

Big Data Management and the Application of Reliability Engineering for the Optimization of Critical Components

Rita Anne Andrade¹, Bruno Mendonça² and Raissa Branco³

1. Mechanical Engineer

2. Electrical Engineer

3. Automation Engineer

Hydro Bauxite & Alumina, Paragominas, Brazil

Corresponding author: rita.andrade.barreto@hydro.com

<https://doi.org/10.71659/icsoba2025-bx016>

Abstract

Hydro Bauxite & Alumina operates two pump station dedicated to the transportation of slurry that feeds the Alunorte refinery. This system utilizes positive displacement triplex diaphragm pumps, powered by 2 000 hp motors, and play a critical role in ore transportation. Given the strategic importance of the positive displacement pumps for process continuity and the significant amount of equipment in operation, rigorous monitoring of the service life of critical components is essential. The integration of Big Data management and reliability engineering applications enable the acquisition of accurate and reliable data, supporting the development of optimal maintenance strategies. Such monitoring ensures operational reliability by enabling the timely replacement of components as they reach the end of their service life. Furthermore, it aims to optimize replacement planning, thereby avoiding overlapping maintenance activities that could lead to workload peaks and increased operational costs. To achieve these objectives, a monitoring system based on the concept of Residual Life was implemented. This approach evaluates the remaining service life of each critical component, providing a clear metric to optimize the maintenance interval. As a result, operational continuity is ensured with greater efficiency and predictability, reinforcing process reliability, and optimizing resource utilization.

Keywords: Big Data, Reliability Engineering, Critical Component.

1. Introduction

Hydro Bauxite & Alumina, a key component in Hydro's global strategy, began operations in March 2007. Since then, the mine, located on the Miltônia 3 Plateau in northeastern Pará, has handled around 16 million tonnes of ore per year, 11 million of which is bauxite. After extraction, the ore undergoes a rigorous beneficiation process—including crushing, grinding, and classification. The processing is carried out using a wet method, and in the end, the bauxite, in slurry form, is ready for transport.

The beneficiated bauxite is pumped through a 244 km slurry pipeline—the first in the world to transport this type of ore—on a route connecting the Paragominas Mine to the Alunorte refinery in Barcarena, where the bauxite is transformed into alumina, an essential raw material for aluminium production. This pioneering solution eliminates heavy vehicle traffic, resulting in lower CO₂ emissions and a safer, more efficient transport system, reinforcing our commitment to sustainability and innovation in the aluminium industry.

Along its route, the pipeline crosses seven municipalities and four major rivers. The system operates at a maximum pressure of 90 bars, with two pumping stations: one at Paragominas Mining (MPSA) in Paragominas, equipped with seven pumps, and another in Tomé-Açu, equipped with six pumps. With a flow rate between 1 800 and 1 950 m³/h, these stations ensure a continuous and safe flow of the slurry. The bauxite is transported using triple positive displacement pumps, model TZPM 2000 – single-acting piston and diaphragm pumps with three

cylinders, driven by a crankshaft and designed to operate with mineral slurries under high pressure.

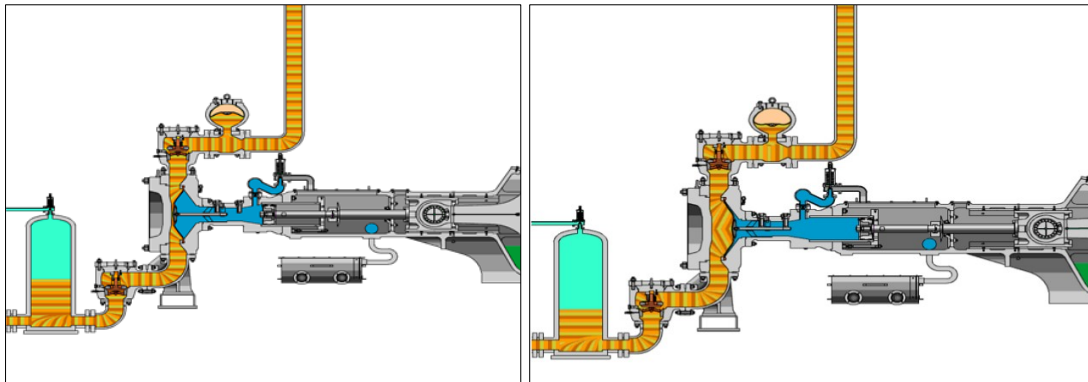


Figure 1. Representation of the Diaphragm Filling and Emptying.

