



Detección de Fallas en Equipos Rotativos mediante la Combinación de Técnicas Espectrales y Análisis Modal con Machine Learning

Emanuel Orzusa^{a,1}, Prof. Dr. Ing. Diego Stalder^{b,2}, Prof. Dr. Ing. Carlos Méndez^{c,3}, Prof. MSc. Ing. Gustavo Martínez^{d,4}, Prof. Ing. Orlando Benítez^{e,5}

^aAlumno, Facultad de Ingeniería, Universidad Nacional de Asunción, San Lorenzo, Paraguay

^bAsesor de Ingeniería Mecánica, Facultad de Ingeniería, Universidad Nacional de Asunción, Paraguay

^cAsesor Externo, Facultad de Ciencias Químicas, Universidad Nacional de Asunción, Paraguay

^dAsesor de Ingeniería Mecánica, Facultad de Ingeniería, Universidad Nacional de Asunción, Paraguay

^eAsesor de Ingeniería Mecánica, Facultad de Ingeniería, Universidad Nacional de Asunción, Paraguay

PALABRAS CLAVE

Diagnóstico de fallas
Vibraciones mecánicas
FFT
HODMD
Aprendizaje automático

RESUMEN

Este trabajo presenta una metodología de diagnóstico de fallas en equipos rotativos basada en el procesamiento por ventanas de señales de vibración y la comparación de dos enfoques complementarios. El Pipeline A construye, por condición, una biblioteca de modos obtenidos por Higher-Order Dynamic Mode Decomposition (HODMD) y clasifica nuevas ventanas mediante métricas de similitud modal y penalización en frecuencia. El Pipeline B transforma cada ventana en un vector de características que combina picos espectrales (FFT), descriptores por bandas tipo Box/STFT y rasgos modales derivados de HODMD, y entrena clasificadores supervisados con optimización de hiperparámetros.

KEYWORDS

Fault diagnosis
Mechanical vibrations
FFT
HODMD
Machine learning

ABSTRACT

This work presents a fault diagnosis methodology for rotating machinery based on windowed processing of vibration signals and the comparison of two complementary approaches. Pipeline A builds, for each condition, a library of modes obtained through Higher-Order Dynamic Mode Decomposition (HODMD) and classifies new windows using modal similarity metrics and a frequency-penalty term. Pipeline B transforms each window into a feature vector that combines spectral peaks (FFT), band-based descriptors (Box/STFT), and HODMD-derived modal features, and trains supervised classifiers with hyperparameter optimization.

1. INTRODUCCIÓN

El mantenimiento basado en condición (CBM) utiliza variables medidas del activo (por ejemplo, vibraciones) para sustentar decisiones de intervención según su estado, en lugar de intervenir únicamente por tiempo o después de la falla. En este marco, las metodologías de diagnóstico basadas en datos pueden integrarse como componentes técnicos dentro de una estrategia de gestión de activos físicos alineada con ISO 55000.

En equipos rotativos, la relación directa entre el comportamiento dinámico y las señales de vibración permite plantear el problema como un diagnóstico multiclase a nivel de ventanas

temporales. Este trabajo propone una metodología que transforma señales crudas en representaciones compactas y comparables mediante una cadena común de ventaneo y preprocesamiento, análisis espectral/tiempo-frecuencia (FFT) y descriptores Box/STFT y extracción modal (HODMD). Sobre esa base se comparan dos enfoques paralelos: (i) un diagnóstico físicamente interpretable mediante una biblioteca de modos HODMD y medidas de similitud (*Pipeline A*) y (ii) un diagnóstico supervisado mediante aprendizaje automático sobre un vector unificado de características espectrales y modales (*Pipeline B*), evaluados bajo la misma información de entrada para asegurar comparabilidad.

2. OBJETIVOS GENERALES Y ESPECÍFICOS

2.1. Objetivo General

Desarrollar y evaluar una metodología para la detección de fallas en equipos rotativos, basada en el análisis de señales de vibración, que combine técnicas espectrales y modales (FFT y HODMD) con modelos de aprendizaje automático, comparando dos pipelines de diagnóstico sobre un dataset seleccionado

¹Alumno de la carrera de Ingeniería Mecánica de Facultad de Ingeniería de la UNA, San Lorenzo, Paraguay (email: eorzusa@fiuna.edu.py)

²Asesor de la Facultad de Ingeniería de la UNA, San Lorenzo Paraguay (email: stalderdiego@gmail.com)

³Asesor de la Facultad de Ciencias Químicas de la UNA. San Lorenzo, Paraguay (email: carlmendezg@gmail.com)

⁴Asesor de la Facultad de Ingeniería de la UNA, San Lorenzo, Paraguay (email: gmartinez@ing.una.py)

⁵Asesor de la Facultad de Ingeniería de la UNA, San Lorenzo, Paraguay (email: odbenitez@fiuna.edu.py)

e incorporando una validación experimental preliminar en un banco de pruebas físico.

2.2. Objetivos Específicos

- **Recopilación y Preprocesamiento de Datos:** Recopilar y preprocesar datos de vibraciones provenientes de repositorios públicos para establecer parámetros iniciales en la aplicación de las técnicas FFT y HODMD.
- **Extracción de Características Espectrales y Modelado ML:** Desarrollar un método para la extracción de características espectrales mediante FFT y utilizar dichas características para desarrollar y optimizar modelos de ML (incluyendo la optimización de hiperparámetros) que permitan detectar fallas.
- **Extracción y Optimización de Características Espacio-Temporales con HODMD:** Aplicar HODMD para extraer características espacio-temporales de las señales de vibración y optimizar sus parámetros para maximizar la calidad de la información.
- **Evaluación y Comparación de Rendimientos por familia de características:** Evaluar el rendimiento individual de las familias de características de orígenes espectrales y espaciales para su posterior comparativa.
- **Validación experimental preliminar:** Aplicar los pipelines de diagnóstico desarrollados (A y B) a datos adquiridos en un banco de pruebas de vibraciones previamente modificado, bajo un conjunto reducido de condiciones controladas. El objetivo es analizar el comportamiento y la aplicabilidad práctica de la metodología en un entorno experimental real, contrastando los resultados con los obtenidos en el dataset VAT (*Vibration, acoustic, temperature, and motor current dataset of rotating machine under varying operating conditions for fault diagnosis*) [1].

