



UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

METODOLOGÍA PARA EL DIAGNÓSTICO Y PRONÓSTICO DE FALLAS EN EL  
MOTOR DE TRACCIÓN DE EQUIPOS DE TRANSPORTE DE MATERIAL MINERO

TESIS PARA OPTAR AL GRADO DE  
MAGÍSTER EN CIENCIAS DE LA INGENIERÍA, MENCIÓN ELÉCTRICA

MEMORIA PARA OPTAR AL TÍTULO DE  
INGENIERO CIVIL ELÉCTRICO

SEBASTIÁN EMMANUEL SERIA RÍOS

PROFESOR GUÍA:  
MARCOS EDUARDO ORCHARD CONCHA

MIEMBROS DE LA COMISIÓN:  
JUAN CARLOS AGÜERO  
JAVIER RUIZ DEL SOLAR SAN MARTÍN

SANTIAGO DE CHILE  
2018

RESUMEN DE LA TESIS PARA OPTAR  
AL TÍTULO DE MAGÍSTER EN CIENCIAS DE LA INGENIERÍA, MENCIÓN ELÉCTRICA  
E INGENIERO CIVIL ELÉCTRICO  
POR: SEBASTIÁN EMMANUEL SERIA RÍOS  
FECHA: 2018  
PROF. GUÍA: MARCOS EDUARDO ORCHARD CONCHA

## METODOLOGÍA PARA EL DIAGNÓSTICO Y PRONÓSTICO DE FALLAS EN EL MOTOR DE TRACCIÓN DE EQUIPOS DE TRANSPORTE DE MATERIAL MINERO

Este trabajo de tesis fue realizado en el marco del proyecto de la Dirección de Tecnología de Komatsu, consistente en el diagnóstico y pronóstico de fallas para el motor de tracción de los equipos de transporte de material minero, Komatsu 960E. El modo de falla que se estudia es la fisura, ruptura y posterior falla catastrófica por desgaste cíclico del perno pasador del eje de giro del piñón, componente que forma parte del sistema planetario del motor. De esta forma, y tomando en cuenta el grado de severidad operacional que se desarrolla en faena, se estima el valor de vida útil remanente para el componente en cuestión, lo que permite establecer mantenimiento predictivo basado en la condición (CBM).

Actualmente, con la convergencia y conectividad de sistemas, como por ejemplo, el internet de las cosas (IoT), la obtención de datos sobre procesos de diversa índole, posee una mayor cobertura producto de la disponibilidad y utilización de múltiples sensores. Además, la utilización de plataformas de computo distribuidas con acceso mediante internet, denominadas comúnmente nube, permite procesar con mayor facilidad esta gran cantidad de datos que se genera constantemente, obteniendo así, un flujo de datos desde los equipos a la nube, lo que permite obtener información sobre los procesos. Este escenario es propicio para generar políticas de mantenimiento basados en la condición y disminuir costos debido a fallas imprevistas o mantenimientos realizados sin ser necesarios.

La metodología consistió en obtener datos operacionales de equipos en faena, y junto a antecedentes de análisis de causa raíz del modo de falla, utilizar las variables monitoreadas que se relacionan con el desgaste cíclico del componente. Esto permitió definir métricas de daño acumulado. Posteriormente, analizando el comportamiento típico se construyó una cadena de Markov, la cual se utilizó para estimar el valor futuro de daño asociado al componente.

De esta forma, se concluyó un método para estimar la vida remanente del componente basados en los historiales de falla, junto a estimaciones de desgaste operacional que indican el nivel de daño acumulado estimado. Se obtuvo como resultado el acierto en el pronóstico de falla para un equipo, con un rango de tiempo de falla de una semana y aproximadamente un mes de anticipación mediante simulaciones, luego de analizar y comparar 4 métricas de daño sobre el componente en cuestión. Para el resto de los equipos, se planteó el pronóstico de falla para la última fecha de obtención de datos.

Dado que cada vez es más frecuente obtener la realimentación de todos los equipos operando en una planta, esta metodología supone un gran punto de partida para estimar y obtener indicios del estado de componentes, muchas veces modelados como caja negra, utilizando herramientas de aprendizaje de máquinas sobre datos masivos, constituidos por las mediciones de los equipos en faena.



*A mi familia y amigos.*

# Agradecimientos

Quisiera agradecer especialmente a dos personas; a mi madre y a mi padre, por el incondicional amor, trabajo y dedicación durante todos estos años.

A la Karen y a la Ninia por ser las mejores hermanas.

A la hermandad y a los cartoneros por darme muchos momentos de felicidad en la U.

Al profe Marcos por su gran voluntad e invaluable guía en mi paso por el DIE.

A la dirección de tecnología de Komatsu por la oportunidad y apoyo.

A los profesores Javier Ruiz del Solar y Juan Carlos Agüero por ayudarme a corregir la tesis.

Y finalmente a todos los músicos que fueron parte del soundtrack día a día, mes a mes y año a año.

# Tabla de Contenido

<b>Índice de Tablas</b>	<b>vii</b>
<b>Índice de Ilustraciones</b>	<b>viii</b>
<b>1. Introducción</b>	<b>1</b>
1.1. Objetivos . . . . .	2
1.2. Hipótesis . . . . .	2
1.3. Metodología . . . . .	3
1.4. Organización . . . . .	4
<b>2. Contexto y Antecedentes</b>	<b>5</b>
2.1. Diagnóstico y Pronóstico de Fallas . . . . .	6
2.2. Mecánica de la fractura por fatiga cíclica . . . . .	7
2.3. Cadena de Markov . . . . .	12
2.4. Rainflow Counting . . . . .	13
2.5. Clustering . . . . .	14
2.5.1. DBSCAN . . . . .	16
2.5.2. K-Means . . . . .	17
<b>3. Desarrollo</b>	<b>19</b>
3.1. Esquema General . . . . .	19
3.1.1. Módulo de modelamiento . . . . .	20
3.1.2. Módulo de pronóstico . . . . .	21
3.2. Estructura de datos adquiridos y utilizados . . . . .	21
3.2.1. Estados de operación . . . . .	22
3.2.2. Parámetros físicos . . . . .	23
3.3. Extracción de características . . . . .	23
3.3.1. DBSCAN para identificación de ciclos de operación . . . . .	27
3.3.2. Rango de variación mínimo - máximo . . . . .	28
3.3.3. Longitud de Arco . . . . .	29
3.3.4. Rainflow Counting . . . . .	30
3.4. Umbral de Falla . . . . .	32
3.4.1. Cadena de Markov . . . . .	33
3.4.2. Simulaciones . . . . .	38
3.5. Diagnóstico de Falla . . . . .	39
3.5.1. Definición de categorías para los datos medidos (clusters) . . . . .	40
3.5.2. Definición de modelos . . . . .	43

