# Comprehensive clustering approach for managing maintenance in large fleet of assets

J. Izquierdo

*Ikerlan Technology Centre, Control Monitoring and O&M Area, 20500 Gipuzkoa,Spain Department of Industrial Management I, School of Engineering, University of Seville, Sevilla, Spain*
*E-mail: jizquierdo@ikerlan.es*

A. Crespo

*Department of Industrial Management I, School of Engineering, University of Seville, Sevilla, Spain*

J. Uribetxebarria

*Ikerlan Technology Centre, Control Monitoring and O&M Area, 20500 Gipuzkoa,Spain*

A. Erguido

*Ikerlan Technology Centre, Control Monitoring and O&M Area, 20500 Gipuzkoa,Spain Department of Industrial Management I, School of Engineering, University of Seville, Sevilla, Spain*

The maintenance management of large fleets of assets which include several technical solutions operating in different operational contexts has been a recurrent research topic in the literature. Current approaches to establishing fleet maintenance plans are primarily criticality-based, considering failures consequences and assets reliability; the reliability model is often supported by the idea of pooling data from similar pieces of equipment. In spite of the capability to reduce the population offered by data-pooling, its criteria may still lead to a quite large number of segments. Therefore, it results in an equally large amount of maintenance plans along with their inherent operational and administrative difficulties. It is the purpose of the paper to introduce a novel and comprehensive approach; it integrates statistical methods and clustering algorithms to render a fleet segmentation which allows better customization of maintenance plans involving fewer efforts. The approach is summarized in a decision chart which collects the logic behind the use of every algorithm, tool and technique.

*Keywords*: Maintenance, Fleet of assets, Reliability, Clustering, Operational context, Proportional hazards

## 1. Introduction

### 1.1. *Motivation and research*

It is broadly accepted in the literature that managing maintenance in a large fleet of assets is not a trivial issue, the asset management discipline provides valuable aid when facing such challenge. In order to successfully manage the assets' maintenance, it is essential to understand their failure behaviours and the uncertainties influencing them. Accordingly, the techniques of reliability engineering are of use to deal with the difficulties of managing the maintenance of fleets involving several technical solutions operating under different working conditions. This paper addresses the maintenance management in such a context.

Nowadays, with the increasing implementation of Information Technology (IT) solutions along with the capacities of Big Data algorithms, new opportunities arise to optimize the maintenance in large fleets of assets. To this aim, this paper contributes to the asset management discipline with a comprehensive approach for optimizing the maintenance of fleets. The proposed framework enables customized maintenance plans through a holistic process. In this process, statistical concepts and a clustering algorithm are proposed in order to overcome two of the main difficulties in the maintenance management of fleets:

(i) Assessment of the impact of operational context on failures. The assets in the fleet operate not only in different geographical locations but under heterogeneous working conditions which, besides, might change. Therefore, it is important to asses the impact of the operational context variables on the asset's reliability so the maintenance plans can be defined and customized accordingly.
(ii) Assets clustering based on their similarities. The technical characteristics of the assets also condition their failure behaviours. In a large fleet of assets, it is not surprising to find a considerable amount of different technical solutions. To face this issue, the k-medoids al-

*Proceedings of the 29th European Safety and Reliability Conference.*
*Edited by* Michael Beer and Enrico Zio

*Published by* Research Publishing, Singapore.
ISBN: 978-981-11-2724-3; doi:10.3850/978-981-11-2724-3_0094-cd

gorithm is proposed in a Spectral Clustering approach.

The research in the paper describes the aspects regarding the information needed to support the mathematical models. It also lays the foundations to establish an asset management strategy based on customized maintenance plans regarding assets' criticality, operational context and technical characteristics.

### 1.2. *Related works and limitations*

Were the assets of the fleet to have a maintenance plan based on a time-to-failure model, it is essential to select an adequate model to such purpose; in Louit et al. (2009) a practical procedure to be followed when making the choice of the model is proposed. The modeling of the reliability of the assets in the fleet involves combining data, in Louit et al. (2009) as well as in Stamatelatos et al. (2002) several conditions to be met by the equipment subject to data pooling are stated. Nonetheless, following these criteria lead to a number of groups of failure data which still involve considerable management complexities.