3. MARCO TEÓRICO

El diagnóstico de fallas en maquinaria rotativa se enmarca en el mantenimiento basado en condición y la gestión de activos, donde las decisiones se apoyan en mediciones objetivas (por ejemplo, vibraciones) y en modelos de análisis que permiten identificar estados anómalos con anticipación [2, 3]. Esta sección resume: (i) el rol del mantenimiento y la gestión de activos, (ii) los equipos rotativos y sus fallas típicas, (iii) el monitoreo de condición por vibraciones con técnicas espectrales, y (iv) el análisis modal basado en datos y el uso de modelos de aprendizaje automático para clasificación de fallas [4, 5, 6, 7].

3.1. Mantenimiento, mantenimiento basado en condición y gestión de activos

El mantenimiento industrial busca asegurar disponibilidad, confiabilidad y seguridad, minimizando costos y riesgos asociados a fallas imprevistas [2]. La tendencia histórica pasa de esquemas correctivos y preventivos hacia enfoques predictivos, donde se interviene según el estado medido del activo, y hacia enfoques prescriptivos, donde se recomiendan acciones considerando costo, riesgo y operación [3]. En este contexto, el mantenimiento basado en condición (CBM) es un pilar: la decisión se fundamenta en variables de condición (vibraciones, temperatura, lubricación, corrientes, etc.), con énfasis particular en vibraciones para máquinas rotativas [4, 5].

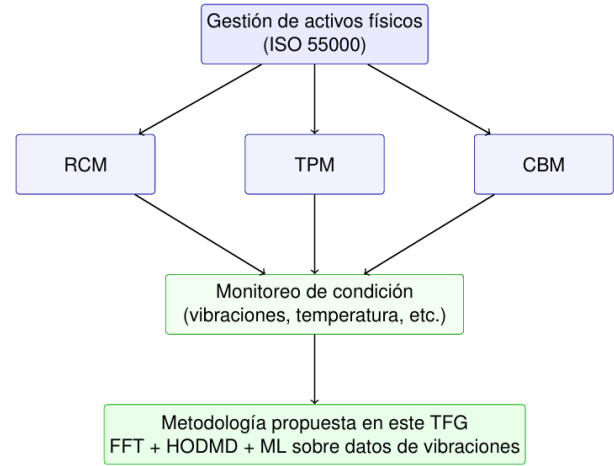


Fig. 1. Relación entre gestión y filosofías de mantenimiento.

A nivel organizacional, la familia ISO 55000 propone un marco de gestión de activos orientado al ciclo de vida, alineando mantenimiento con objetivos estratégicos, riesgo e inversiones [8]. Metodologías de diagnóstico basadas en señales y datos pueden interpretarse como herramientas técnicas específicas dentro de esa estrategia: aportan evidencia para priorizar intervenciones, justificar instrumentación y reducir incertidumbre en la toma de decisiones [8].

La Figura 1 indica la relación entre gestión de activos físicos (ISO 55000), filosofías de mantenimiento (RCM, TPM, CBM), monitoreo de condición y metodología de diagnóstico propuesta.

3.2. Equipos rotativos y modos de falla relevantes

Los equipos rotativos (motores, bombas, ventiladores, compresores, etc.) son críticos por su presencia transversal en procesos industriales y porque muchas fallas alteran directamente la respuesta vibratoria medida en apoyos y estructuras [4, 5]. En trenes típicos motor-acople-carga, la alineación, la rigidez de montaje y el estado de rodamientos condicionan tanto los niveles como las firmas espectrales [4].

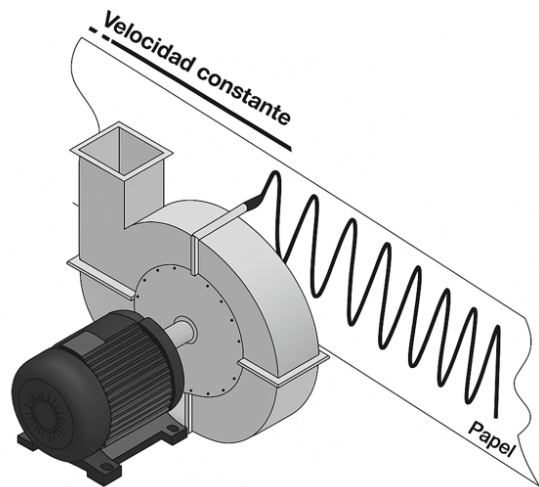


Fig. 2: Registro conceptual de vibraciones mediante marcador sobre banda de papel.

Entre los modos de falla más frecuentes y diagnosticables por vibraciones se incluyen: desbalance (componente dominante a $1 \times \text{RPM}$), desalineación (energía en $1 \times$, $2 \times$ y armónicos), holguras mecánicas (armónicos, subarmónicos y modulación),

defectos en rodamientos (componentes en frecuencias características como BPFI/BPFO y bandas laterales), problemas en la máquina accionada (frecuencia de paso de palas y ruido de banda ancha), resonancias estructurales (amplificación cerca de frecuencias naturales) y ciertos defectos eléctricos con firmas acopladas a vibración/corriente [4, 5, 3]. Este catálogo es importante porque define qué patrones buscar y qué características extraer para un diagnóstico reproducible.