3.5.3. Resultados de los modelos ajustados . . . . .	44
3.6. Pronóstico de Falla . . . . .	46
<b>4. Resultados</b>	<b>51</b>
4.0.1. Mean rainflow counting . . . . .	53
<b>5. Conclusión</b>	<b>61</b>
<b>6. Bibliografía</b>	<b>64</b>

# Índice de Tablas

3.1.	Equipos y rangos de tiempo en el que fueron monitoreados. . . . .	28
3.2.	Promedio de todos los equipos monitoreados por cada métrica. . . . .	32
3.3.	Mediana de todos los equipos monitoreados por cada métrica. . . . .	32
3.4.	Equipos y la cantidad de ciclos identificados y considerados dejando de lado datos anómalos. . . . .	35
3.5.	Promedio de $DOPC_{dmm}$ en cada uno de los estados determinados mediante K-means. Valores en $[Nm]$ . . . . .	36
3.6.	Promedio de $DOPC_{rfc}$ en cada uno de los estados determinados mediante K-means. Valores en $[Nm]$ . . . . .	36
3.7.	Promedio de $DOPC_{la}$ en cada uno de los estados determinados mediante K-means. Valores en $[Nm \cdot Hrs]$ . . . . .	37
3.8.	Error cuadrático medio normalizado entre el DAT y la aproximación de modelos lineales por categoría para métrica distancia mínimo-máximo. . . . .	48
3.9.	Error cuadrático medio normalizado entre el DAT y la aproximación de modelos lineales por categoría para métrica rainflow counting. . . . .	48
3.10.	Error cuadrático medio normalizado entre el DAT y la aproximación de modelos lineales por categoría para métrica longitud de arco. . . . .	49
4.1.	Fecha del último recambio de motor para la ventana de tiempo observada con los datos de eventos recolectados. . . . .	51
4.2.	Valores estimados de umbral de falla. . . . .	51
4.3.	Registro de falla posterior a la fecha de fin de obtención de datos. Este caso se utiliza como prueba de las proyecciones y comparativa entre las diferentes métricas planteadas. . . . .	52
4.4.	Fechas de falla predichas. . . . .	52
4.5.	Valor promedio de mean rainflow counting por motor de todos los equipos. . . . .	54
4.6.	Mediana de mean rainflow counting por motor de todos los equipos. . . . .	54
4.7.	Promedio de $DOPC_{mrfc}$ en cada uno de los estados determinados mediante K-means. Valores en $[Nm]^2$ . . . . .	56
4.8.	Error cuadrático medio normalizado entre el DAT y la aproximación de modelos lineales por categoría para métrica mean rainflow counting. . . . .	56
4.9.	Valores de umbral de falla estimados usando mean rainflow counting. . . . .	57
4.10.	fecha de falla predicha mediante mean rainflow counting. . . . .	58



# Índice de Ilustraciones

2.1. Esquema de CBM para sistemas genéricos [3]. . . . .	7
2.2. Ilustración del desarrollo de grietas en materiales [15]. . . . .	9
2.3. Relación obtenida para la tasa de crecimiento de la grieta por ciclo dado una carga externa $\Delta K$ [21]. . . . .	10
2.4. Efecto de “cierre de grietas” [15]. . . . .	11
2.5. Curva S-N [21]. . . . .	12
2.6. Obtención de puntos peaks para aplicar el algoritmo de rainflow counting. . . . .	14
2.7. Algoritmo Rainflow Counting [1]. . . . .	15
2.8. Ejemplo de extracción de ciclo para el contador [1]. . . . .	15
3.1. Esquema general de la metodología propuesta para el diagnóstico y pronóstico de fallas en el motor de tracción. . . . .	20
3.2. Secuencia temporal de la variable medida, torque de motor de tracción $\tau$ . . . . .	24
3.3. Comparación temporal de variables que indican el estado de operación del equipo para descarga de material. . . . .	25
3.4. Acercamiento al instante de tiempo en que se cambia a marcha atrás. . . . .	25
3.5. Comportamiento sistemático a través de ciclos de operación. . . . .	26
3.6. Extracción de característica; distancia mínimo-máximo. . . . .	29
3.7. Extracción de característica; longitud de arco. . . . .	30
3.8. Histograma de DOPC para la métrica distancia mínimo-máximo. . . . .	31
3.9. Histograma de DOPC para la métrica rainflow counting. . . . .	31
3.10. Histograma de DOPC para la métrica longitud de arco. . . . .	32
3.11. Resumen longitud de arco para la aplicación de un equipo. . . . .	33
3.12. Resumen rainflow counting para la aplicación de un equipo. . . . .	34
3.13. Resumen distancia mínimo-máximo para la aplicación de un equipo. . . . .	34
3.14. Estimación de umbral de falla para métrica longitud de arco. . . . .	38
3.15. Estimación de umbral de falla para métrica rainflow counting. . . . .	38
3.16. Estimación de umbral de falla para métrica distancia mínimo-máximo. . . . .	39
3.17. Relación de longitud de arco y carga-tiempo empleado por ciclo. . . . .	40
3.18. Relación de rainflow counting y carga-tiempo empleado por ciclo. . . . .	41
3.19. Relación de distancia mínimo-máximo y carga-tiempo empleado por ciclo. . . . .	41
3.20. Distribución espacial de zonas de carga y descarga. . . . .	42
3.21. Agrupación y clasificación por conjunto de zonas de carga y descarga. . . . .	42
3.22. Tiempo empleado por ciclo de todos los equipos monitoreados. . . . .	43
3.23. Aproximación de modelos lineales por categorías del DOPC. . . . .	45
3.24. Aproximación de modelos lineales por categorías del DOPC. . . . .	45

3.25. Aproximación de modelos lineales por categorías del DOPC. . . . .	46
3.26. Aproximación del DAT de métrica longitud de arco por los modelos lineales por categoría. . . . .	46
3.27. Aproximación del DAT de métrica longitud de arco por los modelos lineales por categoría. . . . .	47
3.28. Aproximación del DAT de métrica longitud de arco por los modelos lineales por categoría. . . . .	47
3.29. Simulación de comportamiento para el crecimiento de DAT por distancia mínimo-máximo. . . . .	49
3.30. Simulación de comportamiento para el crecimiento de DAT por longitud de arco. . . . .	50
3.31. Simulación de comportamiento para el crecimiento de DAT por rainflow counting. . . . .	50
4.1. Pronóstico de falla utilizando métrica longitud de arco. . . . .	52
4.2. Pronóstico de falla utilizando métrica distancia mínimo-máximo. . . . .	53
4.3. Pronóstico de falla utilizando métrica rainflow counting. . . . .	53
4.4. Histograma de $DOPC_{mrfc}$ por motor de tracción. . . . .	54
4.5. Resumen mean rainflow counting para la aplicación de equipo 76. . . . .	55
4.6. Simulación de comportamiento para el crecimiento de DAT por mean rainflow counting. . . . .	55
4.7. Estimación de umbral de falla para métrica mean rainflow counting. . . . .	57
4.8. Pronóstico de falla utilizando mean rainflow counting. . . . .	57
4.9. MT1 de equipo 57. . . . .	58
4.10. MT2 de equipo 65. . . . .	59
4.11. MT1 de equipo 71. . . . .	59
4.12. MT2 de equipo 72. . . . .	60
4.13. MT2 de equipo 77. . . . .	60

