UNIVERSIDAD NACIONAL JORGE BASADRE GROHMANN

Facultad de Ingeniería

Escuela Profesional de Ingeniería Mecánica

DESARROLLO DE UN MODELO PREDICTIVO PARA ESTIMAR EL TIEMPO DE VIDA ÚTIL REMANENTE EN BOMBAS CENTRÍFUGAS MEDIANTE ANÁLISIS DE COMPONENTES PRINCIPALES

TESIS

Presentada por:

Bach. Luis Carlos Martin Aragon Salas

Para optar el Título Profesional de:

INGENIERO MECÁNICO

TACNA-PERÚ

2026

UNIVERSIDAD NACIONAL JORGE BASADRE GROHMANN

Facultad de Ingeniería

Escuela Profesional de Ingeniería Mecánica

DESARROLLO DE UN MODELO PREDICTIVO PARA ESTIMAR EL TIEMPO DE VIDA ÚTIL REMANENTE EN BOMBAS CENTRÍFUGAS MEDIANTE ANÁLISIS DE COMPONENTES PRINCIPALES

Tesis sustentada, y aprobada el 08 de mayo del 2026, siendo el jurado calificador:

PRESIDENTE: ----------
Mtro. Reynaldo Clemente Telles Ríos

SECRETARIO: ----------
Mgr. Wilson García Mamani

VOCAL: ----------
Dr. Williams Sergio Almanza Quispe

ASESOR: ----------
Dr. Williams Sergio Almanza Quispe

CERTIFICADO DE SIMILITUD

Yo, **Williams Sergio Almanza Quispe**, Asesor de Tesis declara que la Tesis titulada: **"DESARROLLO DE UN MODELO PREDICTIVO PARA ESTIMAR EL TIEMPO DE VIDA ÚTIL REMANENTE EN BOMBAS CENTRÍFUGAS MEDIANTE ANÁLISIS DE COMPONENTES PRINCIPALES",** presentada por el Bachiller **Luis Carlos Martin Aragon Salas,** para optar el título profesional de Ingeniero Mecánico.

Habiendo cumplido con lo establecido en el reglamento de originalidad y de similitud de trabajo de investigación y producción, intelectual, considerando que según la revisión, evaluación y análisis realizado a través del software de similitud textual – Turnitin – cuenta con el nivel de similitud permitido cuyo porcentaje es 7 %. Por lo que, **CERTIFICO LA SIMILARIDAD** de la Tesis está de acuerdo al nivel **PERMITIDO,** para continuar con los trámites correspondientes y para su publicación en el repositorio institucional.

Se emite el presente certificado con fines de continuar con los trámites respectivos para su obtención del Título Profesional de Ingeniero Mecánico.

Tacna, 12 junio de 2026



.....
Dr. Williams Sergio Almanza Quispe
Asesor de Tesis
DNI 44369780
ORCID: 0000-0003-0812-7834



.....
Luis Carlos Martin Aragon Salas
DNI. 72487663

DEDICATORIA

A la memoria de mis amados abuelos, quienes, aunque ya no me acompañan físicamente, dejaron en mí una huella imborrable de amor, sabiduría y esfuerzo. Sus valores, enseñanzas y ejemplo de vida han sido faros que han guiado mi camino y han inspirado cada uno de mis logros. Su recuerdo permanece vivo en mi corazón y en cada paso que doy hacia mis metas.

A mi madre, por su amor incondicional, su paciencia infinita y su constante sacrificio. Su fortaleza, entrega y apoyo han sido la base sobre la cual he construido este logro académico. Gracias por enseñarme a perseverar ante las dificultades y a creer en el poder del trabajo y la dedicación.

Esta tesis es un testimonio del esfuerzo compartido, del legado de quienes me formaron con su ejemplo y del amor que trasciende el tiempo y la distancia. A ellos dedico este logro, con profunda gratitud y respeto.

AGRADECIMIENTOS

Deseo expresar mi más sincero agradecimiento a todas las personas e instituciones que hicieron posible la culminación de este trabajo de investigación.

En primer lugar, a mi familia, por su apoyo incondicional, comprensión y motivación constante durante todo el proceso de formación profesional. Su confianza y aliento fueron el motor que me impulsó a seguir adelante incluso en los momentos más difíciles.

A mi asesor, por su valiosa orientación, dedicación y paciencia, cuyas observaciones y recomendaciones fueron fundamentales para el desarrollo y la calidad de este trabajo.

A la Escuela Profesional de Ingeniería Mecánica y a la Universidad Nacional Jorge Basadre Grohmann, por brindarme la oportunidad de formarme académica y profesionalmente, así como por los recursos y el acompañamiento que hicieron posible mi desarrollo como ingeniero.

A mis compañeros y amigos, con quienes compartí años de aprendizaje, esfuerzo y experiencias inolvidables. Su compañerismo, colaboración y espíritu de equipo fueron esenciales para superar los retos académicos y personales de esta etapa.

A todos ellos, mi más profundo agradecimiento y reconocimiento por haber sido parte de este logro que marca el cierre de una importante etapa de mi vida profesional.

ÍNDICE DE CONTENIDO

DEDICATORIA ii
AGRADECIMIENTOS iii
ÍNDICE DE CONTENIDO iv
ÍNDICE DE TABLAS vii
ÍNDICE DE FIGURAS viii
RESUMEN x
ABSTRACT xi
INTRODUCCIÓN xii
CAPÍTULO I PLANTEAMIENTO DEL PROBLEMA 13
1.1. Realidad problemática 13
1.2. Formulación del problema 14
1.3. Objetivos 15
1.3.1. Objetivo General 15
1.3.2. Objetivos Específicos 15
1.4. Justificación e importancia de la investigación 15
1.4.1. Fundamento que motiva la investigación 15
1.4.2. Importancia del tema de investigación 16
1.5. Limitaciones y viabilidad 17
1.5.1. Limitaciones 17
1.5.2. Viabilidad del estudio 18
1.6. Formulación de hipótesis 19
1.6.1. Hipótesis principal 19
1.6.2. Hipótesis específicas 19
1.7. Variables 20
1.7.1. Definición conceptual y operacional de las variables 20
1.7.2. Operacionalización de las variables 20

CAPÍTULO II Marco Teórico 22
2.1. Antecedentes del trabajo de investigación 22
2.1.1. Antecedentes internacionales 22
2.1.2. Antecedentes nacionales 24
2.2. Bases teóricas sobre el trabajo de investigación 26
2.2.1. Teoría de Bombas Centrífugas 26
2.2.2. Vida útil de las bombas centrífugas 34
2.2.3. Indicador de condición o salud 35
2.2.4. Teoría de Mantenimiento 36
2.2.5. Mantenimiento Predictivo en Bombas Centrifugas 38
2.2.6. Análisis de componentes principales (PCA) 41
2.2.7. Modelos predictivos 43
2.2.7.2. Análisis de Componentes Principales 50
CAPÍTULO III Marco metodológico 52
3.1. Tipo de investigación 52
3.2. Nivel de investigación 52
3.3. Diseño de investigación 52
3.4. Método de Investigación 53
3.5. Población y muestra 54
3.5.1. Población 54
3.5.2. Muestra 54
3.6. Técnicas e instrumentos de recolección de datos 54
3.6.1. Técnicas e instrumentos de recolección y análisis de datos para el primer objetivo específico 55
3.6.2. Técnicas e instrumentos de recolección y análisis de datos para el segundo objetivo específico 56

3.6.3. Técnicas e instrumentos de recolección y análisis de datos para el tercer objetivo específico 58
3.7. Análisis estadístico 60
3.7.1. Análisis estadístico multivariado con PCA 60
3.7.2. Validación estadística del modelo predictivo 62
3.8. Interpretación y relevancia de los resultados del modelo predictivo 63
3.9. Consideraciones del modelo predictivo 64
CAPÍTULO IV RESULTADOS Y DISCUSIÓN 66
4.1. Análisis del comportamiento operativo de bombas centrífugas 66
4.2. Desarrollo del modelo predictivo 77
4.3. Métricas de modelo predictivo 88
4.4. Contraste de hipótesis 96
4.5. Discusión 99
CONCLUSIONES 103
RECOMENDACIONES 104
Bibliografía 106
ANEXOS 110

ÍNDICE DE TABLAS

Tabla 1 Cuadro de operacionalización de variables 20
Tabla 3 Catálogo técnico de bombas centrífugas instaladas.....67
Tabla 3 Métricas de desempeño del modelo PCA–RNA.....93
Tabla 4 Importancia de los componentes principales (ΔMAE).....95

ÍNDICE DE FIGURAS

Figura 1 Bomba Centrífuga 29
Figura 2 Bomba Centrífuga con Difusor y Caja en Espiral 31
Figura 3 Componentes de una bomba centrífuga 32
Figura 4 Bombas Centrifugas por tipo de impulsor 33
Figura 5 Varianza PCA 68
Figura 6 Comportamiento temporal de la vibración RMS en la bomba B01 69
Figura 7 Gradiente térmico entre el rodamiento y el ambiente en la bomba B01 70
Figura 8 Comportamiento temporal del caudal normalizado en la bomba B01 70
Figura 9 Variación temporal de la presión normalizada en la bomba B01 71
Figura 10 Comportamiento temporal de la vibración RMS en la bomba B02 72
Figura 11 Gradiente térmico entre el rodamiento y el ambiente en la bomba B02 72
Figura 12 Comportamiento temporal del caudal normalizado en la bomba B02 73
Figura 13 Variación temporal de la presión normalizada en la bomba B02 74
Figura 14 Comportamiento temporal de la vibración RMS en la bomba B03 75
Figura 15 Gradiente térmico entre el rodamiento y el ambiente en la bomba B03 75
Figura 16 Comportamiento temporal del caudal normalizado en la bomba B03 76
Figura 17 Variación temporal de la presión normalizada en la bomba B03 77
Figura 18 Diagrama de flujo del proceso modelo predictivo 78
Figura 19 Distribución de variables operativas en el sistema de bombeo 79
Figura 20 Sedimentación (Scree Plot) del Análisis de Componentes Principales 81
Figura 21 Distribución y análisis de normalidad de los residuales en el conjunto de prueba 82
Figura 22 Intervalo de predicción P10–P90 del RUL estimado en el conjunto de prueba 83
Figura 23 Curva de probabilidad acumulada de falla del sistema de bombeo (escenario no ideal) 84
Figura 24 Curva de probabilidad de falla del sistema de bombeo – Escenario optimista 85
Figura 25 Curva de probabilidad acumulada de falla del sistema de bombeo – Escenario realista 86
Figura 26 Curva de probabilidad de falla del sistema de bombeo – Escenario pesimista 87

Figura 27 Correlación entre el RUL real y el RUL predicho en el conjunto de entrenamiento.....90
Figura 28 Correlación entre el RUL real y el RUL predicho en la fase de validación..91
Figura 29 Correlación entre el RUL real y el RUL predicho en el conjunto de prueba 92

RESUMEN

El estudio tuvo como propósito desarrollar un modelo predictivo híbrido basado en el Análisis de Componentes Principales (PCA) y Redes Neuronales Artificiales (RNA) para estimar el tiempo de vida útil remanente (RUL) de bombas centrífugas empleadas en entornos industriales. El enfoque metodológico comprendió tres etapas: análisis del comportamiento operativo, diseño del modelo predictivo y evaluación del desempeño mediante métricas de precisión. En la fase de análisis, se examinaron las variables vibración RMS, caudal, presión y gradiente térmico, evidenciando estabilidad general y tendencias que reflejaron los procesos de degradación progresiva. A través del PCA, se logró reducir la dimensionalidad del conjunto de datos, concentrando más del 90 % de la varianza total en cuatro componentes principales, de los cuales el PC3 resultó ser el más influyente en la predicción del RUL, al relacionarse con factores térmicos y vibracionales. El modelo fue implementado mediante una red neuronal tipo perceptrón multicapa (MLP), entrenada con el algoritmo Adam y función de activación ReLU. En la fase de entrenamiento, se obtuvo un $R^2 = 0.9213$, con MAE = 16.93 h y RMSE = 20.75 h, evidenciando un adecuado ajuste a los datos históricos. En la validación, el desempeño disminuyó ($R^2 = 0.1483, MAE = 27.69\ h, RMSE = 33.18\ h$), reflejando sensibilidad a la variabilidad operativa. En el conjunto de prueba, el modelo alcanzó un desempeño intermedio ($R^2 = 0.5203, MAE = 17.45\ h, RMSE = 22.02\ h$), logrando capturar la tendencia general del RUL frente a datos no vistos. Los resultados demostraron que el modelo PCA–RNA es viable para estimaciones iniciales del RUL, permitiendo anticipar la degradación de las bombas y apoyar la planificación del mantenimiento predictivo. Su aplicación contribuye a optimizar tiempos de intervención, reducir costos operativos y aumentar la confiabilidad en sistemas industriales.

Palabras Clave: Análisis de Componentes Principales, Bombas Centrífugas, Mantenimiento Predictivo, Redes Neuronales Artificiales

ABSTRACT

The purpose of the study was to develop a hybrid predictive model based on Principal Component Analysis (PCA) and Artificial Neural Networks (ANN) to estimate the remaining useful life (RUL) of centrifugal pumps used in industrial environments. The methodological approach comprised three stages: analysis of operational behavior, design of the predictive model, and performance evaluation using accuracy metrics. In the analysis phase, the variables RMS vibration, flow rate, pressure, and thermal gradient were examined, revealing overall stability and trends that reflected progressive degradation processes. Through PCA, the dimensionality of the dataset was reduced, concentrating more than 90% of the total variance in four principal components, of which PC3 proved to be the most influential in predicting RUL, as it was related to thermal and vibrational factors. The model was implemented using a multilayer perceptron (MLP) neural network, trained with the Adam algorithm and ReLU activation function. In the training phase, an $R^2 = 0.9213$ was obtained, with MAE = 16.93 h and RMSE = 20.75 h, demonstrating an adequate fit to the historical data. In the validation, performance decreased ($R^2 = 0.1483$, MAE = 27.69 h, RMSE = 33.18 h), reflecting sensitivity to operational variability. In the test set, the model achieved intermediate performance ($R^2 = 0.5203$, MAE = 17.45 h, RMSE = 22.02 h), capturing the general trend of the RUL against unseen data. The results demonstrated that the PCA–RNA model is viable for initial RUL estimates, allowing for the anticipation of pump degradation and supporting predictive maintenance planning. Its application contributes to optimizing intervention times, reducing operating costs, and increasing reliability in industrial systems.

Keywords: Principal Component Analysis, Centrifugal Pumps, Predictive Maintenance, Artificial Neural Networks

INTRODUCCIÓN

El mantenimiento predictivo ha adquirido una relevancia creciente en la industria moderna debido a su capacidad para anticipar fallas y optimizar la gestión de activos críticos. En este contexto, las bombas centrífugas constituyen uno de los equipos más utilizados en procesos industriales, donde su desempeño continuo resulta esencial para garantizar la eficiencia operativa y la seguridad de los sistemas industriales. La estimación del tiempo de vida útil remanente (RUL, por sus siglas en inglés) permite identificar con anticipación el estado de degradación de los equipos, facilitando la toma de decisiones estratégicas y la reducción de costos derivados de paradas no programadas.

En el presente estudio se propone el desarrollo de un modelo predictivo híbrido que combina el Análisis de Componentes Principales (PCA) y las Redes Neuronales Artificiales (RNA) para estimar el RUL de bombas centrífugas. El PCA permite reducir la dimensionalidad de los datos operativos, concentrando la información relevante y eliminando redundancias, mientras que la RNA modela las relaciones no lineales entre las variables dinámicas del sistema, como vibración RMS, presión, caudal y gradiente térmico. Esta integración busca mejorar la capacidad de predicción del deterioro y ofrecer una herramienta confiable para la planificación del mantenimiento.

La investigación se estructura en tres etapas principales. En la primera, se realiza el análisis del comportamiento operativo de tres bombas centrífugas sometidas a mantenimiento preventivo, identificando sus patrones de estabilidad y posibles indicadores de degradación. En la segunda etapa, se desarrolla el modelo PCA–RNA, aplicando técnicas de reducción de dimensionalidad y entrenamiento de red neuronal con optimización Adam y función de activación ReLU, validando su rendimiento mediante escenarios de simulación (optimista, realista y pesimista). Finalmente, se evalúa su desempeño mediante métricas de error y precisión ($MAE, RMSE, MAPE, sMAPE\ y\ R^2$) para determinar su grado de ajuste y capacidad de generalización.

CAPÍTULO I

PLANTEAMIENTO DEL PROBLEMA

1.1. Realidad problemática

Las bombas centrífugas constituyen aproximadamente el 62 % del mercado global de bombas industriales, el cual fue valuado en USD 39,8 mil millones en 2024 y se proyecta que alcance los USD 50,5 mil millones para 2030 (Grand View Research, 2024). En el sector minero, su operación continua resulta esencial para mantener la eficiencia de los procesos; una falla inesperada puede generar interrupciones en la producción, pérdidas económicas y costos elevados de reparación. De hecho, se estima que los paros inesperados pueden representar hasta el 70 % del costo del ciclo de vida total de una bomba industrial (LEWA, 2025).

La experiencia en minas del sur del Perú confirma esta problemática: la implementación de bombas con sensores y mantenimiento predictivo permitió un incremento del 18 % en la operatividad, una reducción del 30 % en el uso de repuestos y la eliminación total de reparaciones de emergencia (Strategic Market Research, 2023). De manera similar, otras minas que adoptaron bombas inteligentes reportaron una disminución del 22 % en el tiempo de inactividad (downtime) y un ahorro energético anual del 15 % (Market Growth Reports, 2025). Estas cifras evidencian la relevancia de contar con estrategias de mantenimiento predictivo que permitan anticipar fallas y reducir interrupciones costosas.

No obstante, en muchas operaciones el mantenimiento aún se realiza mediante programas preventivos basados en cronogramas fijos o inspecciones periódicas, los cuales no siempre detectan fallas a tiempo. Esto puede generar intervenciones innecesarias o, en el peor de los casos, fallos inesperados que afectan la disponibilidad del equipo. Además, los métodos actuales de monitoreo, centrados en variables como presión, caudal y vibraciones, no siempre logran anticipar el desgaste de componentes críticos, como impulsores y sellos mecánicos, lo que dificulta la planificación del mantenimiento y el control de costos asociados (Llihua, 2024).

Para abordar esta situación, se plantea el desarrollo de un modelo predictivo basado en el Análisis de Componentes Principales o en inglés Principal Component Analysis (PCA), que permita identificar patrones de comportamiento en los datos operativos de las bombas. La aplicación de técnicas de aprendizaje automático facilitará la detección temprana de posibles fallas y permitirá optimizar los planos de mantenimiento en función del estado real del equipo. Esto contribuirá a reducir paradas inesperadas y mejorar la gestión de los recursos disponibles. El desarrollo de este modelo requiere recopilar y analizar datos relevantes, definir los parámetros clave de operación y seleccionar los algoritmos adecuados para generar predicciones confiables. La implementación de este sistema permitirá una gestión más eficiente del mantenimiento de las bombas centrífugas, asegurando su funcionamiento adecuado y prolongando su vida útil.

Diversos estudios recientes evidencian que el uso de modelos predictivos basados en análisis de datos y aprendizaje automático permite estimar de manera confiable la vida útil remanente (RUL) de bombas centrífugas y sus componentes críticos (Zhou et al., 2025). Investigaciones aplicadas han demostrado que la integración de sensores, técnicas de reducción de dimensionalidad y redes neuronales mejora la detección temprana de procesos de degradación y reduce la incertidumbre en la planificación del mantenimiento (Chen et al., 2022; Zhu et al., 2024). Estos enfoques superan las limitaciones de los métodos tradicionales al considerar simultáneamente múltiples variables operativas, respaldando la necesidad de desarrollar modelos predictivos orientados a optimizar la gestión del mantenimiento en sistemas de bombeo industriales.

1.2. Formulación del problema

¿Cómo puede un modelo predictivo basado en Análisis de Componentes Principales (PCA) estimar el tiempo de vida útil remanente en bombas centrífugas?

1.3. Objetivos

1.3.1. Objetivo General

Desarrollar un modelo predictivo basado en el Análisis de Componentes Principales (PCA) para estimar el tiempo de vida útil remanente de bombas centrífugas

1.3.2. Objetivos Específicos

a) Analizar el comportamiento operativo de bombas centrífugas sometidas a mantenimiento preventivo para estimar el tiempo de vida útil remanente.
b) Desarrollar un modelo predictivo basado en Análisis de Componentes Principales (PCA) y redes neuronales para estimar el tiempo de vida útil remanente.
c) Aplicar y evaluar el modelo predictivo mediante métricas de precisión y error.

1.4. Justificación e importancia de la investigación

1.4.1. Fundamento que motiva la investigación

Las bombas centrífugas son equipos utilizados en numerosos procesos industriales, incluyendo sectores como la minería, la industria petroquímica y el tratamiento de aguas. Su correcta operación es esencial para garantizar la eficiencia y continuidad de los procesos productivos. Sin embargo, la falta de métodos precisos para la estimación de su vida útil remanente representa un desafío, ya que las fallas inesperadas pueden generar costos elevados de reparación, pérdida de producción y riesgos operativos. El mantenimiento tradicional, basado en revisiones periódicas, no siempre es suficiente para prevenir fallas, ya que no toma en cuenta el estado real de los componentes de la bomba. En este contexto, el desarrollo de un modelo predictivo utilizando técnicas de Análisis de Componentes Principales (PCA) y redes neuronales ofrece una solución que permite analizar datos operativos históricos para predecir con mayor precisión el desgaste de los componentes críticos. Este enfoque no solo mejorará la gestión del mantenimiento y optimizará la disponibilidad de los equipos, sino que también reducirá el impacto ambiental de manera significativa. Al disminuir las intervenciones innecesarias, se evitará el uso excesivo de lubricantes, repuestos y otros insumos

industriales que, en muchos casos, generan desechos peligrosos. Asimismo, prolongar la vida útil de los componentes permitirá reducir la demanda de fabricación y transporte de piezas de reemplazo, actividades que suelen estar asociadas a emisiones de gases de efecto invernadero. Por otro lado, al minimizar las fallas no programadas, se evitarán paradas repentinas que pueden ocasionar derrames de fluidos y un consumo energético mayor durante los arranques de emergencia, contribuyendo así a procesos industriales más sostenibles y con menor huella de carbono.

1.4.2. Importancia del tema de investigación

La investigación es relevante porque propone una solución basada en inteligencia artificial para la estimación del tiempo de vida útil de las bombas centrífugas, un aspecto crítico en la industria. El desarrollo de un modelo predictivo basado en Análisis de Componentes Principales (PCA) y redes neuronales artificiales (RNA) representa un aporte significativo al campo del mantenimiento industrial, en particular al diagnóstico y gestión del estado de bombas centrífugas, ampliamente utilizadas en minería, petróleo, agua y manufactura.

Los beneficios concretos que se obtendrán con la aplicación de este modelo son los siguientes:

- **Anticipación de fallas críticas**: El modelo permite detectar patrones de deterioro antes de que ocurran fallas severas, mejorando la planificación del mantenimiento y evitando paradas inesperadas.
- **Optimización de costos**: Al identificar el momento óptimo para intervenir una bomba, se evitan mantenimientos innecesarios y costosas fallas correctivas, lo que genera un uso más eficiente de los recursos técnicos, logísticos y económicos.
- **Mejora de la confiabilidad y disponibilidad operativa**: La estimación del tiempo de vida útil remanente o en inglés Remaining Useful Life (RUL) facilita la continuidad de los procesos, disminuye los riesgos de paradas no programadas y mejora los indicadores de disponibilidad de equipos.
- **Aplicación flexible y escalable**: El modelo desarrollado en un entorno de simulación puede adaptarse posteriormente a entornos industriales reales, siempre que se cuente con sensores adecuados para la recopilación de datos.

- **Soporte a la toma de decisiones basada en datos**: La integración de algoritmos predictivos con análisis estadístico permite contar con herramientas objetivas para el mantenimiento predictivo, alineadas a la transformación digital de la industria.
- **Transferencia tecnológica y académica**: La investigación sienta las bases para la implementación de modelos similares en otros equipos rotativos y procesos industriales, además de servir como caso de estudio para futuras investigaciones en inteligencia artificial aplicada al mantenimiento.

Además, este estudio contribuye al avance del mantenimiento predictivo en el sector industrial, promoviendo el uso de herramientas analíticas que facilitan la toma de decisiones basadas en datos. Esto no solo aumenta la eficiencia operativa, sino que también prolonga la vida útil de los equipos, garantizando un funcionamiento más seguro y confiable en entornos de alta exigencia.

1.5. Limitaciones y viabilidad

1.5.1. Limitaciones

Se presenta las siguientes limitaciones:

- **Variabilidad en las condiciones operativas:** Las bombas centrífugas pueden operar en una amplia gama de entornos industriales (minería, petroquímica, tratamiento de aguas) y bajo distintas condiciones de carga, temperatura, presión y tipo de fluido. Esta diversidad puede dificultar la generalización del modelo predictivo propuesto, ya que ha sido entrenado con datos específicos de ciertos entornos, lo que podría limitar su aplicabilidad en otros contextos operativos.

- **Disponibilidad y calidad de los datos históricos**: La eficacia del modelo predictivo depende en gran medida de la calidad y cantidad de los datos operativos utilizados durante su entrenamiento. Sin embargo, en algunos casos, los registros históricos pueden ser incompletos, inconsistentes o no estar disponibles, lo que podría afectar la precisión del modelo.

- **Dependencia de los rangos operativos de entrenamiento:** El modelo predictivo presenta una capacidad de generalización condicionada a que las bombas centrífugas analizadas operen dentro de rangos de caudal, presión, temperatura y vibración similares a los considerados durante el entrenamiento. La aplicación del modelo en equipos con configuraciones o condiciones operativas distintas puede requerir procesos de recalibración o reentrenamiento para mantener niveles adecuados de precisión.

1.5.2. Viabilidad del estudio

1.5.2.1. Viabilidad técnica

El estudio es técnicamente viable, ya que se cuenta con herramientas y metodologías comprobadas para el desarrollo de modelos predictivos basados en Análisis de Componentes Principales (PCA) y redes neuronales. Se dispone de software especializado como MATLAB, que permite la recopilación, procesamiento y análisis de datos operativos de las bombas centrífugas. No obstante, es importante considerar que la calidad y disponibilidad de los datos históricos puede variar entre distintos entornos industriales, lo cual podría requerir una etapa previa de validación o depuración de datos para garantizar un entrenamiento adecuado del modelo. Asimismo, aunque el modelo se entrena en condiciones específicas, su desarrollo establece una base sólida para futuras adaptaciones a otros escenarios operativos.

1.5.2.2. Viabilidad económica

El desarrollo de este estudio es económicamente viable, ya que se basa en herramientas de software de código y recursos disponibles en entornos industriales para la recopilación de datos. La inversión principal se centrará en la adquisición de equipos de monitoreo, la implementación del modelo y la capacitación del personal para su uso e interpretación. A largo plazo, la aplicación del modelo predictivo permitirá optimizar los costos operativos al reducir el mantenimiento innecesario y prevenir fallas inesperadas.

1.5.2.3. Viabilidad Metodológica

La metodología planteada en esta investigación es viable, ya que se basa en el uso de técnicas comprobadas para el análisis de datos operativos y el desarrollo de modelos predictivos. En este caso, se empleará un modelo compuesto por el Análisis de Componentes Principales (PCA) y una Red Neuronal Artificial (RNA) de tipo perceptrón multicapa (MLP). Este modelo permite estimar el tiempo de vida útil remanente de bombas centrífugas a partir de datos históricos como caudal, presión, temperatura y vibración.

El PCA será utilizado para reducir la cantidad de variables, conservando la mayor parte de la información relevante. Esto facilita el entrenamiento del modelo y mejora su rendimiento. Posteriormente, la red neuronal procesará estos datos para realizar las predicciones, permitiendo analizar el comportamiento del equipo en función del desgaste. El proceso se desarrollará en un entorno de simulación utilizando el software MATLAB R2024b, lo que permitirá organizar, procesar y evaluar los datos de forma adecuada. La secuencia metodológica propuesta es coherente con los objetivos de la investigación y permite aplicar herramientas estadísticas y computacionales accesibles y conocidas.

1.6. Formulación de hipótesis

1.6.1. Hipótesis principal

El desarrollo de un modelo predictivo basado en el Análisis de Componentes Principales (PCA) permite estimar el tiempo de vida útil remanente de las bombas centrífugas con un nivel de precisión aceptable, a partir del análisis de variables operativas como vibración, caudal, presión y gradiente térmico, contribuyendo a la planificación del mantenimiento predictivo.

1.6.2. Hipótesis específicas

a) El análisis del comportamiento operativo de bombas centrífugas sometidas a mantenimiento preventivo permite identificar patrones de funcionamiento asociados al desgaste, que contribuyen a estimar su tiempo de vida útil remanente.
b) El desarrollo de un modelo predictivo basado en Análisis de Componentes Principales (PCA) y redes neuronales permite representar eficazmente los

patrones de deterioro, mejorando la estimación del tiempo de vida útil remanente de las bombas centrífugas.

c) La aplicación del modelo predictivo y su evaluación mediante métricas estadísticas permite validar su efectividad para predecir el tiempo de vida útil remanente en bombas centrífugas.

1.7. Variables

1.7.1. Definición conceptual y operacional de las variables

a) Variable Independiente: Condiciones operativas de las bombas centrífugas

Las condiciones operativas corresponden a los parámetros físicos medidos durante el funcionamiento de las bombas centrífugas, los cuales influyen directamente en su comportamiento y proceso de degradación.

b) Variable Dependiente: Estimación del tiempo de vida útil remanente de las bombas centrífugas

La estimación del tiempo de vida útil remanente se refiere a la predicción del período durante el cual una bomba centrífuga puede continuar operando de manera eficiente antes de que se presenten fallas críticas o de rendimiento. Esta estimación se basa en el análisis de datos operativos y de desgaste de componentes clave, permitiendo optimizar la planificación del mantenimiento y reducir los tiempos de inactividad no planificados.

1.7.2. Operacionalización de las variables

Tabla 1

Cuadro de operacionalización de variables

VARIABLES	DEFINICION OPERACIONAL	DIMENSIÓN	INDICADORES	UNIDAD DE MEDIDA
Condiciones operativas de las bombas centrífugas	Las condiciones operativas corresponden a los parámetros físicos medidos durante el funcionamiento de las bombas centrífugas, los cuales influyen directamente	Comportamiento operativo	**Parámetros operativos (caudal, presión, vibración, temperatura)**	Valores técnicos (bar, l/s, mm/s, °C)
			Frecuencia de fallas registradas	N° de fallas / periodo

	en su comportamiento y proceso de degradación.		**Intervalos entre mantenimientos preventivos**	Días u horas de operación
Estimación del tiempo de vida útil remanente de las bombas centrífugas	La estimación del tiempo de vida útil remanente de las bombas centrífugas se define como el proceso de cálculo basado en el análisis de datos operativos históricos, utilizando modelos predictivos para determinar el período restante de funcionamiento eficiente antes de la ocurrencia de fallas críticas.	Predicción del tiempo de vida útil	**Precisión de la predicción**	(%)
			Errores en la predicción	(horas)
			Confiabilidad del modelo	(%)
			Tendencia del desgaste	(por día)
		Desarrollo del modelo predictivo	**Análisis de Componentes Principales (PCA)**	(%)
			Implementación de redes neuronales	(cantidad)
			Integración de datos operativos históricos	(número de datos)
		Evaluación del modelo	**Error medio cuadrático (MSE)**	(días²)
			Coeficiente de determinación (R^2)	(rango de 0 a 1)

Nota. Elaboración Propia

CAPÍTULO II

MARCO TEÓRICO

2.1. Antecedentes del trabajo de investigación

2.1.1. Antecedentes internacionales

Chicaiza y González (2021) en su tesis “Sistema de Monitoreo en Nube Para el Diagnóstico Predictivo de Fallos de Motores Eléctricos”, tuvieron como propósito fue evaluar la condición de los motores eléctricos en una planta cementera ubicada en Guayaquil, utilizando información histórica para anticipar posibles fallas. La metodología aplicada se basó en el uso de aprendizaje automático supervisado, entrenando un modelo predictivo con datos previos de pruebas de mantenimiento. Se desarrolló un sistema de monitoreo basado en la nube, capaz de clasificar el estado de los motores y generar informes de mantenimiento. Los resultados alcanzaron una precisión del 94,5 % en la predicción del estado de los motores, lo que permitió optimizar la eficiencia operativa y disminuir los costos de mantenimiento.