In order to facilitate asset management in the stated context, similarity-based approaches have been proposed in the literature Cannarile et al. (2018); Baraldi et al. (2015); Cannarile et al. (2015). It is especially interesting for the approach later presented, the work of Cannarile et al. (2018) and Cannarile et al. (2015) where a spectral clustering approach is proposed. Their work is focused on clustering assets in order to address maintenance optimization of all the assets belonging to the same cluster. A similar approach can be found in Jiang et al. (2013) but instead, the k-medoids clustering algorithm is proposed, which shows better results for probability distribution similarities and has proven to perform better on large data sets according to Velmurugan and Santhanam (2010).

It is also important to consider that every asset in the fleet is operating under different operational conditions which usually change in the life span of the assets. The assumption of constant external factors leads to inaccurate estimations of reliability, therefore it is necessary to directly address the influence of a changing working environment on failure behaviours Peng et al. (2016). This is a topic which has attracted a considerable amount of scientific research, e.g. Okaro and Tao (2016); Lin et al. (2016); Tang et al. (2014). In the related literature, the Proportional Hazards Model has been widely utilized for analysing the effects of the operational context variables Izquierdo et al. (Izquierdo et al., 2018); Xie et al. (2017); Li et al. (2015); Zhiguo et al. (2010); Bendell et al. (1986).

All the mentioned works are designed to support and enhance the maintenance management of the assets in the fleet because maintenance decisions entail important organizational considerations. The definition of the maintenance plans should be conditioned by the chosen asset management strategy which is directly derived from the organizational goals Crespo et al. (2009). Criticality analysis ensures that the strategic demands are met by the maintenance management program and that the assets are properly prioritized Crespo et al. (2016), therefore aligning the operation of the fleet with the and strategic objectives.

### 1.3. *Overview*

Given the importance of the reviewed research in the context of fleet maintenance management, a process is proposed to comprehensively gather the contributions of the works in a novel methodology. The first stage in the methodology consists of foundations to collect the information needed as well as its initial treatment in order to establish the proper Reliability-Maintainability-Availability (RAM) database. Taking as a reference the work proposed in Louit et al. (2009); Stamatelatos et al. (2002) data pooling techniques are applied in the next stage of the methodology, then a valid Proportional Hazards Model is fit to every set of data resulting from combining assets' failures. Once the operational context information has been separated from the baseline reliability of the assets it is possible to cluster them by the spectral clustering approach. Finally, all the potential of the analytical work is framed into a business context, the last stage of the methodology proposes a procedure to align the maintenance management of the fleet with the organizational goals.

The methodology is proposed in section 2 where the whole schema is introduced at first and then every stage is further explained in their corresponding subsections. Then in section 3, the application of the analytical part of the methodology is presented with a case study example. Finally, section 4 summarizes the benefits provided by the application of the proposed methodology as well as those lines of research which are considered worthy of further investigation.

## 2. Proposed approach

The cross-functional process is summarized in the decision chart represented in Figure 1. It has been divided into three stages according to the main aspects covered: data processing, mathematical modelling and maintenance management.

This framework strives to translate the efforts invested in information harvesting, regarding fleets of assets, into specific maintenance actions. In order to realize value from the data, several mathematical algorithms are proposed which are selected to tackle specific difficulties. In the last stage, the criticality analysis provides the align-