# Capítulo 1

## Introducción

Este trabajo de tesis consiste en el desarrollo de una metodología para el diagnóstico y pronóstico de fallas en el motor de tracción de camiones mineros de extracción, utilizados para el transporte de material en la aplicación de una mina en específico. La metodología está construida a partir del planteamiento y formulación de hipótesis para la extracción de características, que indican el deterioro de estado de salud del componente tras análisis de antecedentes, adquisición de variables reales monitoreadas durante faena, el diseño y ejecución de simulaciones.

El modo de falla estudiado corresponde a la rotura de pernos de sujeción de los piñones del motor. Esto genera un efecto retroalimentado en el desgaste de las dentaduras del sistema de manera súbita, lo que es difícil de detectar a tiempo dado el escenario de monitoreo existente. El resultado es la pérdida de tiempos efectivos de operación en desmedro de la productividad y costos por mantenimiento no planificado, entre otros.

Por todo esto, el motivo de generar una metodología de este tipo surge de la aspiración por implementar políticas de mantenimiento predictivo y así disminuir pérdidas por reparaciones imprevistas. De esta forma, el cálculo de la vida útil remanente (VUR) para el componente es el resultado de este trabajo.

Actualmente es extensivo el empleo de técnicas de análisis de grandes cantidades de datos debido al avance en eficiencia y capacidades computacionales a la par de la conectividad. Como resultado es posible observar un auge en el desarrollo de procedimientos interdisciplinarios enfocados en la obtención de información a partir de mediciones o datos, como lo es por ejemplo, la creación de modelos de sistemas con un enfoque de caja negra o el aprendizaje de máquinas a partir de observaciones. El campo del mantenimiento predictivo junto al diagnóstico y pronóstico de fallas no están excluidos de la utilización de estas tecnologías y utilizan enfoques híbridos entre modelos físicos, fenomenológicos y los de aprendizaje a partir de mediciones.

En este sentido, el camión de extracción minero no es la excepción, ya que cuenta con múltiples sensores que permiten monitorear la operación durante el proceso de traslado de material minero, generando a su vez, grandes cantidades de datos. Sin embargo, muchas variables

no cuentan con sensores que midan de manera directa el fenómeno físico. Esto implica el desafío de relacionar y crear modelos que permitan estimar dichas variables. Un ejemplo de esto, corresponde a la tensión mecánica a la que se somete la estructura principal del equipo, que no dispone de sensores dedicados a medirla directamente, sin embargo, se realizan aproximaciones mediante las diferencias en las presiones de suspensiones.

El caso abordado presenta gran variabilidad dependiendo de las condiciones de entorno durante faena y es clara la tendencia de falla prematura en los motores. Esto hecho se atribuye principalmente a condiciones agresivas de operación, las que se reflejan comparando el desempeño del equipo en la mina en cuestión contra el de aplicaciones en otras partes del mundo. Es por esto que la tesis estudia estas características con el fin de, además de estimar la VUR del sistema, identificar prácticas operacionales y/o de entorno que impactan significativamente en la reducción de vida útil del componente. De esta forma, considerando el carácter súbito de la falla dadas las estrategias de monitoreo implementadas actualmente -medición de concentración de material particulado en el aceite con un tiempo de muestreo superior al que desarrolla la evolución de la falla- es de interés el análisis de la evolución del estado de salud del componente y relacionarla con variables que se obtienen de manera directa mediante la conexión de sistemas, sin la necesidad de intervención del equipo, aumentando la eficiencia del diagnóstico del equipo.

## 1.1. Objetivos

El objetivo de este trabajo es estimar la vida útil remanente del componente, planteando un marco de trabajo basado en los procedimientos de la literatura del diagnóstico y pronóstico de fallas, abarcando desde la adquisición y filtrado de datos, selección de características para estimar el estado de salud del componente para, posteriormente, modelar el comportamiento operacional del equipo y pronosticar la fractura del perno.

La expansión de esta metodología para otros sistemas dentro del mismo camión de extracción minero es uno de los objetivos que motivan este estudio, dada la generalidad del concepto de daño acumulado, no obstante, siempre considerando los modos de falla particulares de cada caso.

## 1.2. Hipótesis

La hipótesis que se plantea es que la falla del componente ocurre cuando se alcanza un valor empírico en el “daño acumulado”, que corresponde a una métrica calculada en función de las variables monitoreadas y, al mismo tiempo, se considera una variable aleatoria. El “daño acumulado”, es reflejo de la severidad operacional y además, es aproximadamente el mismo valor para las diferentes muestras de componentes. Por consiguiente, la fractura del perno se caracteriza por una distribución de probabilidad normal sobre el daño acumulado. El valor del daño acumulado en que se maximiza la probabilidad de fractura, se denomina umbral de falla (UDF) y se construye a partir de métricas calculadas del torque de motor, aplicado

durante operación en faena, ponderando en mayor medida su variación en el tiempo debido a los antecedentes de fatiga cíclica. El umbral de falla por ende, se estima utilizando historiales de fallas pasados simulando el comportamiento de daño acumulado asociado a dichos casos y se utiliza para estimar la vida remanente útil, en este estudio, de seis equipos de los cuales uno se conoce la fecha de falla real. Finalmente la validación de esta hipótesis va a estar dada por la diferencia entre las fechas de predicción de falla con el caso real y el análisis de los otros cinco casos de los que no se conoce fecha de falla.

Los principales supuestos consideran: la correlación de deformación y tensión mecánica sobre el componente con la variable monitoreada torque, y, contabilizar el daño acumulado sólo en condiciones más severas de aplicación, atribuidos al estado de operación con carga. La motivación de este último supuesto obedece a no sobrecargar el sistema de monitoreo y adquisición de datos. Luego las métricas de daño acumulado, obtenidas con dichas consideraciones, se validan en base al desempeño de predicción de falla.

### 1.3. Metodología

La metodología empleada para la obtención de resultados inicia con el análisis de antecedentes y estudios previos que se realizaron para explicar y encontrar las causas que provocan la falla. A partir de esto se formulan supuestos que sustentan y justifican la obtención de datos de los equipos al tiempo que operan. En paralelo, se definen los criterios para la posterior extracción de características inspirados en los mismos supuestos, concibiéndose una métrica para aproximar el daño acumulado por el componente. Dicho valor, junto a su proyección futura que considera un comportamiento sistemático para las operaciones, se utiliza directamente para estimar la VUR.