Vargas (2021) en su trabajo fin de grado “Detección de anomalías en un motor eléctrico mediante inteligencia computacional”, tuvo como objetivo desarrollar una metodología para la identificación de fallos en un motor eléctrico trifásico de corriente alterna, utilizando técnicas de Inteligencia Computacional y Aprendizaje Automático. La metodología aplicada se fundamentó en la implementación de Redes Neuronales Artificiales (RNA), Sistemas Neurodifusos y un proceso de toma de decisiones basado en OWA-RIM. Los datos empleados como entrada abarcaron variables como la tensión, la corriente y el par motor, recopilados a través de diferentes sistemas de medición. Los hallazgos demostraron que el sistema de detección fue capaz de identificar con precisión tanto los modos de operación como los fallos del motor, contribuyendo a una detección temprana más eficiente de anomalías ya la optimización del mantenimiento preventivo en entornos industriales.

Alarcón (2022) en su tesis “Desarrollo e implementación de modelos de aprendizaje de máquinas para identificación de fallas en el monitoreo de la condición de activos en planta Celulosa Nueva Aldea”, planteó como objetivo diseñar e implementar

modelos de aprendizaje automático para detectar fallas. en el monitoreo de activos en la Planta de Celulosa Nueva Aldea. La metodología aplicada se centró en la selección de turbomáquinas y prensas de desplazamiento como objetos de estudio, empleando algoritmos de detección de anomalías y redes neuronales para anticipar eventos de alto torque. Los modelos desarrollados fueron validados y evaluados utilizando datos de monitoreo obtenidos de la plataforma Wedge. Los resultados indicaron que los modelos fueron capaces de identificar anomalías y prevenir eventos con cierto nivel de precisión; Sin embargo, se detectaron limitaciones relacionadas con la capacidad de generalización y la presencia de sobreajuste en algunos casos.

Acosta (2023) en su tesis *"Predicción de la vida útil remanente de la batería de un vehículo de movilidad sostenible"*, tuvieron como objetivo predecir la vida útil remanente (RUL) de baterías de ion-litio en bicicletas eléctricas mediante modelos de machine learning. Se aplicó la metodología CRISP-DM, que incluyó comprensión del problema, análisis y limpieza de datos, preparación de variables y construcción de modelos de regresión: Gradient Boost Regressor (GBR), AdaBoost Regressor (ABR) y Support Vector Regressor (SVR). Se realizaron búsquedas de hiperparámetros con Grid Search CV y validación cruzada para evaluar el desempeño mediante métricas como RMSE, MAE, MAPE y R^2. Los resultados mostraron que GBR obtuvo el mejor desempeño, con $R^2 = 0,99$ y RMSE = 6,15, demostrando alta precisión en la predicción del RUL.

Villacrés et al. (2024) en su artículo *"Predicción de vida útil remanente en rodamiento aplicando Machine Learning: Una revisión Sistemática de Literatura"*, tuvieron como objetivo realizar una revisión sistemática de literatura (SLR) para identificar los principales problemas en la predicción de la vida útil remanente (RUL) en rodamientos y los modelos de Machine Learning más utilizados. La metodología aplicada fue el protocolo PRISMA y Kitchenham, con búsqueda en bases de datos científicas (Web of Science, IEEE Xplore, ScienceDirect, entre otras) y criterios de inclusión y exclusión definidos. Se seleccionaron 35 artículos publicados entre 2018-2021. Los resultados mostraron tres problemas clave: la extracción de características, la identificación del inicio del estado de degradación y la falta de modelos generalizables.

Además, el 80 % de estudios utilizó el dataset PRONOSTIA, y los modelos más aplicados fueron de Deep Learning, evaluados con métricas como RMSE, MAE y MAPE

2.1.2. Antecedentes nacionales

Albornoz (2021) en su tesis "Aplicación del aprendizaje automático supervisado en el mantenimiento predictivo de motores eléctricos de inducción en las empresas mineras del Perú", se propuso como objetivo la implementación de técnicas de aprendizaje automático supervisado para mejorar el mantenimiento. predictivo de motores de inducción en el sector minero peruano. La metodología empleada consistió en la comparación de dos enfoques de mantenimiento predictivo: uno fundamentado en métodos convencionales y otro basado en algoritmos de aprendizaje automático supervisado. Se realizaron mediciones del aislamiento del bobinado, así como análisis de vibraciones mecánicas y del comportamiento de la corriente en el estator. Los resultados demostraron que la aplicación del aprendizaje automático supervisado mejoró la eficiencia y confiabilidad del mantenimiento predictivo, facilitando una predicción más precisa de fallas y optimizando los tiempos operativos de los motores.

Adauto (2021) en su tesis titulada "Aplicación de la inteligencia artificial en la detección de fallas en motores eléctricos de corriente continua de imán permanente", se planteó como finalidad utilizar inteligencia artificial para detectar fallas en este tipo de motores eléctricos. Para ello, se realizaron a cabo ensayos en un entorno controlado, recopilando datos sobre el aislamiento de los devanados, las vibraciones mecánicas y el comportamiento de las corrientes. Los resultados evidenciaron que la implementación de inteligencia artificial incrementó de manera notable la confiabilidad y la eficiencia del mantenimiento predictivo, minimizando la ocurrencia de fallas imprevistas y mejorando el rendimiento de los motores eléctricos en aplicación.

Orihuela (2024) en su tesis "Algoritmo predictivo como herramienta de mantenimiento para la detección anticipada de fallas de funcionamiento y operación en sistemas fotovoltaicos autónomos domiciliarios rurales", tuvo como propósito analizar el impacto de un algoritmo predictivo en la identificación temprana de fallas en el funcionamiento y operación de sistemas fotovoltaicos autónomos domiciliarios rurales (SFADR). La metodología aplicada contempló la recopilación de datos meteorológicos a

través de una estación meteorológica y un piranómetro, además de la simulación de escenarios operativos y de fallas en los sistemas fotovoltaicos. Los datos obtenidos fueron procesados utilizando modelos de aprendizaje automático, cuya validación se llevó a cabo mediante una matriz de confusión. Los resultados evidenciaron que el algoritmo predictivo alcanzó una precisión del 95,7 % en la detección de estados normales y del 99,1 % en la identificación de condiciones de falla, lo que contribuyó a mejorar la gestión del mantenimiento preventivo y garantizar el rendimiento óptimo de los sistemas fotovoltaicos.

Quijano y Crisostomo (2023) en su investigación *"Modelo predictivo para el mantenimiento preventivo de computadoras"*, tuvieron como objetivo desarrollar un modelo predictivo que permita anticipar fallas en computadoras y laptops para extender su vida útil y reducir pérdidas operativas. La metodología empleada consistió en una revisión sistemática de literatura siguiendo a Kitchenham y Charters (2007), recopilando artículos de bases de datos como Scopus, IEEE Xplore y Scielo, aplicando criterios de inclusión y exclusión, y clasificando la información en cinco áreas: técnicas, problemas, tecnologías, algoritmos y validación. Los resultados muestran que las técnicas de Machine Learning, como redes neuronales, árboles de decisión y Naive Bayes, combinadas con minería de datos y detección de anomalías, alcanzaron altos niveles de precisión, permitiendo predecir fallas y optimizar los planes de mantenimiento.

Recuay (2024) en su tesis *"Modelo estructural para determinar el tiempo de vida remanente en pavimentos flexibles"*, tuvo como objetivo desarrollar un modelo estructural para estimar el tiempo de vida remanente de pavimentos flexibles en la carretera 22B, desde Tarma hasta La Merced. La metodología empleada fue de tipo aplicada, con nivel explicativo y diseño no experimental, utilizando el método científico. Se recopilaron datos mediante observación directa, análisis documental y conteo vehicular, procesados con la distribución Weibull y la normativa AASHTO para determinar cargas acumuladas y requeridas. Los resultados indicaron que el tramo I presentó una vida remanente del 77,06 %, el tramo II del 52,65 %, el tramo III del 73,88 %, el tramo IV del 58,52 % y el tramo V del 51,02 %, con un promedio global de 62,63 % en toda la vía

2.2. Bases teóricas sobre el trabajo de investigación

2.2.1. Teoría de Bombas Centrífugas

A. Definición de Bombas Centrífugas

Una bomba centrífuga es una turbomáquina utilizada en sistemas que controlan el flujo de líquidos dentro de un circuito cerrado. Su función principal es convertir la energía proporcionada por un motor eléctrico o una turbina en energía cinética o velocidad, que luego se transforma en energía de presión aplicada al fluido en circulación. Este proceso se lleva a cabo dentro de los conductos que conforman un sistema hidráulico, ampliamente empleado en la industria para el transporte o distribución de fluidos, contribuyendo al mantenimiento de un flujo constante en distintos sistemas hidráulicos (Soto, 2021).

La conversión de energía en una bomba ocurre gracias a dos componentes fundamentales de su diseño: el impulsor y el difusor o voluta. El impulsor, que gira a alta velocidad, se encarga de transferir la energía mecánica de la máquina al fluido en forma de energía cinética. En cambio, el difusor o espiral, al ser una parte fija, permite que esa energía cinética se convierta en presión. Cabe resaltar que la bomba no genera presión por sí sola; esta surge como resultado de la resistencia que ofrece el sistema al paso del fluido. Este principio es clave para el rendimiento adecuado de la bomba (Soto, 2021).

B. Clasificación de Bombas Centrífugas

Las bombas centrífugas pueden clasificarse según diferentes criterios técnicos. Uno de ellos es el número de etapas, que se refiere a la cantidad de impulsores instalados en serie dentro de la bomba. Estos impulsores permiten aumentar gradualmente la presión del fluido en cada fase del proceso.

- **Bombas monoetapa:** Incorporan un solo impulsor, lo que les confiere un diseño más simple y compacto. Se utilizan en aplicaciones donde se requiere una presión moderada y no se necesita superar grandes alturas. Son comunes en sistemas de agua potable, riego y procesos industriales básicos.

- **Bombas multietapa:** Disponen de dos o más impulsores conectados en serie, lo que permite incrementar progresivamente la presión del fluido. Son adecuadas para aplicaciones que demandan alta presión, como en calderas, abastecimiento de agua en edificaciones elevadas o procesos industriales exigentes (Fernández, 2024).

Otro criterio utilizado es la dirección del flujo, que determina el trayecto que sigue el fluido al desplazarse dentro de la bomba. Según este parámetro, se reconocen tres configuraciones principales:

- **Radial:** El fluido se mueve a través del impulsor y la carcasa en dirección perpendicular al eje de rotación. Estas bombas son altamente eficientes y versátiles, siendo las más empleadas por su buen rendimiento en una amplia gama de aplicaciones.
- **Axial:** En este diseño, el fluido circula en dirección paralela al eje de rotación. Se caracterizan por su eficiencia y su capacidad para manejar grandes volúmenes de líquido, especialmente en sistemas con baja altura de elevación.
- **Mixto:** Representa una configuración intermedia que combina características de los flujos radial y axial, permitiendo un equilibrio entre presión generada y caudal, según los requerimientos del sistema (Soto, 2021).

Asimismo, las bombas centrífugas también pueden clasificarse de acuerdo con la orientación del eje de rotación, lo que influye en la forma de instalación y en el espacio requerido para su operación:

- **Horizontales:** El eje de rotación está dispuesto de forma paralela al plano de entrada del fluido. Este tipo de bomba es común por su facilidad de instalación y mantenimiento.
- **Verticales:** El eje se encuentra en una posición perpendicular y vertical, lo cual es útil cuando se dispone de espacio limitado en planta o se necesita succionar fluido desde niveles inferiores (Soto, 2021).

C. Elementos de Bombas Centrifugas

A continuación, se describen los componentes principales que hacen posible el funcionamiento de una bomba centrífuga:

La figura 1, presenta una bomba centrífuga, un dispositivo utilizado comúnmente en aplicaciones industriales y domésticas para el transporte de líquidos. A continuación, se explican los componentes esenciales que permiten el correcto funcionamiento de una bomba centrífuga, cada uno con un papel específico en la conversión y transmisión de energía hidráulica:

- **Motor eléctrico:** Ubicado en la parte posterior de la bomba, este componente es responsable de convertir la energía eléctrica en energía mecánica rotacional. A través de un eje, transmite este movimiento al impulsor para generar el flujo del líquido. Suele contar con aletas de ventilación que permiten disipar el calor generado durante la operación continua, evitando el sobrecalentamiento y garantizando un desempeño eficiente y seguro. Este tipo de motor, generalmente asincrónico, es ampliamente utilizado por su fiabilidad, bajo mantenimiento y facilidad de integración con sistemas de automatización (Paredes, 2019).
- **Carcasa (cuerpo de bomba):** De forma cilíndrica o helicoidal (voluta), esta estructura metálica en este caso de color verde cumple una doble función: proteger los componentes internos de la bomba y canalizar el fluido. Al envolver al impulsor, guía el flujo desde la entrada (succión) hacia la salida (descarga), permitiendo la transformación de la energía cinética generada por el impulsor en presión. Este proceso es clave para alcanzar un bombeo eficiente y controlar el caudal con estabilidad. Además, su diseño busca minimizar las pérdidas de carga por fricción interna.

- **Conexiones de entrada y salida:** Estas boquillas o bocas de conexión permiten el ingreso y egreso del fluido. El líquido accede generalmente por la entrada lateral izquierda (succión), siendo acelerado por el impulsor y expulsado con mayor presión por la parte superior o lateral (descarga). Su orientación y diámetro son definidos según el diseño hidráulico y el requerimiento de caudal, asegurando un flujo continuo y uniforme en aplicaciones industriales o domésticas.

- **Base de apoyo:** La bomba se monta sobre una estructura sólida inferior que proporciona soporte físico y estabilidad durante su funcionamiento. Esta base absorbe las vibraciones mecánicas generadas por el motor y el giro del impulsor, lo que contribuye a evitar desplazamientos, daños estructurales o fallos por fatiga mecánica. Una instalación correcta sobre una base nivelada y firme es fundamental para prolongar la vida útil del equipo y mantener su rendimiento óptimo (Paredes, 2019).

Los componentes de la bomba centrífuga trabajan de manera coordinada para permitir su operación eficiente. El fluido ingresa a través de la carcasa y se dirige hacia los impulsores, los cuales aumentan su velocidad. Luego, el fluido es expulsado por la salida con una presión incrementada. Este mecanismo es posible gracias a la geometría de la carcasa y a la energía transmitida por el movimiento rotatorio de los impulsores (Paredes, 2019).

Figura 1

Bomba Centrífuga



Nota. Tomado de (Paredes, 2019).

D. Componentes de la Bomba Centrífuga

A continuación, se describen los elementos que conforman una bomba centrífuga:

- **Corona Directriz:** Esta parte se encarga de orientar el flujo del líquido hacia el impulsor (también conocido como rodete). Su función principal es encauzar el fluido para que ingrese de manera uniforme y controlada a las paletas del rodete.

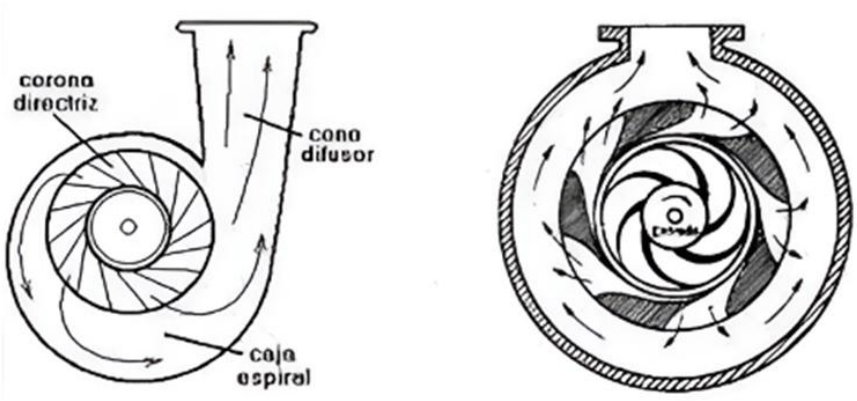
- **El Difusor:** El difusor representa una pieza clave dentro de una bomba centrífuga, ya que cumple una función esencial en la conversión de energía dentro del sistema. Su tarea principal consiste en transformar la energía cinética generada por el rodete en presión, lo que incrementa la eficacia y el desempeño del equipo (Caballero et al., 2021). Una vez que el impulsor acelera el fluido, elevando su velocidad, el difusor actúa reduciendo dicha velocidad. Esta estructura se compone de álabes inmóviles que rodean al rodete, diseñados específicamente para disminuir la velocidad del líquido al atravesarlos. Al realizar esta acción, el difusor convierte la energía del movimiento en presión, facilitando una expulsión del fluido con mayor potencia y eficiencia. El empleo de difusores resulta especialmente relevante en contextos donde se requieren presiones elevadas o en bombas de múltiples etapas, donde es necesario aumentar la presión sin modificar considerablemente la velocidad del flujo. Además, al distribuir el líquido de manera más pareja, el difusor contribuye a minimizar las pérdidas energéticas y a mantener la estabilidad operativa del sistema (Caballero et al., 2021). El difusor está formado por álabes fijos y divergentes. Conforme la carcasa se ensancha, la velocidad del agua disminuye, favoreciendo así la conversión de energía cinética en presión y mejorando el rendimiento de la bomba. Según su estructura y configuración, las bombas centrífugas pueden dividirse en dos clases:

    - ✓ **De tipo helicoidal:** el revestimiento adopta una forma espiral que envuelve al rodete, permitiendo que el área por la que se desplaza el fluido se amplíe progresivamente hasta alcanzar la tubería de descarga.

    - ✓ **De tipo turbina:** el cuerpo de la bomba incorpora difusores estáticos organizados de tal manera que el espacio de paso se expande de forma gradual hacia la salida (Caballero et al., 2021).

Figura 2

Bomba Centrífuga con Difusor y Caja en Espiral



Nota. Tomado de (Caballero et al., 2021).

- **Carcasa en Espiral (Voluta):** La carcasa en espiral, conocida también como voluta, es una estructura fija que envuelve al impulsor y cumple la función de recolectar el fluido impulsado por las paletas. Su forma helicoidal está diseñada estratégicamente para desacelerar progresivamente la velocidad del líquido a medida que se desplaza por su interior. Esta reducción de velocidad permite transformar la energía cinética del fluido en presión, un proceso clave para lograr un bombeo eficiente (Caballero et al., 2021). Además, la forma de la voluta ayuda a distribuir de manera uniforme la presión alrededor del impulsor, lo que reduce esfuerzos mecánicos y mejora la estabilidad operativa de la bomba.

- **Dirección del Flujo:** Las flechas indicadas en el esquema reflejan el trayecto que sigue el fluido dentro del sistema. El líquido ingresa axialmente por la succión, ubicada en el centro del impulsor, y es empujado hacia la periferia como resultado del giro de las paletas. Este movimiento es generado por la fuerza centrífuga, la cual incrementa tanto la velocidad como la presión del fluido (Caballero et al., 2021).

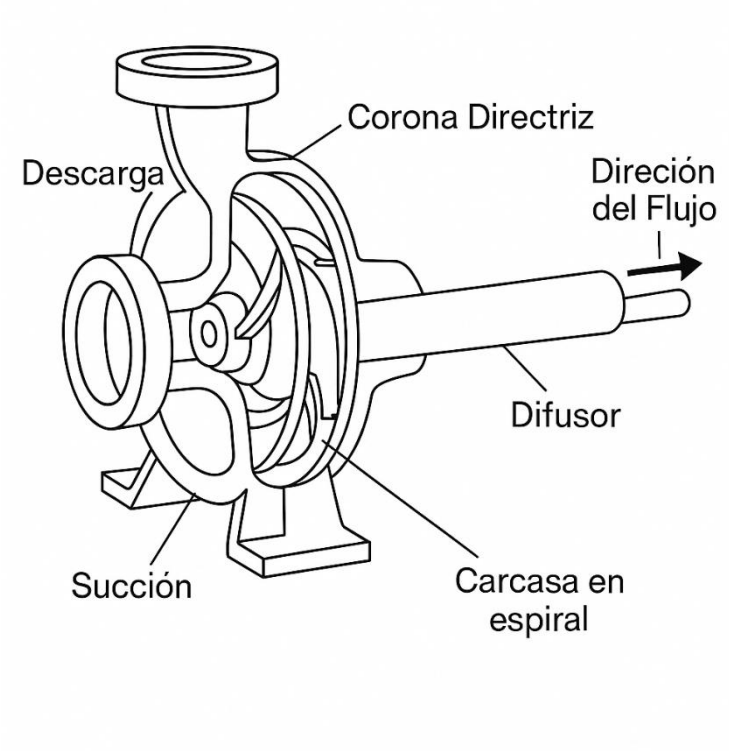
- **Impulsor (Rodete):** El impulsor representa el elemento principal de movimiento dentro de una bomba centrífuga. Su rotación imprime energía cinética al fluido, impulsándolo hacia el borde externo de la carcasa. La configuración geométrica

de sus álabes (número, curvatura, inclinación y grosor) influye directamente en la capacidad de caudal, presión generada y eficiencia global del equipo (Caballero et al., 2021). Por ello, el diseño del impulsor se adapta a los requerimientos específicos de cada aplicación, ya sea para líquidos limpios, cargados o corrosivos.

- **Árbol de Transmisión (Eje de la Bomba):** El eje es una barra metálica de sección circular, cuyo tamaño y resistencia están calculados para soportar las cargas generadas por la rotación del impulsor. Este componente se acopla firmemente al rodete y transmite la energía mecánica que proviene del motor, permitiendo así su movimiento. La alineación, rigidez y materiales del eje son factores determinantes para evitar fallas mecánicas y garantizar una transmisión eficiente del torque (Caballero et al., 2021). Además, su correcto montaje y balanceo son fundamentales para minimizar vibraciones y alargar la vida útil de la bomba.

Figura 3

Componentes de una bomba centrífuga



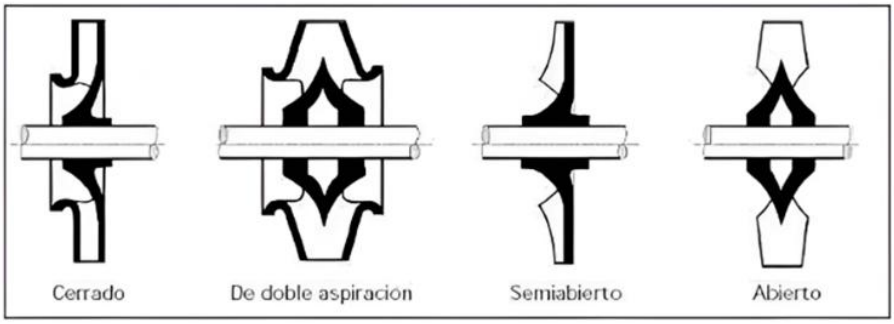
Nota. Tomado de (Caballero et al., 2021).

E. Clasificación de Impulsores en Bombas Centrífugas

Se identifican cuatro clases principales de impulsores empleados en bombas centrífugas, cada uno diseñado con propiedades particulares según el tipo de aplicación. En la figura 4 se detallan las variedades de impulsores y sus respectivas características.

Figura 4

Bombas Centrifugas por tipo de impulsor



Nota. Tomado de (Caballero et al., 2021).

Según Caballero et al. (2021) cada uno de estos tipos de impulsores posee una configuración particular que incide directamente en su desempeño y en los usos para los que resulta más apropiado:

A. Impulsor Cerrado:

- Este diseño está completamente encerrado por dos discos laterales, dejando únicamente el extremo externo de las paletas al descubierto.
- Es el modelo más eficiente para el bombeo de líquidos, ya que minimiza la recirculación interna del fluido, optimizando así el funcionamiento de la bomba.
- Se emplea principalmente en procesos donde se trabaja con líquidos claros, ya que la presencia de partículas sólidas podría obstruir el espacio entre los discos.

B. Impulsor de Doble Succión:

- Este tipo de impulsor permite la entrada del fluido desde ambos extremos del eje, lo que contribuye a equilibrar las fuerzas axiales y mejora la eficiencia en bombas de gran tamaño.
- Es habitual en sistemas que requieren transportar grandes volúmenes de agua u otros líquidos a caudales elevados.

- Al permitir el ingreso del fluido por ambos lados, se disminuye el desgaste mecánico y se prolonga la vida útil del equipo.

C. Impulsor Semiabierto:

- Este diseño cuenta con un solo disco en uno de sus lados, mientras que las paletas quedan visibles en la otra cara.
- Representa un punto medio entre los impulsores cerrados y abiertos, siendo útil para líquidos que contienen pequeñas cantidades de sólidos o presentan viscosidad moderada.
- Ofrece mayor facilidad para su limpieza y mantenimiento en comparación con el cerrado, aunque su eficiencia hidráulica es ligeramente inferior.

D. Impulsor Abierto:

- Este impulsor carece de discos que recubran las paletas, por lo que estas se encuentran completamente expuestas.
- Aunque su rendimiento es menor frente a otros tipos, es especialmente adecuado para fluidos que contienen sólidos en suspensión, lodos o líquidos densos.
- Su estructura simple facilita la limpieza y el mantenimiento, lo que lo convierte en una opción ideal en aplicaciones donde la accesibilidad y el bajo mantenimiento son fundamentales.

2.2.2. Vida útil de las bombas centrífugas

El arranque de una bomba centrífuga puede influir significativamente en su vida útil. Los métodos de arranque tradicionales, como el arranque directo o el uso de configuraciones estrella-triángulo en los motores que las impulsan, pueden generar picos elevados de corriente y presión. Estos picos provocan esfuerzos mecánicos e hidráulicos que pueden acelerar el desgaste de componentes como sellos mecánicos, cojinetes e impulsores, reduciendo así la durabilidad de la bomba. En contraste, el uso de variadores de velocidad permite un arranque más progresivo y controlado, disminuyendo los impactos hidráulicos y mejorando la confiabilidad y longevidad del equipo. La vida operativa de una bomba centrífuga está estrechamente relacionada con el control adecuado de variables como la presión, el caudal y la velocidad de rotación. Un monitoreo

constante de estos parámetros, especialmente durante los procesos de arranque y parada, ayuda a mitigar golpes de ariete y vibraciones excesivas, reduciendo así los esfuerzos mecánicos que afectan la estructura y el rendimiento de la bomba a lo largo del tiempo (Carrillo y Rojas, 2013).

La eficiencia de una bomba centrífuga se ve afectada por diversos factores operativos y mecánicos, entre los que destacan el funcionamiento fuera del punto de máxima eficiencia, el desgaste del impulsor y de los elementos de sellado, la desalineación del eje y la presencia de vibraciones excesivas. Asimismo, fenómenos como la cavitación, las pérdidas hidráulicas en la succión y descarga, la variación de la viscosidad del fluido y una lubricación deficiente influyen negativamente en el rendimiento del equipo. Estas condiciones no solo reducen la eficiencia hidráulica y mecánica, sino que también aceleran el deterioro de los componentes, acortando la vida útil de la bomba (Carrillo y Rojas, 2013).

2.2.3. Indicador de condición o salud

Un indicador de condición basado en algoritmos de aprendizaje supervisado permite monitorear el estado operativo de motores eléctricos en equipos críticos. Este método analiza datos históricos y en tiempo real, como vibraciones, temperatura, corriente y ciclos de operación, para estimar la vida útil restante del motor. Algoritmos como Random Forest, Support Vector Machines (SVM) o redes neuronales analizan estas variables, identificando patrones que permiten anticipar posibles fallas. La implementación de este tipo de indicador optimiza la planificación del mantenimiento, reduciendo costos y minimizando los tiempos de inactividad, al tiempo que mejora la confiabilidad del sistema y asegura un funcionamiento continuo en industrias exigentes como la minería y la fabricación pesada (Carrillo y Rojas, 2013).

El uso de indicadores de condición mediante algoritmos de aprendizaje supervisado proporciona una evaluación continua del estado de los motores eléctricos, ofreciendo una estimación precisa de su tiempo de operación restante. Esta metodología incorpora técnicas avanzadas de aprendizaje automático, como redes neuronales profundas, Random Forest y SVM, para analizar parámetros operativos como vibraciones, temperatura, corriente, frecuencia de fallas y desgaste de los componentes (Silva y Elías, 2017). Al entrenar modelos predictivos con datos tanto históricos como en

tiempo real, se pueden detectar tendencias que predicen posibles averías, permitiendo un mantenimiento predictivo más eficiente. Estos indicadores incrementan la confiabilidad de los equipos al minimizar fallas inesperadas, optimizar los costos operativos y mejorar la eficiencia del sistema. En industrias como la minería, donde las máquinas operan en condiciones extremas, esta herramienta resulta crucial para garantizar la continuidad operativa y prolongar la vida útil de los activos. Además, plataformas como Matlab facilitan el desarrollo de estos modelos, proporcionando herramientas de simulación, análisis de datos y algoritmos avanzados en un entorno unificado (Silva y Elías, 2017).

En el contexto de la presente investigación, los modos de falla no se definen como eventos discretos de falla catastrófica, sino como procesos progresivos de degradación asociados al deterioro de los componentes críticos de la bomba centrífuga. Esta degradación se mide de manera indirecta mediante la evolución temporal de variables operativas, tales como vibración, temperatura, presión y caudal, las cuales actúan como indicadores del estado de salud del equipo. El incremento sostenido de vibraciones, el aumento térmico anómalo o la pérdida gradual de eficiencia hidráulica son interpretados como señales del avance de los modos de falla, permitiendo caracterizar el deterioro antes de la ocurrencia de una falla funcional. Este enfoque es coherente con los principios del mantenimiento predictivo basado en condición, donde la degradación se evalúa de forma continua y no binaria.

2.2.4. Teoría de Mantenimiento

A. Definición de Mantenimiento

El mantenimiento comprende un conjunto de acciones técnicas, operativas y organizativas orientadas a conservar o restablecer el funcionamiento adecuado de equipos, máquinas o sistemas. Su finalidad es asegurar una operación segura, continua y eficiente, mediante la prevención, detección o corrección de posibles fallas que puedan afectar la productividad o causar interrupciones en los procesos industriales (Santillan, 2017).