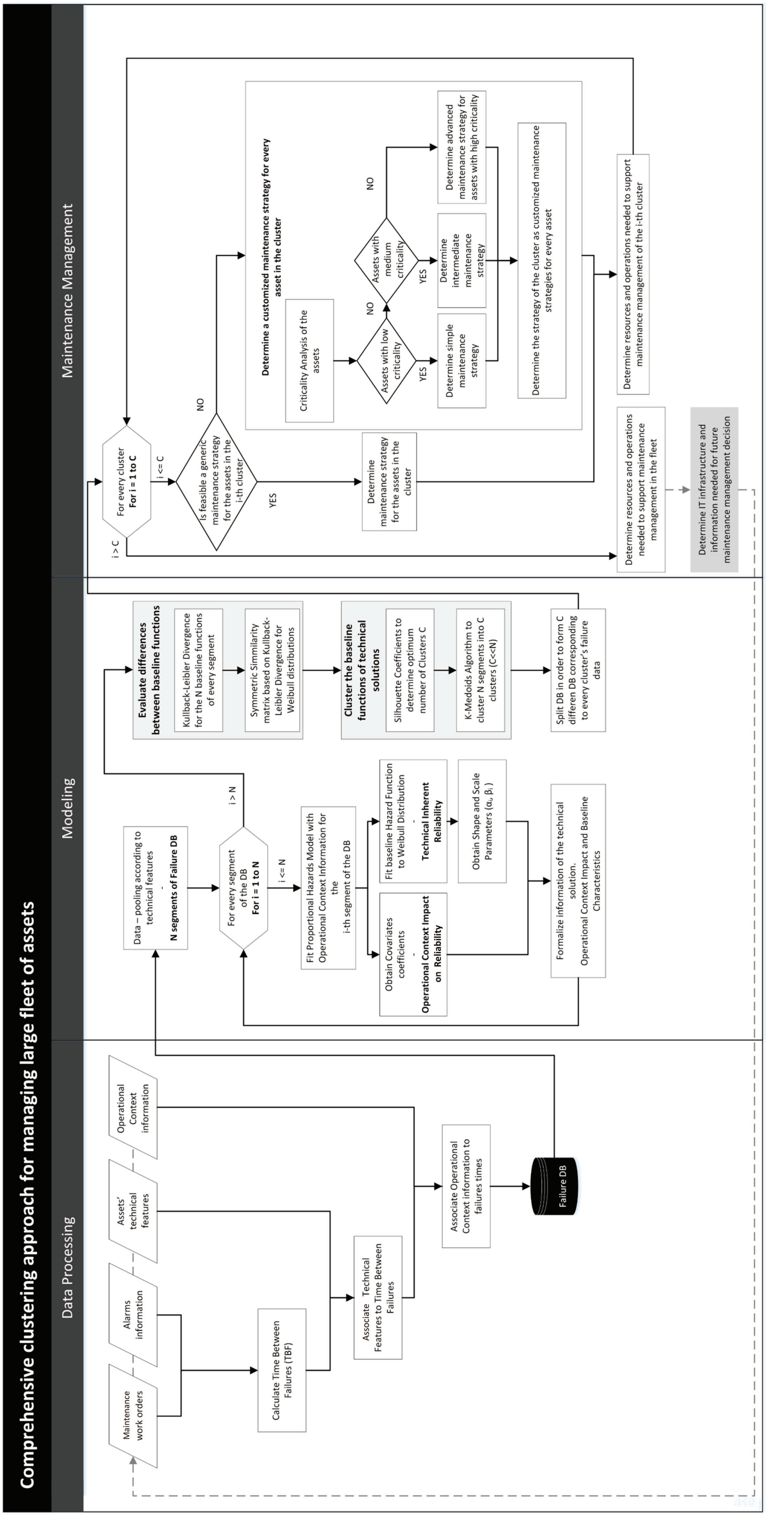


Fig. 1. Decision chart of the proposed approach

ment of the maintenance actions with the business objectives and they are considered the needs regarding infrastructure and resources to support the defined maintenance plans. It is important to notice that the authors do not conceive the approach as a static process, but as a cycle based on fleet data exploitation. According to this point of view, in the last stage of the framework, there is a step proposing to review the IT infrastructure required to fulfil maintenance management objectives. In the followings subsections, each of the three stages is thoroughly explained.

### 2.1. *Data processing*

This first stage of the approach defines the information structure of the failure database taking into account the challenges to overcome when modelling reliability. The principles stated in related literature Hameed et al. (2011); Louit et al. (2009) are integrated into the approach as well as ideas extrapolated from ISO14424 (2006). The RAM database is to be built in an organized and structured way, to such purpose, different types of information are represented in the Data Processing stage of Figure 1.

- **Maintenance Data.** This information, often found in GMAO systems, is mainly based on insights regarding failure behaviour, and maintenance works. It is essential to identify maintenance tasks in equipment and components at the defined indenture level. Another key aspect is the cost of the maintenance actions in terms of downtime and resources employed.
- **State Information.** It consists of measurements and alarms, generally, coming from condition monitoring technology or SCADA alarms. The states shown by these sources should match the information contained in the maintenance data, however, the real practice shows that it is not usually as straight forward. Therefore it is essential to identify the patterns in the state information which matches the maintenance data in order to calculate the RAM indicators.
- **Assets technical features.** Mainly the design characteristics but also the manufacturer's data and equipment attributes. The different features of the assets in the fleet may entail different inherent reliabilities. Thus, this information will provide its value when applying the data pooling techniques. The characterization of the different components in the asset should be consistent with the indenture level of the maintenance data in order to link technical features to failures modes.
- **Operational Context Information.** Data of the variables regarding the operation of the asset and its environment. In order to fully characterize the failure events of the assets is essential to describe their operation with parameters that may influence failure times. Therefore, this information will enable more advanced reliability models and thus more accurate maintenance strategies.