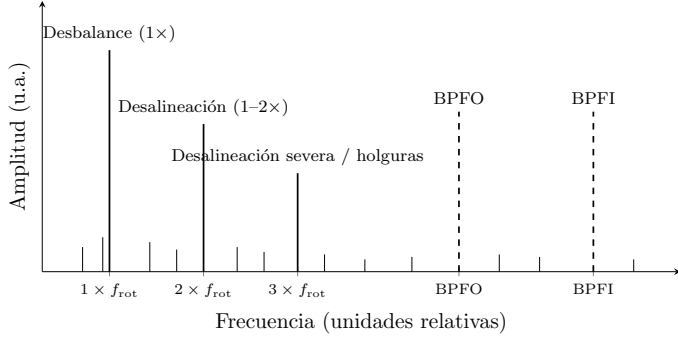


Fig. 3: Espectro conceptual de velocidad vibratoria con patrones típicos de desbalance, desalineación, holguras y defectos en rodamientos.

3.3. Monitoreo de condición por vibraciones: adquisición, severidad y análisis en frecuencia

El monitoreo por vibraciones se basa en medir aceleración/velocidad/desplazamiento con sensores (típicamente acelerómetros) instalados en carcasas de rodamientos o puntos estructuralmente representativos, seguido de procesamiento para obtener indicadores y apoyar decisiones [4, 5]. En máquinas con velocidad aproximadamente constante, el análisis en frecuencia es central: múltiples fallas se reflejan como picos en órdenes ($1\times$, $2\times$, etc.) y como componentes asociadas a fenómenos periódicos (paso de palas, engranajes, frecuencias características de rodamientos) [5].

3.4. Análisis modal basado en datos: DMD/HODMD y similitud modal

Además del espectro clásico, el análisis modal basado en datos permite describir la dinámica como combinación de modos espacio-temporales con frecuencia y amortiguamiento, estimados directamente desde mediciones multicanal, sin requerir un modelo físico explícito del sistema [9, 6]. En este marco, DMD aproxima la evolución entre “instantáneas” mediante un operador lineal reducido, y HODMD extiende el enfoque incorporando incrustaciones temporales (retardos), lo que mejora la separación de frecuencias próximas y la robustez frente a ruido en registros finitos [6].

En el análisis modal clásico, la respuesta de un sistema lineal de múltiples grados de libertad se expresa como superposición de modos propios espaciales y coordenadas modales temporales. Para

$$\mathbf{M}\ddot{\mathbf{q}}(t) + \mathbf{C}\dot{\mathbf{q}}(t) + \mathbf{K}\mathbf{q}(t) = \mathbf{f}(t), \quad (1)$$

el problema de autovalores

$$(\mathbf{K} - \omega_r^2 \mathbf{M})\boldsymbol{\phi}_r = \mathbf{0} \quad (2)$$

proporciona frecuencias naturales ω_r y modos propios $\boldsymbol{\phi}_r$ que permiten reconstruir la respuesta como combinación modal [4].

DMD y HODMD trasladan esta idea al ámbito *basado en datos*, sin requerir las matrices $\mathbf{M}$, $\mathbf{C}$ y $\mathbf{K}$. Siguiendo [6], se

busca representar un vector de observables $\mathbf{v}(t) \in \mathbb{C}^J$ (por ejemplo, aceleraciones en J sensores) como

$$\mathbf{v}(t) \approx \sum_{n=1}^N a_n \mathbf{u}_n \exp[(\delta_n + i\omega_n)(t - t_1)], \quad t_1 \leq t \leq t_1 + T, \quad (3)$$

donde $\mathbf{u}_n$ son modos espaciales, a_n amplitudes iniciales, δ_n tasas de crecimiento/decaimiento y ω_n frecuencias angulares. Con $\lambda_n = \delta_n + i\omega_n$:

$$\mathbf{v}(t) \approx \sum_{n=1}^N a_n \mathbf{u}_n e^{\lambda_n(t-t_1)}. \quad (4)$$

En sistemas reales, autovalores y modos aparecen en pares conjugados complejos, lo que permite reconstruir oscilaciones reales a partir de combinaciones de modos conjugados [6]. El problema central es identificar, a partir de $\{\mathbf{v}(t_k)\}$, los conjuntos $\{\lambda_n, \mathbf{u}_n, a_n\}_{n=1}^N$ que aproximan la dinámica observada.

Para comparar modos espaciales y cuantificar su similitud se emplean criterios tipo MAC (valores en $[0, 1]$), útiles para agrupar modos coherentes, evaluar repetibilidad y asociar modos observados a patrones representativos [4]. En diagnóstico, esto habilita medidas objetivas de “parecido modal” complementarias a la comparación espectral.

3.5. Aprendizaje automático para diagnóstico de fallas: modelos y evaluación

El aprendizaje automático (ML) permite construir clasificadores que asignan estados de máquina a partir de vectores de características extraídas de señales (por ejemplo, descriptores espectrales y/o modales), aprendiendo la relación entre datos y etiquetas a partir de ejemplos [7, 3]. En aplicaciones supervisadas multiclase, es habitual utilizar modelos clásicos como regresión logística (línea base), árboles y ensambles (Random Forest, boosting), SVM y kNN, seleccionando hiperparámetros mediante validación para controlar sobreajuste y maximizar generalización [7, 3].

La evaluación no debe reducirse a exactitud global cuando hay desbalance de clases o cuando ciertos errores son más costosos. Por ello se emplean métricas por clase y promedios (por ejemplo, F1 macro o exactitud balanceada), además del análisis detallado de la matriz de confusión para identificar confusiones sistemáticas entre fallas con firmas similares [7, 3]. Este marco define cómo comparar modelos entrenados con diferentes conjuntos de características y cómo interpretar su desempeño en términos de errores relevantes para mantenimiento.