De esta forma, se puede dividir el trabajo en las siguientes fases e hitos de importancia:

- **Estudio de antecedentes:** investigación del modo de falla que afecta al componente y trabajos de investigación ya realizados en Komatsu.
- **Revisión del estado del arte:** búsqueda de literatura relacionada con el modo de falla, procedimientos para el diagnóstico y pronóstico de fallas como marco de trabajo y técnicas de procesamiento de datos.
- **Adquisición y filtrado de datos:** obtención de datos, estructuración y preprocesamiento para quitar valores anómalos
- **Extracción de características y definición de daño acumulado:** procesamiento de datos para obtener la métrica que se utilizará para caracterizar el estado de salud del componente.
- **Clustering:** identificación de condiciones de operación y entrenamiento de modelos que estiman daño acumulado a partir de dichas condiciones
- **Gestión de datos faltantes con modelos por condición de operación:** estimación del valor de daño acumulado pasado desde el último recambio de componente por cada equipo.
- **Simulaciones y proyección de comportamiento futuro:** modelamiento mediante

cadena de Markov del daño acumulado para cada equipo.

- **Obtención de parámetros para estimación de la VUR usando historial de fallas:** utilizando las realizaciones de los equipos que se tiene documentación de las horas de operación antes de la fractura, se estima el umbral de falla (UDF) de la hipótesis, calculado como el valor medio de los valores de daño acumulado para cada caso cuando ocurrió la falla.
- **Validación y eventual corrección de parámetros a partir de nuevos casos de falla identificados:** verificación con el caso real de falla. La corrección considera re-calcular el UDF tomando el valor medio al nuevo conjunto aumentado, el cual es resultado de agregar los nuevos casos de falla por fractura. Este paso queda planteado como trabajo futuro.

Las limitaciones presentes del sistema significan una variable importante para la precisión del modelo. Esto significa la pérdida de datos monitoreados cuando el equipo se encuentra en operación y la sobrecarga del sistema cuando hay una gran cantidad de datos cargándose en la aplicación de gestión de descarga de datos. Por esta razón, en la metodología, los puntos de modelamiento de daño acumulado en función de condiciones de entorno se hace indispensable para completar datos faltantes.

Por otro lado, el carácter iterativo del último punto de la metodología fue pensado para diseñar un sistema que se vuelve experto a medida que nuevas fallas van sucediendo. Finalmente este método fue codificado de manera modular de forma tal que podría agregarse nuevas variables a las definiciones de métricas, pensando que el sistema que almacena datos se robustece y deja de ser un factor restrictivo debido a la gran cantidad de datos que gestiona.

## 1.4. Organización

A continuación, este trabajo se divide en 4 capítulos, organizados de la siguiente manera: en el capítulo 2, contexto y antecedentes, se explica la situación en particular y marco teórico, que está pensado para sustentar y explicar el trasfondo de los supuestos más importantes. Luego, el capítulo 3, desarrollo, explica y presenta de forma detallada la metodología empleada. Los resultados se muestran en el capítulo 4 y finalmente se discute los aspectos más relevantes en el capítulo 5 de conclusiones.

# Capítulo 2

## Contexto y Antecedentes

En este capítulo se expone la composición del sistema en estudio junto a los antecedentes generales que describe el modo de falla del componente crítico. Por otra parte, se aborda los principales conceptos teóricos que sustentan las hipótesis planteadas y que se utilizan durante el desarrollo de la tesis.

El detalle del diseño del motor de tracción se omite debido a que es información protegida por Komatsu, por lo que se presentan los aspectos relevantes para comprender el modo de falla del equipo sin graficar su estructura. El equipo cuenta con dos motores de tracción que corresponden a motores-rueda eléctricos. Estos transmiten la energía convirtiéndola en movimiento a través de un sistema de engranajes planetario constituido por un eje solar y piñones planetarios, denominados “carriers”, los que están sujetos en su eje de giro por dos pernos pasantes. Los carriers se dividen en dos grupos que se clasifican según el régimen de velocidad de giro, considerando su posición en el sistema de transmisión; están los carriers de alta velocidad y los de baja velocidad. El modo de falla estudiado es el agrietamiento debido a fatiga cíclica del perno pasante del carrier de baja velocidad y posterior fractura, hecho que desencadena el desgaste en las dentaduras de engranajes debido a la circulación de restos de perno al interior del motor-rueda, lo que finalmente provoca una falla catastrófica en el motor de tracción y por consiguiente, no es posible la movilidad de manera autónoma.

El estudio de causas del modo de falla, en primer lugar relaciona la deformación de la estructura completa que contiene los engranajes planetarios con el esfuerzo que genera el motor en términos del torque aplicado. Esta tesis desarrolla una metodología para predecir la falla catastrófica del componente dado el perfil de operación del equipo y su incidencia en el proceso de fatiga cíclica. El modelado de los efectos estructurales corresponde al enfoque de caja negra y se comienza el estudio con la revisión de las metodologías utilizadas en el diagnóstico y pronóstico de fallas de sistemas.

## 2.1. Diagnóstico y Pronóstico de Fallas

En esta sección se presenta el desarrollo de los tópicos y técnicas empleadas para el diagnóstico y pronóstico de fallas (PHM). PHM se enmarca en la aplicación de técnicas para el mantenimiento basado en la condición (CBM). Estas técnicas utilizan datos e información del sistema extraídos mediante sensores, donde la finalidad es determinar la condición del sistema ante la aparición de eventos identificados como comportamientos anómalos según el desempeño esperado. Dependiendo del sistema a considerar, una falla es definida como la desviación no permitida de al menos una propiedad o variable característica de dicho sistema. También se entiende como la interrupción permanente en el funcionamiento del sistema bajo las condiciones y requerimientos de operación [11] [3]. La evolución de estos eventos es utilizada para programar reparaciones, anticipándose a la falla. Idealmente, esta predicción debe ser capaz de realizarse en una ventana de tiempo favorable en relación a temas logísticos y económicos. Así, se asegura disponibilidad del sistema y la correspondiente reducción de costos por mantenimiento. De esta forma, CBM se conforma como una metodología que engloba el proceso desde la adquisición de datos del sistema, su procesamiento y la toma de decisiones, donde cada categoría tiene asociado estudios y herramientas que son aplicables dependiendo del fenómeno estudiado. Por lo tanto, es común que CBM involucre procesos tales como: estudio de la física del desgaste mecánico de componentes, estrategias de monitoreo, algoritmos para simulación, detección y predicción de comportamientos, diseño de experimentos con sus respectivos protocolos de pruebas y cálculo de métricas para validación. De ahí que PHM es parte de CBM, correspondiendo a la metodología que predice el comportamiento futuro del sistema. Esto permite estimar la vida útil remanente (VUR), entendida como el tiempo futuro en que el sistema opera cumpliendo los requerimientos para los que fue diseñado antes de que se produzca la falla y la funcionalidad se vea afectada. Este paradigma de mantenimiento se diferencia a los enfoques correctivos y preventivos, ya que, en el primer caso existe una reparación posterior a la falla, presentándose contratiempos en la disponibilidad del sistema asociados a la falla imprevista. En el segundo caso, se realiza un mantenimiento fijo independiente del estado del sistema, lo que conlleva ineficiencias ante la eventual revisión de sistemas y reemplazo de componentes aún cuando pueden seguir operando con un margen de tiempo significativo.