The piston fits into the cylinder rod, which operates inside a lined cylinder – both of which are highly resistant to wear. In the event that solids penetrate the drive fluid compartment due to a diaphragm rupture, no immediate damage will occur [1].

The piston has two sets of seals, as shown in Figure 2: one that seals the pressure side (020) and another that prevents the suction of air and oil (21). Between the two sets of seals, a ring (012) is placed to support the gravitational load of the piston and the piston rod [1].

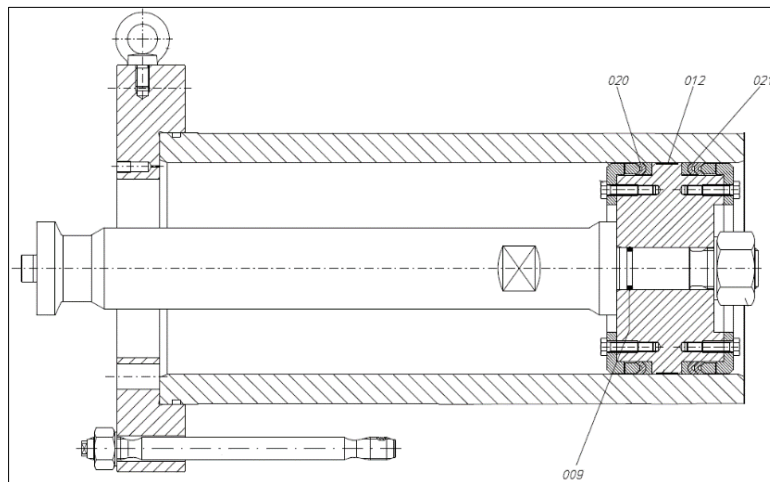


Figure 2. Piston seal rings in the Piston Unit.

The manufacturer's manual states that both the diaphragms and the piston seal sets have a service life of 8 000 hours, and their replacement is recommended after this period [1].

A diaphragm can fail due to excessive pressure in the drive fluid, excessive hammering caused by high pulsations, lack of drive fluid, contaminants in the slurry, or fatigue. Any of these possible causes can allow slurry to pass into the drive fluid section, leading to abrasive wear of the components and consequently greater damage to the pump's operation [1].

The piston seals, in turn, may fail due to severe localized wear around the circumference, grooves on the sealing rings, a rough texture on the rings' surface, and wear on the outer diameter.

Hydro Bauxite & Alumina uses PIMS (Plant Information Management System), based on the PI System, which enables data collection from various sources within an industrial plant, such as PLC, DCS, SCADA and other systems. It stores this data, allowing for the visualization of both historical and real-time information, identification of trends, analysis of asset and process performance, and support for decision-making based on asset behaviour.

The reliability engineering initiative presented in this paper began in 2021 with the application of classical techniques. In 2024, the activities were expanded with the use of modern asset management methods, with the aim to optimize critical components of the bauxite pumping system.

2. Methodology

2.1 Data Collection and Preparation

The initial maintenance strategy for managing the critical assets of the positive displacement pumps at Hydro Bauxite & Alumina employed a time-based scheduled maintenance, using the SAP Plant Maintenance module from the SAP ERP system, and focused on asset maintenance management. In this model, the replacement of critical components was carried out at predefined intervals, regardless of actual operating conditions or expected remaining service life.

Although the implementation of SAP brought significant benefits – such as reduced operating costs, increased efficiency, improved safety indicators, and extended asset life – the operational performance of the critical items in the slurry pipeline did not reach the desired level of optimization [2, 3].