Providing that every database is available and the data in it fulfil certain quality standards, it is possible to create a single source of information containing the failure information. This single Failure Data Base (DB) consists of the TBF data from the different failure modes of the assets in the fleet, each one of them associated not only with its corresponding equipment and technical features, but with the operational context variables values as well.

### 2.2. *Modelling*

Having the failure data properly organized, the next course of action in the proposed process is to divide it into several segments in order of begin addressing the maintenance problem. To do so, data-pooling propose combining data taking into account technical features (same design, same hardware, same function, etc.). Therefore, in the database the TBF data are grouped according to the technical characteristics, rendering a segmented failure database according to exclusively technical solutions.

The data associated with every technical solution is susceptible to being fit into a proper reliability model, in the presented framework the proposed one is the Proportional Hazards Model (PHM) Cox (1972). Its most common expression is the hazard function in Equation 1,where $h(t, X_1, ..., X_k)$ represents the hazard rate at time t for an asset with covariates $\mathbf{X} = (X_1, ..., X_k)$, $h_o(t)$ represents the baseline hazard function and $\boldsymbol{\omega} = (\omega_1, ..., \omega_k)$ is the vector of the parameters of the model which describe the effects of each of the covariates.

$$h(t, X_1, ..., X_k) = h_0(t) exp\left(\sum_{j=1}^{k} \omega_j X_j\right) \quad (1)$$

It is proposed the variation of the model known as Weibull PHM, where the baseline hazard $h_o(t)$ follows a Weibull distribution being $\alpha$ the scale parameter (characteristic life) and $\beta$ the shape parameter. Thus, every segment of the database will be fit into a proper Weibull PHM. This will provide for every technical solution in the fleet two kinds of information: (i) the technical inherent reliability collected in the shape and scale parameters and (ii) the effect of the operational context on that reliability collected in the coefficients of the exponential part of the model. Expressing the failure behaviour of the solutions in such way allows for decoupling the technical reliability information and the operational context variations.

Considering that every technical solution have a defined Weibull probability distribution it is possible to address the clustering approach following the work in Cannarile et al. (2018, 2015). It is possible to define the similarity between two distributions according to Equation 2, where $d_{KL}^{sym}$ is the Kullback-Leibler Divergence defined by Equations 3 and 4 being $\gamma \approx 0.5772$ the Euler-Mascheroni constant and $\Gamma = \int_{-\infty}^{+\infty} t^{z-1} e^{-t} dt \; z \geq 0$ the gamma function.

$$w_{ij} = \frac{1}{1 + d_{KL}^{sym}} \quad (2)$$

$$d_{KL}^{sym} = \frac{1}{2}(d_{KL}(\mu_i \| \mu_j) + d_{KL}(\mu_j \| \mu_i)) \quad (3)$$

$$\mathrm{d}_{KL}(\mu_i \| \mu_j) = \log\left(\frac{\beta_i \alpha_j^{\beta_j}}{\beta_j \alpha_i^{\beta_i}}\right) - (\beta_i - \beta_j)\left(\log(\alpha_i) - \frac{\gamma}{\beta_i}\right) + \left(\frac{\alpha_i}{\alpha_j}\right)^{\beta_j} \Gamma\left(\frac{\beta_j}{\beta_i} + 1\right) - 1 \quad (4)$$

It is feasible to define a similarity matrix $\overline{\overline{W}}$ in which each element is $w_{ij}$. This similarity matrix is the first step in the process of Spectral Clustering, the process is explained here but the reader is referred to Cannarile et al. (2018); Baraldi et al. (2015) for further details.