B. Clasificación de Mantenimiento

El mantenimiento puede agruparse en tres categorías principales, de acuerdo con el momento en que se llevan a cabo las intervenciones y el enfoque adoptado:

- **Mantenimiento correctivo:** Consiste en realizar reparaciones una vez que el equipo ha fallado. Este tipo de mantenimiento es reactivo y suele implicar mayores costos y tiempos de detención del sistema.
- **Mantenimiento preventivo:** Se basa en realizar intervenciones programadas según un calendario o nivel de uso del equipo, con el objetivo de evitar posibles fallas. Incluye revisiones periódicas, ajustes y el cambio anticipado de piezas con desgaste.
- **Mantenimiento predictivo:** Se fundamenta en el monitoreo de la condición real del equipo, utilizando sensores, pruebas técnicas y análisis de datos para anticipar posibles fallas. Esto permite intervenir en el momento más oportuno, extendiendo la vida útil de los componentes y minimizando los periodos de inactividad (Santillan, 2017).

C. Mantenimiento Preventivo

Es un tipo de mantenimiento que, como su nombre lo sugiere, busca minimizar la probabilidad de fallas o averías, mediante la ejecución de un conjunto de acciones orientadas a ese propósito. Esto permite prolongar la vida útil de los equipos y mejorar el rendimiento operativo al disminuir los denominados tiempos de inactividad, causados por fallos imprevistos (Albornoz, 2021).

Las actividades que normalmente se llevan a cabo en un mantenimiento preventivo incluyen:

- Desarme del sistema.
- Sustitución de componentes con alta probabilidad de fallo o que presenten un grado notable de desgaste.
- Reensamblaje, pruebas y validación del sistema.
- Elaboración de una base de datos histórica, con el registro detallado de las fallas ocurridas (Albornoz, 2021).

D. Mantenimiento Predictivo

El mantenimiento predictivo se basa en la supervisión continua y el análisis detallado del estado operativo del transformador, con el propósito de detectar fallos potenciales antes de que ocurran. A través de herramientas avanzadas como el análisis de

gases disueltos y sensores especializados, es posible identificar anomalías de manera temprana, lo que permite planificar intervenciones oportunas. A diferencia del mantenimiento preventivo, este enfoque no sigue un calendario fijo, sino que se adapta a las condiciones reales del equipo. Aunque su implementación requiere inversiones en sistemas de monitoreo y capacitación del personal, los beneficios incluyen la optimización del rendimiento del transformador, la reducción de paradas inesperadas y una mayor disponibilidad operativa (Carrillo y Rojas, 2013).

Este enfoque se centra en la vigilancia constante del equipo durante su funcionamiento normal, mediante el uso de mediciones en tiempo real y análisis de datos históricos obtenidos de inspecciones regulares. Con esta información, se desarrollan estrategias para reemplazar o mejorar componentes específicos que presenten signos de deterioro. De esta manera, se previenen fallos, se optimiza el rendimiento y se prolonga la vida útil del equipo a través de intervenciones dirigidas a los elementos críticos (Silva y Elías, 2017).

El mantenimiento predictivo se distingue de otras estrategias, como el mantenimiento preventivo, al priorizar el estado actual del equipo en lugar de seguir intervalos de tiempo predeterminados. Este tipo de mantenimiento permite realizar ajustes dinámicos en el plan de mantenimiento, optimizando los recursos disponibles y reduciendo los costos asociados a fallos imprevistos. Gracias a sistemas de monitoreo continuo y diagnóstico avanzado, se obtiene una visión integral del desempeño del transformador, lo que contribuye a incrementar la eficiencia y la confiabilidad operativa a largo plazo (Aranda, 2024).

2.2.5. Mantenimiento Predictivo en Bombas Centrifugas

El mantenimiento predictivo es una estrategia basada en la supervisión continua o periódica del estado operativo de un equipo, con el fin de anticiparse a posibles fallos antes de que ocurran. En el caso de bombas centrífugas, esta técnica se aplica para detectar anomalías, reducir tiempos de inactividad no planificados, y alargar la vida útil de sus componentes, optimizando el funcionamiento general del sistema. A diferencia del mantenimiento correctivo (que actúa tras el fallo) o del preventivo (que sigue un cronograma fijo), el predictivo actúa en función del comportamiento real del equipo, usando herramientas de monitoreo, análisis de datos y criterios de falla (Fuentes, 2025).

A. Objetivos del mantenimiento predictivo en bombas centrífugas

- Identificar signos tempranos de desgaste o desalineación.
- Evitar fallas inesperadas en sistemas críticos.
- Reducir costos operativos mediante intervenciones dirigidas y oportunas.
- Optimizar el tiempo de servicio sin afectar la continuidad de operación.
- Mejorar la eficiencia energética y el rendimiento hidráulico del equipo.

B. Parámetros monitoreados comúnmente

Según Fuentes (2025), para implementar un sistema predictivo eficaz en bombas centrífugas, se evalúan parámetros clave que reflejan su estado operativo:

- **Vibraciones:** Un incremento en los niveles de vibración puede ser síntoma de desbalance del impulsor, desalineación del eje, aflojamiento de piezas o cavitación. Se utilizan sensores acelerómetros para medir la frecuencia y amplitud de las vibraciones.
- **Temperatura:** Se supervisa la temperatura del motor, cojinetes y sellos mecánicos. Un aumento inusual puede indicar fricción excesiva, falta de lubricación o sobrecarga.
- **Presión y caudal:** Variaciones fuera del rango nominal pueden reflejar obstrucciones en las líneas, desgaste del impulsor o pérdida de eficiencia hidráulica.
- **Consumo eléctrico:** Cambios en la corriente absorbida por el motor eléctrico permiten inferir sobrecargas mecánicas o bloqueos internos.
- **Nivel de ruido:** Sonidos anómalos pueden ser indicio de desequilibrios, cavitación o fallas en cojinetes. El análisis acústico complementa la detección por vibración.
- **Inspección visual y análisis de aceite/lubricante:** Se busca evidencia de fugas, desgaste de piezas o contaminación. El análisis de partículas en lubricantes revela el estado interno de los componentes móviles.

C. Técnicas utilizadas en el mantenimiento predictivo

- **Análisis de vibraciones:** Herramienta esencial para evaluar el estado dinámico de la bomba. Permite detectar fallas como desbalance, holgura, cavitación o fallos en rodamientos.

- **Termografía infrarroja:** Permite visualizar puntos de sobrecalentamiento en componentes eléctricos y mecánicos. Muy útil en motores eléctricos y rodamientos.
- **Ultrasonido:** Ayuda a detectar fugas, cavitación o desgaste prematuro mediante sonidos de alta frecuencia.
- **Análisis de aceites y lubricantes:** Se examina la viscosidad, presencia de metales y partículas contaminantes para anticipar desgaste interno de piezas críticas.
- **Monitoreo remoto o en línea (online monitoring):** Con sensores conectados a sistemas SCADA o plataformas IoT, se pueden registrar y analizar datos en tiempo real, generando alertas automáticas ante condiciones anómalas.

D. Ventajas del mantenimiento predictivo

- **Disminución de paradas inesperadas:** Al actuar antes de que ocurra una avería, se garantiza mayor disponibilidad operativa.
- **Reducción de costos por reparaciones mayores:** Detectar una falla en etapa temprana evita daños colaterales en otros componentes.
- **Mayor vida útil de la bomba:** Al mantener condiciones óptimas de operación, se minimiza el desgaste prematuro.
- **Optimización del inventario de repuestos:** Se puede planificar la reposición de piezas con base en el estado real y no por intervalos fijos.
- **Mejora en la seguridad operativa:** Minimiza riesgos de fallas catastróficas que puedan generar accidentes o pérdidas de producción.

E. Aplicación práctica en una bomba centrífuga

En entornos industriales, una de las implementaciones más efectivas del mantenimiento predictivo en bombas centrífugas consiste en la instalación de sensores especializados de vibración y temperatura en puntos estratégicos del equipo, como el eje de rotación, la carcasa, los rodamientos y los sellos mecánicos. Estos sensores recogen datos en tiempo real y los transmiten a un sistema digital de monitoreo generalmente un software SCADA, una plataforma IoT o un sistema de gestión de mantenimiento computarizado (CMMS) para su análisis y seguimiento (Fuentes, 2025).

¿Cómo funciona?

1. **Detección temprana de fallos mecánicos:** Las vibraciones excesivas suelen ser los primeros indicadores de problemas como desbalance del impulsor, desalineación del eje, aflojamiento de componentes, desgaste en cojinetes o cavitación. Los sensores acelerómetros miden la frecuencia y la amplitud de estas vibraciones. Si los valores superan los límites preestablecidos por el fabricante o por estándares internacionales (como ISO 10816), el sistema emite una alerta automática, recomendando una inspección inmediata antes de que se produzca una falla grave.
2. **Control térmico del equipo:** La temperatura anormal en ciertas zonas del equipo puede deberse a fricción, lubricación deficiente, bloqueo interno o sobrecarga del motor. Los sensores térmicos como los termopares o sensores infrarrojos permiten detectar estos cambios y correlacionarlos con el comportamiento general del sistema.

Un caso recurrente de degradación mecánica en bombas centrífugas es el desgaste excesivo del eje en la zona de los prensaestopas, el cual suele estar asociado a una desalineación del eje, un ajuste inadecuado del sistema de sellado o una lubricación deficiente. Estas condiciones generan fricción anormal y vibraciones elevadas, provocando un deterioro acelerado de la superficie del eje. Asimismo, la operación fuera del punto óptimo de diseño y la presencia de vibraciones persistentes contribuyen a intensificar este tipo de daño, pudiendo derivar en fugas, fallas prematuras del sello y reducción de la vida útil del conjunto rotativo.

3. **Integración con el monitoreo remoto:** Gracias al avance del internet industrial de las cosas (IoT), los datos recogidos pueden visualizarse en plataformas web o móviles, lo que permite a los técnicos recibir notificaciones incluso sin estar físicamente presentes en la planta. Esto acelera la toma de decisiones y facilita la programación de intervenciones de mantenimiento solo cuando realmente se necesitan, evitando gastos innecesarios por paradas programadas sin justificación.

2.2.6. Análisis de componentes principales (PCA)

El Análisis de Componentes Principales (PCA, por sus siglas en inglés) es una técnica estadística utilizada para reducir la dimensionalidad de un conjunto de datos,

manteniendo la mayor cantidad de información posible. Su objetivo principal es transformar un conjunto de variables posiblemente correlacionadas en un nuevo conjunto de variables no correlacionadas, llamadas componentes principales, las cuales se ordenan de acuerdo con la cantidad de varianza explicada. Esta técnica es ampliamente utilizada en el análisis de grandes volúmenes de datos, facilitando la interpretación, visualización y reducción del ruido en los datos. El PCA se basa en la descomposición de la matriz de covarianza o correlación de los datos originales, identificando las direcciones en las que la variabilidad es máxima. Aplicado en diversas áreas como la ingeniería, la economía y la biomedicina, el PCA permite optimizar modelos predictivos, simplificar la representación de los datos y mejorar la eficiencia computacional en tareas de clasificación, regresión y segmentación (Li et al., 2018).

En el contexto de la presente investigación, la aplicación del Análisis de Componentes Principales (PCA) adquiere especial relevancia debido a la naturaleza multivariada y correlacionada de las variables operativas caudal, presión, vibración y temperatura registradas durante la operación. Estos parámetros no son independientes: por ejemplo, un incremento en la vibración puede estar asociado a variaciones en la presión o al desgaste mecánico de componentes internos, lo que genera colinealidad estadística y redundancia de información en los datos crudos (Jolliffe y Cadima, 2016).

El PCA permite proyectar las variables originales en un espacio ortogonal donde los nuevos ejes (componentes principales) son no correlacionados y están ordenados según la varianza explicada. Así, la primera componente concentra la mayor información del sistema relacionada, por ejemplo, con la degradación progresiva de la bomba mientras que las siguientes capturan variaciones secundarias o ruido experimental. Esta transformación ofrece ventajas críticas:

1. **Detección temprana de fallas**: Al concentrar la variabilidad relevante en pocos componentes, se facilita la identificación de patrones anómalos asociados a desgaste, cavitación o desbalance dinámico, fenómenos descritos por Li et al. (2018) en sistemas de bombeo.
2. **Reducción de ruido y dimensionalidad**: La compresión de información elimina redundancias en señales de alta frecuencia o picos espurios en sensores, mejorando la calidad estadística del conjunto de entrenamiento (Shlens, 2014).

3. **Eficiencia computacional**: Al disminuir la cantidad de variables de entrada, el PCA reduce el costo de entrenamiento y mejora la convergencia de algoritmos como redes neuronales artificiales y modelos de aprendizaje profundo, según reportes recientes en mantenimiento predictivo industrial (Shlens, 2014).

Adicionalmente, la visualización de los primeros componentes facilita la interpretación física del comportamiento de la bomba a lo largo del tiempo, permitiendo distinguir con claridad etapas de degradación o periodos de operación normal. Por ello, el PCA se integra como una fase previa esencial antes del modelado predictivo, asegurando mayor precisión, robustez y capacidad de generalización del sistema final para la estimación del RUL en entornos industriales complejos.

2.2.7. Modelos predictivos

2.2.7.1.Redes Neuronales Artificiales

Las Redes Neuronales Artificiales (RNA) son sistemas computacionales inspirados en el funcionamiento del sistema nervioso humano, especialmente en las conexiones entre las neuronas biológicas. Su diseño imita la forma en que el cerebro procesa, aprende y adapta la información, permitiendo a las máquinas reconocer patrones complejos, establecer relaciones entre datos y realizar predicciones. Estas redes constituyen una de las herramientas más potentes en el campo de la inteligencia artificial y el aprendizaje automático, siendo aplicadas en tareas como la clasificación de datos, regresión, reconocimiento de voz e imágenes, diagnóstico médico, procesamiento del lenguaje natural (NLP), y predicción de series temporales (Noya y Sutizal, 2023).

Una característica fundamental de las RNA es su capacidad de aprendizaje supervisado y no supervisado, lo que les permite ajustarse a distintas formas de entrenamiento según la disponibilidad de etiquetas en los datos. Las redes neuronales modernas, particularmente las redes profundas (Deep Learning), han demostrado un rendimiento sobresaliente en entornos complejos, superando en muchos casos a los algoritmos tradicionales (Noya y Sutizal, 2023).

A. Arquitectura de una Red Neuronal Artificial

Una red neuronal artificial se organiza en distintas capas de nodos o neuronas artificiales, conectadas entre sí por pesos sinápticos que determinan la influencia de una neurona sobre otra. Según Noya y Sutizal (2023) estas capas pueden clasificarse en tres grupos principales:

1. **Capa de entrada**:
    - Es la primera capa del sistema, responsable de recibir los datos brutos provenientes del entorno o del conjunto de entrenamiento.
    - No realiza procesamiento alguno, su función consiste en canalizar la información hacia las neuronas de las capas subsiguientes.
    - Los datos pueden ser de distinta naturaleza: imágenes, señales eléctricas, registros numéricos, texto, etc.
2. **Capas ocultas**:
    - Constituyen el núcleo de procesamiento de la red, donde se llevan a cabo transformaciones complejas de los datos.
    - Cada neurona en estas capas realiza un cálculo ponderado de las entradas recibidas, seguido por la aplicación de una función de activación (como ReLU, Sigmoid o Tanh), que introduce no linealidad al sistema.
    - Una red puede contar con una o múltiples capas ocultas. Cuando se integran varias, se habla de redes neuronales profundas (DNN), las cuales son especialmente útiles para problemas de alta complejidad o con grandes volúmenes de datos.
3. **Capa de salida**:
    - Es la capa encargada de proporcionar el resultado final del modelo, ya sea una predicción de clase (en clasificación) o un valor numérico (en regresión).
    - El número de neuronas en esta capa depende directamente del tipo de problema. Por ejemplo, una red para clasificar entre tres categorías tendrá tres neuronas en su capa de salida.
    - En esta etapa también puede aplicarse una función de activación específica, en clasificación multiclase, para interpretar los resultados como probabilidades.

Cabe señalar que el rendimiento de una red neuronal depende de múltiples factores, entre ellos: la calidad y cantidad de los datos de entrada, el número de capas y neuronas, las funciones de activación elegidas, el algoritmo de optimización (como Adam o SGD) y los criterios de regularización (como Dropout o L2). Gracias a estos componentes, las RNA pueden ser entrenadas para detectar patrones complejos, realizar inferencias precisas y adaptarse dinámicamente a nuevas condiciones o datos.

B. Funcionamiento General de una Neurona Artificial

En una Red Neuronal Artificial (RNA), cada neurona funciona como una unidad computacional que procesa múltiples señales de entrada. Estas entradas están asociadas a coeficientes ponderados conocidos como pesos, que determinan la influencia de cada dato en el cálculo de la salida. A este valor ponderado se le suma un término adicional denominado sesgo (bias), cuya función es permitir el desplazamiento del resultado, incluso cuando las entradas sean nulas, mejorando así la capacidad de adaptación de la neurona al patrón de datos (Noya y Sutizal, 2023).

El comportamiento de cada neurona se describe mediante una función matemática que sintetiza el procesamiento interno:

$$y = f\left(\sum_{i=1}^{n} w_i x_i + b\right)$$

Donde:

- y es la salida de la neurona.
- w_i son los pesos asociados a cada entrada x_i
- b es el sesgo o bias.
- f es la función de activación, que introduce no linealidad al sistema.

Este mecanismo imita el comportamiento de las neuronas biológicas, donde los estímulos recibidos son procesados y conducidos hacia una respuesta si se alcanza un cierto umbral. En las RNA, esta respuesta se regula a través de funciones como ReLU,

sigmoide o tangente hiperbólica, esenciales para evitar que la red se limite a representar únicamente relaciones lineales.

C. Proceso de Entrenamiento

El entrenamiento de una red neuronal artificial es el proceso mediante el cual la red aprende a realizar tareas específicas, como clasificación o predicción, ajustando sus parámetros internos (pesos y sesgos) a partir de un conjunto de datos. Este ajuste busca minimizar la diferencia entre las salidas generadas por la red y los valores deseados, es decir, reducir el error de predicción. Cuanto menor es este error, mayor es la capacidad de la red para generalizar y responder correctamente a datos nuevos (Noya y Sutizal, 2023).

Según Noya y Sutizal (2023) el proceso de entrenamiento se desarrolla a través de las siguientes fases fundamentales:

1. **Propagación hacia adelante (Forward Propagation):**
    - Los datos se introducen en la capa de entrada y se transmiten hacia las capas ocultas, donde cada neurona realiza un cálculo y produce una salida. Finalmente, se obtiene una predicción en la capa de salida.
    - Este paso implica la evaluación de la función de activación en cada nodo y el cómputo de la salida global del modelo.
2. **Cálculo del error:**
    - Se cuantifica la diferencia entre la salida estimada por la red y el valor real esperado. Para ello, se emplean funciones de costo como el Error Cuadrático Medio (MSE), el entropía cruzada, entre otras, dependiendo del tipo de problema (regresión o clasificación).
3. **Retropropagación del error (Backward Propagation):**
    - A través del algoritmo de retropropagación, se calculan los gradientes del error respecto a cada peso y sesgo, utilizando el método de la derivada parcial en sentido inverso por toda la red.
    - Este proceso aplica la regla de la cadena para distribuir el error desde la capa de salida hasta las capas anteriores.

4. **Ajuste de parámetros (Actualización de Pesos y Sesgos):**
    - Se actualizan los valores de los pesos y sesgos utilizando un algoritmo de optimización, siendo el descenso por gradiente estocástico (SGD) uno de los más comunes.
    - Existen también variantes más avanzadas como Adam, RMSProp o Adagrad, que aceleran la convergencia y mejoran la estabilidad del aprendizaje.
5. **Iteraciones o épocas:**
    - Este ciclo completo se repite a lo largo de varias épocas, que consisten en recorridos completos del conjunto de entrenamiento. Durante cada época, los parámetros se ajustan progresivamente hasta que la red alcanza un rendimiento satisfactorio.
    - El número de épocas, la tasa de aprendizaje y el tamaño del *batch* son hiperparámetros críticos que influyen en la eficiencia del aprendizaje y en la capacidad de evitar problemas como el sobreajuste.

D. Clasificación de Redes Neuronales Artificiales

Según lo planteado por López y Fernández (2008) en su libro "Las redes neuronales artificiales", existen diversas arquitecturas de redes neuronales artificiales (RNA), diseñadas para resolver distintos tipos de problemas dependiendo de la naturaleza de los datos y los objetivos del análisis. A continuación, se describen las variantes más representativas y sus respectivos ámbitos de aplicación:

a) Redes Neuronales de Propagación Directa (Feedforward Neural Networks – FFNN)

- Este tipo de red representa la forma más elemental de una RNA. En ella, la información fluye exclusivamente en un único sentido: desde la capa de entrada, pasando por una o varias capas ocultas, hasta llegar a la capa de salida.
- Las FFNN no poseen conexiones cíclicas ni mecanismos de retroalimentación, por lo que son adecuadas para resolver problemas estáticos, donde las entradas no dependen del orden ni del contexto, como la clasificación de imágenes simples o la predicción de resultados numéricos.

- Destacan por su estructura sencilla, lo que facilita su desarrollo, entrenamiento y análisis. Son especialmente útiles cuando se dispone de datos estructurados y etiquetados.
- A pesar de su facilidad de implementación, no son apropiadas para datos secuenciales o dependientes del tiempo, dado que carecen de mecanismos de memoria.

b) Redes Neuronales Recurrentes (Recurrent Neural Networks – RNN)

- Las RNN están diseñadas para gestionar información secuencial o temporal, como texto, audio, datos financieros o series cronológicas. Esto es posible gracias a las conexiones recurrentes entre las neuronas, que les permiten mantener información del pasado.
- A diferencia de las redes feedforward, las RNN tienen la capacidad de conservar un "estado interno" que sirve de contexto, haciendo que las predicciones actuales puedan depender de entradas anteriores.
- Esta arquitectura es fundamental en aplicaciones donde el orden o la dependencia entre datos consecutivos es relevante, como en traducción automática, generación de texto, análisis de sentimiento y reconocimiento de voz.
- Uno de los principales retos al entrenar RNNs es el problema del desvanecimiento o explosión del gradiente, lo cual puede limitar el aprendizaje en secuencias extensas.
- Para superar estas limitaciones, se han desarrollado variantes como las LSTM (Long Short-Term Memory) y las GRU (Gated Recurrent Units), que incorporan mecanismos de control para retener o descartar información a través del tiempo, mejorando la estabilidad del entrenamiento.

c) Redes Neuronales Convolucionales (Convolutional Neural Networks – CNN)

- Las CNN son especialmente eficaces para procesar datos estructurados en forma de matriz o grilla, como imágenes, mapas de calor o señales tridimensionales.
- Utilizan filtros o kernels convolucionales que escanean la entrada para detectar características espaciales locales, como bordes, texturas, colores o formas. Esta capacidad de extraer patrones espaciales las hace ideales para tareas visuales.

- Su estructura suele incluir capas de convolución, capas de agrupamiento (pooling) y capas densas totalmente conectadas, que permiten refinar las características extraídas y realizar la clasificación final.
- Son ampliamente utilizadas en visión artificial, como en clasificación de imágenes, detección de objetos, reconocimiento facial, diagnóstico por imágenes médicas y conducción autónoma.
- Más allá del procesamiento de imágenes, las CNN también han sido adaptadas para procesamiento de texto y análisis de señales acústicas, extendiendo su uso a aplicaciones como la detección de emociones en voz y análisis de espectrogramas.

d) Redes Neuronales Profundas (Deep Neural Networks – DNN)

- Las DNN se caracterizan por tener múltiples capas ocultas entre la entrada y la salida, lo que permite que estas redes aprendan representaciones jerárquicas y complejas de los datos.
- Esta profundidad estructural les otorga la capacidad de identificar patrones abstractos y relaciones no evidentes entre variables, siendo fundamentales para resolver problemas complejos de clasificación, predicción o segmentación.
- Constituyen la base del aprendizaje profundo (Deep Learning), campo que ha transformado múltiples industrias gracias a su efectividad en tareas como procesamiento del lenguaje natural (NLP), generación de contenido, reconocimiento de voz y sistemas de recomendación.
- Las DNN requieren una gran cantidad de datos etiquetados y recursos computacionales significativos, como unidades gráficas de procesamiento (GPU), para ser entrenadas de manera eficiente.
- No obstante, su capacidad para superar modelos tradicionales en precisión y adaptabilidad las ha posicionado como una de las tecnologías más poderosas y disruptivas en la actualidad, permitiendo desarrollos como chatbots inteligentes, sistemas de diagnóstico automático y modelos generativos como GANs.

El uso de redes neuronales artificiales se fundamenta en la necesidad de modelar relaciones no lineales entre las variables operativas de las bombas centrífugas y su tiempo de vida útil remanente. En este tipo de equipos, el deterioro no depende de un solo

parámetro, sino de la interacción conjunta entre la vibración, el caudal, la presión y el gradiente térmico, los cuales pueden variar de manera progresiva durante la operación.

En ese sentido, las redes neuronales artificiales permiten procesar datos históricos y reconocer patrones asociados al comportamiento operativo y al desgaste del equipo. A diferencia de los métodos convencionales, este enfoque tiene la capacidad de aprender a partir de los datos y ajustar sus predicciones conforme identifica tendencias de degradación. Por ello, su aplicación resulta pertinente para estimar el tiempo de vida útil remanente de las bombas centrífugas (Noya y Sutizal, 2023). Asimismo, la integración de las redes neuronales con el Análisis de Componentes Principales permite mejorar el tratamiento de la información, ya que el PCA reduce la dimensionalidad de los datos y conserva los componentes más representativos. Posteriormente, la red neuronal utiliza dichos componentes para estimar el RUL, favoreciendo un modelo más ordenado y con menor redundancia en las variables de entrada.

2.2.7.2.Análisis de Componentes Principales

El Análisis de Componentes Principales (PCA) es una técnica estadística perteneciente al campo del análisis multivariante cuya finalidad es reducir la dimensionalidad de conjuntos de datos complejos manteniendo la mayor parte de la información relevante. PCA transforma un conjunto de variables originales, que pueden estar altamente correlacionadas, en un nuevo conjunto de variables no correlacionadas llamadas componentes principales. Cada componente principal representa una combinación lineal de las variables originales, ordenadas de manera que el primer componente captura la mayor parte de la varianza del conjunto de datos, el segundo la mayor varianza residual bajo la restricción de ser ortogonal al primero, y así sucesivamente. Esta técnica es ampliamente utilizada como método de preprocesamiento en modelos predictivos para mejorar el rendimiento y reducir la complejidad del modelo (IBM, s.f.).

Matemáticamente, dado un conjunto de datos centrado (las medias de cada variable restadas), representado por la matriz $X \in \mathbb{R}^{n \times p}$, donde n es el número de observaciones y p el número de variables, el PCA se fundamenta en el análisis de la matriz de covarianza:

$$Q = \frac{1}{n-1} X^T X$$

Los componentes principales se obtienen como los autovectores (vectores propios) de la matriz de covarianza Q, y los valores propios asociados representan la cantidad de varianza explicada por cada componente:

$$Q w_k = \lambda_k w_k$$

donde w_k es el k-ésimo vector propio y λ_k su correspondiente valor propio. La suma acumulada de los valores propios λ_k indica la cantidad de varianza total explicada por los primeros componentes principales, lo que permite seleccionar un subconjunto de componentes que retienen la mayor parte de la información original del conjunto de datos. En modelos predictivos, como los usados para estimar la vida útil remanente (RUL) de equipos industriales, PCA se emplea para reducir la dimensionalidad de los datos de entrada antes de aplicar algoritmos de aprendizaje automático, como redes neuronales. Esto reduce la multicolinealidad entre las variables, mitiga el ruido y ayuda a evitar el sobreajuste, permitiendo al modelo enfocar su entrenamiento en las características que aportan mayor información relevante a la predicción (Tsallis et al., 2025)

CAPÍTULO III

MARCO METODOLÓGICO

3.1. Tipo de investigación

La investigación corresponde al tipo aplicado, ya que, según Hernández et al. (2014), su propósito es generar conocimientos que puedan ser utilizados para la resolución de problemas específicos en el ámbito del mantenimiento industrial, como la estimación del tiempo de vida útil remanente de bombas centrífugas. Además, se enmarca dentro del enfoque cuantitativo, porque se basa en la recopilación y análisis de datos operativos históricos para desarrollar y validar el modelo predictivo.

3.2. Nivel de investigación

La investigación corresponde al nivel explicativo, de acuerdo con Hernández et al. (2014), dado que busca analizar la relación entre variables como las condiciones operativas de las bombas centrífugas y su vida útil remanente. Este nivel permite comprender los factores que influyen en el desgaste del equipo y su comportamiento a lo largo del tiempo, contribuyendo a una mejor planificación del mantenimiento. En este caso, el estudio busca identificar y explicar los factores que influyen en el deterioro de los equipos y cómo estos pueden ser modelados matemáticamente para predecir el tiempo restante de funcionamiento eficiente.

3.3. Diseño de investigación

El diseño de investigación que se adopta en este estudio es no experimental, longitudinal y explicativo, en concordancia con lo propuesto por Hernández et al. (2014).

Se trata de un diseño no experimental porque no se manipularán deliberadamente las variables independientes, sino que se observarán y analizarán datos operativos históricos ya existentes o simulados de bombas centrífugas, tal como se obtendrán en el entorno de simulación mediante el software MATLAB R2024b. No se aplicará ningún tratamiento a un grupo experimental o control, sino que se realizará un análisis observacional retrospectivo de registros técnicos. Asimismo, el diseño es de tipo longitudinal ya que el análisis de datos se realiza considerando el comportamiento de las bombas a lo largo del tiempo, permitiendo estudiar su evolución operativa, desgaste

progresivo y frecuencia de fallas, lo cual es fundamental para estimar el tiempo de vida útil remanente (RUL). Esta característica temporal es clave para aplicar técnicas como el Análisis de Componentes Principales (PCA) y el entrenamiento de modelos de predicción con redes neuronales artificiales.

Finalmente, el diseño se clasifica como explicativo, ya que busca no solo describir el comportamiento de las variables operativas (como caudal, presión, temperatura y vibración), sino también explicar cómo se relacionan entre sí y cómo estas relaciones permiten predecir el deterioro del equipo. Este nivel de diseño tiene como propósito establecer vínculos causales entre los patrones operativos detectados y la estimación de la vida útil remanente, a través de un modelo predictivo validado con métricas estadísticas como el Error Medio Cuadrático (MSE) y el Coeficiente de Determinación (R^2)..

3.4. Método de Investigación

La presente investigación se sustenta en el método hipotético-deductivo, el cual parte de la formulación de una hipótesis principal y específicas que serán contrastadas mediante la observación sistemática y el análisis de datos cuantitativos. Este método permite explicar fenómenos a partir de la deducción lógica de consecuencias empíricas verificables, lo que resulta adecuado para el desarrollo y validación de un modelo predictivo basado en técnicas estadísticas y computacionales (Hernández et al., 2014).