La evaluación de modelos de clasificación exige métricas coherentes con la aplicación. En diagnóstico de fallas, no basta con maximizar la proporción global de aciertos, ya que ciertos errores (no detectar una falla severa) son mucho más críticos que otros [3].

Sea una clasificación multi-clase con K estados y matriz de confusión $\mathbf{N} = [N_{ij}]$, donde N_{ij} es el número de muestras de clase verdadera i clasificadas como j . En el caso binario, la Figura 4 ilustra los conteos básicos: verdaderos positivos (TP), falsos positivos (FP), falsos negativos (FN) y verdaderos negativos (TN).

A partir de la matriz de confusión se definen, entre otras, las siguientes métricas:

- Exactitud (*accuracy*):

$$\text{Accuracy} = \frac{1}{N} \sum_{i=1}^K N_{ii}, \quad N = \sum_{i=1}^K \sum_{j=1}^K N_{ij},$$

		Matriz de Confusión	
		Clases reales	
Clases predichas	Predicho Positivo	Real Positivo Verdadero Positivo (VP)	Real Negativo Falso Positivo (FP)
	Predicho Negativo	Falso Negativo (FN)	Verdadero Negativo (VN)

Fig. 4: Matriz de confusión binaria: clases reales vs. predichas y sus cuatro resultados básicos. **Fuente:** Elaboración propia.

equivalente en el caso binario a $\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FP} + \text{FN} + \text{TN})$. Es simple pero puede resultar engañosa si las clases están desbalanceadas.

- Precisión, *recall* y F1 por clase (caso binario, para la clase positiva):

$$\text{Precisión} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}},$$

$$\text{F1} = 2 \frac{\text{Precisión} \times \text{Recall}}{\text{Precisión} + \text{Recall}}.$$

En el caso multi-clase estas métricas se definen por clase usando un esquema uno-contra-todos [10].

- Promedios macro, micro y ponderado [7]: para $K > 2$ se calcula precisión, *recall* y F1 por clase y luego se promedian:
 - *micro-promedio*: agrega primero TP, FP y FN de todas las clases; da más peso a clases frecuentes;
 - *macro-promedio*: media aritmética de las métricas por clase; otorga igual peso a cada clase y es útil cuando se desea tratar fallas raras y estados frecuentes con la misma importancia;
 - *promedio ponderado*: media ponderada por el tamaño de cada clase.

En escenarios desbalanceados, los promedios *macro* suelen ser más informativos que la exactitud global [3].

- Exactitud balanceada (*balanced accuracy*):

$$\text{BalancedAccuracy} = \frac{1}{K} \sum_{i=1}^K \text{Recall}_i,$$

equivalente al *recall macro*; resulta útil cuando las distribuciones de clase son muy desiguales.

- Curvas ROC y AUC: si el clasificador proporciona puntuaciones continuas o probabilidades por clase, pueden construirse curvas ROC (*Receiver Operating Characteristic*) que muestran la tasa de verdaderos positivos frente a la tasa de falsos positivos al variar el umbral. El área bajo la curva (AUC) resume la capacidad discriminativa (1: ideal, 0,5: azar) [10]. Existen extensiones al caso multi-clase.

La Figura 5 ilustra de forma conceptual el espacio ROC, la diagonal correspondiente a un clasificador aleatorio y dos curvas ROC asociadas a modelos con distinta capacidad de discriminación.

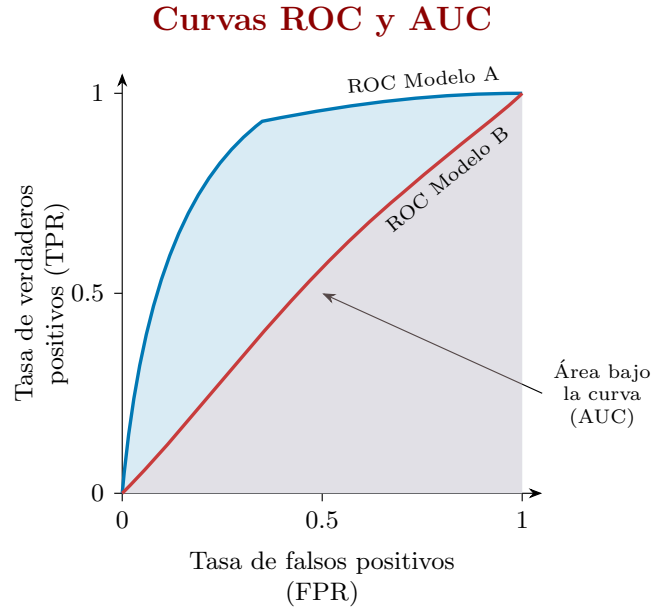


Fig. 5: Curvas ROC conceptuales para dos clasificadores binarios y la recta asociada a un clasificador aleatorio. **Fuente:** Elaboración propia a partir de [10, 7].

4. METODOLOGÍA Y DISEÑO DEL PIPELINE DE DIAGNÓSTICO

Este trabajo propone y evalúa una metodología integrada para el diagnóstico de fallas en máquinas rotativas a partir de señales de vibración multicanal, combinando análisis espectral (FFT y PSD), descriptores tiempo-frecuencia (Box/STFT), descomposición modal mediante Higher-Order Dynamic Mode Decomposition (HODMD) y modelos supervisados de aprendizaje automático [1, 11, 5, 12, 9, 6]. El diagnóstico se formula como un problema de clasificación multiclase, donde cada ventana de vibración debe asignarse a una condición de estado dentro de un conjunto discreto de clases (condición normal y distintos tipos/severidades de falla).