From the similarity matrix $\overline{\overline{W}}$ of size $[N, N]$ it is possible to construct the graph $G = (V, E)$ where each of the nodes $v_i$ represents the baseline Weibull distribution of certain technical solution in the fleet, and each edge $e_{ij}$ is the similarity between two distributions based on the Kullback-Leibler Divergence. Having defined such graph, the problem is to find the partition of the graph such that the weights of the edges are small and large for intercluster and intracluster connections correspondingly. To such aim, the spectral clustering algorithm is proposed, therefore it is necessary to compute the normalized Graph Laplacian Matrix. From $\overline{\overline{W}}$ the degree matrix $\overline{\overline{D}}$ is calculated, it is a diagonal matrix whose diagonal entries are defined according to Equation 5, and then the normalized Graph Laplacian matrix is calculated as described in Equation 2.2, where $\overline{\overline{L}} = \overline{\overline{D}} - \overline{\overline{W}}$ and $\overline{\overline{I}}$ is the identity matrix of size $[N, N]$.

$$d_i = \sum_{j=1}^{N} w_{ij}, \quad i = 1, ..., N \quad (5)$$

$$L_{sym} = \overline{\overline{D}}^{-1/2} \overline{\overline{L}}\, \overline{\overline{D}}^{-1/2} = \overline{\overline{I}} - \overline{\overline{D}}^{-1/2} \overline{\overline{W}}\, \overline{\overline{D}}^{-1/2} \quad (6)$$

In order to extract the information of the graph, the C smallest eigenvalues $\lambda_1, ..., \lambda_C$ are selected along with their corresponding eigenvectors $\bar{u}_1, ..., \bar{u}_C$, being C the desired number of clusters. The relevant information is considered by transforming matrix $\overline{\overline{W}}$ into a reduced matrix $\overline{\overline{U}}$ of size $[N, C]$. The columns of $\overline{\overline{U}}$ are the C eigenvectors $\bar{u}_1, ..., \bar{u}_C$ which contain the information regarding the similarities among the i-th baseline distribution and the others. It has been proven that it is possible to enhance cluster properties of the data by normalizing the rows of matrix $\overline{\overline{U}}$ and forming matrix $\overline{\overline{T}}$ Von Luxburg (2007), where every element is computed following Equation 7.

$$t_{ic} = \frac{u_{ic}}{\left(\sum_{c=1}^{C} u_{ic}^2\right)^{0.5}}, \; i = 1, ..., N \,, \; c = 1, ..., C \quad (7)$$

Once the matrix $\overline{\overline{T}}$ is obtained, a k-medoids algorithm is proposed as an unsupervised clustering to partition the data set into C clusters which has proven to perform better for large data sets Velmurugan and Santhanam (2010).

### 2.3. *Maintenance strategy*

The ultimate goal of the previous procedure is the improvement in maintenance management. To such aim, the approach proposes certain guidelines for defining a maintenance strategy which enables the exploitation of the value behind the mathematical modelling. It is mainly based on the criticality analysis of the assets, their prioritization ensures alignment between the strategic demands and the maintenance management programs.

On the one hand, for every cluster identified in the fleet, it is important to initially assess its characteristics since it is possible that a general maintenance strategy for the cluster can address operational and strategic needs of every asset in it. However, while this might be true for clusters whose assets have similar business impact and behaviours, it is important to consider that in most of the cases heterogeneity in strategic needs or failure frequencies will be the most common reality within the clusters of a fleet. To deal with this issue, the approach considers a customized maintenance strategy for every asset in the cluster depending on its criticality. Assets with low, medium and high criticality would have either a simple, intermediate or advanced maintenance strategy correspondingly.

The basic maintenance strategies to which the authors make reference are based on a considerable load of corrective maintenance, periodic inspections or calendar-based preventive actions. The maintenance plans derived from these strategies do not take into account the operational context of the assets, they are designed based on average values of the cluster like the Mean Time Between Failures (MTBF), Mean Down Time (MDT) or the Mean Time To Repair (MTTR).

However, for assets with medium criticality, the maintenance actions might be defined according to some basic reliability models such as Exponential or Weibull distribution. It is also possible to consider as intermediate maintenance plans the ones based on simple cost models build on information regarding the distribution of the failures. The cost models would set the optimum preventive maintenance intervals given several assumptions and simplifications.