En primer lugar, se identifican los parámetros operativos relevantes de las bombas centrífugas (caudal, presión, temperatura, vibración), los cuales se recopilan desde registros históricos de funcionamiento. A partir de estos datos, se realiza una depuración y procesamiento utilizando herramientas de análisis multivariado, específicamente el Análisis de Componentes Principales (PCA), que permite reducir la dimensionalidad del conjunto de datos y resaltar las variables más representativas. Posteriormente, se emplea una Red Neuronal Artificial (RNA) como técnica de modelado predictivo, entrenada con los datos procesados para generar estimaciones del tiempo de vida útil remanente de las bombas. El desempeño del modelo se evalúa utilizando métricas estadísticas como el Error Medio Cuadrático (MSE) y el coeficiente de determinación (R^2), lo cual permite validar su precisión y capacidad de generalización.

3.5. Población y muestra

3.5.1. Población

La población de estudio está constituida por el conjunto de registros operativos históricos de bombas centrífugas utilizadas en procesos industriales, en los cuales se dispone de variables técnicas relevantes para el mantenimiento predictivo, tales como vibración, caudal, presión y temperatura. Esta población comprende equipos que operan bajo condiciones continuas y que cuentan con sistemas de monitoreo que permiten el análisis de su comportamiento operativo a lo largo del tiempo. Conforme a Hernández et al. (2014), la población corresponde al total de unidades o elementos sobre los cuales se desea realizar inferencias en función de los objetivos de la investigación.

3.5.2. Muestra

La muestra estuvo conformada por tres bombas centrífugas, seleccionadas mediante un muestreo no probabilístico por conveniencia técnica. El criterio de selección se basó en la disponibilidad y calidad de los registros históricos de datos operativos, condición indispensable para el desarrollo y validación del modelo predictivo PCA–RNA. Las bombas seleccionadas fueron: B01 (KSB Etanorm C 050-125 C10), B02 (Gorman-Rupp 63E31-B) y B03 (INOXPA DIN-FOOD 01.110.33.01), las cuales presentan información operativa suficiente y continua para el análisis del comportamiento, la reducción de dimensionalidad mediante PCA y la estimación del tiempo de vida útil remanente.

3.6. Técnicas e instrumentos de recolección de datos

Para desarrollar la presente investigación, se emplearán técnicas de recolección y análisis de datos alineadas a cada objetivo específico. Estas combinan métodos estadísticos y de aprendizaje automático para estimar con mayor precisión la vida útil remanente de las bombas centrífugas. Para la adquisición de los datos operativos se consideran sensores industriales comúnmente utilizados en sistemas de mantenimiento predictivo de bombas centrífugas. Entre ellos se incluyen sensores de vibración, empleados para identificar desbalance, desalineación y desgaste mecánico en componentes rotativos; sensores de temperatura, utilizados para monitorear el

comportamiento térmico de rodamientos, sellos mecánicos y motor; sensores de presión, que permiten evaluar la eficiencia hidráulica del sistema; y sensores de caudal, orientados a analizar el desempeño operativo y detectar pérdidas de eficiencia.

3.6.1. Técnicas e instrumentos de recolección y análisis de datos para el primer objetivo específico

La caracterización del sistema de bombeo consiste en definir las condiciones técnicas y operativas bajo las cuales se desarrolla el modelo predictivo. Se consideran:

- Variables hidráulicas
- Variables mecánicas
- Variables eléctricas

Objetivo Específico 1:

Analizar el comportamiento operativo de bombas centrífugas sometidas a mantenimiento preventivo para estimar el tiempo de vida útil remanente.

A. Técnicas e instrumentos de recolección de datos:

a) Recolección de datos: Se aplicó una revisión detallada de documentos técnicos y registros operativos con la finalidad de identificar los principales parámetros que influyen en el deterioro de las bombas. Se consideró variables como caudal, presión, vibración y temperatura, las cuales serán analizadas en función de su comportamiento a lo largo del tiempo.

Instrumentos utilizados:

- Informes de mantenimiento preventivo (fallas, reparaciones, intervenciones periódicas).
- Registros de operación continua con datos históricos.
- Fichas técnicas del fabricante con valores nominales y límites de funcionamiento.

B. Análisis estadístico:

Mediante técnicas de estadística descriptiva (media, desviación estándar, coeficiente de variación, tendencias), se evaluará la variabilidad de los datos para identificar la pertinencia de su uso en modelos predictivos. Esto permitirá luego aplicar el PCA con una base sólida.

Software: MATLAB R2024b.

3.6.2. Técnicas e instrumentos de recolección y análisis de datos para el segundo objetivo específico

Objetivo Específico 2:

Desarrollar un modelo predictivo basado en Análisis de Componentes Principales (PCA) y redes neuronales para estimar el tiempo de vida útil remanente.

A. Técnicas e instrumentos de recolección de datos:

a) **Técnica:** Integración y depuración de datos históricos operativos.
b) **Instrumento:** Conjunto de datos previamente organizados.

B. Técnicas de análisis y procesamiento de datos:

a) **Técnicas:**
   - Análisis de Componentes Principales (PCA) para reducción de dimensionalidad.
   - Redes Neuronales Artificiales (RNA) para entrenamiento del modelo.
b) **Software:** Se utilizará MATLAB R2024b para procesar y analizar la información recopilada para el desarrollo del modelo predictivo.

El modelo predictivo que se utiliza en esta investigación es una Red Neuronal Artificial (RNA) del tipo Perceptrón Multicapa (Multilayer Perceptron, MLP), una arquitectura ampliamente utilizada en problemas de regresión no lineal. Esta red consiste en:

- Una capa de entrada que recibe los componentes principales obtenidos tras aplicar el Análisis de Componentes Principales (PCA).
- Una o más capas ocultas, donde cada neurona aplica funciones de activación como ReLU o sigmoide para modelar relaciones complejas entre las variables.
- Una capa de salida con una única neurona que estima el tiempo de vida útil remanente (RUL) como valor continuo.

La red será entrenada mediante el algoritmo backpropagation, utilizando el optimizador Adam y un conjunto de datos de entrenamiento previamente normalizado y

reducido por PCA. Durante el entrenamiento, se ajustarán los pesos y sesgos internos de la red para minimizar el Error Medio Cuadrático (MSE) entre los valores reales y las predicciones. Se seleccionarán de manera iterativa los hiperparámetros más adecuados (número de neuronas, tasa de aprendizaje, número de épocas) con base en el desempeño observado en un conjunto de validación. Finalmente, el modelo entrenado será aplicado a nuevos datos de prueba para evaluar su capacidad predictiva.

C. Preparación de datos para el modelo

Los datos históricos obtenidos serán preprocesados, normalizados y depurados, a fin de garantizar que no existan valores atípicos o vacíos que afecten el entrenamiento del modelo.

- Se aplica el Análisis de Componentes Principales (PCA), lo que permite reducir la cantidad de variables conservando la mayor proporción de información posible.
- Se seleccionan los componentes principales que expliquen más del 90 % de la varianza, optimizando así la entrada de información para la red neuronal y reduciendo el riesgo de sobreajuste.

Además, la construcción de curvas características del sistema experimental sigue el procedimiento: Se registran presión, caudal y potencia en distintos puntos de operación, desde caudal cero hasta el máximo permitido.

D. Construcción del modelo

Se emplea una Red Neuronal Artificial del tipo perceptrón multicapa (MLP), entrenada con los componentes principales.

- Se seleccionan hiperparámetros como número de capas ocultas, número de neuronas, función de activación (ReLU o sigmoide), tasa de aprendizaje y algoritmo de optimización (Adam).
- El entrenamiento se realiza mediante backpropagation, ajustando iterativamente los pesos y sesgos.

E. Descripción del sistema del banco experimental

El sistema experimental de bombeo se encuentra configurado con instrumentación capaz de medir en tiempo real los parámetros operativos críticos: caudal

(Q), presión (P), vibración (V) y temperatura (T). Estos datos constituyen la base para el desarrollo del algoritmo del modelo predictivo orientado a estimar el tiempo de vida útil remanente (RUL) de las bombas centrífugas. Los valores capturados serán procesados en el software MATLAB R2024b, donde se aplica la siguiente metodología:

1. **Adquisición de datos:**
   Los registros de caudal, presión, vibración y temperatura se almacenan en formato compatible con MATLAB para su procesamiento automatizado.
2. **Preprocesamiento y limpieza:**
   - Filtrado de datos ruidosos y tratamiento de valores atípicos.
   - Normalización de cada parámetro para evitar escalas desiguales en el modelo.
3. **Aplicación de Análisis de Componentes Principales (PCA):**
   - Los cuatro parámetros serán transformados en componentes principales, reduciendo la dimensionalidad del sistema y concentrando la información relevante en un menor número de variables.
   - Esta reducción mejora la eficiencia computacional y elimina redundancias entre datos altamente correlacionados, como presión y caudal.
4. **Entrenamiento del modelo predictivo:**
   - Los componentes principales servirán como entradas para una Red Neuronal Artificial (RNA) de tipo perceptrón multicapa (MLP).
   - El modelo se entrena para correlacionar las condiciones operativas con la tendencia de desgaste y la estimación del RUL.
5. **Validación y predicción:**
   - Se aplican métricas estadísticas como Error Medio Cuadrático (MSE) y Coeficiente de Determinación (R^2) para evaluar el rendimiento del modelo.
   - Los resultados permitieron generar curvas de predicción del tiempo de vida útil remanente en distintas condiciones de operación.

3.6.3. Técnicas e instrumentos de recolección y análisis de datos para el tercer objetivo específico

Objetivo Específico 3:

Aplicar y evaluar el modelo predictivo mediante métricas de precisión y error.

A. Técnicas e instrumentos de recolección de datos:

a) Técnica:

- Evaluación del modelo predictivo: Mediante métricas como el error medio cuadrático (MSE), coeficiente de determinación (R^2) y precisión de las predicciones.

b) Instrumento:

- Conjunto de datos de prueba (valores históricos no usados en el entrenamiento).
- Resultados de salida generados por el modelo predictivo tras su aplicación.

B. Técnicas de análisis y procesamiento de datos:

a) Técnicas:

Evaluación mediante métricas estadísticas:

- Error Medio Cuadrático (MSE)
- Coeficiente de determinación (R^2)

b) Software: MATLAB R2024b para simulación y evaluación del modelo predictivo.

C. Evaluación del modelo:

Una vez entrenado, el modelo es validado con un conjunto de datos no utilizados durante el entrenamiento.

- Se analizan métricas como Error Medio Cuadrático (MSE), coeficiente de determinación (R^2), Error Absoluto Medio (MAE) y precisión.
- Se generan gráficos de dispersión entre valores iniciales y predichos, y se evalúa la dispersión del error.

D. Interpretación y relevancia:

Los resultados obtenidos del modelo permiten estimar con alto grado de confiabilidad la vida útil remanente de las bombas. Esto representa un aporte valioso para la gestión del mantenimiento, ya que permite programar intervenciones con anticipación,

reduciendo el riesgo de fallas imprevistas y optimizando los recursos disponibles.

3.7. Análisis estadístico

El análisis estadístico constituye una etapa fundamental en la presente investigación, ya que permite extraer, procesar e interpretar de manera objetiva la información operativa histórica de las bombas centrífugas. Este análisis no solo proporciona una comprensión preliminar de los datos, sino que también sirve de base para la reducción de dimensionalidad, la selección de variables significativas y la validación del modelo predictivo. Para cumplir con los objetivos específicos del estudio, se emplearán dos tipos de análisis estadístico: descriptivo e inferencial, integrados con herramientas de aprendizaje automático.

A. Análisis estadístico descriptivo

El análisis descriptivo permite caracterizar el comportamiento de las variables operativas (caudal, presión, vibración, temperatura, entre otras) mediante medidas como:

- **Media aritmética**: para estimar el valor promedio de cada parámetro.
- **Desviación estándar** y **varianza**: para evaluar la dispersión de los datos.
- **Coeficiente de variación**: para comparar la variabilidad relativa entre variables.
- **Análisis de correlación**: para identificar relaciones entre variables y determinar posibles redundancias.

Este análisis preliminar facilita la detección de valores atípicos, datos faltantes o inconsistencias, lo cual es esencial antes de aplicar técnicas más avanzadas como el Análisis de Componentes Principales (PCA).

3.7.1. Análisis estadístico multivariado con PCA

Para esta investigación se emplea el Análisis de Componentes Principales (PCA, por sus siglas en inglés) como técnica multivariada destinada a reducir la dimensionalidad del conjunto de datos operativos históricos de las bombas centrífugas, conservando la mayor cantidad posible de información relevante. Esta técnica permite identificar patrones ocultos, disminuir la redundancia de variables correlacionadas y facilitar el entrenamiento eficiente del modelo predictivo.

El procedimiento consistirá en:

- Estandarización de las variables operativas (caudal, presión, temperatura y vibración) mediante normalización z-score, con el fin de garantizar que todas contribuyan de manera equitativa al análisis.
- Cálculo de la matriz de correlación entre las variables, para evidenciar las relaciones lineales existentes y justificar el uso del PCA.
- Cálculo de autovalores y autovectores, que permitirán generar los componentes principales como combinaciones lineales de las variables originales.
- Selección de componentes principales mediante el criterio de varianza explicada acumulada, priorizando aquellos que representen conjuntamente al menos el 90 % de la varianza total del sistema.
- Análisis gráfico de sedimentación (scree plot), para identificar el punto de corte óptimo en función del cambio de pendiente en los autovalores.
- Interpretación de las cargas factoriales de cada componente, para asociar los patrones hallados con las variables originales más influyentes (por ejemplo, si un componente representa mayoritariamente el comportamiento de la vibración o del caudal).
- Proyección de los datos originales en el nuevo espacio reducido, que servirá como entrada para el entrenamiento del modelo de redes neuronales artificiales.

El número de componentes principales retenidos se definió mediante un criterio combinado que consideró la varianza explicada acumulada, el análisis del gráfico de sedimentación y la interpretabilidad física de los componentes. Se seleccionaron aquellos que explicaron conjuntamente al menos el 90% de la varianza total, identificando mediante el scree plot el punto a partir del cual la incorporación de nuevos componentes aportaba una ganancia marginal de información. Este criterio permitió equilibrar la reducción de la dimensionalidad con la estabilidad del modelo y la adecuada representación de los procesos de degradación. Asimismo, el uso del Análisis de Componentes Principales resulta apropiado debido a la naturaleza multivariada y correlacionada de las variables operativas, ya que integra información de caudal, presión, vibración y temperatura, reduce la redundancia y el ruido, y mejora la eficiencia computacional del modelo frente a enfoques univariados basados en señales individuales.

3.7.2. Validación estadística del modelo predictivo

Una vez entrenado el modelo con los datos procesados mediante Análisis de Componentes Principales (PCA), se procede a su validación y análisis estadístico a lo largo de las distintas fases del desarrollo. El proceso estará compuesto por cuatro etapas: recopilación y preprocesamiento de datos, aplicación del PCA, entrenamiento del modelo predictivo, y validación estadística. A continuación, se detallan:

a) Recopilación y preprocesamiento de datos:

Se recopilan datos operativos históricos o simulados de bombas centrífugas, incluyendo variables como caudal, presión, vibración y temperatura. Esta información será sometida a las siguientes acciones:

- Detección y eliminación de valores atípicos.
- Imputación de datos faltantes.
- Normalización mediante z-score para homogeneizar la escala de los datos.
- Visualización inicial mediante diagramas de líneas y diagramas de dispersión para analizar la evolución de cada variable a lo largo del tiempo.

b) Aplicación del Análisis de Componentes Principales (PCA):

Se aplica PCA para reducir la dimensionalidad del conjunto de datos, transformando las variables originales en un conjunto de componentes principales no correlacionados. Para ello:

- Se calculan autovalores y varianza explicada acumulada.
- Se seleccionan los componentes que acumulen al menos el 90 % de la varianza total.
- Se interpretan las cargas factoriales para identificar qué variables dominan cada componente.
- Se generan gráficos de sedimentación y diagramas de barras para visualizar el aporte de cada componente.

c) Desarrollo del modelo predictivo:

Los componentes principales seleccionados se utilizan como entrada de una Red Neuronal Artificial (RNA) tipo perceptrón multicapa (MLP), entrenada con los siguientes criterios:

- Uso de función de activación ReLU o sigmoide.
- Algoritmo de optimización Adam.
- Tasa de aprendizaje y número de épocas definidos empíricamente.
- Entrenamiento y validación cruzada con conjuntos separados.

Durante esta etapa se utilizarán gráficos de líneas para visualizar la evolución del error por época, así como histogramas para observar la distribución de los errores y los residuos del modelo.

d) Validación y ajuste del modelo:

El desempeño del modelo se evalúa aplicando las siguientes métricas:

- **Error Medio Cuadrático (MSE)**: cuantifica la desviación promedio al cuadrado entre los valores iniciales y los predichos.
- **Error Absoluto Medio (MAE)**: mide el promedio de las diferencias absolutas entre las predicciones y los datos iniciales.
- **Coeficiente de Determinación (R^2)**: indica el grado de ajuste del modelo predictivo a los datos observados.
- **Porcentaje de error relativo**: expresa la relación entre el error absoluto y el valor real, en forma porcentual.

3.8. Interpretación y relevancia de los resultados del modelo predictivo

La interpretación de los resultados obtenidos con el modelo predictivo desarrollado en esta investigación se realiza en función del desempeño alcanzado en la estimación del tiempo de vida útil remanente (RUL) de bombas centrífugas. Para ello, se utilizan métricas de evaluación como el Error Medio Cuadrático (MSE), el Coeficiente de Determinación (R^2), el Error Absoluto Medio (MAE) y gráficos de dispersión que comparen los valores de registros iniciales y predichos. La relevancia de estos resultados radica en su capacidad para evidenciar el grado de precisión del modelo y su utilidad como herramienta de apoyo en la toma de decisiones relacionadas con el mantenimiento predictivo. Por ejemplo, un bajo MSE y un R^2 cercano a 1 indicarán una alta concordancia

entre las predicciones del modelo y los datos iniciales, validando su aplicación en escenarios de operación.

Además, el modelo permite identificar tendencias de deterioro antes de que ocurran fallas críticas, contribuyendo a:

- Optimizar la planificación del mantenimiento, programando intervenciones preventivas más efectivas.
- Reducir tiempos de inactividad no planificados.
- Prolongar la vida útil de los componentes de las bombas.
- Minimizar los costos operativos asociados a fallas inesperadas.
- Mejorar la confiabilidad del sistema en contextos industriales exigentes como minería, petróleo o tratamiento de aguas.

En términos académicos y técnicos, los resultados del modelo demouestran la viabilidad de aplicar herramientas de análisis multivariado (PCA) y redes neuronales artificiales en la estimación del RUL, marcando un aporte significativo en el campo del mantenimiento inteligente de equipos rotativos.

3.9. Consideraciones del modelo predictivo

En el desarrollo del modelo predictivo para estimar el tiempo de vida útil remanente en bombas centrífugas se han adoptado una serie de consideraciones metodológicas, técnicas y computacionales que aseguran la coherencia y validez del proceso de modelado. Estas consideraciones son las siguientes:

- **Naturaleza de los datos**: Se emplean datos operativos históricos (simulados) que contengan registros de variables como caudal, presión, temperatura y vibración, en intervalos temporales regulares. Se asume que dichos datos representan con fidelidad el comportamiento del equipo durante su ciclo de vida.
- **Normalización previa al análisis**: Para garantizar la comparabilidad entre variables con diferentes unidades y escalas, todos los datos serán normalizados utilizando el método z-score antes de aplicar el Análisis de Componentes Principales (PCA).
- **Reducción de dimensionalidad**: El PCA permite seleccionar únicamente los componentes principales que expliquen al menos el 90 % de la varianza total del

sistema, con el fin de reducir la complejidad del modelo y mejorar la eficiencia del entrenamiento de la red neuronal.

- **Tipo de red neuronal**: Se utiliza una red neuronal artificial (RNA) del tipo perceptrón multicapa (MLP), debido a su capacidad para resolver problemas de regresión no lineales y modelar relaciones complejas entre variables reducidas y la variable de salida (tiempo de vida útil estimado).
- **Criterios de entrenamiento**: El modelo es entrenado utilizando el algoritmo de retropropagación, con funciones de activación ReLU o sigmoide, tasa de aprendizaje adecuada y validación cruzada para evitar sobreajuste. La selección de hiperparámetros (capas, neuronas, épocas) se hará en función del rendimiento en un conjunto de validación.
- **Evaluación del rendimiento**: La precisión del modelo es evaluada mediante métricas como el Error Medio Cuadrático (MSE), el Error Absoluto Medio (MAE), el Coeficiente de Determinación (R^2) y análisis gráfico de dispersión entre valores iniciales y predichos.
- **Entorno de implementación**: Todo el proceso es ejecutado en un entorno de simulación con MATLAB R2024b, utilizando sus librerías de Machine Learning y Deep Learning para el procesamiento, visualización, entrenamiento y validación del modelo.

El modelo predictivo presenta robustez frente a variaciones en el comportamiento operativo de las bombas centrífugas, independientemente de su marca, siempre que operen dentro de rangos similares de caudal, presión, temperatura y vibración a los considerados en el entrenamiento. Sin embargo, su capacidad de generalización depende de la cantidad y calidad de los datos históricos disponibles y se ve limitada ante condiciones de operación no representadas, por lo que su aplicación en otros sistemas requiere procesos de recalibración o reentrenamiento.

CAPÍTULO IV

RESULTADOS Y DISCUSIÓN

4.1. Análisis del comportamiento operativo de bombas centrífugas

El presente apartado describe el análisis descriptivo y temporal del comportamiento operativo de las bombas centrífugas objeto de estudio, con el fin de identificar patrones de operación, signos tempranos de degradación y condiciones anómalas que sirvan de insumo al desarrollo del modelo predictivo. Para ello se emplea el catálogo técnico de las unidades instaladas (características hidráulicas, materiales, tipos de sello y parámetros nominales) y las series temporales de variables operativas caudal, presión, vibración RMS, temperaturas y corriente recolectadas durante el periodo evaluado. Los resultados incluyen la visualización de tendencias temporales, la cuantificación de variabilidad y la localización de eventos fuera de rango que pudieran asociarse a desgaste, cavitación o problemas en sellos y rodamientos, lo que posibilita una primera clasificación del estado de cada bomba. Este análisis exploratorio se realiza en el entorno de simulación y procesamiento establecido para la investigación, y prepara las transformaciones y la selección de variables que posteriormente serán sometidas a Análisis de Componentes Principales (PCA) y modelado mediante redes neuronales para la estimación del tiempo de vida útil remanente.

La Tabla 3 presenta el catálogo técnico de las tres bombas centrífugas analizadas en el sistema. Estas unidades presentan configuraciones y capacidades hidráulicas distintas, lo que permite comparar su desempeño operativo bajo diferentes condiciones de caudal y altura manométrica. La bomba B01 (KSB Etanorm C 050-125 C10) está diseñada para manejar mayores volúmenes de flujo con eficiencia del 77.1 % y una potencia nominal de 4.01 kW, orientada a procesos de recirculación o impulsión de caudal medio. La B02 (Gorman-Rupp 63E31-B) destaca por su velocidad de operación de 3600 rpm y su mayor presión máxima de 12 bar, adecuada para operaciones de impulsión en líneas de descarga extendidas. Por su parte, la B03 (INOXPA DIN-FOOD 01.110.33.01) corresponde a una bomba sanitaria fabricada íntegramente en acero inoxidable AISI 316, empleada en sistemas donde se requiere compatibilidad alimentaria o resistencia a la corrosión.

Tabla 2

Catálogo técnico de bombas centrífugas instaladas

id_bomba	B01	B02	B03
Marca	KSB	Gorman-Rupp	INOXPA
Modelo	Etanorm C 050-125 C10	63E31-B	DIN-FOOD 01.110.33.01
Q_nom_m3h	76	30	20
H_nom_m	15	20	10
rpm	2929	3600	3000
Potencia_kW	4.01	5	2.5
Diam_succion_mm	65	102	50
Diam_descarga_mm	50	76	40
Material_carcasa	Hierro fundido	Hierro dúctil	AISI 316
Material_impulsor	Acero aleado	Aluminio	AISI 316
Tipo_sello	Sello mecánico simple	Sello mecánico doble	Sello sanitario
Eficiencia_pct	77.1	73	75
NPSHr_m	2.5	3	2
PresionMax_bar	10	12	8

Nota. Elaboración Propia

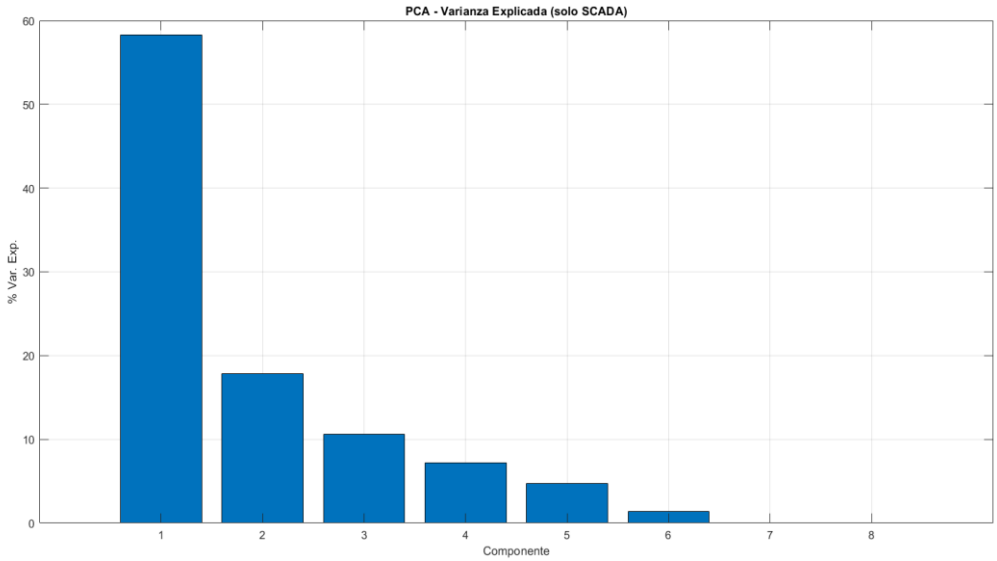
l Análisis de Componentes Principales (PCA) se aplicó a un conjunto de 36 variables operativas obtenidas a partir de señales temporales del sistema SCADA de las bombas centrífugas. Previamente a la aplicación del PCA, las variables fueron sometidas a un proceso de normalización mediante el método z-score, con el objetivo de eliminar el efecto de las diferencias de escala y magnitud entre las señales medidas (caudal, presión, vibración, temperatura, entre otras). La normalización z-score se realizó transformando cada variable para que presente media cero y desviación estándar unitaria, garantizando que todas las señales aporten de manera equilibrada al análisis multivariado. Este procedimiento resulta especialmente adecuado para el PCA, dado que evita que variables con mayores rangos numéricos dominen la estructura de varianza y distorsionen la identificación de los componentes principales. En consecuencia, todas las figuras y análisis asociados al PCA se basan exclusivamente en datos normalizados mediante z-score. Una vez normalizadas las variables, se ejecutó el PCA sobre la matriz de datos estandarizada. Los resultados indican que el primer componente principal (PC1) explica

aproximadamente el 58.3 % de la varianza total, mientras que la combinación de PC1 y PC2 acumula cerca del 76.1 %, evidenciando que los dos primeros componentes concentran la mayor parte de la información estructural del sistema. Este nivel de varianza acumulada permite una representación reducida del comportamiento dinámico de las bombas, preservando las relaciones fundamentales entre las variables operativas. En este contexto, el espacio bidimensional definido por PC1 y PC2, mostrado en las proyecciones gráficas correspondientes, se emplea exclusivamente con fines exploratorios y de visualización, facilitando la identificación de patrones generales de operación y posibles transiciones entre estados normales y anómalos. No obstante, para el desarrollo del modelo predictivo PCA–RNA, se emplearon cuatro componentes principales (PC1–PC4), seleccionados en función del criterio de varianza acumulada superior al 90 %, conforme a lo detallado en la sección metodológica.

La Figura 5 presenta la distribución de la varianza explicada por cada componente principal, obtenida a partir del PCA aplicado a las variables del sistema SCADA previamente normalizadas mediante z-score. Se observa que el PC1 concentra cerca del 58 % de la varianza, seguido del PC2 con aproximadamente 18 %, mientras que las componentes restantes aportan porcentajes decrecientes.

Figura 5

Varianza PCA



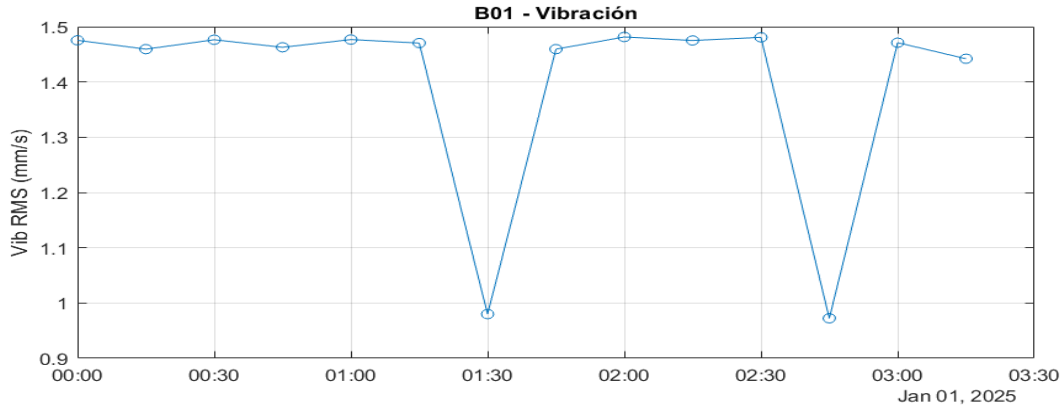
Nota. Elaboración Propia

Los criterios de severidad de vibración empleados como referencia en la presente investigación se sustentan en la norma ISO 10816, la cual establece límites de vibración global para máquinas rotativas industriales. Cabe precisar que dichos criterios se utilizan únicamente como marco de referencia técnica para la interpretación del estado vibracional, sin constituir un análisis normativo exhaustivo de aceptación o rechazo del equipo.

La Figura 6 muestra la evolución temporal de la vibración RMS registrada en la bomba B01 (Ver Anexo N°2), durante un intervalo de operación continua. Se observa un comportamiento mayormente estable, con valores promedio cercanos a 1.45 mm/s, lo cual se encuentra dentro del rango aceptable de funcionamiento según las normas ISO 10816 para maquinaria rotativa. Sin embargo, se identifican dos descensos bruscos en la señal, próximos a las 01:30 y 03:00 horas, donde la vibración cae por debajo de 1.0 mm/s, indicando posibles interrupciones momentáneas del registro o periodos de parada del equipo. Este análisis permite caracterizar el patrón vibracional típico de la bomba y constituye una referencia para la detección de desviaciones futuras.