La estrategia central consiste en segmentar los registros en ventanas temporales de duración fija, tratándolas como instancias de decisión. En la evaluación cuantitativa principal se emplea el dataset VAT [1], con dos acelerómetros biaxiales (cuatro canales: YS1, ZS1, YS2, ZS2), tres niveles de carga (0, 2 y 4 Nm) y escenarios asociados a condición y severidad. Se utiliza un conjunto de 45 escenarios y 900 ventanas (20 ventanas de 10 s por escenario), manteniendo metadatos y trazabilidad mediante un manifiesto global (CSV) que vincula cada ventana con su origen y etiqueta.

El flujo de procesamiento se organiza en dos *pipelines* paralelos con etapas iniciales comunes y reglas de decisión distintas: (A) *Pipeline A*, basado en una biblioteca de modos HODMD y medidas de similitud modal (enfoque interpretable); (B) *Pipeline B*, basado en vectores de características (FFT/PSD, Box/STFT y HODMD) y clasificadores supervisados [2, 3, 10, 7].

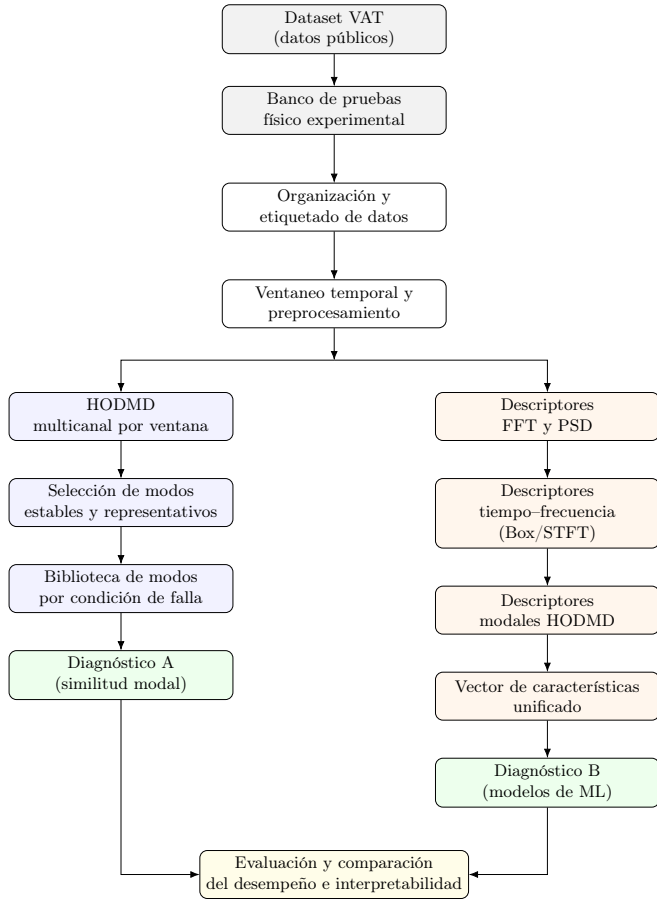


Fig. 6. Esquema general de la metodología propuesta

La Figura 6 resume esquemáticamente estos bloques comunes y la bifurcación hacia los dos pipelines.

En el preprocesamiento por ventana se aplica eliminación de media y *detrend* lineal; para el análisis espectral se utiliza ventana de Hann y corrección por ganancia coherente [13]. Para HODMD se realiza, además, una normalización tipo *z-score* por canal dentro de la ventana, buscando homogeneizar escalas entre sensores. No se impone un filtrado temporal agresivo; la selección de información se concentra principalmente mediante recorte y selección de bandas en el dominio frecuencial.

Las características espectrales clásicas se obtienen calculando el espectro unilateral de amplitud por FFT en la banda $[10 \text{ Hz}, f_s/2]$. Para condensar el espectro, se seleccionan $N_{\text{FFT}} = 5$ picos por canal usando detección por prominencia relativa y separación mínima en frecuencia, con un ranking dominado por amplitud y un refuerzo armónico acotado [5]. Esto genera, por ventana, pares frecuencia–amplitud (freq/amp) para cada pico y canal.

Como complemento tiempo–frecuencia, se calcula un espectrograma tipo STFT (modo PSD) y se construye un espectro Box mediante la mediana temporal por frecuencia, aportando robustez frente a transitorios aislados. Sobre este espectro Box se seleccionan $M_{\text{BOX}} = 5$ picos por canal, almacenando frecuencia y amplitud (en dB), junto con metadatos de resolución. En el caso VAT, los descriptores FFT y Box/STFT aportan un bloque espectral de 80 variables por ventana (40 de FFT + 40 de Box).

Para la extracción modal, HODMD se aplica por ventana multicanal con una búsqueda en rejilla de parámetros: orden de *embedding* $d \in \{8, 12, 16, 20, 24, 28\}$ y rango SVD solici-

tado en un conjunto discreto, usando un rango efectivo adaptativo según las dimensiones de los datos [12, 6]. Los modos se filtran por estabilidad (autovalores cercanos al círculo unidad y límites en la parte real continua) y por banda de frecuencia $[10 \text{ Hz}, f_s/2]$. Se seleccionan exactamente $K = 5$ modos por ventana, aplicando supresión no máxima en frecuencia para evitar redundancias. Las formas modales complejas se alinean en fase, se reducen al espacio físico de sensores y se fijan en representación real, preservando consistencia de signo.

Las características HODMD se codifican como: frecuencias modales, amplitudes, normas de las formas modales y componentes por canal (forma modal espacial), manteniendo el orden por frecuencia ascendente. Estas familias modales se utilizan tanto para el diagnóstico directo del *Pipeline A* como como entrada para modelos supervisados en el *Pipeline B*.