In response to this situation, raw operational data was collected from the positive displacement pumps. The data was extracted from multiple sources: PIMS systems, SAP records, and instrumentation signals from the PLCs of the positive displacement pumps, as shown in Table 1.

Table 1. Evaluated Variables

Maintenance Variables	Operation Variables	Process Variables Bauxite Slurry
Mean time between failure (h)	Instantaneous Flow Rate (m ³ /h)	Density (g/cm ²)
Mean time to Repair (h)	Pressure (kgf/cm ²)	Viscosity
Physical Availability (%)	Cumulative Production (t)	Solid Percentage (%)
Maintenance Cost	Physical Utilization (%)	pH
Change History	Power (kW)	Granulometry
Strokes (S/m)		
Run time (h)		
Motor Current (A)		

Based on this information, a data-driven approach was initiated to improve the understanding of the operational context in which the critical assets are an integral part. It was verified whether the main process variables were operating within their design limits. The results will be discussed in the next chapter.

A comprehensive analysis of the historical data for critical process variables was carried out, using Big Data management practices and Data Science techniques to ensure the adequacy and reliability of the information. The study aimed to verify the compliance of operational parameters

with the criteria established in the client-supplier agreement, to support decision-making regarding the acceptance of increased maintenance risk in pursuit of cost reduction through extended replacement intervals for critical components [4].

The results demonstrated that all analyzed items were within the agreed limits: the solids content ranged from 50.5 to 52.5 %, corresponding to a density between 1.451 and 1.478 g/cm³; the percentage of particles smaller than 10 microns remained at or below 26.3 %; silica and alumina levels showed controlled variations of ± 0.5 % and ± 1.0 %, respectively, compared to the monthly forecast; daily production fluctuated within a ± 5.0 % range of the projected value; mass balance deviation remained within ± 3.0 %; the passing fraction through mesh #325 ranged from 40 to 43 %, while mesh #32 showed performance above 98.7 %; and finally, turbidity stayed within the 50 mg/L limit.

Regarding operational data, both maintenance and operational targets were met, with physical availability exceeding 95 % and utilization above 99 %.

These results confirm that the process is under control and in compliance with the defined and agreed quality standards and parameters.

2.2 Risk Management

Given that the process has been demonstrated to be stable – both from an operational standpoint and in terms of statistical process control – it is reasonable and strategically advantageous to consider adopting more advanced maintenance practices aimed at optimizing costs and increasing operational efficiency. Operational stability indicates that the equipment is functioning within expected parameters, without significant fluctuations or recurring failures. At the same time, process stability, evidenced by quality indicators, repeatability, and statistical control, shows that the system as a whole is predictable and robust [5–7].

This scenario significantly reduces the risks associated with adopting maintenance approaches that aim for greater economic efficiency, such as predictive maintenance based on data analysis, extending intervals between scheduled interventions, or even introducing condition-based maintenance strategies. By reducing unnecessary interventions and focusing resources only where there are real indications of imminent failure, it is possible to minimize production downtime and labor and parts costs without compromising process reliability, as shown in Figure 3. Furthermore, this paradigm shift can free up technical capacity for higher-value activities, such as continuous improvement and operational innovation [5, 7].