Figura 6

Comportamiento temporal de la vibración RMS en la bomba B01

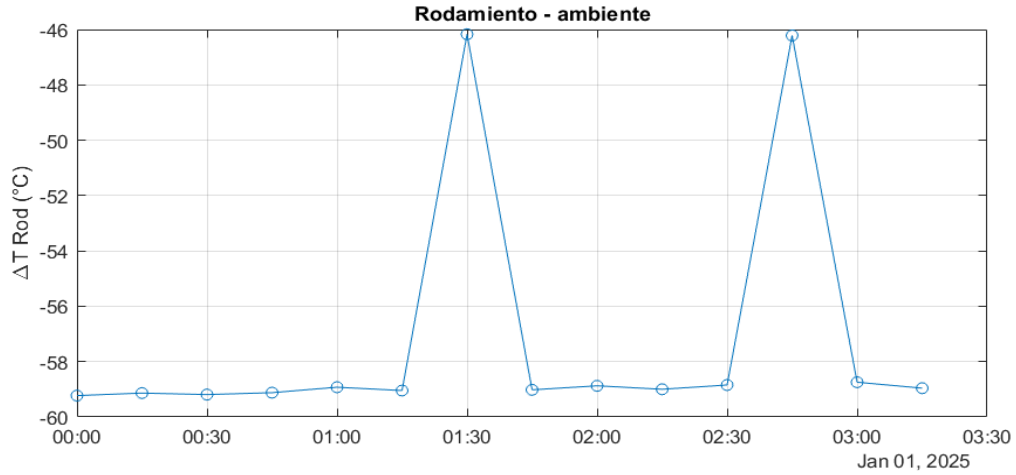


Nota. Elaboración Propia a partir del Anexo N°2.

La Figura 7 ilustra el comportamiento del gradiente térmico (ΔT) entre la temperatura del rodamiento y la temperatura ambiente en la bomba B01 a lo largo del periodo monitoreado. Se aprecia que, en condiciones normales de operación, el diferencial térmico se mantiene estable alrededor de -59 °C, lo que indica un equilibrio térmico adecuado entre los componentes mecánicos y el entorno.

Figura 7

Gradiente térmico entre el rodamiento y el ambiente en la bomba B01

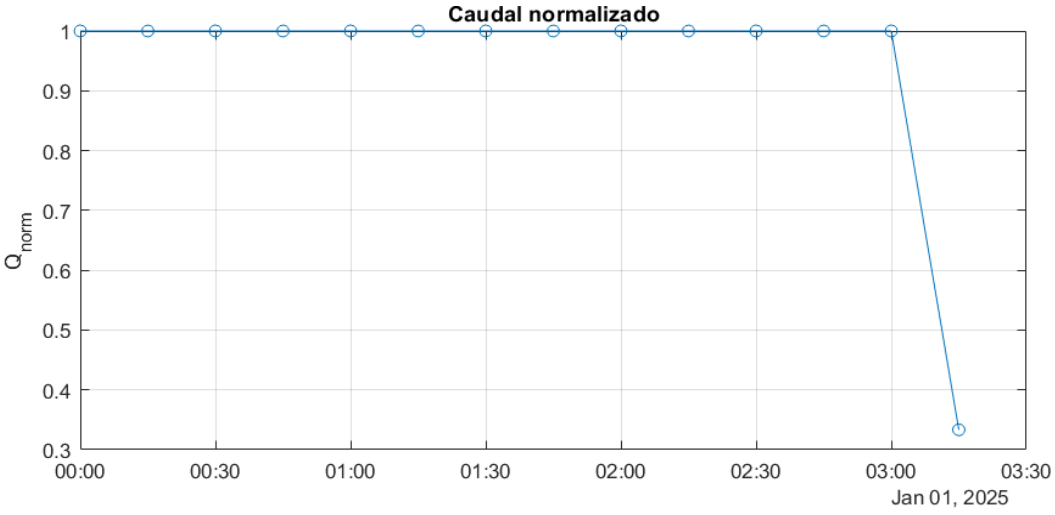


Nota. Elaboración Propia a partir del Anexo N°2.

La Figura 8 muestra la variación temporal del caudal normalizado (Q_norm) en la bomba B01 durante el periodo de observación. En general, el flujo se mantiene estable y cercano a su valor nominal ($Q_norm \approx 1$) a lo largo del intervalo de funcionamiento, lo que indica un desempeño hidráulico constante y sin pérdidas de rendimiento significativas. Sin embargo, se evidencia una caída abrupta en torno a las 03:10 horas, momento en el que el caudal desciende hasta aproximadamente 0.35, sugiriendo una detención momentánea o una disminución de la carga hidráulica.

Figura 8

Comportamiento temporal del caudal normalizado en la bomba B01

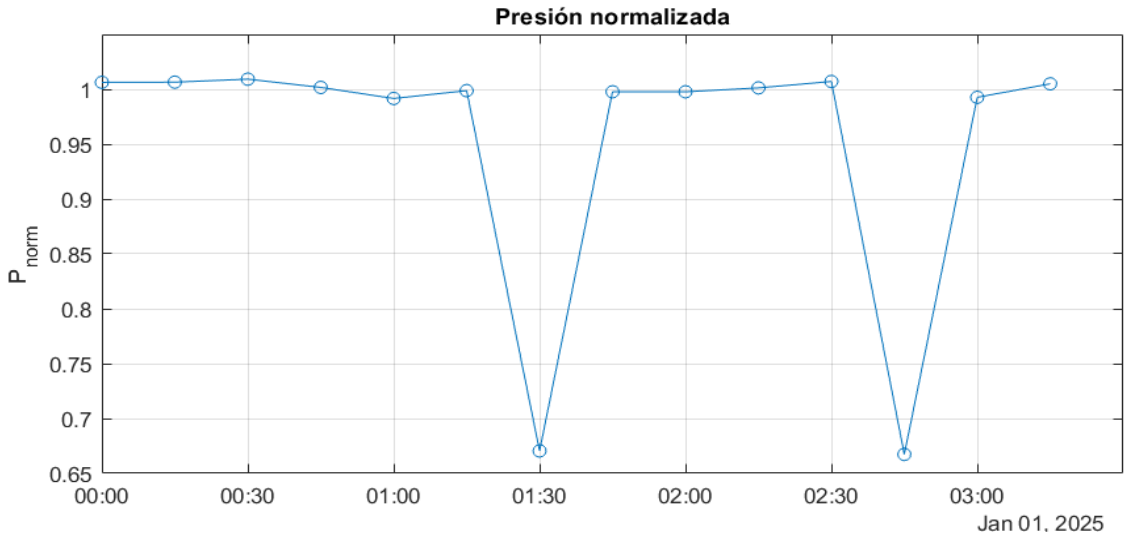


Nota. Elaboración Propia a partir del Anexo N°2.

La Figura 9 representa el comportamiento temporal de la presión normalizada (P_norm) en la bomba B01 a lo largo del periodo monitoreado. En condiciones normales, la presión se mantiene estable con valores próximos a la unidad, reflejando un funcionamiento hidráulico equilibrado y dentro de los parámetros nominales del sistema. No obstante, se observan dos caídas pronunciadas alrededor de las 01:30 y 02:45 horas, donde la presión desciende hasta valores cercanos a 0.68, coincidiendo con las anomalías detectadas en el caudal y la vibración. Estos descensos podrían estar asociados a eventos transitorios como paradas breves o pérdidas momentáneas de carga, lo cual sugiere una correlación entre las variables de presión, caudal y vibración.

Figura 9

Variación temporal de la presión normalizada en la bomba B01

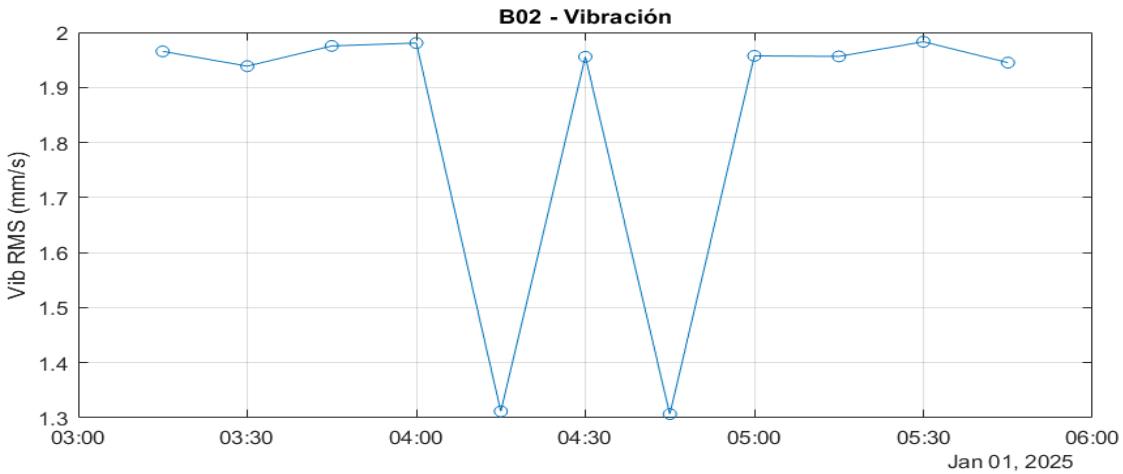


Nota. Elaboración Propia a partir del Anexo N°2.

La Figura 10 presenta la evolución temporal de la vibración RMS registrada en la bomba B02 durante el periodo de monitoreo. En la mayor parte del intervalo, los valores se mantienen estables alrededor de 1.95 mm/s, lo que refleja un desempeño mecánico adecuado dentro de los límites permisibles establecidos por la norma ISO 10816 para equipos rotativos. Sin embargo, se identifican dos descensos abruptos en torno a las 04:15 y 05:00 horas, donde la amplitud vibracional cae hasta aproximadamente 1.3 mm/s, lo cual sugiere interrupciones breves o variaciones en la carga del sistema. Estas fluctuaciones podrían estar relacionadas con transiciones operativas, cambios de régimen hidráulico o eventos de parada parcial del motor.

Figura 10

Comportamiento temporal de la vibración RMS en la bomba B02

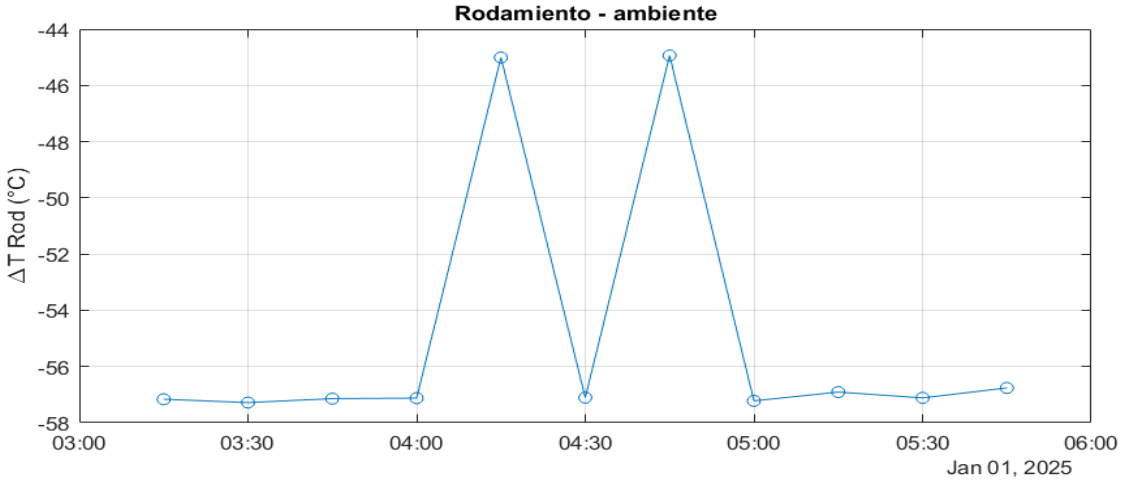


Nota. Elaboración Propia a partir del Anexo N°2.

La Figura 11 muestra la evolución del gradiente térmico (ΔT) entre la temperatura del rodamiento y la del ambiente en la bomba B02. Se aprecia que la mayor parte del tiempo el diferencial térmico se mantiene constante, con valores próximos a -57 °C, lo que sugiere un funcionamiento térmico estable y una adecuada disipación de calor. Sin embargo, se evidencian dos picos notables alrededor de las 04:10 y 04:50 horas, donde la diferencia se reduce abruptamente hasta -45 °C.

Figura 11

Gradiente térmico entre el rodamiento y el ambiente en la bomba B02

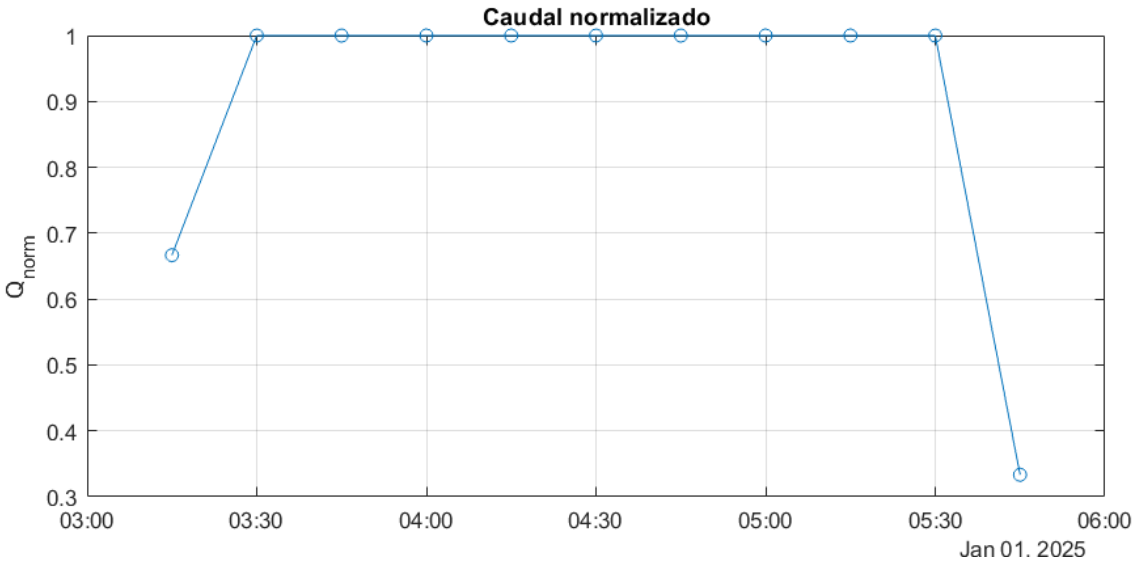


Nota. Elaboración Propia a partir del Anexo N°2.

La Figura 12 presenta la variación temporal del caudal normalizado (Q_norm) en la bomba B02 durante el periodo de operación registrado. En la mayor parte del intervalo, el caudal se mantiene estable y cercano a su valor nominal ($Q_norm \approx 1$), lo que evidencia una adecuada estabilidad hidráulica del sistema. Sin embargo, se identifican dos eventos de caída pronunciada: el primero al inicio del registro, donde el caudal parte de un valor reducido ($Q_norm \approx 0.68$) antes de estabilizarse, y el segundo alrededor de las 05:40 horas, cuando desciende abruptamente hasta 0.35, indicando una posible interrupción de flujo o una parada momentánea del equipo. Estas variaciones coinciden temporalmente con los descensos observados en las señales de vibración y temperatura del rodamiento, sugiriendo la ocurrencia de eventos transitorios comunes.

Figura 12

Comportamiento temporal del caudal normalizado en la bomba B02



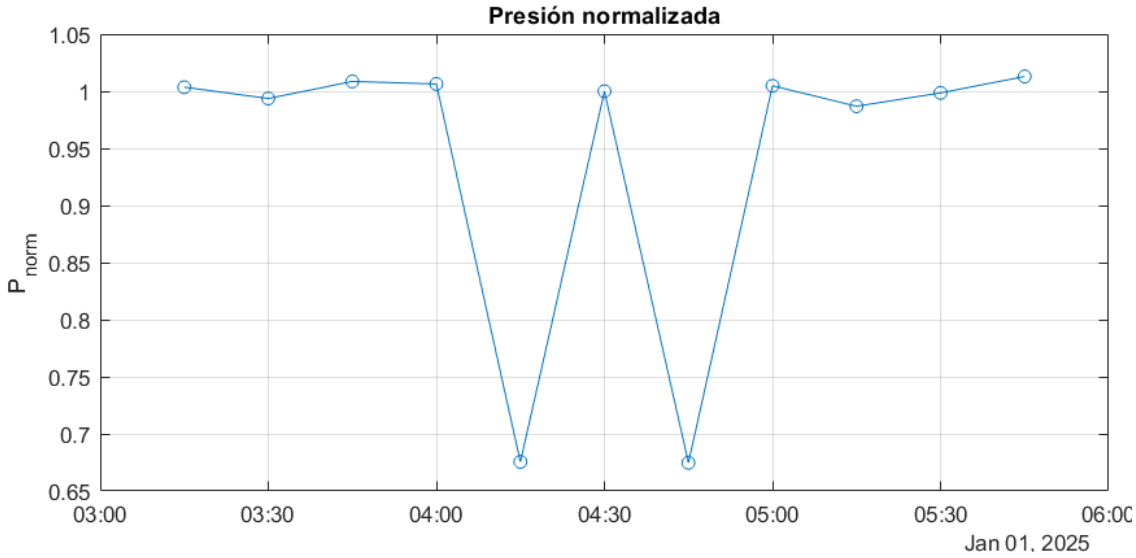
Nota. Elaboración Propia a partir del Anexo N°2.

La Figura 13 muestra la evolución de la presión normalizada (P_norm) en la bomba B02 durante el periodo de monitoreo. En condiciones normales, la presión se mantiene cercana a la unidad, reflejando un régimen hidráulico estable y un adecuado desempeño del sistema. No obstante, se observan dos caídas marcadas alrededor de las 04:15 y 04:50 horas, donde la presión desciende hasta valores aproximados de 0.68, evidenciando interrupciones temporales o disminuciones súbitas de carga en la línea de impulsión.

Estas variaciones coinciden con los descensos simultáneos en el caudal y las fluctuaciones térmicas del rodamiento, lo cual sugiere la presencia de eventos transitorios asociados a paradas breves o a cambios de operación en la bomba.

Figura 13

Variación temporal de la presión normalizada en la bomba B02



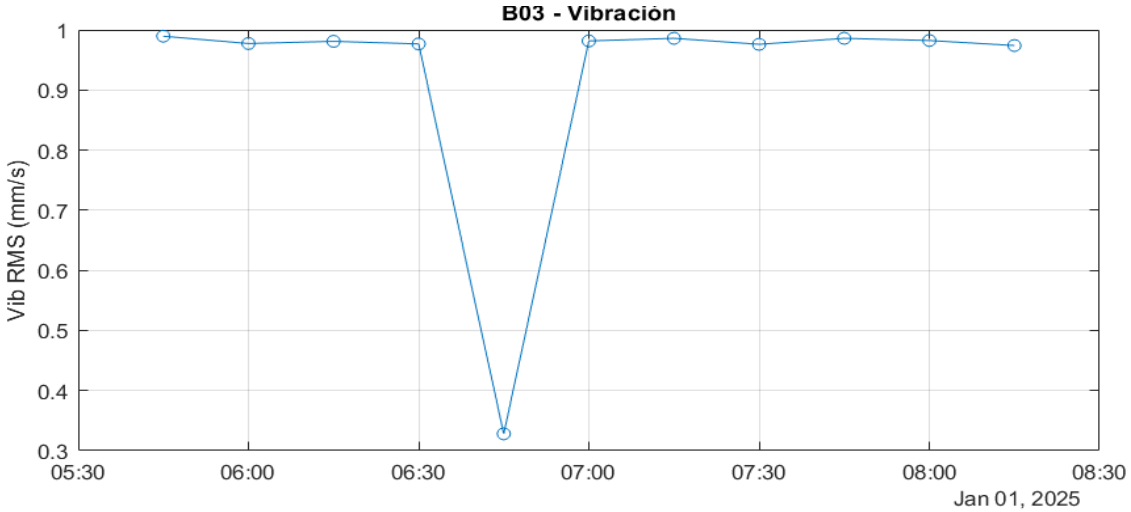
Nota. Elaboración Propia a partir del Anexo N°2.

La Figura 14 presenta la evolución temporal de la vibración RMS registrada en la bomba B03 durante el intervalo de monitoreo. En la mayor parte del periodo, los valores se mantienen cercanos a 1.0 mm/s, lo que indica un funcionamiento estable y dentro de los niveles permisibles establecidos por la norma ISO 10816 para equipos de pequeña potencia. Sin embargo, se evidencia un descenso abrupto alrededor de las 06:40 horas, donde la amplitud vibracional cae a aproximadamente 0.35 mm/s, sugiriendo una posible interrupción temporal o una reducción momentánea en la velocidad de operación.

Tras este evento, la señal recupera su nivel habitual sin indicios de incremento sostenido en la vibración, lo que descarta la presencia de desequilibrio o desalineación significativa. Este comportamiento confirma un régimen dinámico estable en la bomba B03 y refuerza su papel como referencia comparativa frente a las otras unidades del sistema, particularmente en la fase de modelado predictivo y análisis de degradación.

Figura 14

Comportamiento temporal de la vibración RMS en la bomba B03

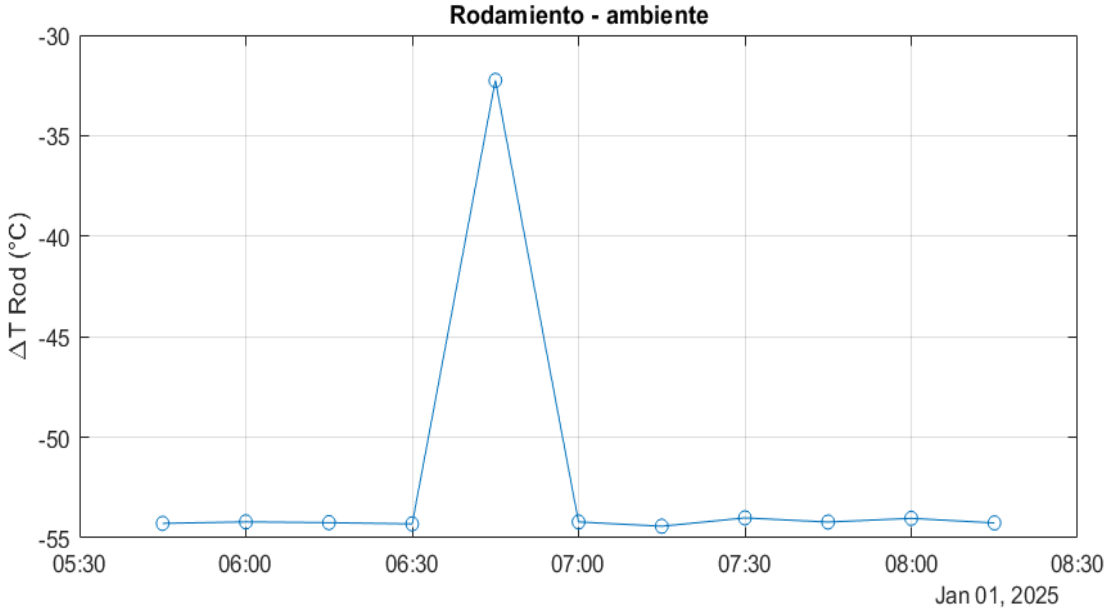


Nota. Elaboración Propia a partir del Anexo N°2.

La Figura 15 muestra la variación temporal del gradiente térmico (ΔT) entre la temperatura del rodamiento y la temperatura ambiente en la bomba B03.

Figura 15

Gradiente térmico entre el rodamiento y el ambiente en la bomba B03



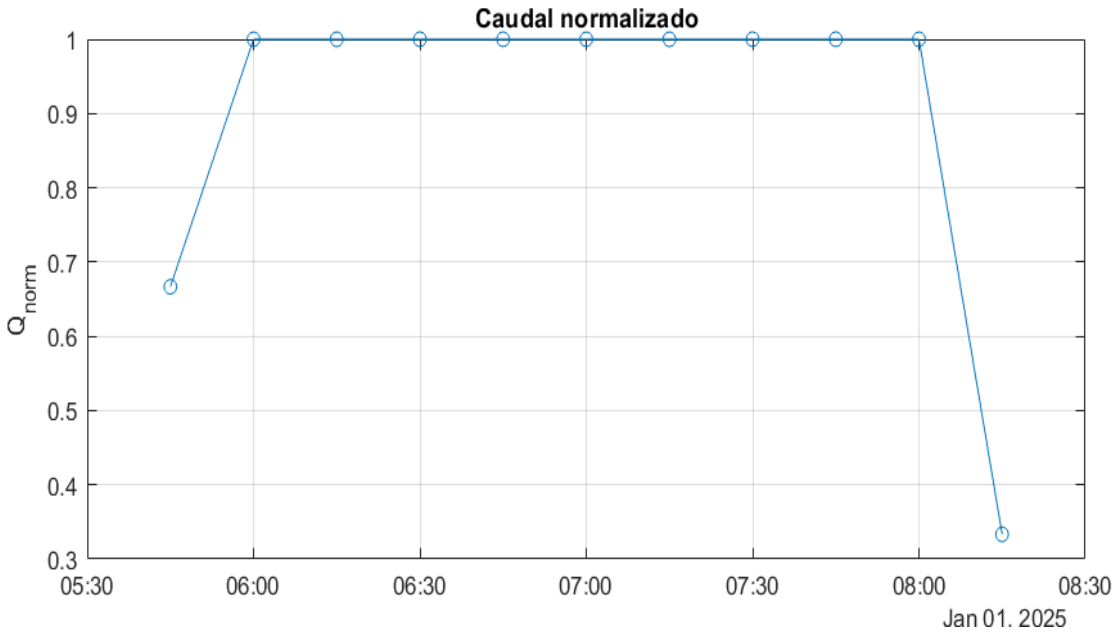
Nota. Elaboración Propia a partir del Anexo N°2.

Durante la mayor parte del registro, el diferencial se mantiene estable en torno a -54 °C, reflejando un comportamiento térmico uniforme y una disipación de calor adecuada en los componentes mecánicos. No obstante, se observa un pico pronunciado cerca de las 06:40 horas, donde la diferencia se reduce hasta aproximadamente -33 °C, coincidiendo con el descenso de vibración detectado en el mismo intervalo temporal. Este comportamiento sugiere la ocurrencia de un evento transitorio, posiblemente asociado a una reducción temporal de carga o a un breve cese operativo del equipo.

La Figura 16 representa la evolución temporal del caudal normalizado (Q_norm) en la bomba B03. Se observa que el flujo mantiene una tendencia estable y cercana al valor nominal ($Q_norm \approx 1$) durante la mayor parte del periodo de monitoreo, indicando un desempeño hidráulico constante y sin fluctuaciones significativas. No obstante, se registran dos eventos atípicos: uno al inicio del intervalo, donde el caudal parte de un valor reducido ($Q_norm \approx 0.68$), y otro hacia las 08:10 horas, en el cual se presenta una caída abrupta hasta 0.35, sugiriendo un cese momentáneo de operación o una pérdida temporal de carga.

Figura 16

Comportamiento temporal del caudal normalizado en la bomba B03

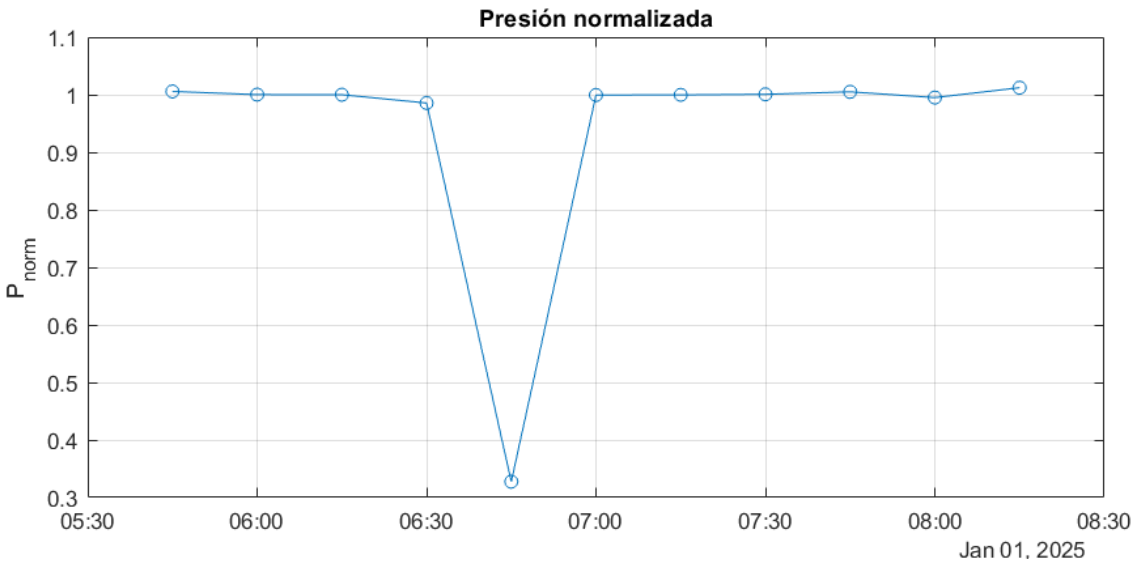


Nota. Elaboración Propia a partir del Anexo N°2.

La Figura 17 muestra el comportamiento temporal de la presión normalizada (P_norm) en la bomba B03 durante el periodo de observación. Los valores permanecen estables y próximos a la unidad, lo que refleja un funcionamiento hidráulico eficiente y acorde a las condiciones nominales de diseño. Sin embargo, se identifica un descenso abrupto alrededor de las 06:40 horas, donde la presión cae hasta 0.35, coincidiendo con las variaciones observadas en el caudal y la vibración. Este evento sugiere una interrupción momentánea o una reducción repentina en la carga hidráulica del sistema, posiblemente originada por una parada parcial o un cambio de régimen operativo.

Figura 17

Variación temporal de la presión normalizada en la bomba B03



Nota. Elaboración Propia a partir del Anexo N°2.