For critical assets with important an impact on the business, it is required to adopt advanced maintenance strategies. For these assets are considered state-of-the-art models which take into account the operational context enabling a customized maintenance plan which specifically addresses the needs of each one of them. Some examples of these models would be the dynamic opportunistic maintenance proposed in Erguido et al. (2017) or the dynamic reliability modelling based on artificial neural networks in Izquierdo et al. (Izquierdo et al.).

Having defined the maintenance strategy for every one of the assets, it is important to consider the resources and infrastructure needed to support the maintenance activities derived for the defined plans. To such aim, the proposed process considers the definition of the operations needs for every one of the clusters and after that a general definition of the resources and operations for the whole fleet. Finally, in order to continuously improve the management of the fleet and allow for more advanced maintenance strategies and better performance of the mathematical modelling, it has been considered the definition of the IT infrastructure to support the data module. By these means, the whole approach is aligned with a continuous improvement perspective which considers new investments in IT technologies, programs for improvement of data quality, alternative data structures, implementation of new information systems, etc. These considerations lead to the further development of the company within the asset management discipline and support the strives for better management of large fleets of assets.

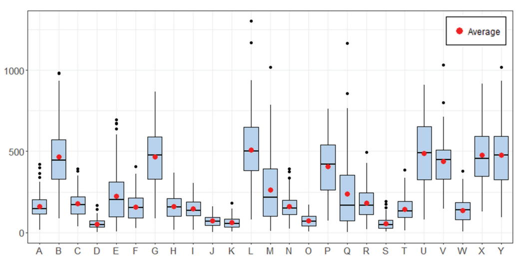


Fig. 2. TBF distribution for every technical solution

## 3. Case Study

To test the suitability of the mathematical procedures proposed in the modelling module of the framework, a case study is presented here. It consists of a practical exemplification of the application of every one of the algorithms considered in the module. In order to prove the validity of the proposed approach, a database has been simulated from real-field data. It consists of multiple failures of a certain failure mode of Heating Ventilating and Air Conditioning systems for rail solutions (HVAC systems). The failures correspond to different technical solutions and have been simulated under, and therefore are associated to, specific conditions of temperature and relative humidity.

According to the data pooling techniques previously introduced, the data failure can be segmented into smaller datasets considering exclusively the technical features of the HVAC system in which the failure occurred. This data pooling technique leads to a segmentation consisting of data sets which correspond to 25 different technical solutions. In Figure 2, where the distribution of the TBFs of each of the technical solutions have been represented in a box-plot, it can be seen the heterogeneity in the TBFs due to both the differences in the technical solutions and the operational context.

Following the steps of the second module of the framework, in order to render a reliability model that explains the influence of the operational context in the failure times, a proportional hazards model has been fit for each of the technical solutions. The PHM of each one of them considers the impact of temperature variations as well as the impact of relative humidity. In order to characterize the baseline hazard functions, they have been adjusted to a two-parameter Weibull distribution and hence characterized by a shape and scale parameter. The reliability function given by the parameters of the baseline hazard of each technical solution can be seen in Figure 3; since the functions are free of the bias in the failure probability provided by the operational context variables, it is possible to already appreciate certain similitude among some of the graphs.

Given that the obtained baseline functions contain the information regarding the reliability of the technical solution it is now possible to cluster them. To such aim, as stated in the previous section, a spectral clustering approach is proposed based on the Kullback-Leibler Divergence. It is calculated for each pair of technical solutions and then a similarity matrix is constructed. From that information, the silhouette coefficients for 1 to 10 number of clusters have been assessed, ranking the solution with 4 clusters as the best with a coefficient of 0.7812, followed by the solution consisting of 3 clusters.