El propósito de desarrollar e implementar distintos niveles de diagnóstico y pronóstico de fallas en sistemas complejos que involucran interacción de sub-sistemas y maquinarias relativamente costosas está dado principalmente por maximizar el tiempo de operación de los sistemas y minimizar el costo por mantención y operación de los equipos [23]. Por esta razón, es crítico contar con la habilidad de detectar fallas en un tiempo prudente para gestionar de mejor manera la información y posterior toma de decisiones.

En la Figura 2.1 se esquematiza el proceso de CBM/PHM en general. En primer lugar se tiene el sistema genérico, el cual podría ser un avión, equipo de transporte de material minero, batería de litio, etc. De este se obtiene datos de variables físicas relevantes del sistema mediante el uso de sensores (quedando definidas ciertas características como frecuencia de muestreo, precisión del sensor, etc). Dichos datos sin procesar muestran cómo se está comportando el sistema, sin embargo esta información a priori no es indicativa de la vida útil remanente del sistema. Para este propósito, los datos obtenidos se preprocesan y se les extrae características



que son útiles para identificar tendencias que den cuenta de la evolución del sistema ante la eventual falla. En el preprocesamiento, normalmente se utilizan herramientas para quitar datos anómalos, filtrar, cambiar la estructura de los datos a una forma más conveniente, etc. Esto propicia el inicio de la etapa de extracción de características. La gran importancia de este último paso radica en que la característica seleccionada debe ser naturalmente proporcional a un indicador de cómo se está comportando el sistema en relación a su estado de salud. Si la tendencia del sistema es a decaer en su estado de salud, es deseable construir una característica que siga aquella tendencia. Por esta razón, es de vital importancia realizar un estudio de cómo falla y cuáles son las causas que desencadenan el mal funcionamiento del sistema. El principio utilizado para este propósito es FMECA (Failure Modes and Effects Criticality Analysis), el cual es un proceso de estudio que caracteriza los modos de falla y la combinación de estos que afectan al sistema. Para cada uno de estos, se identifican síntomas sobre variables medidas, mediante el uso de sensores adecuados, o métricas calculadas, para identificar y aislar patrones de comportamiento relacionados a las fallas. Este proceso corresponde a “clasificación de fallas” de la Figura 2.1, y a continuación, se tiene la etapa de pronóstico. Dicha característica que se obtiene bajo el análisis de FMECA, se extrapola en el futuro para obtener una estimación de la VUR. Los algoritmos empleados para realizar la predicción varían dependiendo del sistema. Para el caso de este estudio se utiliza un modelamiento de comportamiento futuro con cadena de Markov. Históricamente, hasta este punto, una vez que se obtenía la característica, se realizaba el diagnóstico del sistema y se tomaban decisiones en cuanto a dicho estado. El paso que marca la diferencia al mantenimiento predictivo o CBM, es la utilización de información obtenida del sistema (diagnóstico, extracción de característica y otros) para predecir la evolución futura de la característica y estimación de la VUR.

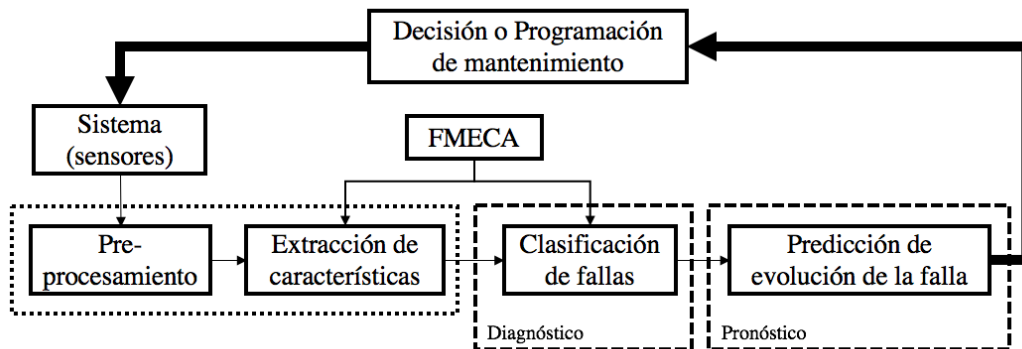


Figura 2.1: Esquema de CBM para sistemas genéricos [3].

La efectividad de los enfoques analíticos, conocimiento experto y basado en datos depende de la calidad y tipos de modelos disponibles, la calidad y cantidad de datos disponibles [3].

## 2.2. Mecánica de la fractura por fatiga cíclica

En esta sección se describe los principales mecanismos que rigen el fenómeno de fatiga de materiales y que motiva la elección de esta metodología, en consideración del modo de falla

identificado en los antecedentes.

El concepto de fatiga de materiales es considerado como el proceso físico de debilitamiento del componente al aplicar esfuerzos cíclicos. Estos esfuerzos pueden ser de magnitud inferior al límite máximo o tensión de fluencia en el cual el material se deforma de manera irreversible pasando de la zona elástica a la zona plástica. Esa es una de las razones para comprender cómo se generan grietas y posterior fractura total en materiales. Las nociones de este fenómeno corresponden a la acumulación de esfuerzos al interior del material debido a fatiga prolongada en el tiempo, lo que resulta en un estado susceptible a desarrollar micro grietas que posteriormente se expanden hasta llegar a la fractura del material. Por ende, la fatiga de materiales se refiere a un fenómeno por el cual la rotura de los materiales bajo cargas dinámicas cíclicas se produce más fácilmente que con cargas estáticas.

Históricamente el estudio de fatiga de materiales se vio impulsado como resultado de la revolución industrial del siglo XIX, tras la masiva utilización de maquinarias y, desde entonces se consideraba y postulaba que la fractura se producía luego de aplicar carga o esfuerzo al sistema con atributos cíclicos, lo cual es resultado de investigación y supuestos de August Wöhler, quien fue ingeniero ferroviario. Un sólo esfuerzo de la misma magnitud no produciría fractura ni daño [21], [20]. Este hecho se corroboró posteriormente en el transcurso del siglo XX.