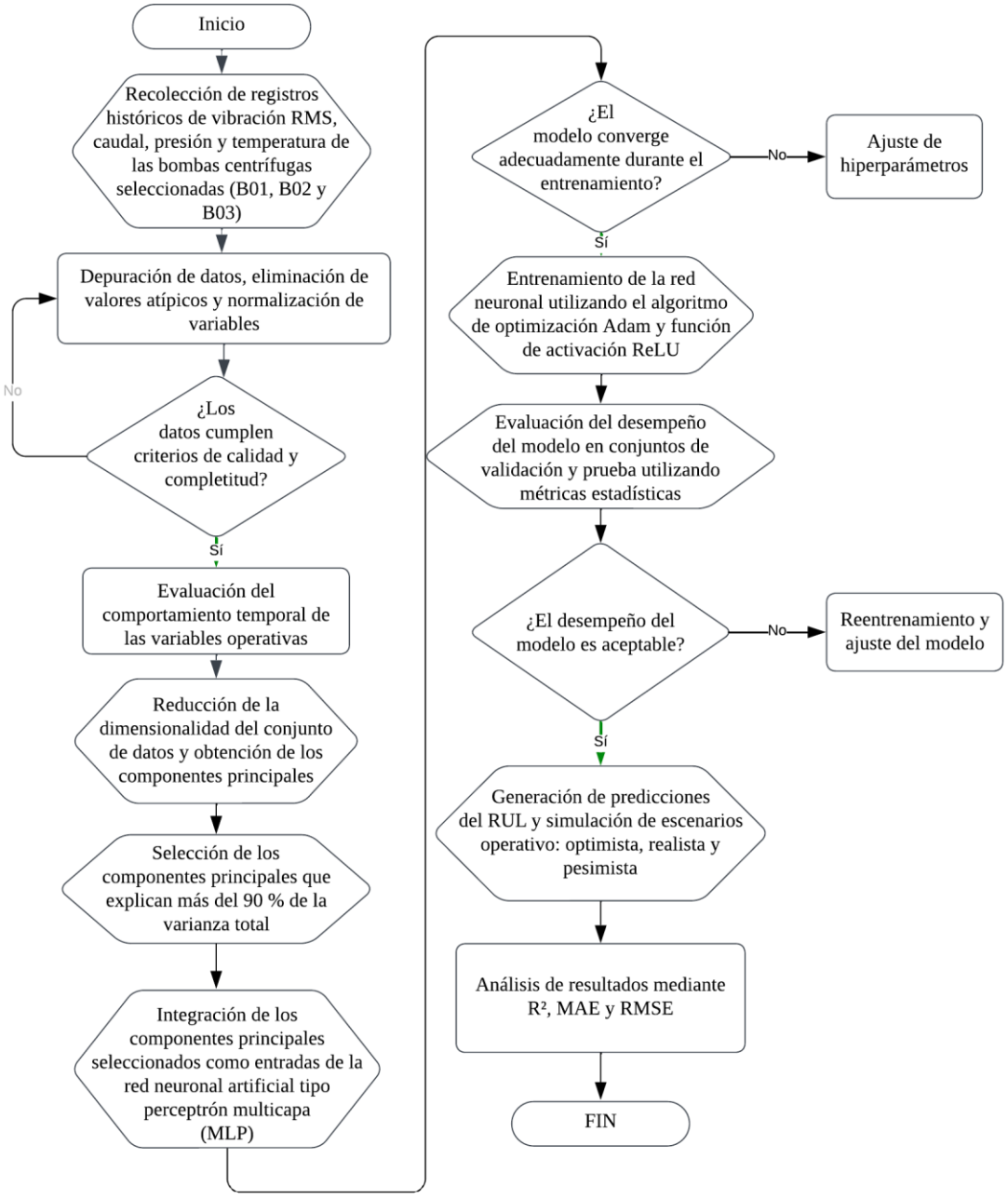
4.2. Desarrollo del modelo predictivo

En este apartado se describe el proceso de construcción, entrenamiento y validación del modelo predictivo destinado a estimar el comportamiento de degradación y la vida útil remanente de las bombas centrífugas analizadas. El desarrollo del modelo se basa en la información procesada de las variables operativas obtenidas del sistema SCADA caudal, presión, temperatura y vibración RMS, las cuales fueron previamente normalizadas y reducidas mediante el Análisis de Componentes Principales (PCA), garantizando la eliminación de redundancias y la conservación de la mayor parte de la varianza del sistema.

La metodología contempla la aplicación de técnicas de aprendizaje supervisado, específicamente redes neuronales artificiales (RNA) optimizadas mediante validación cruzada, con el propósito de modelar relaciones no lineales entre los indicadores de condición y el estado funcional de los equipos (Ver Figura 18).

Figura 18

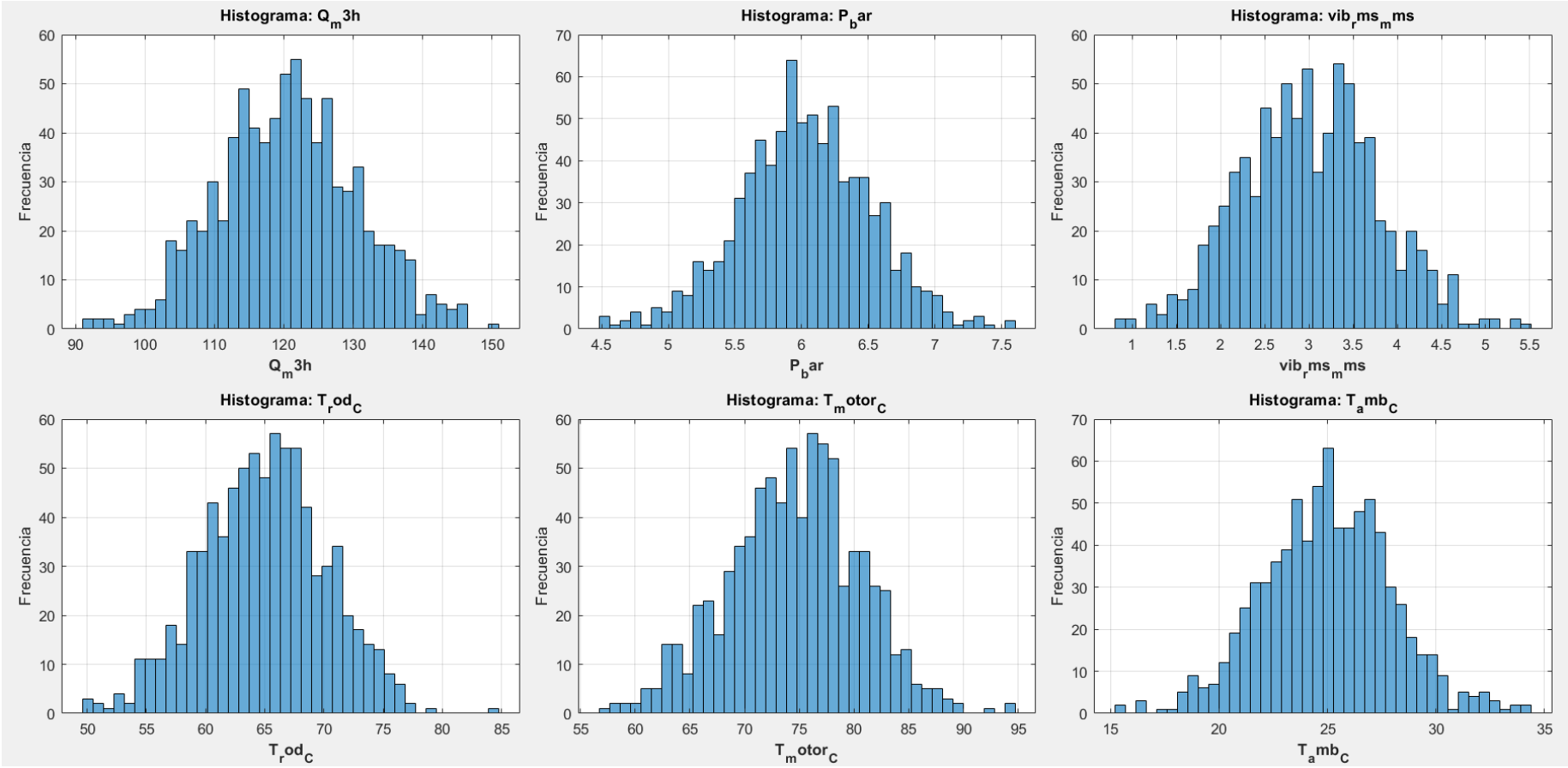
Diagrama de flujo del proceso modelo predictivo



Nota. Elaboración propia

Figura 19

Distribución de variables operativas en el sistema de bombeo



Nota. Elaboración Propia

La Figura 19 muestra la distribución estadística de las principales variables operativas registradas en el sistema de bombeo, incluyendo caudal (Q_{m3h}), presión (P_b), vibración RMS (vib_rms), temperatura de rodamiento (T_rod), temperatura del motor (T_motor) y temperatura ambiente (T_amb). Los histogramas permiten visualizar la frecuencia de ocurrencia de los valores medidos, facilitando la identificación de patrones de comportamiento y la detección de posibles outliers en las señales adquiridas.

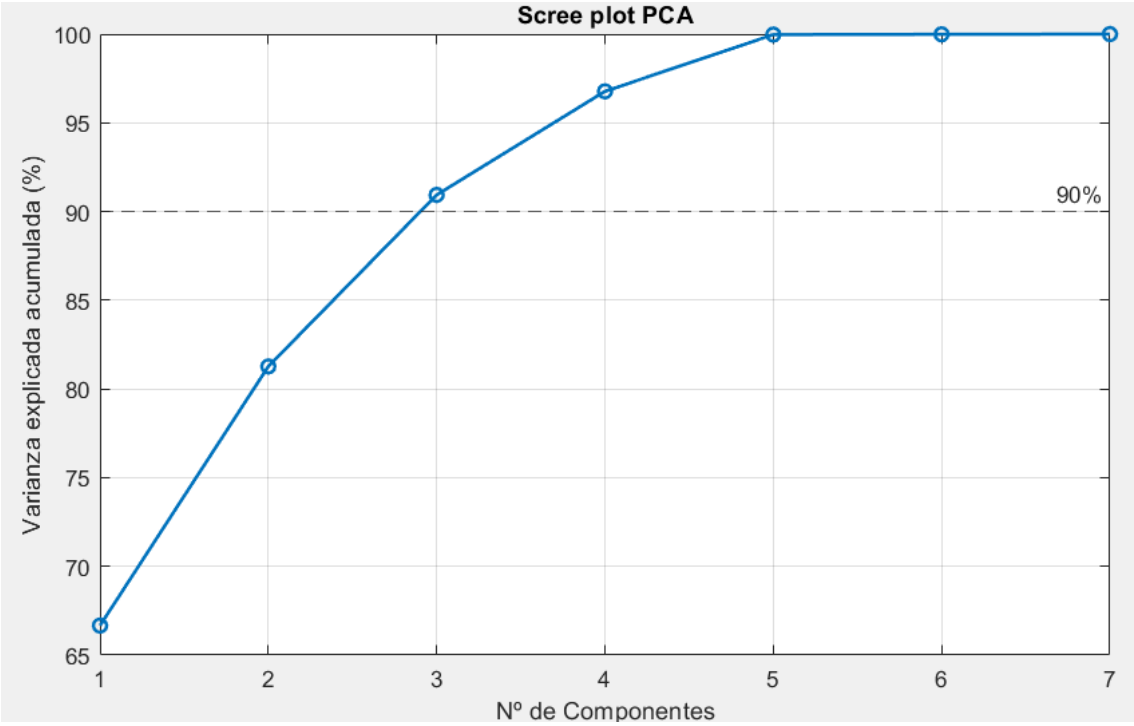
Se observa que las variables hidráulicas caudal y presión presentan distribuciones discretas, lo cual responde a las diferentes condiciones de operación en los ciclos de arranque, régimen y parada de las bombas. Por otro lado, las temperaturas del motor y del rodamiento exhiben concentraciones elevadas en rangos específicos, evidenciando estabilidad térmica durante la mayor parte del tiempo de operación. La variable de vibración RMS muestra una distribución más dispersa, indicando la presencia de fluctuaciones mecánicas leves propias del sistema rotativo.

La Figura 20 muestra la gráfica de sedimentación (Scree Plot) obtenida a partir del Análisis de Componentes Principales (PCA) aplicado a las variables operativas del sistema de bombeo. En el eje horizontal se representa el número de componentes principales y en el eje vertical la varianza explicada acumulada (%). Se observa que las cuatro primeras componentes explican más del 90 % de la varianza total, lo que evidencia que estas concentran la mayor parte de la información relevante del sistema. A partir de la quinta componente, la ganancia de varianza se vuelve marginal, lo que justifica la reducción dimensional al considerar únicamente las primeras componentes en el proceso de modelado.

Este resultado respalda la aplicación del PCA como técnica previa de extracción de características, permitiendo eliminar redundancias y minimizar el ruido en los datos sin comprometer la representatividad del conjunto. Dicho enfoque facilita la eficiencia computacional y mejora la capacidad de generalización del modelo predictivo durante la etapa de entrenamiento.

Figura 20

Sedimentación (Scree Plot) del Análisis de Componentes Principales



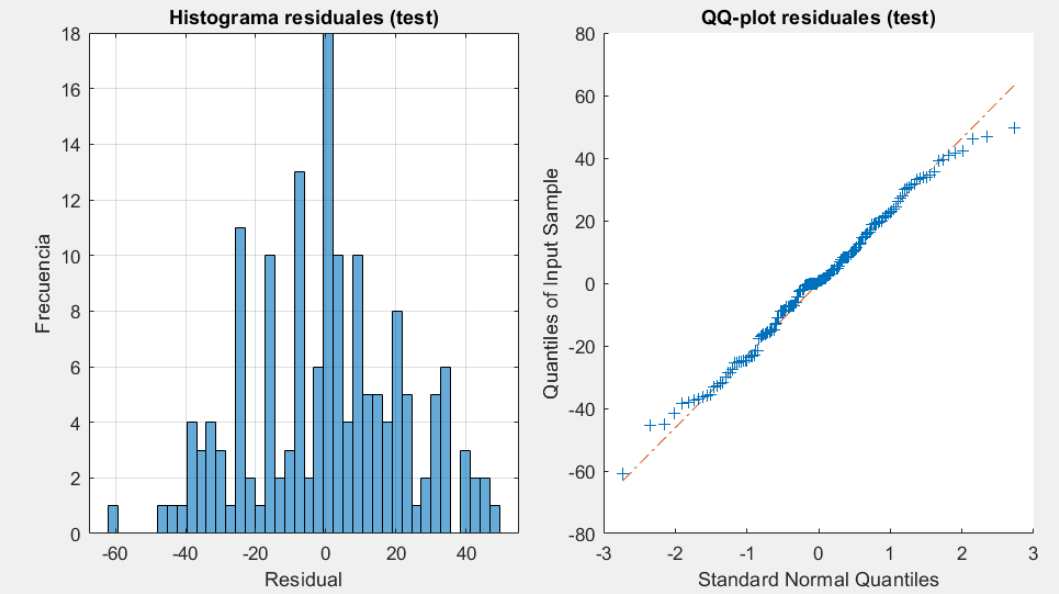
Nota. Elaboración Propia

La Figura 21 muestra el análisis estadístico de los residuales obtenidos en el conjunto de prueba del modelo predictivo PCA-RNA. En el gráfico izquierdo se presenta el histograma de los residuales, donde se observa una distribución aproximadamente centrada alrededor de cero, sin un desplazamiento sistemático hacia valores positivos o negativos, lo que indica la ausencia de sesgo significativo en las predicciones del modelo.

Por su parte, el gráfico QQ-plot (lado derecho) compara los cuantiles empíricos de los residuales con los cuantiles de una distribución normal teórica. La alineación predominante de los puntos a lo largo de la línea diagonal sugiere que los residuales presentan un comportamiento cercano a la normalidad, especialmente en la región central de la distribución. No obstante, se aprecian ligeras desviaciones en los cuantiles extremos, lo cual es habitual en modelos aplicados a sistemas industriales reales y puede asociarse a condiciones operativas transitorias, cambios de régimen de carga o eventos de degradación no lineal del sistema de bombeo.

Figura 21

Distribución y análisis de normalidad de los residuales en el conjunto de prueba



Nota. Elaboración Propia

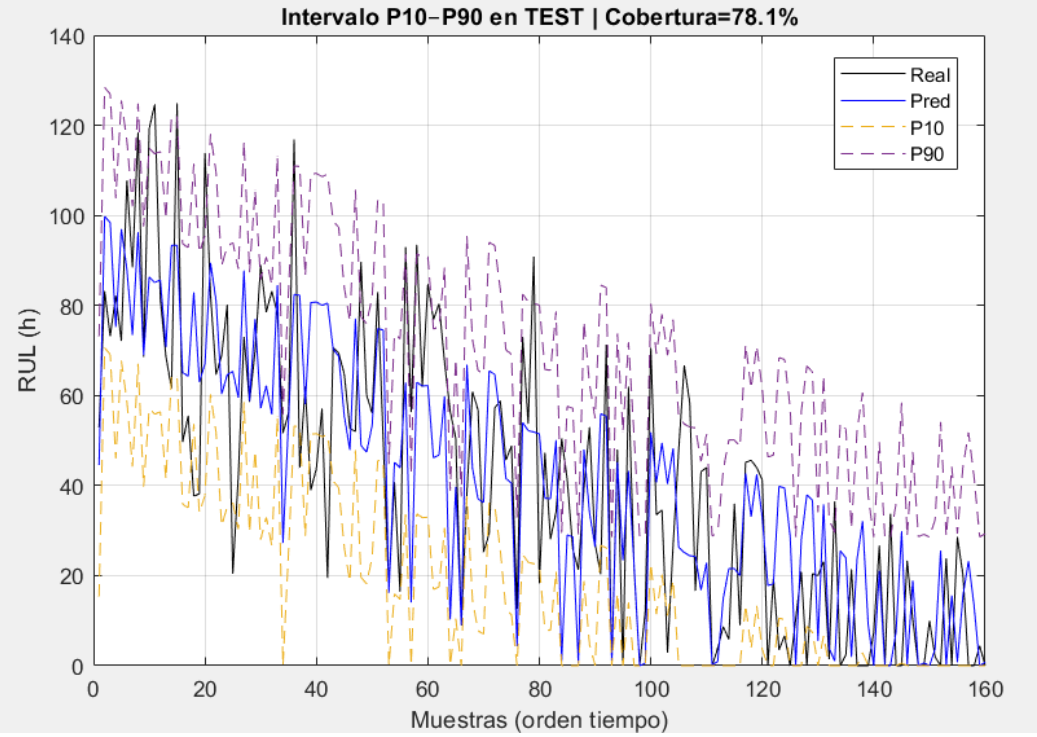
La Figura 22 muestra el intervalo de predicción P10–P90 del tiempo de vida útil remanente (RUL) estimado por el modelo PCA–RNA en el conjunto de prueba. En la figura se presenta la comparación entre el RUL real (línea negra), el RUL predicho por el modelo (línea azul) y los límites inferiores (P10) y superiores (P90) del intervalo predictivo, representados por líneas discontinuas. El intervalo P10-P90 delimita el rango dentro del cual se espera que se encuentre el valor real del RUL con una probabilidad del 80 %, permitiendo cuantificar la incertidumbre asociada a las predicciones del modelo. La cobertura observada del 78,1 % indica que la mayoría de los valores reales del RUL se encuentran dentro de dicho intervalo, lo que evidencia una capacidad razonable del modelo para capturar la variabilidad del proceso de degradación en condiciones de prueba.

Se observa que, durante las primeras muestras, correspondientes a estados de mayor RUL, el intervalo presenta una amplitud relativamente mayor, asociada a una mayor dispersión de las variables operativas y a la incertidumbre propia de etapas tempranas de degradación. A medida que avanza el tiempo y el sistema se aproxima al fin de su vida útil, el intervalo tiende a estrecharse, reflejando una reducción de la incertidumbre en la estimación del RUL y una mayor consistencia entre los valores

predichos y reales. Las desviaciones puntuales del RUL real fuera del intervalo predictivo se asocian principalmente a transiciones operativas abruptas, cambios en las condiciones de carga o efectos no lineales del proceso de degradación que no son completamente capturados por el modelo. No obstante, el comportamiento global del intervalo predictivo confirma que el modelo PCA–RNA proporciona una estimación del RUL acompañada de una medida cuantitativa de incertidumbre, lo que resulta relevante para la toma de decisiones en estrategias de mantenimiento predictivo.

Figura 22

Intervalo de predicción P10–P90 del RUL estimado en el conjunto de prueba



Nota. Elaboración Propia

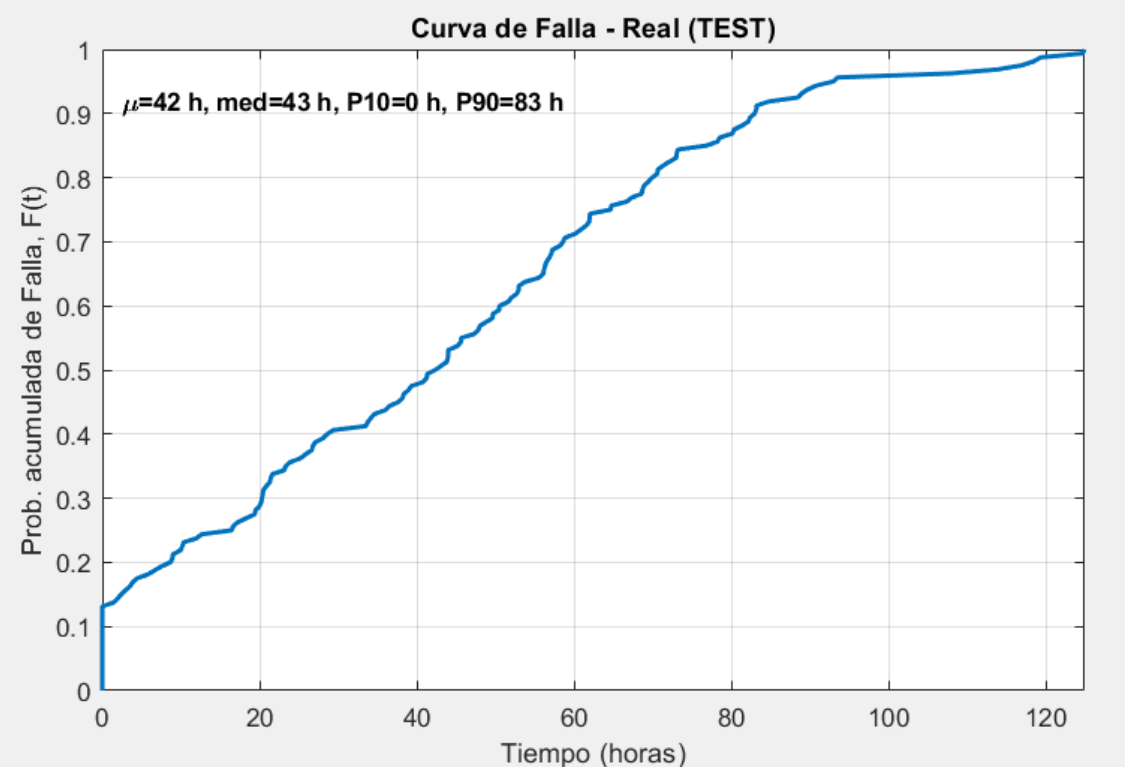
La Figura 23 muestra la curva de probabilidad acumulada de falla, $F(t)$, del sistema de bombeo bajo un escenario no ideal, representando la evolución temporal del riesgo de falla en función del tiempo de operación. Esta curva expresa la probabilidad de que el sistema haya experimentado una falla antes o en un instante determinado, constituyéndose en una herramienta para analizar el comportamiento de degradación del equipo a lo largo de su ciclo operativo.

En la gráfica se identifican los principales estadísticos descriptivos del tiempo hasta la falla (TTF): una media de $\mu = 42$horas y una mediana de 43 horas, lo que evidencia una distribución aproximadamente simétrica del proceso de falla. Asimismo, los percentiles P10 = 0 h y P90 = 83 h delimitan el intervalo en el que ocurre el 80 % de los eventos de falla, reflejando la dispersión temporal del fenómeno bajo condiciones reales de operación.

Se observa que la probabilidad acumulada de falla incrementa de forma progresiva desde las primeras horas de funcionamiento, alcanzando valores cercanos al 50% alrededor de las 40-45 horas, y superando el 90% aproximadamente a las 80 horas de operación. Este comportamiento indica que el sistema presenta un proceso de degradación continuo, sin una etapa prolongada de operación libre de fallas, característica de equipos sometidos a desgaste acumulativo.

Figura 23

Curva de probabilidad acumulada de falla del sistema de bombeo (escenario no ideal)



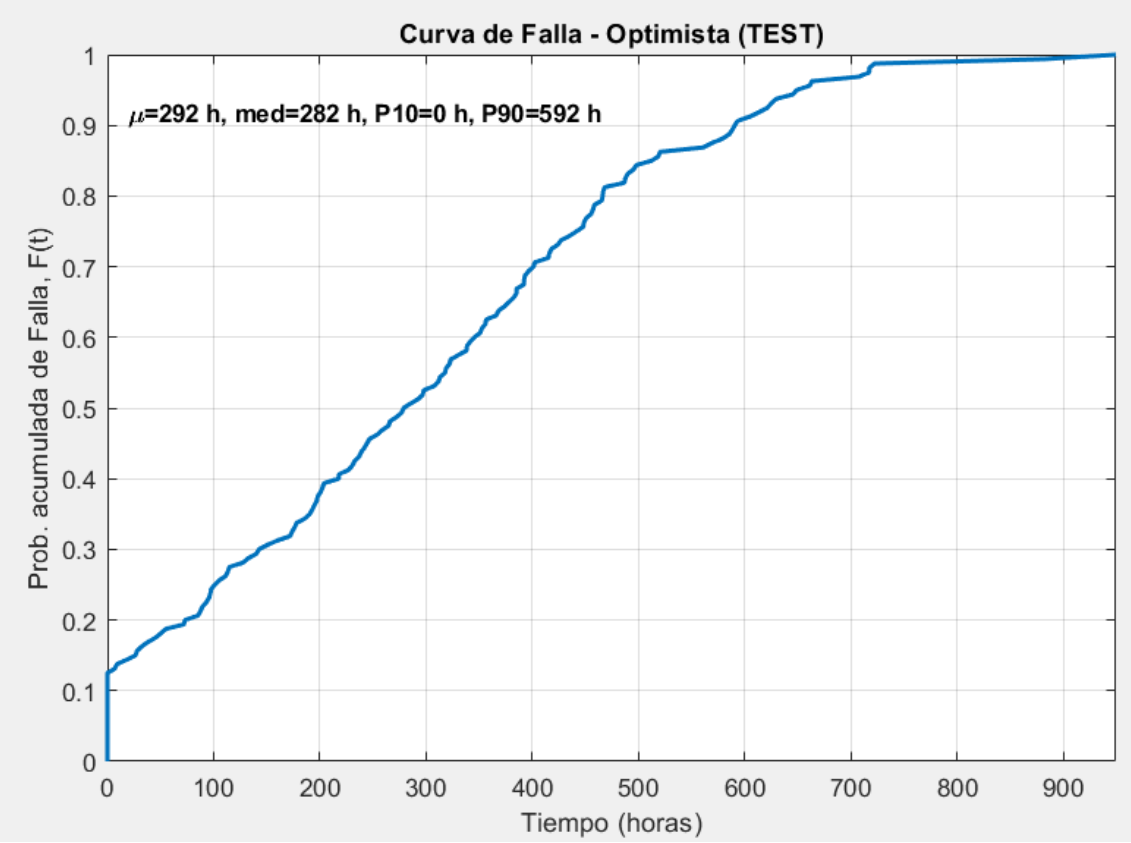
Nota. Elaboración Propia

La Figura 24 muestra la curva de probabilidad acumulada de falla, $F(t)$, del sistema de bombeo bajo el escenario optimista, el cual representa condiciones operativas favorables caracterizadas por un funcionamiento estable del equipo y una degradación controlada de sus componentes. Esta curva describe la evolución temporal del riesgo de falla, expresando la probabilidad de que el sistema falle antes o en un determinado tiempo de operación.

En este escenario, los estadísticos descriptivos del tiempo hasta la falla (TTF) presentan una media de $\mu = 292$ horas y una mediana de 282 horas, lo que evidencia una vida útil considerablemente extendida en comparación con los escenarios real y realista. Asimismo, los percentiles P10 = 0 h y P90 = 592 h delimitan el intervalo en el cual ocurre el 80 % de los eventos de falla, reflejando una amplia dispersión temporal y una mayor variabilidad en la ocurrencia de fallas bajo condiciones ideales de operación.

Figura 24

Curva de probabilidad de falla del sistema de bombeo – Escenario optimista



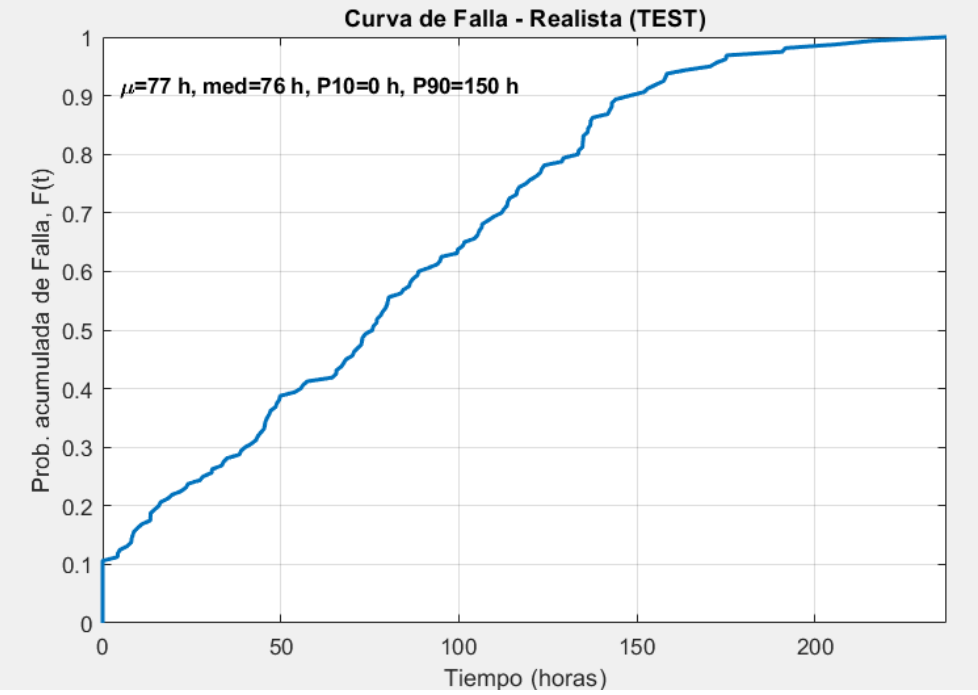
Nota. Elaboración Propia

Desde el punto de vista temporal, la curva presenta un crecimiento progresivo de la probabilidad acumulada de falla, alcanzando valores cercanos al 50% alrededor de las 280-300 horas, y aproximándose al 100% a partir de las 700-800 horas de operación. La pendiente relativamente suave de la curva indica un proceso de degradación gradual, coherente con un régimen de operación optimizado, en el que los efectos de vibración, carga mecánica y estrés térmico se mantienen dentro de rangos controlados. En términos físicos, este escenario representa el mejor desempeño esperable del sistema de bombeo dentro del rango de operación analizado, constituyéndose en una referencia superior para la evaluación comparativa del modelo predictivo desarrollado.

La Figura 25 muestra la curva de probabilidad acumulada de falla, $F(t)$, correspondiente al escenario realista, el cual representa condiciones normales de operación del sistema de bombeo, con presencia moderada de variaciones térmicas, vibracionales y de presión propias del funcionamiento habitual del equipo. Esta curva describe la evolución temporal del riesgo de falla, expresando la probabilidad de que el sistema falle antes o en un determinado tiempo de operación.