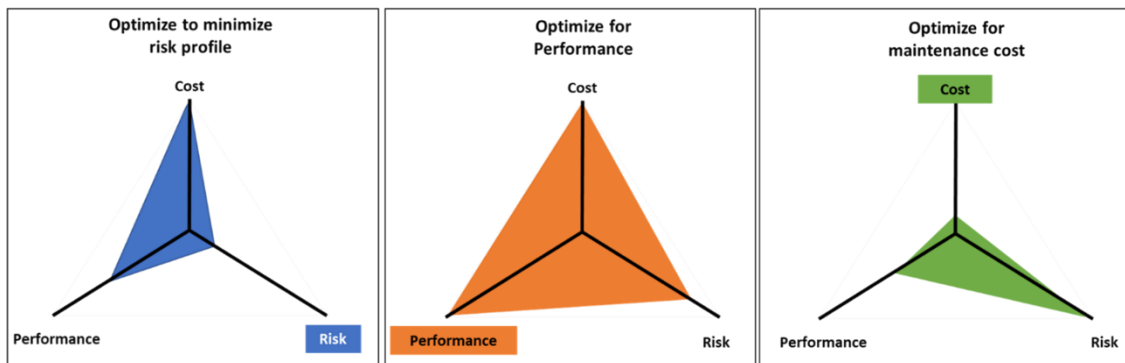


Figure 3. Maintenance Management

With the confirmation of the system's proper operational condition for the pumping system, we identified the opportunity to adopt new management strategies for critical components, accepting higher operational risk levels in order to reduce costs without significantly compromising system performance.

To enable this approach, classical reliability engineering techniques were applied, combined with modern methodologies based on reliability engineering modeling. This allowed for accurate simulation of asset behavior and failure prediction across various operational scenarios.

3. Classical Reliability Engineering Techniques

Reliability Engineering (RE) is the field of engineering that studies the ability of systems and components to function without failure over a specified period and under defined conditions. It combines statistical analysis, failure modeling, and maintenance strategies to increase the safety and efficiency of products and processes [8]. Its goal is to predict, prevent, and mitigate failures in order to optimize costs and ensure performance [8, 9].

Within RE, Life Data Analysis is a statistical technique used to model the time until failure of products or systems. It helps estimate reliability, service life, and failure rates using test or field data. The analysis typically utilizes probabilistic distributions such as Weibull, Lognormal, and Exponential. The results guide decisions regarding warranty, maintenance, and design improvement. It is essential in reliability engineering to predict the behavior of products over time [10, 11].

The methodological approach for Life Data Analysis begins with the definition of the item under analysis, as shown in Figure 4, which in this study refers to the piston sealing elements and the pump diaphragm. After this identification, the objective of the analysis is defined. For this study, it consists of characterizing the service life behavior of the components in question, aiming to support the definition of the most appropriate maintenance strategy based on empirical performance and failure data, and not solely on the manufacturer's recommendation [8, 11].

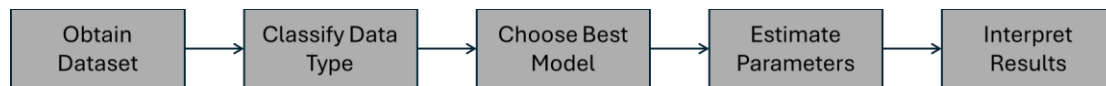


Figure 4. Life Data Analysis (LDA) Flowchart

As described in Table 1, the Mean time between failure and Mean time to Repair Time data - used as the basis for reliability studies - were obtained through the PI System and further processed and analyzed by the maintenance engineering team. The characteristics of the downtime was categorized into complete data and right-censored data. No grouped data, interval data, or left-censored data were recorded, as shown in Figure 5.

Since the failure data has been processed, the next step is to conduct a goodness-of-fit test for the data against a set of reliability distributions. To do so, the Weibull++ software was used to select the models that best fit the dataset. Four goodness-of-fit tests were performed: Kolmogorov-Smirnov, Correlation Coefficient, Likelihood Value, along with a visual-graphical analysis of the distributions' fit to the proposed curves. To estimate the parameters of the distributions, the Linear Regression method and Maximum Likelihood Estimator were used. All methods allow for the use of complete data, and both left and right-censored data.

Confidence limits are widely used in life data analysis (or survival analysis) to estimate the uncertainty associated with parameters such as failure rate, reliability function, and mean time to failure. They provide an interval within which the true value of the parameter is expected to lie, with a specified confidence level – typically bilateral 90 % in the mining industry.

In practice, these limits help make safer decisions regarding the reliability of products or systems over time, especially when the data is based on samples rather than complete populations. In our

study, we adopted the Fischer Matrix and Beta Binomial methods to estimate the confidence limits.