El tiempo efectivo en que el material opera antes de la fractura por fatiga cíclica está dividido en dos fases: la iniciación de la grieta -también conocida como nucleación de la grieta-, y el posterior crecimiento de la grieta [21]. La Figura 2.2 ilustra el entorno y evolución en que se produce la grieta en el material.

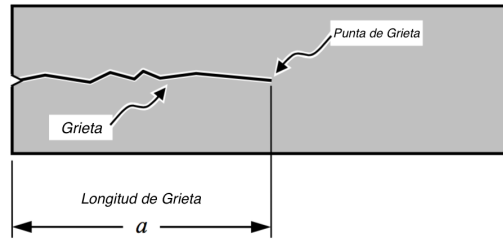
Fue a principios de 1960 cuando se publicaron estudios sobre la relación entre el crecimiento de la grieta  $da/dN$  y el rango de estrés al que se somete el material  $\Delta K$ . Paris, Gomez y Anderson tomaron los resultados propuestos por Irwin en 1957 de su publicación “Analysis of stresses and strains near the end of a crack traversing a plate” y continuaron con el resultado del valor  $K$ , conocido como el factor geométrico de intensidad de tensión [21], [15], [12], en consideración al análisis del campo de tensiones en el entorno de la punta de grieta. Esto último ilustrado en la Figura 2.2b. El resultado más importante de este análisis, es que para materiales con comportamiento lineal y propiedades básicas constantes en el tiempo, la teoría de fractura mecánica elástica lineal (LEFM) caracteriza el crecimiento de la grieta en torno a la punta de grieta en función del factor de tensión  $K$ , como se indica en la Ecuación 2.1.

$$K = S \cdot \sqrt{\pi \cdot a} \cdot f(\text{geometría del cuerpo, forma de la grieta}) \quad (2.1)$$

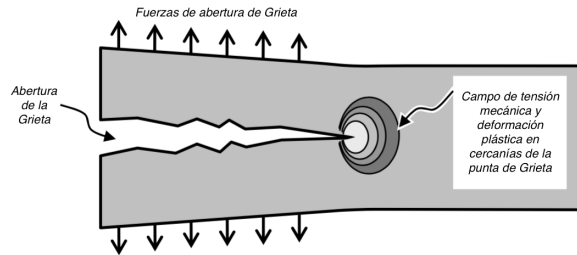
Donde  $S$  es la carga externa que se aplica sobre el componente,  $f$  es un factor que depende de la geometría y forma de la grieta, mientras que  $a$  es el largo de la grieta. Luego las componentes direccionales de tensión sobre la punta de grieta se obtienen mediante la Ecuación 2.2.

$$\sigma = K \cdot f(r, \theta, \text{aplicación de carga externa}) \quad (2.2)$$

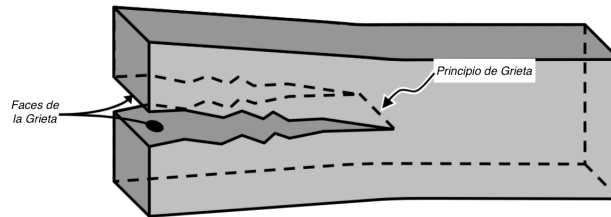
Donde  $f$  es un factor geométrico que depende de  $r$  y  $\theta$ . Estas últimas variables denotan el



(a) longitud de la grieta.



(b) Concentración de tensiones en punta de grieta.



(c) Planos opuestos a la abertura y desarrollo de la grieta.

Figura 2.2: Ilustración del desarrollo de grietas en materiales [15].

sistema de coordenadas empleado en la punta de grieta, que para este caso corresponde a un sistema polar. De esta forma se caracteriza el crecimiento de grieta en torno a la punta de grieta tomando las consideraciones geométricas que describen la fisura.

Las publicaciones de Paris mostraban los resultados como una relación lineal entre  $\log(da/dN)$  y  $\log(\Delta K)$ . Posteriores pruebas sobre el crecimiento de grietas se realizaron y la tendencia de los datos validaban la ecuación de Paris:

$$da/dN = C\Delta K^m \quad (2.3)$$

El elemento de fuerza que lleva al crecimiento de la fractura está presente en  $\Delta K$ . Las constantes  $C$  y  $m$  no se caracterizan fácilmente sólo con las propiedades físicas del material y se obtienen mediante experimentación en gran parte. Sin embargo, para valores grandes y pequeños de  $\Delta K$ , las muestras se alejaban de la tendencia de Paris, lo que conlleva a la definición de tres regiones como se muestra en la Figura 2.3 [22], [21]. Las regiones que destacan por su criticidad son las dadas por  $I$  y  $II$ . Se observa que existe un valor para  $\Delta K$

tal que la tasa de crecimiento de la grieta es tal que la fractura es instantánea. Para el caso  $\Delta K = K_{max}$ , es intuitivo considerando que si se alcanza tal nivel de estrés, la fractura se realiza en condiciones quasi-estáticas.

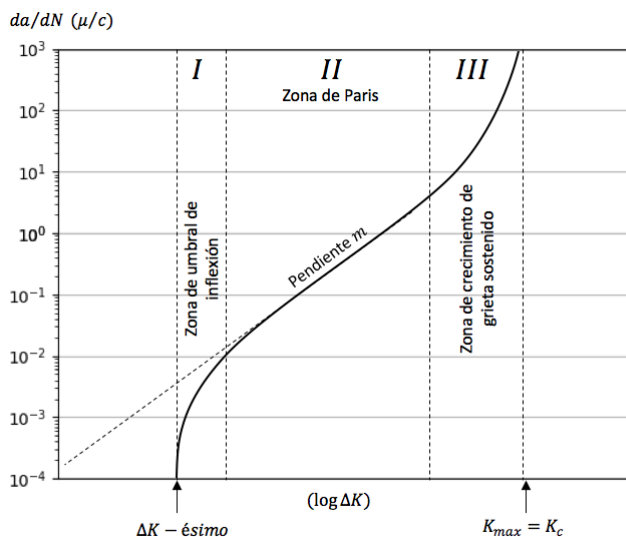


Figura 2.3: Relación obtenida para la tasa de crecimiento de la grieta por ciclo dado una carga externa  $\Delta K$  [21].

Cabe mencionar un importante efecto sobre la grieta cuando está sometida a cargas externas. Este es denominado “cierre de grietas” y es inducido por residuos de tensiones compresivas cuando la carga externa  $S$  experimenta una transición de  $S_{max}$  a  $S_{min}$ . Como se mencionó anteriormente,  $\sigma$  provoca deformaciones plásticas que propagan el crecimiento de la grieta, sin embargo, esta variación desde el valor máximo a mínimo de la carga externa hace que  $\sigma$  tome un valor negativo. Esto se aprecia en la Figura 2.4; se muestra el caso para la dimensión vertical  $y$ . Tras la transición de  $S_{max}$  a  $S_{min}$ , los valores negativos de  $\sigma$  en función de la dimensión  $x$  que se aprecian en el gráfico inferior derecho generan el efecto de cerradura en la grieta, y a menos que, se aplique un valor  $\sigma_{op}$  que supere el umbral que permita reabrir la grieta, esta queda en este estado [15].