Figura 25

Curva de probabilidad acumulada de falla del sistema de bombeo – Escenario realista



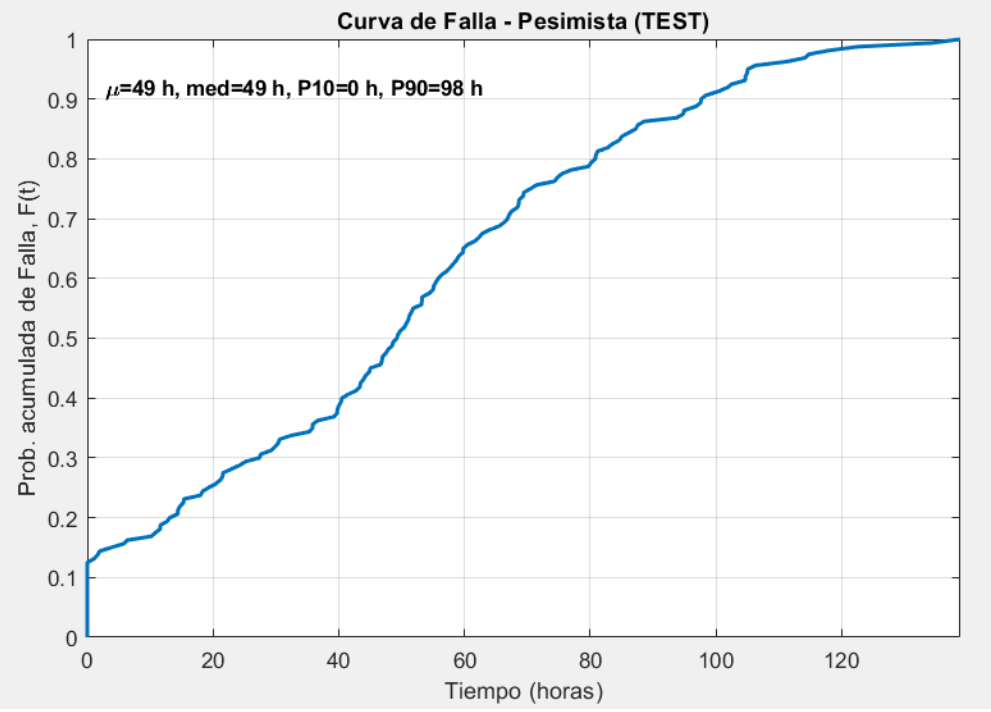
Nota. Elaboración Propia

Los parámetros estadísticos característicos del tiempo hasta la falla (TTF) indican una media de $\mu = 77$horas y una mediana de 76 horas, lo que evidencia una distribución centrada y consistente del proceso de falla. Asimismo, los percentiles P10 = 0 h y P90 = 150 h delimitan el intervalo en el cual ocurre el 80% de los eventos de falla, reflejando una dispersión moderada en los tiempos de ocurrencia de fallas bajo condiciones operativas realistas. Desde un punto de vista cuantitativo, la curva presenta un crecimiento progresivo de la probabilidad acumulada de falla, alcanzando valores cercanos al 50% alrededor de las 70-80 horas, superando el 70% aproximadamente a las 110-120 horas, y aproximándose al 100% a partir de las 180-200 horas de operación.

La Figura 26 representa la curva de probabilidad acumulada de falla, $F(t)$, del sistema de bombeo bajo el escenario pesimista, el cual simula condiciones de operación adversas caracterizadas por incrementos sostenidos de carga térmica, mayores niveles de vibración y un control limitado del mantenimiento. Esta curva describe la evolución temporal del riesgo de falla, expresando la probabilidad de que el sistema falle antes o en un determinado tiempo de operación.

Figura 26

Curva de probabilidad de falla del sistema de bombeo – Escenario pesimista



Nota. Elaboración Propia

En este escenario, los parámetros estadísticos del tiempo hasta la falla (TTF) muestran una media de $\mu = 49$ horas y una mediana de 49 horas, lo que indica una concentración del proceso de falla en intervalos de tiempo relativamente cortos. Asimismo, los percentiles P10 = 0 h y P90 = 98 h delimitan el intervalo en el cual ocurre el 80% de los eventos de falla, evidenciando una dispersión moderada, aunque desplazada hacia tiempos menores en comparación con los escenarios realista y optimista.

Desde el punto de vista temporal, la curva presenta un crecimiento acelerado de la probabilidad acumulada de falla, alcanzando valores cercanos al 50% alrededor de las 45-50 horas, y superando el 90% aproximadamente a las 90-100 horas de operación. Este comportamiento refleja un proceso de degradación más severo, en el que las fallas se manifiestan de forma anticipada como consecuencia del desgaste acelerado de componentes críticos. La pendiente más pronunciada de la curva, en relación con los otros escenarios, sugiere un entorno operativo con mayores exigencias mecánicas y térmicas, donde factores como desalineación, cavitación, lubricación deficiente o sobrecarga hidráulica intensifican el deterioro del equipo. La reducción del TTF y el desplazamiento de la curva hacia tiempos menores confirman que el escenario pesimista representa el límite inferior del desempeño del sistema de bombeo dentro del rango analizado.

En términos comparativos, este escenario constituye un caso de referencia crítica para la evaluación del modelo predictivo, permitiendo analizar su capacidad para identificar condiciones de degradación acelerada y estimar de forma adecuada la reducción del tiempo de vida útil remanente (RUL) bajo contextos operativos desfavorables.

4.3. Métricas de modelo predictivo

En este apartado se presentan los resultados de las métricas de desempeño utilizadas para evaluar la precisión, robustez y capacidad de generalización del modelo predictivo aplicado al diagnóstico del sistema de bombeo. Estas métricas permiten cuantificar la diferencia entre los valores estimados por el modelo y los datos reales obtenidos en las pruebas experimentales, proporcionando una medida objetiva de su eficacia en la predicción del Remaining Useful Life (RUL) y en la detección de condiciones anómalas.

El análisis incluye indicadores estadísticos ampliamente empleados en la evaluación de modelos de aprendizaje automático, tales como el Error Absoluto Medio (MAE), la Raíz del Error Cuadrático Medio (RMSE), el Error Porcentual Absoluto Medio (MAPE), el Error Porcentual Simétrico Medio (sMAPE) y el Coeficiente de Determinación (R^2). Cada una de estas métricas proporciona una perspectiva complementaria sobre la exactitud de las predicciones, la dispersión de los errores y la correlación entre los valores observados y los generados por el modelo.

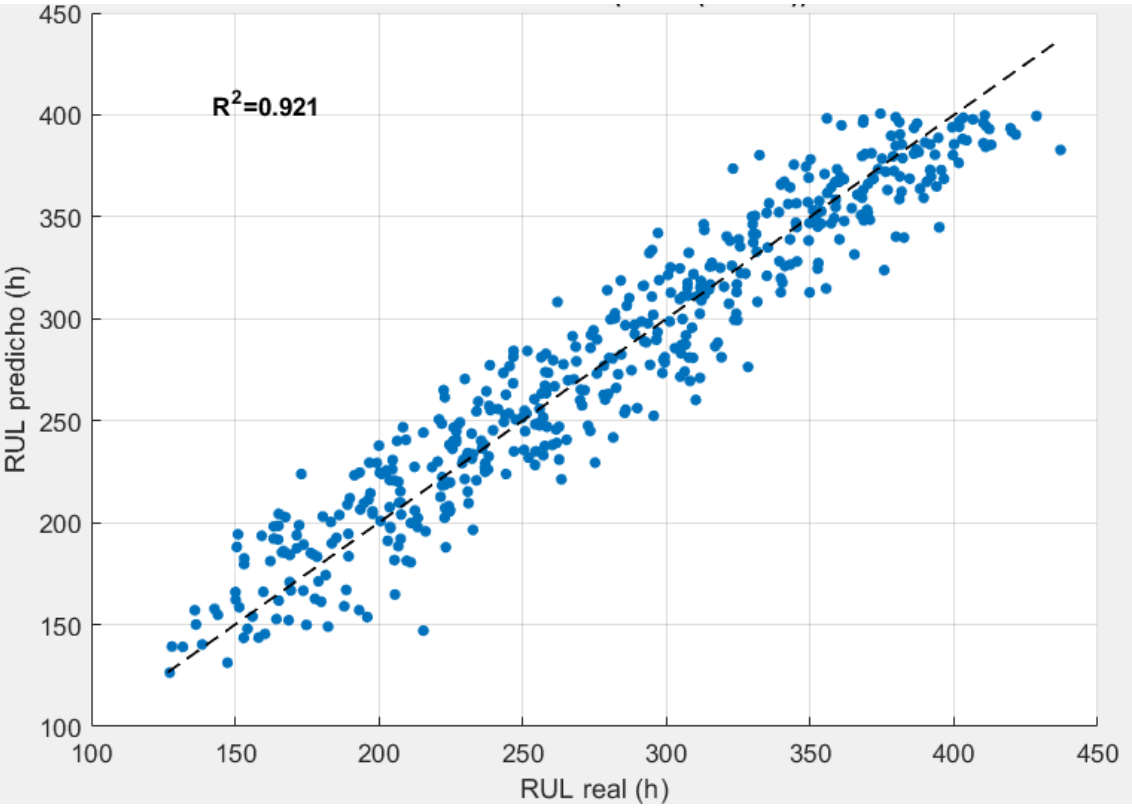
La Figura 27 presenta la correlación entre los valores de vida útil remanente (RUL) reales y los valores predichos por el modelo PCA–RNA en el conjunto de entrenamiento. La línea discontinua representa la relación ideal de igualdad entre el RUL real y el RUL estimado, mientras que los puntos corresponden a las predicciones generadas durante el proceso de aprendizaje del modelo.

Desde un punto de vista cuantitativo, el coeficiente de determinación obtenido es $R^2 = 0.921$, lo que indica que aproximadamente el 92.1% de la variabilidad del RUL real es explicada por las predicciones del modelo en esta etapa. Este resultado evidencia un alto nivel de ajuste durante el entrenamiento y confirma la capacidad del modelo para capturar las relaciones no lineales existentes entre las variables operativas monitoreadas, tales como vibración, presión, temperatura y caudal.

Visualmente, se observa que la mayoría de los puntos se concentran en torno a la línea de referencia, lo que refleja una consistencia predictiva adecuada y una correcta representación del comportamiento de degradación del sistema durante la fase de aprendizaje. No obstante, se aprecian ligeras dispersiones, principalmente en el rango intermedio de RUL (aproximadamente entre 150 y 300 horas), las cuales pueden atribuirse a la variabilidad inherente de los datos operativos y a la naturaleza estocástica del proceso de entrenamiento de la red neuronal.

Figura 27

Correlación entre el RUL real y el RUL predicho en el conjunto de entrenamiento



Nota. Elaboración Propia

La Figura 28 muestra la relación entre los valores reales y los valores predichos de la vida útil remanente (RUL) durante la fase de validación del modelo. En esta etapa se emplearon datos que no formaron parte del entrenamiento, por lo que el análisis permite evaluar la capacidad de generalización del modelo frente a condiciones operativas nuevas.

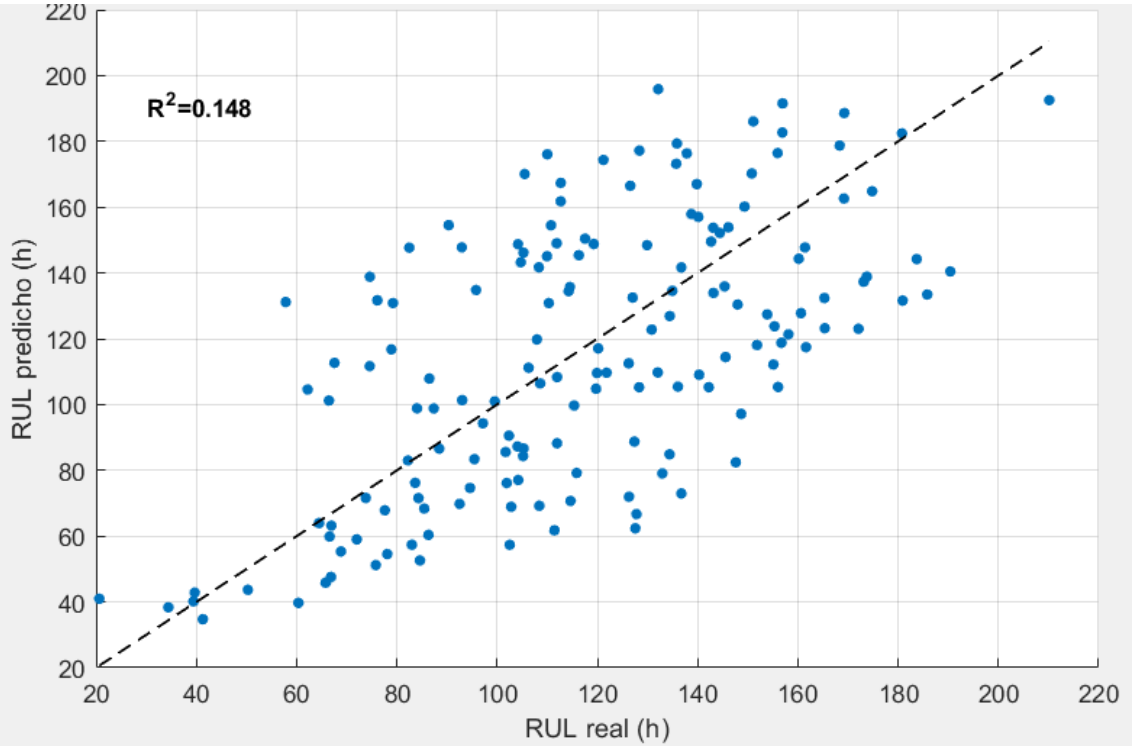
Desde el punto de vista cuantitativo, el coeficiente de determinación obtenido es $R^2 = 0.148$, lo que evidencia una reducción en el nivel de correlación entre los valores reales y los valores estimados de RUL respecto a la fase de entrenamiento. Este comportamiento indica que el modelo presenta mayores dificultades para reproducir con precisión el comportamiento del sistema cuando se enfrenta a datos no vistos previamente, lo cual es consistente con la complejidad inherente de los procesos de degradación en equipos industriales reales.

Desde un contexto físico y operativo real, este comportamiento se explica por la alta variabilidad de las condiciones de operación del sistema de bombeo durante la etapa

de validación. En particular, combinaciones de carga hidráulica, vibración, temperatura y presión distintas a las observadas en el conjunto de entrenamiento generan patrones de degradación más heterogéneos, lo que introduce incertidumbre adicional en la estimación del RUL. En sistemas de bombeo centrífugo, fenómenos como cavitación incipiente, desalineación mecánica, variaciones en la calidad del fluido y cambios transitorios en el régimen de operación influyen de manera no uniforme en la evolución del desgaste, dificultando la extrapolación directa del modelo. Asimismo, la dispersión observada en la nube de puntos, especialmente en los rangos intermedios de RUL, refleja que el proceso de degradación no sigue una trayectoria estrictamente determinista, sino que responde a múltiples factores operativos que pueden variar de un ciclo a otro. En este sentido, el resultado obtenido no debe interpretarse como una falla del modelo, sino como una representación realista del comportamiento del sistema bajo condiciones operativas cambiantes, propias de un entorno industrial.

Figura 28

Correlación entre el RUL real y el RUL predicho en la fase de validación



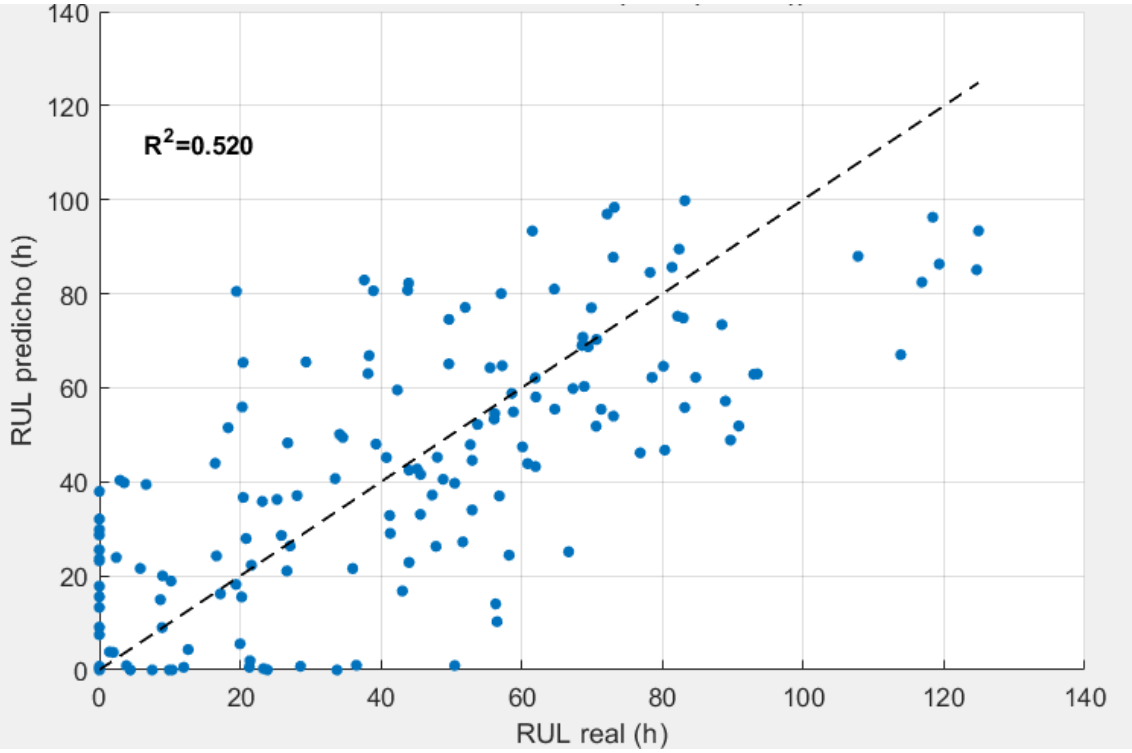
Nota. Elaboración Propia

La Figura 29 muestra la relación entre los valores reales y los valores predichos de la vida útil remanente (RUL) obtenidos en el conjunto de prueba, el cual representa el escenario más exigente para evaluar la capacidad de generalización del modelo predictivo, dado que las muestras empleadas no fueron utilizadas durante las fases de entrenamiento ni validación.

Desde un punto de vista cuantitativo, el coeficiente de determinación obtenido es $R^2 = 0.520$, lo que indica una correlación moderada entre los valores reales y las estimaciones de RUL en el conjunto de prueba. Este resultado evidencia que el modelo logra capturar parcialmente la tendencia general del proceso de degradación, aun cuando se enfrenta a datos completamente nuevos y a condiciones operativas no vistas previamente.

Figura 29

Correlación entre el RUL real y el RUL predicho en el conjunto de prueba



Nota. Elaboración Propia

Desde un contexto físico y operativo real, este comportamiento se explica por la variabilidad inherente al funcionamiento de las bombas centrífugas, especialmente en etapas avanzadas de degradación. En esta fase, fenómenos como cambios en el régimen

de operación, fluctuaciones en la carga hidráulica, variaciones térmicas y alteraciones en los niveles de vibración influyen de manera no uniforme en la evolución del desgaste, generando dispersiones naturales en la relación entre las variables operativas y el RUL real. Esta complejidad se refleja en la dispersión observada en la nube de puntos respecto a la línea de referencia. Asimismo, se observa que las mayores desviaciones se concentran en los rangos bajos e intermedios de RUL, donde el proceso de degradación puede acelerarse o desacelerarse de forma no lineal dependiendo de las condiciones de operación. En este sentido, las diferencias entre los valores reales y predichos no deben interpretarse como un fallo del modelo, sino como una manifestación realista del comportamiento estocástico del sistema bajo condiciones operativas cambiantes.

En términos metodológicos, el desempeño obtenido en el conjunto de prueba confirma que el modelo presenta una capacidad de generalización aceptable, superior a la observada en la fase de validación, y aporta evidencia de que el enfoque PCA–RNA es capaz de proporcionar estimaciones útiles del tiempo de vida útil remanente (RUL), incluso en escenarios de alta incertidumbre. Este resultado refuerza la validez del modelo como herramienta de apoyo al mantenimiento predictivo, al tiempo que pone de relieve la necesidad de continuar fortaleciendo la base de datos y la representación de estados de degradación más diversos.

La Tabla 3 presenta las métricas de desempeño obtenidas para el modelo predictivo híbrido PCA-RNA (Análisis de Componentes Principales y Red Neuronal Artificial) en las fases de entrenamiento, validación y prueba. Los indicadores evaluados MAE, RMSE, MAPE, sMAPE y R^2 permiten cuantificar la precisión y consistencia del modelo en la estimación de la vida útil remanente (RUL) de las bombas centrífugas.

Tabla 3

Métricas de desempeño del modelo PCA–RNA

Conjunto	MAE	RMSE	MAPE (%)	sMAPE (%)	R^2
Entrenamiento	16.9335	20.7543	6.5970	6.5533	0.9213
Validación	27.6906	33.1843	25.5273	24.9273	0.1483
Prueba	17.4512	22.0243	44.4521	71.1225	0.5203

Nota. Elaboración Propia

- En el conjunto de entrenamiento, el modelo alcanza valores reducidos de error, con un MAE de 16.93 h y un RMSE de 20.75 h, lo que indica una adecuada capacidad para aprender los patrones de degradación presentes en los datos históricos. Asimismo, los valores de MAPE (6.60%) y sMAPE (6.5 %) confirman una buena precisión relativa en la estimación del RUL, mientras que el coeficiente de determinación $R^2 = 0.9213$ evidencia que el modelo explica más del 92 % de la variabilidad del RUL real durante esta etapa. Estos resultados reflejan un ajuste satisfactorio del modelo en la fase de aprendizaje.
- En la fase de validación, se observa un incremento en los errores, con un MAE de 27.69 h y un RMSE de 33.18 h, acompañado de valores más elevados de MAPE (25.53%) y sMAPE (24.93%). El R^2 obtenido (0.1483) indica una reducción significativa en la capacidad explicativa del modelo frente a datos no utilizados en el entrenamiento. Desde un punto de vista operativo, este comportamiento es coherente con la variabilidad de las condiciones reales de operación del sistema de bombeo, donde combinaciones de carga, vibración y temperatura distintas a las aprendidas inicialmente introducen mayor incertidumbre en la estimación del RUL.
- Por su parte, en el conjunto de prueba, el modelo presenta un desempeño intermedio, con un MAE de 17.45 h y un RMSE de 22.02 h, valores cercanos a los obtenidos en la fase de entrenamiento. El $R^2 = 0.5203$ indica una correlación moderada entre los valores reales y predichos de RUL, lo cual resulta factible en un contexto operativo real, donde el proceso de degradación de bombas centrífugas está influenciado por múltiples factores variables y no completamente deterministas. Los valores de MAPE (44.45%) y sMAPE (71.12%) se explican principalmente por la presencia de escenarios con vida útil remanente reducida, en los que pequeños errores absolutos generan errores porcentuales elevados. Este comportamiento es común en aplicaciones reales de mantenimiento predictivo y no invalida la utilidad del modelo, ya que este logra capturar la tendencia general del deterioro y apoyar la toma de decisiones de mantenimiento.

En conjunto, los resultados presentados en la Tabla 3 muestran que el modelo PCA–RNA presenta un buen desempeño en la fase de entrenamiento, una sensibilidad significativa a la variabilidad operativa en la validación, y una capacidad de

generalización aceptable en el conjunto de prueba. Este comportamiento es consistente con la naturaleza compleja y no lineal del proceso de degradación de bombas centrífugas en entornos industriales reales, y respalda la utilidad del modelo como herramienta de apoyo para la estimación del tiempo de vida útil remanente (RUL), reconociendo al mismo tiempo las limitaciones inherentes al uso de datos operativos heterogéneos.

La Tabla 4 presenta los resultados del análisis de importancia relativa de los componentes principales, evaluada mediante el método de permutación sobre el error absoluto medio (ΔMAE). Este enfoque consiste en perturbar individualmente cada componente principal y cuantificar el incremento resultante en el error de predicción del modelo, permitiendo identificar cuáles componentes contienen la información más relevante para la estimación del tiempo de vida útil remanente (RUL) dentro del modelo híbrido PCA–RNA. A diferencia del análisis tradicional de varianza explicada del PCA, que prioriza los componentes con mayor aporte a la variabilidad global de los datos, el análisis de permutación sobre el ΔMAE permite evaluar la relevancia predictiva de cada componente en relación directa con la variable objetivo. En este sentido, un mayor valor de ΔMAE indica que el componente perturbado desempeña un rol significativo en la capacidad del modelo para estimar el RUL.

Tabla 4

Importancia de los componentes principales (ΔMAE)

PC	Delta MAE
PC3	104.2
PC1	32.348
PC2	0.028673

Nota. Elaboración Propia

Los resultados muestran que el componente principal PC3 genera el mayor incremento en el error de predicción (ΔMAE = 104.2), evidenciando que este componente concentra información crítica para la estimación del RUL. La perturbación de PC3 provoca un deterioro sustancial del desempeño del modelo, lo que indica que este componente captura patrones latentes directamente asociados al proceso de degradación

progresiva de las bombas centrífugas. En consecuencia, PC3 puede interpretarse como el componente que integra las señales operativas más sensibles al desgaste del equipo, independientemente de su configuración nominal.

El componente PC1 presenta una contribución intermedia (ΔMAE = 32.35), lo que sugiere que, si bien contiene información relevante para el modelo, su influencia sobre la predicción del RUL es secundaria en comparación con PC3. Este comportamiento resulta coherente con el análisis de PCA, dado que los primeros componentes suelen capturar tendencias generales del sistema y condiciones operativas globales, las cuales no siempre se encuentran directamente vinculadas a los mecanismos de deterioro.

Por su parte, el componente PC2 muestra un impacto prácticamente nulo en el error de predicción (ΔMAE = 0.0287), lo que indica que la información que representa posee una baja correlación con la variable objetivo. Este resultado sugiere que PC2 podría estar asociado a variaciones operativas de carácter estable o redundante, que no aportan información significativa al proceso de estimación del RUL.

En conjunto, estos resultados evidencian que la importancia predictiva de los componentes principales no depende exclusivamente de la varianza explicada, sino de su capacidad para representar dinámicas operativas vinculadas al deterioro del equipo. En este contexto, el PC3 se consolida como la variable latente más influyente del modelo, por lo que su interpretación física posiblemente relacionada con combinaciones de variables térmicas, vibracionales y de carga hidráulica resulta relevante para comprender los mecanismos de degradación y optimizar la estructura del modelo predictivo PCA-RNA.

4.4. Contraste de hipótesis

El contraste de las hipótesis planteadas en la presente investigación se realiza de manera indirecta, mediante la verificación del cumplimiento de los objetivos de investigación y el análisis del desempeño del modelo predictivo desarrollado. Dado que el estudio se enmarca en un enfoque cuantitativo aplicado, basado en simulación y

modelamiento predictivo, la validación de las hipótesis no se efectúa a través de pruebas estadísticas clásicas de inferencia (como pruebas t o χ^2), sino mediante métricas de desempeño propias de modelos de regresión y aprendizaje supervisado.

En este contexto, el cumplimiento de los objetivos específicos y general se evalúa mediante indicadores cuantitativos que permiten medir la capacidad del modelo PCA–RNA para estimar el tiempo de vida útil remanente (RUL) de bombas centrífugas, considerando tanto su precisión como su capacidad de ajuste y generalización. Las métricas empleadas para este contraste son el coeficiente de determinación (R^2), el error absoluto medio (MAE) y la raíz del error cuadrático medio (RMSE), calculadas para las fases de entrenamiento, validación y prueba.

a) Contraste de la hipótesis general

La hipótesis general establece que *el desarrollo de un modelo predictivo basado en el Análisis de Componentes Principales (PCA) permitirá estimar con precisión el tiempo de vida útil remanente de las bombas centrífugas*. Esta hipótesis se contrasta directamente con el objetivo general, orientado a desarrollar dicho modelo predictivo.

El contraste se fundamenta en el nivel de ajuste alcanzado por el modelo en las distintas fases de evaluación. En la fase de entrenamiento, el modelo obtuvo un coeficiente de determinación $R^2 = 0.9213$, lo que indica que más del 92 % de la variabilidad del tiempo de vida útil remanente (RUL) es explicada por el modelo a partir de los componentes principales derivados de las variables operativas. Asimismo, se registraron errores reducidos (MAE = 16.93 h y RMSE = 20.75 h), evidenciando un adecuado ajuste a los datos históricos utilizados para el entrenamiento.

En la fase de validación, el modelo presentó un desempeño inferior ($R^2 = 0.1483$, MAE = 27.69 h, RMSE = 33.18 h), reflejando sensibilidad frente a la variabilidad operativa y a datos no utilizados en el entrenamiento. No obstante, en la fase de prueba, el modelo alcanzó un coeficiente de determinación $R^2 = 0.5203$, con errores controlados (MAE = 17.45 h y RMSE = 22.02 h), lo que demuestra su capacidad para capturar la tendencia general del RUL en datos no vistos. Estos resultados permiten aceptar la hipótesis general dentro del alcance del estudio, al evidenciar que el modelo predictivo desarrollado es capaz de estimar el tiempo de vida útil remanente con un nivel de precisión aceptable en un entorno de simulación.

b) Contraste del primer objetivo específico

El primer objetivo específico, orientado a analizar el comportamiento operativo de las bombas centrífugas sometidas a mantenimiento preventivo, se contrasta con la hipótesis que plantea que dicho análisis permite identificar patrones de funcionamiento asociados al desgaste del equipo.

Este contraste se sustenta en los resultados del análisis exploratorio de variables operativas (vibración RMS, caudal, presión y gradiente térmico) y en la aplicación del Análisis de Componentes Principales. Los resultados muestran que más del 90% de la varianza total del sistema se concentra en cuatro componentes principales, destacando el componente PC3 como el más influyente en la predicción del RUL, asociado principalmente a variables térmicas y vibracionales.

c) Contraste del segundo objetivo específico y su hipótesis asociada

El segundo objetivo específico, relacionado con el desarrollo de un modelo predictivo basado en PCA y redes neuronales artificiales, se contrasta con la hipótesis que sostiene que esta integración permite representar eficazmente los patrones de deterioro de las bombas centrífugas.

La validez de esta hipótesis se evidencia mediante los valores de error obtenidos en las distintas fases del modelo. En particular, los valores de MAE inferiores a 20 horas en las fases de entrenamiento y prueba indican que el modelo logra aproximar con razonable precisión el valor real del RUL. Asimismo, el comportamiento del RMSE confirma que los errores más significativos se concentran en la fase de validación, lo cual es consistente con la variabilidad operativa de los datos y no invalida la capacidad predictiva general del modelo.

d) Contraste del tercer objetivo específico

El tercer objetivo específico, orientado a aplicar y evaluar el modelo predictivo mediante métricas de precisión y error, se contrasta con la hipótesis que plantea que dicha evaluación permite validar la efectividad del modelo para la predicción del RUL. El análisis comparativo de las métricas $R^2, MAE\ y\ RMSE$ entre las fases de entrenamiento, validación y prueba permitió evaluar la estabilidad y robustez del modelo. La presencia de un alto ajuste en entrenamiento, junto con un desempeño intermedio y coherente en la fase de prueba, confirma que el modelo es funcional para estimaciones iniciales del RUL

y puede ser utilizado como herramienta de apoyo para la planificación del mantenimiento predictivo.

4.5. Discusión

Los resultados obtenidos con el modelo predictivo PCA–RNA evidencian un comportamiento diferenciado a lo largo de las tres fases de evaluación: entrenamiento, validación y prueba. Durante la fase de entrenamiento, el modelo alcanzó un coeficiente de determinación $R^2 = 0.9213$, acompañado de errores reducidos (MAE = 16.93 h y RMSE = 20.75 h), lo que demuestra una adecuada capacidad de aprendizaje y un correcto ajuste a los datos históricos de operación. Estos valores indican que el modelo logró capturar de manera efectiva las relaciones no lineales existentes entre las variables operativas y la evolución del proceso de degradación del sistema de bombeo.

En la fase de validación, se observa una disminución significativa del desempeño predictivo, reflejada en un $R^2 = 0.1483$, junto con un incremento de los errores (MAE = 27.69 h y RMSE = 33.18 h). Este comportamiento pone de manifiesto una sensibilidad del modelo a la variabilidad operativa, ya que los datos de validación presentan combinaciones de condiciones que no estuvieron plenamente representadas durante el entrenamiento. No obstante, este resultado resulta consistente con el comportamiento esperado en sistemas industriales reales, donde los estados de degradación no siguen trayectorias homogéneas y se ven influenciados por múltiples factores externos.