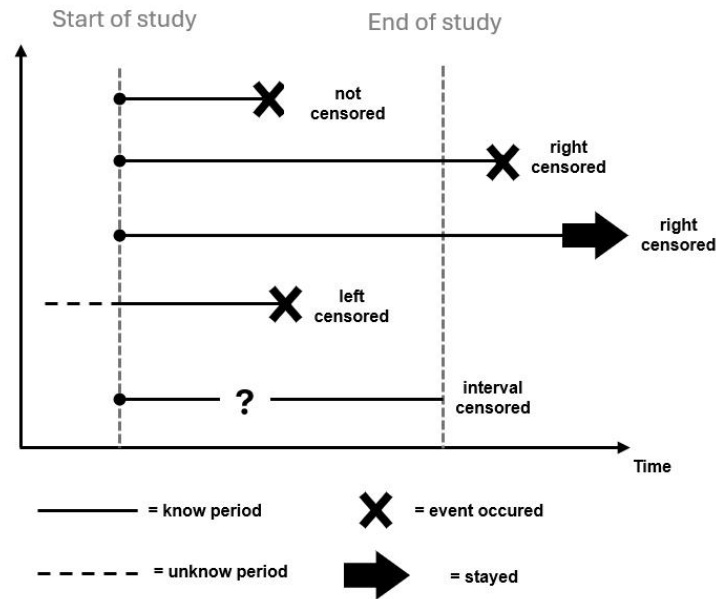


Figure 5. Failure Data Types

Thus, it is necessary to implement life cycle control for these critical components by monitoring their installation dates and predicted replacement schedules, ensuring that the replacement occurs before the component reaches its maximum service life. Initially, the component’s run time was tracked using calendar time, considering 24 hours of operation per day, or 720 hours per month. However, the equipment is subject to both scheduled and unscheduled downtimes, meaning that 100 % utilization is not achieved.

Nevertheless, with a clear definition of the components’ criticality and the possibility of changing the strategy - considering the optimization of component life, which allows for cost reduction, increased pump availability, fewer preventive shutdowns, and consequently greater maintenance team availability - it becomes possible to analyze the life cycle of these components. This approach optimizes the lifespan of components that have been subjected to extended operating times. Figures 6 and 7 present the reliability of these components considering an operational time greater than 8 000 hours.

4. Discussion of the Results

Based on the data obtained, it was possible to perform a goodness-of-fit test using the Weibull++ software, the obtained results are presented in Tables 2, 3, 4 and 5. Weibull++ uses the statistical methods Likelihood Value (LKV), Bayesian Information Criterion (BIC), and Akaike Information Criterion (AIC), comparing them and presenting the most appropriate distributions.

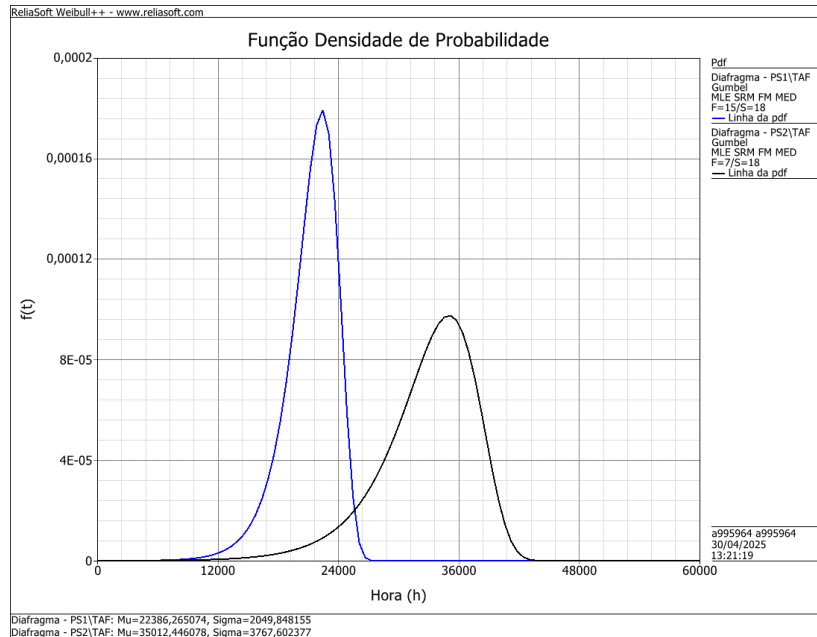


Figure 6. Probability Density Function of Failure Data for PS1 and PS2 Diaphragms

