance (DBM) at the University of Twente, where she is developing hybrid models to predict failure in military systems' components in collaboration with the PrimaVera project.

Richard Loendersloot Loendersloot received a MSc degree in Mechanical Engineering (2001) and a did his PhD research at the University of Twente, on thermoset resin flow processes through textile reinforcements during composite production process Resin Transfer Moulding and obtained his PhD degree in 2006. He worked in an engineering office on high-end FE simulations of a variety mechanical problems to return to the University of Twente in 2008 as assistant professor for Applied Mechanics. His research started to focus on vibration based structural health and condition monitoring, being addressed in both research and education. He became part of the research chair Dynamics Based Maintenance upon its initiation in 2012. His research covers a broad range of applications: from rail infra structure monitoring, to water mains condition inspection and aerospace health monitoring applications, using both structural dynamics and ultrasound methods. He is involved in a number of European and National funded research projects. He became associate professor in 2019, currently in charge of the daily lead of the Dynamics Based Maintenance group.

Annemieke Meghoe obtained her Bachelor of Science degree in Mechanical Engineering at the Anton de Kom University of Suriname. In 2015 she received her Master of Science degree in Mechanical Engineering at the Applied Mechanics department of the University of Twente. She continued as a PhD student in the chair of Dynamics Based Maintenance (DBM) at the University of Twente, where she developed physics-based models for the rail infrastructure in collaboration with Strukton Rail. She obtained her PhD degree in December 2019 and is currently working as an assistant professor in the chair of DBM starting from January 2020. Her research interests include railway infrastructure and predictive maintenance.

Tiedo Tinga is a full professor in dynamics based maintenance at the University of Twente since 2012 and full professor Life Cycle Management at the Netherlands Defence Academy since 2016. He received his PhD degree in mechanics of materials from Eindhoven University in 2009. He is chairing the smart maintenance knowledge center and now leads a number of research projects on developing predictive maintenance concepts, mainly based on physics of failure models, but also following data-driven approaches.