En el *Pipeline A*, se construye una biblioteca de prototipos por clase: para cada clase y cada índice modal k , se agrega robustamente (mediana tras alineación de signo) la forma modal unitaria y se agrega la frecuencia prototipo. La similitud entre una ventana y una clase se calcula combinando: (i) MAC real entre formas modales [14], (ii) ponderación por amplitud y norma mediante exponentes, y (iii) penalización gaussiana en frecuencia desde datos de entrenamiento (MAD y escalado). Para comparar conjuntos de modos se resuelve una asignación óptima modo–modo mediante el algoritmo húngaro [15], obteniendo un puntaje por clase y decidiendo por $\arg \max$.

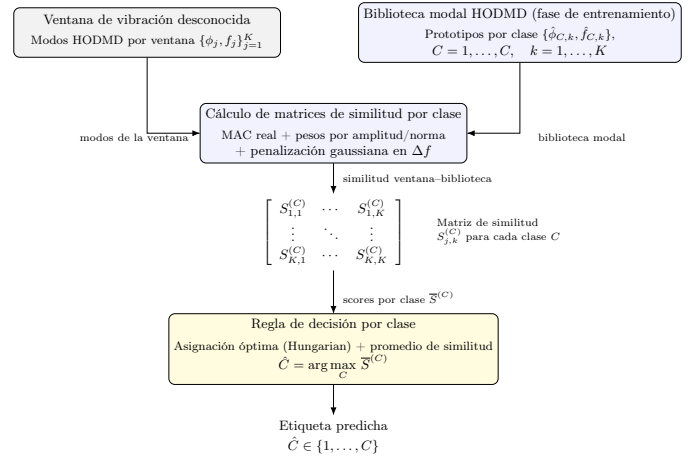


Fig. 7. Esquema detallado del Pipeline A.

La Figura 7 detalla el flujo de cálculo de $S_c(j, k)$, la asignación óptima y el puntaje final $\text{score}(c)$.

En el *Pipeline B*, se construye una matriz unificada de características $\mathbf{X}$ mediante un *join* entre el manifiesto de ventanas y las tablas de características por familia (FFT, Box/STFT, HODMD), asegurando alineación por clave (scenario, ventana_idx). Se evalúan distintos subconjuntos de características (solo FFT, solo Box/STFT, solo HODMD y combinaciones) y se entrenan clasificadores supervisados para el problema multiclase, siguiendo esquemas estandarizados de particionado estratificado y validación cruzada, con métricas como exactitud, F1 macro y matrices de confusión [16].

La Figura 8 muestra el esquema general del *Pipeline B* basado en extracción de características y modelos supervisados.

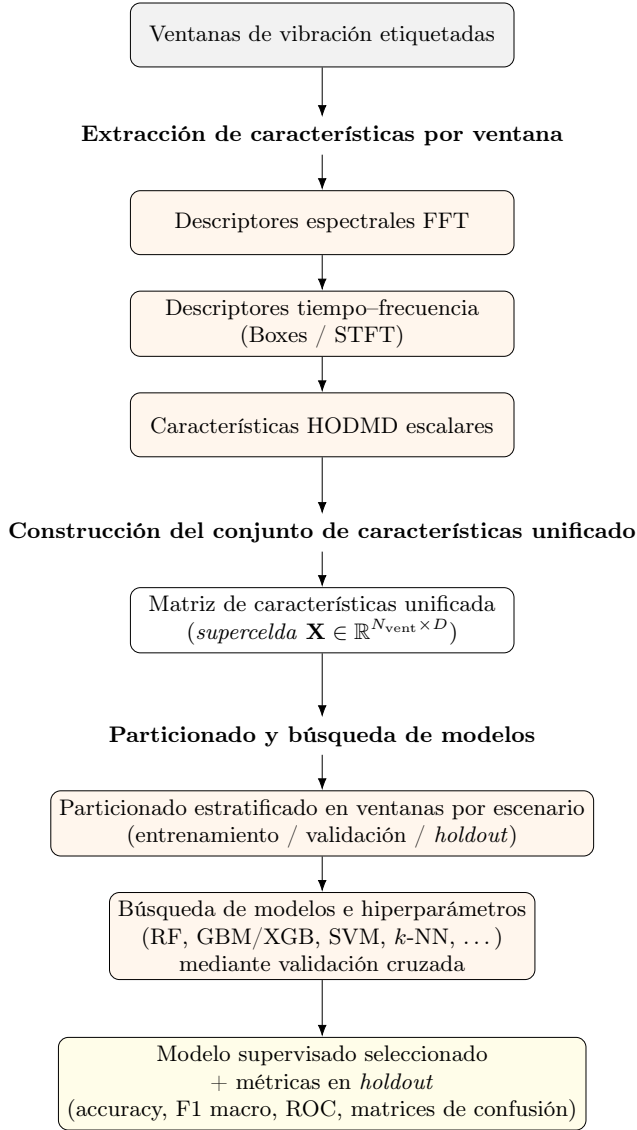


Fig. 8. Esquema general del Pipeline B

4.1. Aplicación al dataset VAT: experimentos y resultados

Se aplicó la metodología propuesta al dataset VAT utilizando exclusivamente señales de vibración del banco experimental, bajo el esquema de 45 escenarios finos definidos por el protocolo original (combinación de condición del rodamiento, carga y velocidad). Para cada escenario se emplearon los cuatro canales de aceleración disponibles (YS1, ZS1, YS2, ZS2), sin descartar ninguno, construyendo así ventanas multicanal de dimensión 4.