Por su parte, en el conjunto de prueba, el modelo muestra un desempeño intermedio, con valores de error (MAE = 17.45 h, RMSE = 22.02 h) cercanos a los obtenidos en el entrenamiento y un $R^2 = 0.5203$, lo que indica una correlación moderada entre los valores reales y los valores predichos de RUL. Desde un contexto físico y operativo real, este nivel de correlación es factible, dado que la degradación de bombas centrífugas está sujeta a fenómenos transitorios, variaciones de carga hidráulica, cambios térmicos y eventos no estacionarios que introducen incertidumbre en la estimación del RUL. En este sentido, el modelo logra capturar la tendencia general del deterioro, aun cuando se enfrenta a datos completamente nuevos. Los valores de MAPE (44.45%) y sMAPE (71.12%) observados en el conjunto de prueba deben interpretarse

considerando que, en escenarios con vida útil remanente reducida, pequeños errores absolutos pueden traducirse en errores porcentuales elevados. Este fenómeno es ampliamente reconocido en aplicaciones reales de mantenimiento predictivo y no invalida la utilidad del modelo, sino que evidencia una limitación inherente a las métricas relativas cuando el RUL se aproxima a valores bajos.

Estos resultados son comparables con los obtenidos por Chicaiza y González (2021), quienes reportaron altos niveles de precisión durante el entrenamiento de modelos supervisados para motores eléctricos, pero una mayor estabilidad en validación al disponer de conjuntos de datos más amplios y balanceados. De manera similar, Vargas (2021) demostró que los modelos basados en redes neuronales y sistemas neurodifusos presentan una alta capacidad de identificación de estados de operación, aunque su desempeño depende fuertemente de la calidad y representatividad de las señales de entrada.

Asimismo, Alarcón (2022) reportó limitaciones en la generalización de modelos aplicados al monitoreo de turbomáquinas bajo condiciones de alta variabilidad operativa, lo que coincide con los hallazgos del presente estudio. Estos antecedentes refuerzan la evidencia de que, si bien las redes neuronales poseen un alto poder de representación, su desempeño práctico está condicionado por la diversidad y cobertura del conjunto de datos utilizado para el entrenamiento.

Por otro lado, Acosta (2023) logró estimaciones altamente precisas del RUL en baterías de ion-litio mediante el uso de técnicas avanzadas de ajuste de hiperparámetros y validación cruzada, lo que demuestra que la optimización sistemática del modelo y la adecuada selección de variables explicativas son factores determinantes para mejorar la precisión predictiva. Estos enfoques podrían incorporarse en futuras iteraciones del modelo PCA-RNA desarrollado. La revisión sistemática de Villacrés et al. (2024) identifica como principales desafíos en la predicción del RUL la extracción de características relevantes, la identificación del inicio del estado de degradación y la limitada generalización de los modelos. Estas dificultades coinciden con las limitaciones observadas en la presente investigación, particularmente en la representación de la

degradación progresiva de las bombas y la estabilidad del modelo ante escenarios operativos no estacionarios.

En relación con la sensibilidad del modelo frente a variaciones operativas, los resultados evidencian que el PCA-RNA responde de manera diferenciada a cambios en variables como caudal, presión, vibración y temperatura, las cuales se reflejan en la variación de los componentes principales. En particular, los componentes asociados a variables vibracionales y térmicas presentan mayor influencia en la estimación del RUL, lo que confirma que el modelo es capaz de captar cambios operativos relevantes vinculados al deterioro físico de las bombas centrífugas. Sin embargo, cuando dichas variaciones exceden los rangos presentes en el conjunto de entrenamiento, se observa una disminución en la precisión, confirmando que el desempeño del modelo está condicionado por el dominio operativo aprendido.

Desde el punto de vista físico, la degradación de componentes como rodamientos, sellos mecánicos e impulsor se manifiesta directamente en las variables analizadas. El desgaste de rodamientos se asocia a incrementos de vibración y temperatura, el deterioro del sello mecánico a variaciones térmicas locales y pérdidas de presión, y el desgaste del impulsor a disminuciones de caudal y eficiencia hidráulica, tal como señalan Villacrés et al. (2024). Estas relaciones permiten interpretar las variables medidas como indicadores indirectos del estado físico del sistema, facilitando la estimación del RUL.

En relación con la hipótesis planteada, los resultados permiten validarla de manera parcial. El objetivo general de estimar la vida útil remanente de bombas centrífugas mediante un modelo PCA-RNA se ve respaldado por el alto nivel de ajuste alcanzado en el entrenamiento y por la capacidad del modelo para capturar tendencias de degradación en el conjunto de prueba. No obstante, la reducción del desempeño en la validación evidencia que la precisión del modelo depende de la disponibilidad de datos representativos y de rangos operativos controlados.

En síntesis, el modelo PCA-RNA desarrollado se encuentra alineado con las tendencias actuales del mantenimiento predictivo basado en inteligencia artificial. Si bien los resultados confirman su viabilidad técnica y aplicabilidad práctica, la comparación

con estudios previos evidencia la necesidad de incorporar mayores volúmenes de datos, estrategias avanzadas de regularización y validación robusta, así como arquitecturas que consideren la dinámica temporal, con el fin de mejorar su precisión, robustez y capacidad de generalización en entornos industriales reales.

CONCLUSIONES

El modelo propuesto logró estimar el RUL de tres bombas centrífugas reduciendo la dimensionalidad de los datos y modelando relaciones no lineales entre variables operativas y degradación. Aunque presenta limitaciones de generalización, su aplicación demuestra el potencial del mantenimiento predictivo basado en inteligencia artificial, contribuyendo a optimizar tiempos de intervención, reducir costos operativos y aumentar la confiabilidad en entornos industriales.

El análisis operativo de las bombas B01, B02 y B03 evidenció un comportamiento estable en las variables críticas vibración RMS, caudal, presión y gradiente térmico, con valores dentro de los límites de la norma ISO 10816. Se identificaron fluctuaciones transitorias asociadas a paradas o cambios hidráulicos, con descensos en caudal y presión (≈ 0.35) y variaciones térmicas hasta -33 °C. Estos patrones reflejan una operación controlada, útil para la detección temprana de degradación y la proyección del RUL con base en datos reales de mantenimiento preventivo.

Se desarrolló un modelo híbrido PCA-RNA (MLP) procesando 36 variables operativas, donde el PCA redujo la dimensionalidad concentrando más del 90 % de la varianza en cuatro componentes. El componente PC3 resultó ser el más influyente ($\Delta MAE = 104.2$), relacionado con variaciones térmicas y vibracionales. Las curvas de probabilidad acumulada de falla demostraron la capacidad del modelo para reproducir escenarios realista, optimista y pesimista, destacando su utilidad para simular condiciones

Las métricas de desempeño (MAE, RMSE, sMAPE y R^2) muestran un alto ajuste en el entrenamiento ($R^2 = 0.9213$) y una reducción del desempeño en validación ($R^2 = 0.1483$), asociada a la variabilidad operativa y a condiciones no vistas durante el aprendizaje. En el conjunto de prueba, el modelo alcanza un desempeño intermedio ($R^2 = 0.5203$), evidenciando que logra capturar la tendencia general del RUL frente a datos nuevos. En conjunto, los resultados confirman que el enfoque PCA-RNA es viable como herramienta inicial para la estimación del deterioro, aunque su desempeño puede fortalecerse mediante regularización, ampliación del conjunto de datos y mejoras en la representación temporal del proceso de degradación.

RECOMENDACIONES

Se recomienda a las áreas de mantenimiento y confiabilidad recopilar un mayor volumen de datos históricos y en tiempo real que abarquen diferentes condiciones de operación variaciones de carga, temperatura ambiente, caudal y presión, para permitir fortalecer el proceso de entrenamiento del modelo y reducir el sobreajuste observado en las fases de validación y prueba. Una base de datos más representativa favorecería la detección de patrones más precisos de degradación, mejorando la capacidad del modelo para generalizar y predecir fallas en entornos reales.

Es conveniente realizar ajustes en la estructura de la red neuronal, aplicando técnicas de regularización (Dropout o Early Stopping), validación cruzada y búsqueda de hiperparámetros, lo que contribuirá a estabilizar los resultados y reducir la dispersión de errores observada en las métricas de validación. Con ello, el modelo podría mantener un equilibrio entre precisión y robustez, evitando la pérdida de rendimiento ante datos no vistos.

Se recomienda a las áreas de mantenimiento y confiabilidad integrar variables adicionales relacionadas con el comportamiento eléctrico del motor (corriente, potencia activa y reactiva, temperatura del devanado) para enriquecer la caracterización del estado operativo. Esto permitiría que el análisis de componentes principales (PCA) capture una mayor proporción de la varianza explicada, mejorando la sensibilidad del modelo ante señales tempranas de deterioro en los sistemas de bombeo.

Se recomienda al equipo de ingeniería de confiabilidad explorar arquitecturas de redes neuronales más sofisticadas, como las redes LSTM (Long Short-Term Memory) o CNN–LSTM híbridas, que permiten procesar secuencias temporales y capturar dependencias dinámicas entre las variables operativas. Estas estructuras podrían superar las limitaciones del modelo actual al manejar mejor la variabilidad temporal, ofreciendo predicciones más precisas y confiables sobre la vida útil remanente (RUL) de los equipos en condiciones industriales complejas.

Finalmente, se recomienda evaluar la integración del modelo predictivo desarrollado dentro de un sistema de gestión de mantenimiento computarizado (CMMS), de modo que las estimaciones de vida útil remanente (RUL) puedan emplearse como soporte para la programación de actividades de mantenimiento. La implementación conceptual podría contemplar la vinculación de los datos de sensores provenientes de sistemas SCADA o plataformas IoT con el CMMS, permitiendo generar alertas tempranas, priorizar órdenes de trabajo y optimizar la planificación del mantenimiento basada en condición.

BIBLIOGRAFÍA

Acosta, D. (2023). Predicción de la vida útil remanente de la batería de un vehículo de movilidad sostenible. Udea.edu.co; Universidad de Antioquia. https://hdl.handle.net/10495/37623

Adauto, R. (2021). Aplicación de la inteligencia artificial en la detección de fallas en los motores eléctricos de corriente continua de imán permanente. Uncp.edu.pe. http://hdl.handle.net/20.500.12894/7313

Alarcón González, L. I. (2022). Desarrollo e implementación de modelos de aprendizaje de máquinas para identificación de fallas en el monitoreo de la condición de activos en planta Celulosa Nueva Aldea. https://repositorio.uchile.cl/handle/2250/189642

Albornoz, G. (2021). Aplicación del aprendizaje automático supervisado en el mantenimiento predictivo de los motores eléctricos de inducción en las empresas mineras del Perú. Uncp.edu.pe. http://hdl.handle.net/20.500.12894/7567

Aranda, J. (2024). Influencia del método Duval en Python del análisis de aceite aislante en el diagnóstico preventivo del transformador de potencia 10/12 MVA - U.M. Raura. Uncp.edu.pe. http://hdl.handle.net/20.500.12894/10992

Caballero A., Llanca, E., Osorio, L. y Saico C. (2021). *Informe N°3: Bombas centrífugas* (Informe de laboratorio, Universidad Nacional Mayor de San Marcos, Laboratorio de Ingeniería Química 1).

Carrillo-Romero, J. A., y Rojas-Boyacá, J. A. (2013). El variador de velocidad como método de arranque ideal para motores eléctricos de inducción. Ingeniería Investigación y Desarrollo, 13(1), 34-39.

Chen, L., Wei, L., Wang, Y., Wang, J., y Li, W. (2022). Monitoring and predictive maintenance of centrifugal pumps based on smart sensors. Sensors, 22(6), 2106. https://doi.org/10.3390/s22062106

Chicaiza, R., y González, H. (2021). Sistema de Monitoreo en Nube Para el Diagnóstico Predictivo de Fallos de Motores Eléctricos [Tesis de Maestría]. Escuela Superior

Politécnica del Litoral, Guayaquil. http://www.dspace.espol.edu.ec/handle/123456789/54265

Contreras, J. (2020). Diseño de un modelo para mantenimiento predictivo en motores de inducción utilizando técnicas de la Industria 4.0. Utp.edu.pe. https://hdl.handle.net/20.500.12867/4275

Fernández González, À. (2024). Diseño, análisis y simulación de una bomba centrífuga de carga para VLCCs (Bachelor's thesis, Universitat Politècnica de Catalunya).

Fuentes, J. (2025). Mantenimiento e instalación de bombas centrifugas en Empresa de Operaciones Marítimas - Puerto de Ilo. Ujcm.edu.pe. https://hdl.handle.net/20.500.12819/3502

Grand View Research. (2024). Centrifugal pump market size, share & trends analysis report [Datos sobre el mercado global y previsiones]. Grand View Research. https://www.grandviewresearch.com/industry-analysis/centrifugal-pump-market-report

Hernández Sampieri, R., Fernández Collado, C., Baptista Lucio, M. (2014). Metodología de la investigación (6° ed.). México: McGraw Hill Interamericana Editores S.A. de C.V.

IBM. (2024). ¿Qué es el análisis de componentes principales (PCA)? IBM. https://www.ibm.com/mx-es/think/topics/principal-component-analysis

Jolliffe, I. T., y Cadima, J. (2016). Principal component analysis: A review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202

LEWA. (2025). Minimum life cycle costs: How low operating costs amortize higher initial costs [Incidencia de tiempos de inactividad en el costo de ciclo de vida]. LEWA. https://www.lewa.com/en-US/consulting/minimum-life-cycle-costs

Li, X., Wei, L., y He, J. (2018). Design and Implementation of Equipment Maintenance Predictive Model Based on Machine Learning. IOP Conference Series Materials

Science and Engineering, 466, 012001–012001. https://doi.org/10.1088/1757-899x/466/1/012001

Llihua, M. (2024). Mejoramiento del diagnóstico de fallas eléctricas de las electrobombas centrífugas en el Hospital Nacional Ramiro Prialé Prialé EsSalud - Huancayo. Uncp.edu.pe. http://hdl.handle.net/20.500.12894/11680

López, R. F., y Fernández, J. M. F. (2008). Las redes neuronales artificiales. Netbiblo.

Market Growth Reports. (2025, julio 28). Mining pumps market size, share & trends, 2033 [Reporte sobre bombas inteligentes y mantenimiento predictivo]. MarketGrowthReports. https://www.marketgrowthreports.com/market-reports/mining-pumps-market-113770

Noya, R., y Sutizal, R. (2023). Soluciones para una clase de programación lineal en dos niveles, vía algoritmos genéticos. Unac.edu.pe. https://hdl.handle.net/20.500.12952/7930

Orihuela Lazo, M. A. (2024). Algoritmo predictivo como herramienta de mantenimiento para la detección anticipada de fallas de funcionamiento y operación en sistemas fotovoltaicos autónomos domiciliarios rurales. http://repositorio.uncp.edu.pe/handle/20.500.12894/11192

Paredes Ego-Aguirre, C. A. (2019). Diseño y Construcción de Un Módulo de Bombas Centrífugas e Implementación de Guías Prácticas para El Estudio de sus Principales Características.

Quijano, M., y Crisostomo, D. (2023). Modelo predictivo para el mantenimiento preventivo de computadoras. Upc.edu.pe. http://hdl.handle.net/10757/667740

Recuay, P. (2021). Modelo estructural para determinar el tiempo de vida remanente en pavimentos flexibles. Upla.edu.pe. https://hdl.handle.net/20.500.12848/2674

Santillan Armas, C. P. (2017). Programa de mantenimiento centrado en confiabilidad para bombas centrifugas horizontales warman 450 mcr en Minera Cerro Corona.

Shlens, J. (2014). A tutorial on principal component analysis. arXiv preprint arXiv:1404.1100. https://arxiv.org/abs/1404.1100

Silva, A., y Elías, M. (2017). Análisis de algoritmos de control predictivo basado en el modelo aplicado al accionamiento hexafásico (Master's thesis, FIUNA).

Soto Zabala, D. F. (2021). Implementación de mantenimiento predictivo para una bomba centrífuga utilizando machine learning.

Strategic Market Research. (2023). Use case: Copper mine in Peru modernizes slurry pumping. Strategic Market Research. https://www.strategicmarketresearch.com/market-report/mining-pumps-market

Tsallis, C., Papageorgas, P., y Munteanu, R. A. (2025). Application-wise review of machine learning-based predictive maintenance: Trends, challenges, and future directions. Applied Sciences, 15(9), 4898. https://doi.org/10.3390/app15094898

Vargas, Á. (2021). Detección de anomalías en un motor eléctrico mediante inteligencia computacional. Uvadoc.uva.es. https://uvadoc.uva.es/handle/10324/47247

Villacrés Parra, S. R., Zavala León, M. A., y Viscaíno Cuzco, M. A. (2024). Predicción de vida útil remanente en rodamiento aplicando Machine Learning: Una revisión Sistemática de Literatura. ConcienciaDigital, 7(3.1), 46–67. https://doi.org/10.33262/concienciadigital.v7i3.1.3120

Zhou, S., Zhang, L., y Yang, X. (2025). Remaining useful life prediction of centrifugal pump bearings based on digital twins. Scientific Reports, 15, 19513. https://doi.org/10.1038/s41598-025-03952-2

Zhu, R., Liu, Y., Zhang, H., y Li, X. (2024). Predicting the remaining useful life of centrifugal pump bearings based on KPCA–LSTM. Energies, 17(16), 4167. https://doi.org/10.3390/en17164167

ANEXOS

Anexo N°1. Matriz de Consistencia

Título: "Desarrollo de un modelo predictivo para estimar el tiempo de vida útil remanente en bombas centrífugas mediante análisis de componentes principales"					
Objetivos	**Hipótesis**	**Variables**	**Dimensiones**	**Indicadores**	**Metodologías**
General	**General**	**Variable Independiente**			
Desarrollar un modelo predictivo basado en el Análisis de Componentes Principales (PCA) para estimar el tiempo de vida útil remanente de bombas centrífugas	El desarrollo de un modelo predictivo basado en el Análisis de Componentes Principales (PCA) permitirá estimar con precisión el tiempo de vida útil remanente de las bombas centrífugas, optimizando la planificación del mantenimiento y reduciendo la ocurrencia de fallas inesperadas.	VI: Condiciones operativas de las bombas centrífugas	**Comportamiento operativo** **Desarrollo del modelo predictivo** **Evaluación del modelo**	▪ Parámetros operativos (caudal, presión, vibración, temperatura) ▪ Frecuencia de fallas registradas ▪ Intervalos entre mantenimientos preventivos ▪ Análisis de Componentes Principales (PCA) ▪ Implementación de redes neuronales ▪ Integración de datos operativos históricos ▪ Error medio cuadrático (MSE) ▪ Coeficiente de determinación (R^2)	La investigación corresponde al tipo aplicado. La investigación corresponde al nivel explicativo. La investigación corresponde a un diseño no experimental y longitudinal. La población de estudio corresponde a todas las bombas centrífugas utilizadas en procesos industriales donde se requiere un monitoreo continuo para la gestión del mantenimiento. La muestra estará conformada por un conjunto representativo de bombas centrífugas seleccionadas mediante un muestreo no probabilístico por conveniencia.
Específicos	**Específicos**	**Variable Dependiente**			
Analizar el comportamiento operativo de bombas centrífugas sometidas a mantenimiento preventivo para estimar el tiempo de vida útil remanente.	El análisis del comportamiento operativo de bombas centrífugas sometidas a mantenimiento preventivo permite identificar patrones de funcionamiento asociados al desgaste, que contribuyen a estimar su tiempo de vida útil remanente.	VD1: Estimación del tiempo de vida útil remanente de las bombas centrífugas			
Desarrollar un modelo predictivo basado en Análisis de Componentes Principales (PCA) y redes neuronales para	El desarrollo de un modelo predictivo basado en Análisis de Componentes Principales (PCA) y redes neuronales permite representar eficazmente				

estimar el tiempo de vida útil remanente.	los patrones de deterioro, mejorando la estimación del tiempo de vida útil remanente de las bombas centrífugas.		**Predicción del tiempo de vida útil**	▪ Precisión de la predicción ▪ Errores en la predicción ▪ Confiabilidad del modelo ▪ Tendencia del desgaste	
Aplicar y evaluar el modelo predictivo mediante métricas de precisión y error.	La aplicación del modelo predictivo y su evaluación mediante métricas estadísticas permite validar su efectividad para predecir el tiempo de vida útil remanente en bombas centrífugas.				

Anexo N°2. Data de las bombas centrifugas

timestamp	id_bomba	Q_m3h	P_bar	vib_rms_mms	T_rod_C	T_motor_C	T_amb_C	I_A	estado_bomba
2025-01-01 00:00:00	B01	76.38	1.48	1.018	45.11	60.29	20.4	10	1
2025-01-02 00:00:00	B01	75.89	1.47	1.039	45.17	60.24	19.55	10	1
2025-01-03 00:00:00	B01	76.49	1.47	1.139	45.33	60.35	20.7	10	1
2025-01-04 00:00:00	B01	77.16	1.47	1.072	45.32	60.25	19.3	10	1
2025-01-05 00:00:00	B01	75.82	1.45	1.106	45.42	60.3	20.29	10	1
2025-01-06 00:00:00	B01	75.82	1.46	1.091	45.29	60.13	21.1	10	1
2025-01-07 00:00:00	B01	77.2	1.47	1.135	45.16	60.26	19.5	10	1
2025-01-08 00:00:00	B01	76.58	1.49	1.098	45.16	60.58	19.72	10	1
2025-01-09 00:00:00	B01	75.64	1.48	1.124	45.17	60.09	20.05	10	1
2025-01-10 00:00:00	B01	76.41	1.45	1.133	46.17	60.24	19.75	10	1
2025-01-11 00:00:00	B01	75.65	1.48	1.11	45.19	60.29	19.22	10	1
2025-01-12 00:00:00	B01	75.65	1.47	1.19	45.36	60.27	20.03	10	1
2025-01-13 00:00:00	B01	76.18	1.46	1.179	45.3	60.05	19.47	10	1
2025-01-14 00:00:00	B01	74.55	1.48	1.16	45.22	60.13	20.24	10	1
2025-01-15 00:00:00	B01	74.69	1.49	1.175	45.12	60.11	19.54	10	1
2025-01-16 00:00:00	B01	75.57	1.49	1.271	45.25	60.28	20.77	10	1
2025-01-17 00:00:00	B01	75.23	1.46	1.188	45.26	60.04	19.61	10	1
2025-01-18 00:00:00	B01	76.24	1.47	1.195	45.1	60.47	19.84	10	1
2025-01-19 00:00:00	B01	75.31	1.48	1.185	45.17	60.12	20.41	10	1
2025-01-20 00:00:00	B01	0	0	0.714	20.88	21.38	19.38	0	0
2025-01-21 00:00:00	B01	77.11	1.46	1.318	45.73	60.23	20.11	10	1
2025-01-22 00:00:00	B01	75.83	1.47	1.249	45.59	60.3	20.65	10	1
2025-01-23 00:00:00	B01	76.05	1.46	1.253	45.24	60.37	19.2	10	1
2025-01-24 00:00:00	B01	74.92	1.45	1.3	45.52	60.19	20.09	10	1

2025-01-25 00:00:00	B01	75.59	1.48	1.259	45.18	60.12	20.13	10	1
2025-01-26 00:00:00	B01	76.08	1.49	1.289	45.37	60.27	20.39	10	1
2025-01-27 00:00:00	B01	75.13	1.47	1.418	45.06	60.2	19.38	10	1
2025-01-28 00:00:00	B01	76.29	1.49	1.291	45.36	60.08	19.34	10	1
2025-01-29 00:00:00	B01	75.54	1.48	1.308	45.26	60.31	20.26	10	1
2025-01-30 00:00:00	B01	75.78	1.46	1.303	45.25	60.51	20.15	10	1
2025-01-31 00:00:00	B01	75.54	1.48	1.461	45.27	60.2	20.13	10	1
2025-02-01 00:00:00	B01	77.41	1.49	1.32	45.11	60.32	20.17	10	1
2025-02-02 00:00:00	B01	75.99	1.47	1.333	45.06	60.13	19.66	10	1
2025-02-03 00:00:00	B01	0	0	1.036	21.62	22.12	20.12	0	0
2025-02-04 00:00:00	B01	76.63	1.43	1.364	45.7	60.12	20.15	10	1
2025-02-05 00:00:00	B01	75.07	1.48	1.383	45.24	60.19	19.64	10	1
2025-02-06 00:00:00	B01	76.16	1.47	1.372	45.66	60.34	20.93	10	1
2025-02-07 00:00:00	B01	74.51	1.47	1.473	45.11	60.12	20.24	10	1
2025-02-08 00:00:00	B01	74.99	1.47	1.481	45.26	60.1	19.4	10	1
2025-02-09 00:00:00	B01	76.15	1.44	1.46	45.32	60.42	20.33	10	1
2025-02-10 00:00:00	B02	29.66	1.96	0.963	43.11	58.05	19.85	8	1
2025-02-11 00:00:00	B02	30.38	1.97	1.008	43.52	58.25	19.63	8	1
2025-02-12 00:00:00	B02	29.99	1.94	1.055	43.26	58.28	20.23	8	1
2025-02-13 00:00:00	B02	29.49	1.9	0.975	43.04	58.25	19.48	8	1
2025-02-14 00:00:00	B02	29.69	1.98	1.071	43.07	58.42	19.32	8	1
2025-02-15 00:00:00	B02	30.4	2	1.122	43.58	58.2	19.72	8	1
2025-02-16 00:00:00	B02	30.13	1.96	1.051	43.3	58.21	19.6	8	1
2025-02-17 00:00:00	B02	29.98	1.97	1.058	43.21	58.25	19.33	8	1
2025-02-18 00:00:00	B02	29.33	1.97	1.055	43.38	58.28	19.84	8	1
2025-02-19 00:00:00	B02	30.42	2	1.083	43.06	58.23	19.17	8	1
2025-02-20 00:00:00	B02	30.56	1.98	1.126	43.06	58.13	20.86	8	1

2025-02-21 00:00:00	B02	30.51	1.97	1.113	43.17	58.18	19.25	8	1
2025-02-22 00:00:00	B02	0	0	0.609	21.18	21.68	19.68	0	0
2025-02-23 00:00:00	B02	30.11	1.96	1.276	43.21	58.16	19.8	8	1
2025-02-24 00:00:00	B02	30.15	1.93	1.098	43.57	58.42	20.84	8	1
2025-02-25 00:00:00	B02	29.54	1.95	1.174	43.31	58.07	19.13	8	1
2025-02-26 00:00:00	B02	30.05	1.98	1.229	43.69	58.33	20.07	8	1
2025-02-27 00:00:00	B02	30.32	1.96	1.157	43.64	58.22	20.52	8	1
2025-02-28 00:00:00	B02	30.23	1.96	1.152	43.57	58.17	19.85	8	1
2025-03-01 00:00:00	B02	0	0	0.693	20.92	21.42	19.42	0	0
2025-03-02 00:00:00	B02	29.74	1.97	1.197	43.31	58.84	19.95	8	1
2025-03-03 00:00:00	B02	30.24	1.93	1.198	43.21	58.25	20.76	8	1
2025-03-04 00:00:00	B02	30.19	1.97	1.271	43.27	58.24	19.69	8	1
2025-03-05 00:00:00	B02	29.32	1.96	1.303	43.55	58.43	20.72	8	1
2025-03-06 00:00:00	B02	29.26	1.93	1.287	43.06	58.13	20.13	8	1
2025-03-07 00:00:00	B02	29.98	1.98	1.322	43.37	58.08	20.62	8	1
2025-03-08 00:00:00	B02	29.82	1.96	1.374	43.08	58.55	20.11	8	1
2025-03-09 00:00:00	B02	29.91	2	1.329	43.16	58.46	19.76	8	1
2025-03-10 00:00:00	B02	29.88	1.99	1.298	43.24	58.34	20.82	8	1
2025-03-11 00:00:00	B02	30.3	1.95	1.348	43.22	58.11	20.11	8	1
2025-03-12 00:00:00	B03	20.31	0.99	0.859	40.01	55.19	20.82	6	1
2025-03-13 00:00:00	B03	20.11	0.99	0.833	40.13	55.06	19.98	6	1
2025-03-14 00:00:00	B03	20.42	0.98	0.893	40.34	55.1	20.04	6	1
2025-03-15 00:00:00	B03	20.28	0.97	0.978	40.06	55.15	19.53	6	1
2025-03-16 00:00:00	B03	19.6	0.98	0.895	40.41	55.13	19.61	6	1
2025-03-17 00:00:00	B03	20.21	0.99	1.022	40.06	55.14	18.9	6	1
2025-03-18 00:00:00	B03	20.14	0.96	0.927	40.33	55.16	20.03	6	1
2025-03-19 00:00:00	B03	19.94	0.99	0.92	40.35	55.3	21.25	6	1

2025-03-20 00:00:00	B03	19.72	1	0.967	40.43	55.34	19.49	6	1
2025-03-21 00:00:00	B03	20.12	0.96	1.101	40.59	55.24	21.49	6	1
2025-03-22 00:00:00	B03	19.58	0.97	1.098	40.19	55.5	20.71	6	1
2025-03-23 00:00:00	B03	0	0	0.578	21.26	21.76	19.76	0	0
2025-03-24 00:00:00	B03	0	0	0.562	21.51	22.01	20.01	0	0
2025-03-25 00:00:00	B03	19.75	0.98	1.108	40.2	55.23	19.93	6	1
2025-03-26 00:00:00	B03	20.24	0.98	1.053	40.35	55.18	20.85	6	1
2025-03-27 00:00:00	B03	19.99	0.97	1.011	40.15	55.44	20.26	6	1
2025-03-28 00:00:00	B03	20.03	0.99	1.04	40.53	55.11	20.59	6	1
2025-03-29 00:00:00	B03	20.18	0.99	1.157	40.25	55.33	20.18	6	1
2025-03-30 00:00:00	B03	20	0.98	1.198	40.11	55.52	19.94	6	1
2025-03-31 00:00:00	B03	20.09	1	1.123	40.58	55.89	20.16	6	1
2025-04-01 00:00:00	B03	20.16	0.98	1.149	40.5	55.1	19.8	6	1
2025-04-02 00:00:00	B03	20.11	0.97	1.177	40.88	55.29	19.27	6	1
2025-04-03 00:00:00	B03	20.06	0.98	1.24	40.52	55.2	20.41	6	1
2025-04-04 00:00:00	B03	20.36	0.98	1.12	40.05	55.52	19.25	6	1
2025-04-05 00:00:00	B03	20.25	0.99	1.197	40.14	55.38	19.87	6	1
2025-04-06 00:00:00	B03	19.98	0.99	1.226	40.27	55.28	19.78	6	1
2025-04-07 00:00:00	B03	19.82	0.99	1.198	40.09	55.34	20.32	6	1
2025-04-08 00:00:00	B03	20.13	0.99	1.203	40.11	55.29	20.78	6	1
2025-04-09 00:00:00	B03	20.04	0.98	1.239	40.21	55.1	20.6	6	1
2025-04-10 00:00:00	B03	20.34	0.97	1.273	40.15	55.53	20.34	6	1