El concepto de límite de fatiga se instaura luego de los experimentos de Wöhler donde obtuvo las curvas  $S-N$ . En los ensayos que este realizara, se sometió a distintos especímenes muestreados a tensión mecánica cíclica. La magnitud del estrés cíclico aplicado,  $S$ , se contrastó con la cantidad de ciclos de longevidad observados,  $N$ , definidos como la cantidad de ciclos aplicados antes que la fractura ocurra en el espécimen. Dicha curva se muestra en la Figura 2.5, donde los valores de  $S$  se presentan en escala lineal y  $N$  en escala logarítmica. La propiedad que satisface  $S$ , es que corresponde a una tensión mecánica cíclica de amplitud constante. De la curva  $S-N$ , se distingue la línea recta inferior que define una cota para la cual la fatiga no ocurre después de una cantidad de ciclos considerablemente alta dado ese nivel de estrés aplicado. Este valor se denomina límite de fatiga y dicha zona se denomina fatiga de alto ciclo. En contraparte, la fatiga de bajo ciclo está presente en la zona superior del gráfico. Además de la diferencia en cuanto a magnitud de esfuerzo aplicado, la fatiga de alto ciclo depende en mayor medida de condiciones superficiales de la pieza como rugosidades o capas constituyentes, a diferencia de la fatiga de bajo ciclo, donde cobra mayor importan-

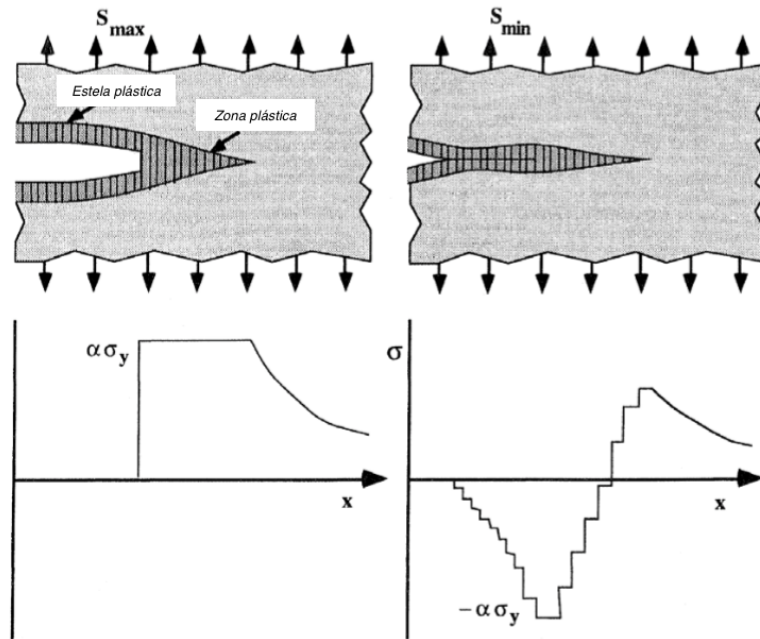


Figura 2.4: Efecto de “cierre de grietas” [15].

cia las deformaciones plásticas que ocurren en la superficie. El límite de fatiga cobra gran importancia en el campo de la ingeniería, ya que, como se ha mencionado, se define como el valor de estrés al que el componente es sometido de tal modo que la longevidad se extiende al infinito -una cantidad de ciclos muy grande relativamente-, y, para valores mayores a este, la vida del componente se acorta a un número finito. Luego los niveles de esfuerzo a los que el componente debe estar sometido por diseño idealmente no superan aquel umbral. Sin embargo, esta relación considera un esfuerzo aplicado cíclico de componente promedio cero. Para el análisis de fatiga con componente promedio mayor que cero se utiliza el diagrama de Goodman, que relaciona los valores de tensión alternante y tensión media para definir los límites de operación en los cuales el componente acorta su longevidad debido a fatiga. Por otro lado, si no hay nucleación de micro grietas, la fractura no ocurre [21]. Sin embargo, la superficie aún está sometida a deslizamientos pero debajo del límite de fatiga. Luego de experimentos en que una muestra entallada se sometía a esfuerzos debajo del límite de fatiga, no se observaba la aparición de macro grietas, sin embargo, sí aparecían micro grietas [7], [21]. Por lo tanto, el límite de fatiga puede ser definido como el umbral en que la amplitud de los esfuerzos ejercidos sobre el componente comienzan a generar macro grietas.

Dado estos antecedentes, se han propuesto metodologías para predecir el daño por fatiga de los componentes. Considerando la aplicación a la cual están expuestos los materiales, uno de los principales desafíos es cuantificar el daño generado en la estructura, tomando en cuenta las variaciones cíclicas de la tensión mecánica, pero esta vez, examinando las amplitudes variables de los esfuerzos ejercidos sobre los componentes. Para esto se distinguen dos categorías: comportamiento por similitud y daño acumulado por fatiga [21]. Para el primer caso se considera que condiciones similares a componentes constituidos de la misma manera producen comportamientos similares y por ende las consecuencias son parecidas. El segundo enfoque relaciona la longevidad con el daño acumulado en función de ciclos acumulados, de forma tal que al aplicar  $n_i$  veces ciclos de amplitud  $S_{a,i}$  sobre una longevidad de  $N_i$ , entonces

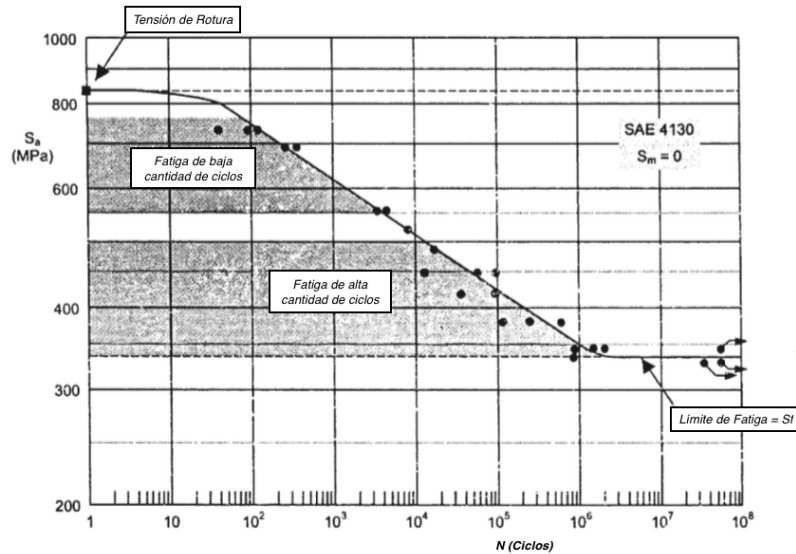


Figura 2.5: Curva S-N [21].

se ha desgastado una proporción de  $n_i/N_i$  de su vida útil antes de la fractura.