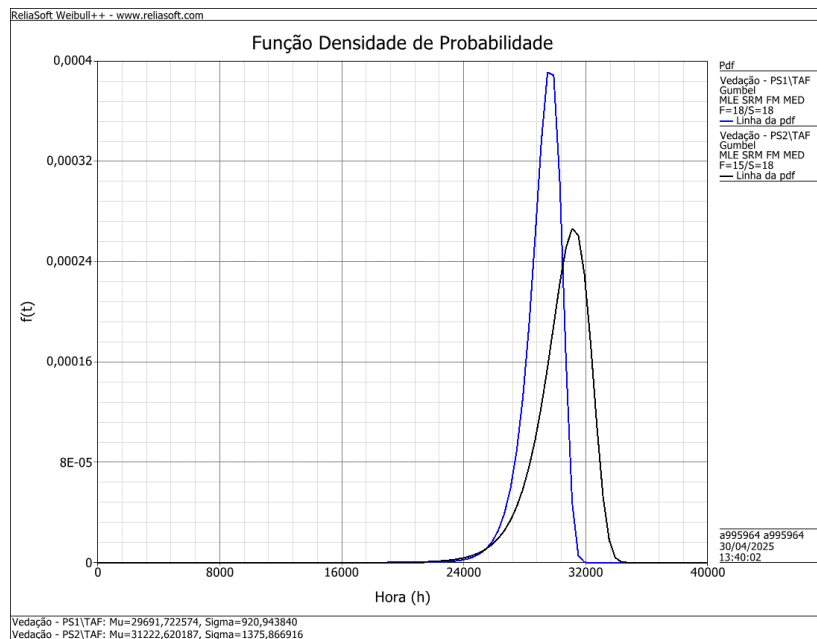


Figure 7. Probability Density Function of Failure Data for PS1 and PS2 Seals

Table 2. Goodness-of-Fit Test Results for PS1 Diaphragm

Distribution	Ranking	LKV	BIC	AIC
Gamma-G	1	-140.4	288.92	286.8
Gumbel	2	-142.55	290.53	289.11
Logistic	3	-145.95	297.32	295.9
Weibull-2P	4	-146.92	299.25	297.84
Normal	5	-149.25	303.91	302.49
Loglogistic	6	-150.93	307.28	305.87

Gamma	7	-153.6	312.62	311.2
Lognormal	8	-156.07	317.55	316.13
Exponential-2P	9	-164.16	333.73	332.32
Exponential-1P	10	-170.35	343.4	342.69

Although the Generalized Gamma distribution yielded the highest result, upon evaluating the graphical representation of the distribution, it was identified that the Gumbel distribution also provided a well-fitted model for the failure data, while also aligning within the established 90 % confidence limits. This was not the case with the Generalized Gamma distribution.

Table 3. Adhesion Test Results of PS1 Seal

Distribution	Ranking	LKV	BIC	AIC
Gamma-G	1	-145.67	300	297.33
Gumbel	2	-152.68	311.14	309.36
Normal	3	-155.07	315.92	314.14
Gamma	4	-155.27	316.32	314.54
Lognormal	5	-155.37	316.52	314.74
Logistic	6	-155.82	317.42	315.64
Loglogistic	7	-156.09	317.97	316.19
Exponential-1P	8	-208.22	419.33	418.44

Similarly to the PS1 diaphragm, a Gumbel distribution was selected instead of a Generalized Gamma distribution due to the better graphical fit to the confidence limits.