El procesamiento se basó en un ventaneo uniforme común a ambos pipelines: a partir de cada archivo .mat se reconstruyó la frecuencia de muestreo efectiva (f_s típicamente $\approx 25,6$ kHz) y se generaron $K = 20$ ventanas por escenario, de duración fija $W = 10$ s, con inicios distribuidos casi uniformemente en $[0, N - W]$. En total, el conjunto de datos quedó conformado por 900 ventanas etiquetadas (45×20), cada una asociada unívocamente a un escenario fino. El preprocesamiento se realizó por ventana y canal: para FFT y BOX/STFT se aplicó centrado (y detrend cuando correspondía) y ventana de Hann antes del cálculo espectral, restringiendo la selección de picos/bandas a $[10 \text{ Hz}, f_s/2]$. Para HODMD, las señales se reordenaron en el orden canónico $\{YS1, ZS1, YS2, ZS2\}$ y se normalizaron por

canal (media cero y desviación estándar unitaria) antes del embebido temporal y la descomposición. En el pipeline supervisado, valores faltantes residuales se manejaron mediante imputación y escalado dentro de los *pipelines* de modelos. Cabe señalar que la partición estratificada se realizó **a nivel de ventanas** (no por grupos/archivos), y se documenta así para consistencia con la implementación reportada.

Resultados del Pipeline A (biblioteca modal HODMD)

Se ejecutó un experimento completo siguiendo las etapas S1–S3 descritas previamente: extracción HODMD por ventana sobre una grilla de dimensiones de embebido $d \in \{8, 12, 16, 20, 24, 28\}$ y truncamientos SVD solicitados $r_{req} \in \{6, 8, 10, 12, 14, 16, 18, 20, 24, 28\}$, con rango efectivo acotado por el número de canales y muestras. Se filtraron modos inestables y se restringió la banda a $[10 \text{ Hz}, f_s/2]$. Por ventana se seleccionaron exactamente $K = 5$ modos dominantes (política *by_norm*) y se almacenaron sus frecuencias, amplitudes, normas y patrones espaciales.

La biblioteca se construyó agrupando ventanas de entrenamiento por clase y por orden modal, alineando el signo de los vectores modales y obteniendo prototipos espaciales robustos (mediana) y frecuencias prototipo (mediana). La clasificación se realizó proyectando cada ventana contra la biblioteca mediante similitud modal y penalización gaussiana en frecuencia. En términos de desempeño por ventana sobre 45 clases, el Pipeline A alcanzó $\text{Acc} \approx 0,79$ y $\text{F1}_{\text{macro}} \approx 0,78$, con baja dispersión entre pliegues. Los errores se concentraron principalmente en confusiones entre escenarios muy próximos (diferencias sutiles de severidad/carga dentro de la misma familia de falla), manteniéndose la mayor parte de la masa en la diagonal.

TABLA 1: Resumen del desempeño del Pipeline A (biblioteca modal HODMD) en el dataset VAT (45 clases).

Esquema	Acc (media)	F1 _{macro}	Precisión _{macro}
5 pliegues ext	$\approx 0,79$	$\approx 0,78$	$\approx 0,83$

Resultados del Pipeline B (diagnóstico supervisado)

Se evaluaron siete combinaciones de familias de características a partir de las mismas 900 ventanas (45 clases): (i) picos FFT por canal, (ii) cajas espectrales tipo BOX/STFT por canal y (iii) descriptores HODMD globales (frecuencias, normas y componentes espaciales ϕ_k por canal) de modos dominantes. En cada familia se seleccionaron $k' = 5$ picos/bandas/modos (fft5A, box5A, hodmd5N). El esquema de evaluación fue un *hold-out* estratificado 80/20 por ventanas (16 entrenamiento y 4 prueba por clase) y una validación cruzada interna estratificada ($k \approx 10$) para optimización de hiperparámetros con Optuna. Se consideraron cinco modelos (RF, GB, XGB, SVM y k-NN), integrados en *pipelines* con imputación y escalado cuando correspondía.

Los resultados mostraron desempeños muy altos en todos los casos. Incluso la configuración más pobre (sólo HODMD) alcanzó $\text{F1}_{\text{macro}} \approx 0,92$ y $\text{AUC} > 0,996$. La mejor combinación fue FFT+Box+HODMD con Random Forest, con $\text{F1}_{\text{macro}} \approx 0,994$ y $\text{AUC} \approx 0,9999$. El análisis cualitativo de errores indicó que las confusiones residuales se concentran principalmente en la estimación de severidad (niveles adyacentes) dentro del mismo torque, sin cambios de torque en los errores dominantes. A nivel interpretativo, FFT y Box tienden a dominar la

separación entre estados y niveles globales de vibración, mientras que HODMD aporta información complementaria útil para discriminar severidades cercanas.

TABLA 2: Resumen de resultados del Pipeline B (mejor modelo por combinación de características) en el dataset VAT.

Familias de características	Mejor modelo	F1 _{macro}	AUC (prueba)
FFT + Box + HODMD	RF	0.994	0.9999
FFT + Box	RF	0.989	1.0000
Box + HODMD	RF	0.989	0.9999
Sólo Box	RF	0.983	0.9999
FFT + HODMD	RF	0.972	0.9997
Sólo FFT	GB	0.949	0.9990
Sólo HODMD	XGB	0.919	0.9969

Discusión integrada (VAT)

En el dataset VAT, ambos pipelines confirman que la información espectral (FFT/Box) y modal (HODMD) captura firmas diagnósticas relevantes. El Pipeline B logra el máximo desempeño cuando se combinan familias y se emplea un clasificador tipo *ensemble* (en particular Random Forest), mientras que el Pipeline A ofrece un enfoque modal e interpretable con desempeño alto, pero con mayor sensibilidad a fronteras finas entre escenarios muy similares. En conjunto, los resultados respaldan el uso de un enfoque multimodal (FFT+Box+HODMD) para maximizar exactitud en clasificación fina, manteniendo la biblioteca HODMD como referencia complementaria para análisis e interpretación de errores.