The next step consists of the application of the

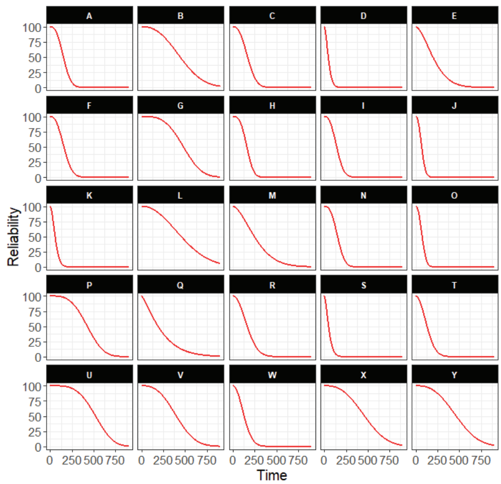


Fig. 3. Baseline reliability of every technical solution

K-Medoids algorithm to obtain 4 clusters out of the initial 25 technical solutions. The rendered solution is represented in Figure 4. In the representation, it can be seen the graph corresponding to the different technical solutions (nodes) connected by edges expressing the similarities among them according to the similarity matrix. The weighted edges rank from 0 to 1, being 0 the least similar and 1 identical failures distributions correspondingly. Along with the graph, the common reliability functions for each one of the clusters is represented. The distributions are characterized by a two-parameter Weibull distribution and they can be customized to certain operational context since the information of the impact of temperature and relative humidity on failure is considered in the exponential part of the Weibull PHM.

It is remarkable, that the proposed approach is capable of detecting clusters with a very different number of nodes within them. In the proposed solution the largest cluster consists of 9 technical solutions while the smallest consists only of 3. It suggests that the algorithms provide an approach valid in fleets in which the different technical solutions are present in different proportions.

Having the fleet cluster into four groups enables to address the maintenance management in a structured and organized way. It is possible to determine whether a common maintenance strategy for the a cluster is possible or otherwise every asset needs a maintenance plan customized to its operational context. However, these considerations belong to the third module of the proposed framework whose application is highly conditioned by the particularities of each company.

It is important to state that the limitations of the solution lies within two aspects of the proposal. Regarding the data there are two main limitations, (i) the proposed solution is limited given the bast amount of required failure data; and (ii) it is also limited by the information regarding the operational context which can be expensive and difficult to obtain in some cases. Considering that those limitations could be overcome the validity of the clustering solution should be subject to the outcome of statistical methods intended to verify that the effect of operational context variables is homogeneous within the clusters.

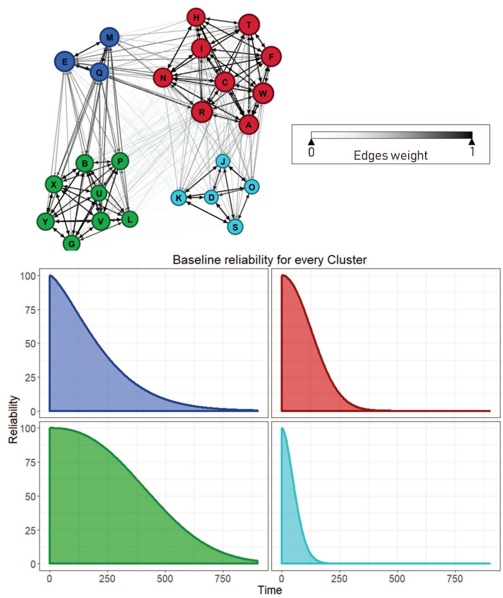


Fig. 4. Baseline reliability clustering results

## 4. Concluding Remarks

The presented framework for managing the maintenance in a large fleet of assets is a comprehensive proposal. This holistic approach considers different aspects of the problem in fleets' maintenance management, it deals with information and database structure, mathematical modelling of failure behaviours and maintenance strategy definition. A remarkable contribution is the explicit treatment of two sources of variability in assets' failures, the presence of different technical solutions for the same failure mode and the heterogeneity in the operational contexts.

The Data Processing module establishes a guide to structure the information coming from several sources of different nature into a database oriented to maintenance management. From a generic point of view, it formalizes the process undergone by practitioners who want to optimize the maintenance of the fleet through reliability engineering. The consideration of information sources of different nature in the database not only enables the mathematical modelling to optimize maintenance, but also provides potential to support other data-driven decision making processes.

The Modelling module has been validated through a case study based on real-field data, the different methods and algorithms proposed are specifically integrated in order to solve specific issues faced when defining the maintenance strategy. The integration of the Proportional Hazards Model with the spectral clustering approach has proven to be useful explaining the variability and grouping the technical solutions in similarity-based clusters taking into account the impact of the operational context. The proposed algorithms have proven their stability against two important aspects, the variation in the TBF distribution due to the operational context and the unequal proportion of certain assets in the fleet.