Table 4. Adhesion Test Results of PS2 Diaphragm

Distribution	Ranking	LKV	BIC	AIC
Gumbel	1	-75.18	154.26	154.37
Gamma-G	2	-75.2	156.23	156.39
Logistic	3	-76.31	156.51	156.61
Weibull-2P	4	-76.98	157.85	157.95
Normal	5	-77.52	158.94	159.05
Loglogistic	6	-78.18	160.25	160.36
Gamma	7	-79.02	161,94	162.05
Lognormal	8	-79.84	163.57	163.67
Exponential-2P	9	-81.96	167.8	167.91
Exponential-1P	10	-86.04	174.03	174.08

For the failure analysis of the PS2 diaphragms, the highest-ranked distribution, Gumbel, was chosen because it represented the failure data both mathematically and graphically in the best way.

Finally, the choice of the Generalized Gamma distribution was made for the PS2 seals, as it was the distribution that best fitted the confidence limits of the evaluated data. The results of the reliability distributions for the diaphragms PS1 and PS2 are presented in Figure 6 and the reliability distributions for the PS1 and PS2 seal are presented in Figure 7.

Table 5. Adhesion Test Results of PS2 Seal

Distribution	Ranking	LKV	BIC	AIC
Gamma-G	1	-128.16	264.45	262.32
Gumbel	2	-133.18	271.77	270.35
Logistic	3	-135.63	276.68	275.26
Normal	4	-135.73	276.88	275.47
Gamma	5	-136.11	277.64	276.22
Loglogistic	6	-136.11	277.64	276.22
Lognormal	7	-136.31	278.04	276.62
Exponential-1P	8	-171.88	346.48	345.77

The data shows that in both cases, the frequency distributions of failures extend beyond the 8 000 hours defined by the manufacturer as the recommended maintenance interval, which allows for changes in the maintenance strategies. Thus, the current maintenance strategy for the PS1 diaphragms is based on the preventive replacement of components upon reaching a previous downtime of 24 000 hours or conducting inspections every 8 000 hours to identify any deviations in the assembly. This approach results in a 3x increase in component performance, a 53 % reduction in costs, and increased availability without impacting the process or operation. As for the PS2 diaphragm, the maintenance strategy is based on extending its operational life to a maximum of 35 000 hours, compared to the 8 000 hours proposed by the manufacturer.

The maintenance strategy for the seals was based on the ability of the items to exceed the 24 000 hours indicated by the manufacturer. Currently, the new strategy is based on replacing the seals in both pumping stations after up to 32 000 hours, resulting in a 38.6 % cost reduction and an extension of the material’s lifespan by approximately 6 000 hours.

Figure 8 shows the current dashboard used to monitor the diaphragms and seals, displaying the dates of upcoming replacements and visually identifying the current status of each item.