5. CONCLUSIONES Y RECOMENDACIONES

En este trabajo se desarrolló y evaluó una metodología de diagnóstico de fallas en equipos rotativos a partir de señales de vibración, combinando descriptores espectrales (FFT y BOX/STFT), descriptores modales (HODMD) y modelos supervisados. La validación principal se realizó sobre el dataset VAT (45 escenarios finos, 900 ventanas multicanal) y se complementó con un banco de pruebas propio (3 condiciones, 30 ventanas).

- Evidencia de viabilidad en VAT. Con 45 clases, el enfoque supervisado (*Pipeline B*) alcanzó desempeño casi perfecto cuando se combinan familias FFT+Box+HODMD (mejor caso: $F1_{\text{macro}} \approx 0,994$, $AUC \approx 0,9999$). En contraste, el enfoque modal puro (*Pipeline A*) logró un desempeño alto pero inferior (típicamente $\text{Acc} \approx 0,79$, $F1_{\text{macro}} \approx 0,78$), con errores concentrados en clases muy cercanas (principalmente severidades/cargas similares).
- Complementariedad de información. FFT y BOX/STFT tienden a dominar la separación “estado normal vs. falla” y diferencias globales de condición, mientras que HODMD aporta información adicional útil para discriminar escenarios muy próximos (variantes finas), además de mayor interpretabilidad física a través de frecuencias y patrones espaciales de modos.
- Transferencia preliminar al banco de pruebas. En un conjunto reducido (30 ventanas), *Pipeline B* separó sin errores las tres clases ($F1_{\text{macro}} = 1,00$ en test), mientras que *Pipeline A* presentó confusiones puntuales entre normal y desalineación (aprox. 28/30 correctas, $\text{Acc} \approx 0,93$). Estos

resultados son indicativos, pero estadísticamente frágiles por el tamaño muestral.

- Recomendación de uso. Si se dispone de datos etiquetados representativos del dominio operativo, *Pipeline B* es la opción preferente por desempeño. *Pipeline A* resulta recomendable cuando la interpretabilidad y la trazabilidad física son prioritarias, o cuando se desea un esquema menos dependiente de un entrenamiento supervisado clásico.

6. TRABAJOS FUTUROS

- Ampliación de fallas y regímenes: incorporar más modos de falla (holguras, engranajes, ejes, defectos progresivos) y explorar variaciones amplias de velocidad/carga, incluyendo transitorios y cambios de carga en tiempo real.
- Diseño experimental y sensado: estudiar sistemáticamente número/ubicación de sensores y enriquecer el sensado con señales adicionales (corriente de motor, temperatura, acústica), evaluando el aporte incremental respecto a vibración.
- Modelos y representación: explorar selección/extracción adicional de características y modelos alternativos (incluyendo aprendizaje profundo), comparando desempeño vs. costo computacional y requerimientos de datos.
- Implementación en línea: desarrollar un prototipo *near real-time* que integre adquisición, ventaneo y diagnóstico, midiendo latencias, estabilidad y requerimientos de hardware.
- Transferencia entre dominios: investigar adaptación de dominio y estrategias de recalibración/reentrenamiento para transferir modelos entrenados en laboratorio a máquinas reales en planta, reduciendo campañas de medición por caso.

7. REFERENCIAS

- [1] W. Jung, S.-H. Kim, S.-H. Yun, J. Bae, and Y.-H. Park, “Vibration, acoustic, temperature, and motor current dataset of rotating machine under varying operating conditions for fault diagnosis,” *Data in Brief*, vol. 48, p. 109049, 2023.
- [2] A. K. S. Jardine, D. Lin, and D. Banjevic, “A review on machinery diagnostics and prognostics implementing condition-based maintenance,” *Mechanical Systems and Signal Processing*, vol. 20, no. 7, pp. 1483–1510, 2006.
- [3] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin, “Applications of machine learning to machine fault diagnosis: A review and roadmap,” *Mechanical Systems and Signal Processing*, vol. 138, p. 106587, 2020.
- [4] E. P. Marín, *Elementos de medición y análisis de vibraciones en máquinas rotatorias*. La Habana, Cuba: Editorial Félix Varela, 2012.
- [5] R. B. Randall and J. Antoni, “Rolling element bearing diagnostics—a tutorial,” *Mechanical Systems and Signal Processing*, vol. 25, no. 2, pp. 485–520, 2011.
- [6] J. Vega and S. Le Clainche, *Higher-Order Dynamic Mode Decomposition*. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 2017.
- [7] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. New York: Springer, 2009.
- [8] *ISO 55000:2014 Asset management – Overview, principles and terminology*, International Organization for Standardization Std., 2014.
- [9] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2016.
- [10] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer, 2006.

- [11] A. V. Oppenheim and R. W. Schaffer, *Discrete-Time Signal Processing*, 3rd ed. Upper Saddle River, NJ: Prentice Hall, 2010.
- [12] P. J. Schmid, "Dynamic mode decomposition of numerical and experimental data," *Journal of Fluid Mechanics*, vol. 656, pp. 5–28, 2010.
- [13] F. J. Harris, "On the use of windows for harmonic analysis with the discrete fourier transform," *Proceedings of the IEEE*, vol. 66, no. 1, pp. 51–83, 1978.
- [14] N. M. M. Maia and J. M. M. Silva, *Theoretical and Experimental Modal Analysis*. Taunton, UK: Research Studies Press, 1997.
- [15] H. A. Taha, *Operations Research: An Introduction*, 10th ed. Harlow, UK: Pearson, 2017.
- [16] D. M. W. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation," *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, 2011.