The Maintenance Management module is presented as the part of the framework intended to translate into organizational value the previous modules. Following the proposed process, the maintenance of the assets is either addressed cluster-wise or customized to the operational context for those assets with high criticality. The criticality analysis provides alignment between the mathematical models describing the reality of the fleet, and the organizational objectives of the company realized by the maintenance strategy; however, it is highly conditioned by the particularities of every organization, so further research, as well as practical implementation, is encouraged. Another important, and worth of further research, aspect of the framework is taking into account the definition of the resources and infrastructure needed to support the operations derived from the maintenance management of the fleet according to the defined maintenance plans.

## References


Baraldi, P. et al. (2015). Unsupervised clustering of vibration signals for identifying anomalous conditions in a nuclear turbine. *Journal of Intelligent & Fuzzy Systems*.

Bendell, A. et al. (1986). Proportional hazards modelling in reliability analysis - an application to brake discs on high speed trains. *Quality and Reliability Engineering International*.

Cannarile, F. et al. (2015). Handling reliability big data: A similarity-based approach for clustering a large fleet of assets. In *Safety and Reliability of Complex Engineered Systems-Proceedings of the 25th ESREL Conference*.

Cannarile, F. et al. (2018). A clustering approach for mining reliability big data for asset management. *Proceedings of the Institution of Mechanical Engineers, Journal of Risk and Reliability*.

Cox, D. R. (1972). Models and life-tables regression. *JR Stat. Soc. Ser. B*.

Crespo, A. et al. (2009). The maintenance management framework: A practical view to maintenance management. *Journal of Quality in Maintenance Engineering*.

Crespo, A. et al. (2016). Criticality analysis for maintenance purposes: A study for complex in-service engineering assets. *Quality and reliability engineering international*.

Erguido, A. et al. (2017). A dynamic opportunistic maintenance model to maximize energy-based availability while reducing the life cycle cost of wind farms. *Renewable Energy*.

Hameed, Z. et al. (2011). Challenges in the reliability and maintainability data collection for offshore wind turbines. *Renewable Energy*.

ISO14424 (2006). *Petroleum, Petrochemical and Natural Gas Industries: Collection and Exchange of Reliability and Maintenance Data for Equipment*. International Organization for Standardization.

Izquierdo, J. et al. Dynamic artificial neural network-based reliability for assessing the impact of operational context on assets' failures. Unpublished.

Izquierdo, J. et al. (2018). Assessing the impact of operational context variables on rolling stock reliability. a real case study. In *Safety and Reliability - Safe societies in a changing world*, Proceedings of ESREL 2018. Taylor & Francis Group, London.

Jiang, B. et al. (2013). Clustering uncertain data based on probability distribution similarity. *IEEE Transactions on Knowledge and Data Engineering*.

Li, L. et al. (2015). Cox-proportional hazards modeling in reliability analysis: A study of electromagnetic relays data. *IEEE Transactions on Components, Packaging and Manufacturing Technology*.

Lin, S. et al. (2016). A failure rate model for traction transformer based on phm considering multiple factors. In *2016 Prognostics and System Health Management Conference (PHM-Chengdu)*.

Louit, D. et al. (2009). A practical procedure for the selection of time-to-failure models based on the assessment of trends in maintenance data. *Reliability Engineering & System Safety*.

Okaro, I. A. and L. Tao (2016). Reliability analysis and optimisation of subsea compression system facing operational covariate stresses. *Reliability Engineering & System Safety*.

Peng, W. et al. (2016). Reliability of complex systems under dynamic conditions: A bayesian multivariate degradation perspective. *Reliability Engineering & System Safety*.

Stamatelatos, M. et al. (2002). Fault tree handbook with aerospace applications. -.

Tang, Z. et al. (2014). Analysis of significant factors on cable failure using the cox proportional hazard model. *IEEE Transactions on Power Delivery*.

Velmurugan, T. and T. Santhanam (2010). Computational complexity between k-means and k-medoids clustering algorithms for normal and uniform distributions of data points. *Journal of computer science*.

Von Luxburg, U. (2007). A tutorial on spectral clustering. *Statistics and computing*.

Xie, Q. et al. (2017). Cox proportional hazards modelling of blockage risk in vitrified clay wastewater pipes. *Urban Water Journal*.

Zhiguo, L. et al. (2010). Change detection in the cox proportional hazards models from different reliability data. *Quality and Reliability Engineering International*.