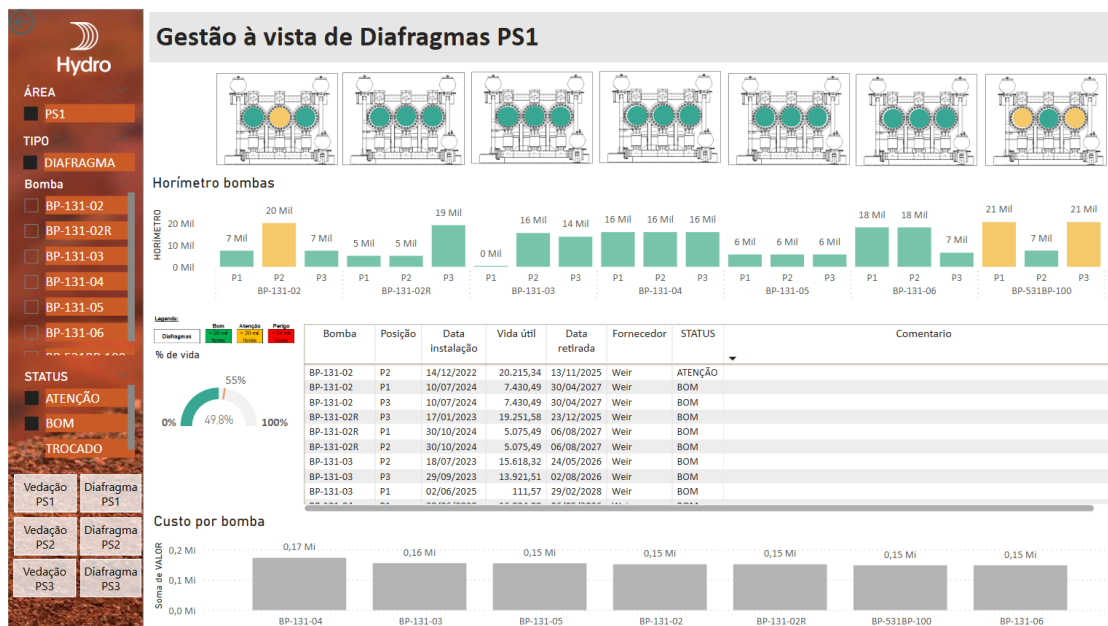


Figure 8. Lifespan control dashboard.

5. Conclusion

Based on the reliability analysis conducted, with data collected over the years, the component's lifespan, which was previously 8 000 hours, has exceeded the expected life by three times without compromising the operational performance of the pump.

The quality and reliability of a structured database are essential for defining operational limits and for the development of effective decision-making strategies, directly impacting the reduction of operational costs and increasing asset availability.

Among the various monitored elements, some critical components – as defined by the manufacturers – have a predefined lifespan, requiring periodic replacement to ensure continuous operation and system reliability. To address this, a critical component lifespan control dashboard was implemented, allowing the technical team to monitor the lifecycle of these items in real-time. Thus, whenever the system indicates that a particular component has reached its operational limit, the replacement is promptly scheduled and executed, preventing unexpected failures and optimizing predictive maintenance.

With the change in the data collection method, it is possible to identify a gain in the lifespan of the components, considering that preventive and corrective pump stops occur over time.

6. References

1. Operation manual of the positive displacement pump GEHO TZPM 2000, Weir Group. <https://www.global.weir/pt-BR/brands/geho-positive-displacement-pumps/>
2. Wulan et al., (2024), Impact of ERP System Implementation on Operational and Financial Efficiency in Manufacturing Industry, *Journal of Economic Education and Entrepreneurship Studies*, 5 (3), 491-501.
3. F. Ghaiti and M. Laghrissi, Implementation of the management system maintenance in SAP, *Proceedings of 2013 International Conference on Industrial Engineering and Systems Management (IESM)*, Agdal, Morocco, 2013, 1-6.
4. B. Balamurugan et al., (2022), *Big Data Analytics for Smart Manufacturing*, CRC Press.
5. D. Thomas and B. Weiss, (2021), Maintenance Costs and Advanced Maintenance Techniques: Survey and Analysis, *International Journal of Prognostics and Health Management*, Vol. 12 No. 1 (2021).
6. John D. Campbell and Andrew K.S. Jardine, *Maintenance Excellence: Optimizing Equipment Life-Cycle Decisions*, CRC Press, 2001.
7. H. Dan et al., Economic optimal model integrating statistical process control and preventive maintenance, *IEEE International Conference of Online Analysis and Computing Science (ICOACS)*, Chongqing, China, 2016, 218-224, DOI: 10.1109/ICOACS.2016.7563083.
8. A. Birolini, *Reliability Engineering: Theory and Practice*, Springer, 2017.
9. J. Pulido, Life data analysis using the competing failure modes technique, *Annual Reliability and Maintainability Symposium (RAMS)*, Palm Harbor, FL, USA, 2015, 1-6, DOI: 10.1109/RAMS.2015.7105142.
10. M. Xu and H. Mao, q-Weibull Distributions: Perspectives and Applications in Reliability Engineering, *IEEE Transactions on Reliability*, (2024): 1-14, DOI: 10.1109/TR.2024.344828.
11. William Q. Meeker and Luis A. Escobar, *Statistical Methods for Reliability Data*, Wiley, 2014, 712 pages.