Dicho lo anterior, los estudios del crecimiento de la fractura se concentraron principalmente en dos aristas en la literatura: el estudio del campo de fuerzas actuando en las cercanías de la fractura y por otro lado, consideraciones de disipación de energía por el cuerpo en su totalidad para caracterizar la falla por fatiga [24]. En este último caso, es común la situación en que no es directa la medición de fuerzas que actúan en la zona afectada, por lo que el enfoque es medir el efecto de manera global y con perspectiva de caja negra, es decir, mediante entradas que se suponen proporcionales a la severidad de la aplicación y salidas que relacionan el grado de estado de salud del componente [15].

En consecuencia, es trascendental la identificación de fuerzas a las que está sometido el componente. En el caso de estudio, estas corresponden a dos tipos de esfuerzo provenientes del motor, las cuales recaen en el componente en cuestión: esfuerzo de régimen constante y de régimen variable, dado un punto de operación.

### 2.3. Cadena de Markov

Gran cantidad de procesos en la realidad ocurren exponiendo una secuencia sistemática de salidas o símbolos, las cuales podemos observar. Dichas salidas obedecen por supuesto a eventuales estados del sistema que no necesariamente somos capaces de definir de forma determinista. Una metodología para determinar un modelo que mejor se ajuste a los datos -u observaciones- consiste en convenir posibles estados ocultos que condicionen dichas mediciones. De esta forma, se construye un sistema estocástico embebido, dado por la distribución de probabilidades de las observaciones y las probabilidades de transición entre los estados ocultos. Esta metodología corresponde a los modelos ocultos de Markov [18].

Este procedimiento fue desarrollado a inicios del siglo *XX* por Andrei Andreyevich Markov (1856-1922), quien estudió secuencias de variables aleatorias donde el valor futuro inmediato de la variable sólo depende del estado actual. Su trabajo fue fundamental para el desarrollo de la teoría de procesos estocásticos [4].

El modelo de cadena de Markov queda definido al identificar 5 parámetros [19]:

1.  $N$ , el número de estados.
2.  $M$ , el número de posibles observaciones.
3. Las probabilidades de transición de estados.

$$a_{ij} = P[q_{t+1} = S_j | q_t = S_i], \quad 1 \leq i, \quad j \leq N \quad (2.4)$$

4. La distribución de probabilidades para las posibles observaciones dado un estado  $j$ ,  $B = b_j(k)$

$$b_j(k) = P[v_k \text{ en } t | q_t = S_j], \quad 1 \leq j \leq N, \quad 1 \leq k \leq M \quad (2.5)$$

5. La distribución de probabilidad para el estado inicial  $\pi = \pi_i$

$$\pi_j = P[q_1 = S_i], \quad 1 \leq i \leq N \quad (2.6)$$

Las aplicaciones comunes hoy en día que utilizan modelación mediante cadenas de Markov son el procesamiento digital de voz, machine learning, y en general diferentes “caminatas aleatorias” (random walks) para generar observaciones de distinta índole y obtener resultados estadísticos.

## 2.4. Rainflow Counting

Rainflow counting es un método que estima la cantidad de ciclos completos realizados por una señal periódica irregular y de amplitud variable. Esta técnica es vastamente utilizada en el estudio de fatiga de materiales [1], [8], [5], donde el objetivo es ser una aproximación numérica del daño acumulado por el componente. Dado que en aplicaciones reales, el componente es sometido a esfuerzos cíclicos irregulares, esta herramienta es de gran interés para la estimación de la cantidad de ciclos netos agregados en un período de tiempo, en este caso, con respecto a la tensión mecánica sobre el material. Gran parte de los resultados que surgieron a partir de ensayos empíricos de laboratorio son consecuencia de aplicar esfuerzos controlados sobre el componente, lo que en aplicaciones reales no se cumple necesariamente. Por esta razón, estimar la acumulación de daño mediante métricas sobre la señal de esfuerzo aplicado y correlacionarla con la longevidad del componente, es valioso para la estimación de vida útil del componente en dicha aplicación [5]. El enfoque es descomponer la señal irregular en una serie de eventos equivalentes a ciclos regulares en el tiempo con amplitud constante [8].

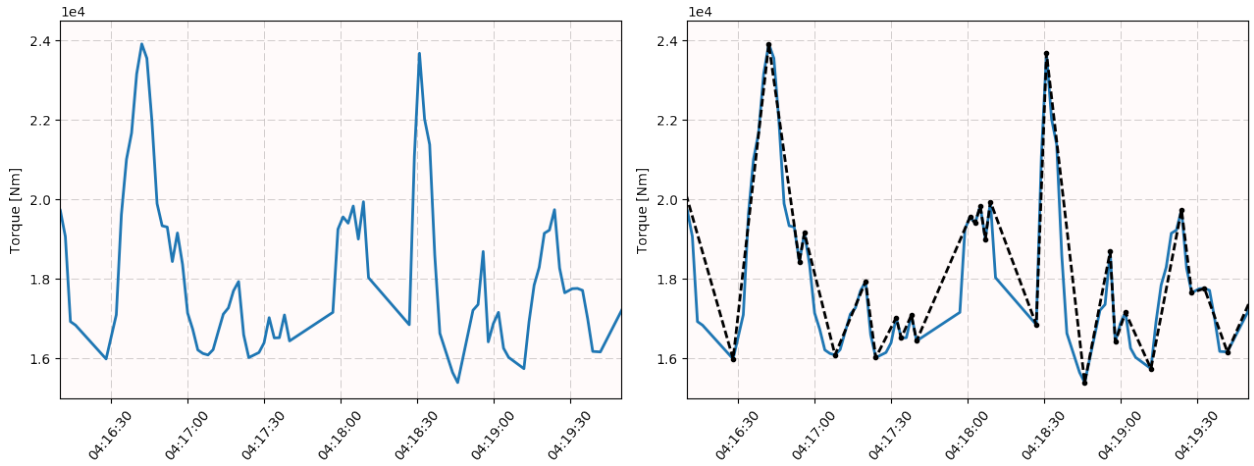
A continuación se detalla el algoritmo utilizado para la extracción de ciclos completos sobre una señal cíclica arbitraria [1].

1. Obtener los peaks de la señal. En la Figura 2.6a se muestra como ejemplo la secuencia temporal de torque aplicado por el motor de tracción del equipo minero. Para dicha

señal se obtiene los extremos identificados por el cambio de signo de la derivada con respecto al tiempo. Estos son ilustrados con línea punteada en la Figura 2.6b.

2. Se crea la sucesión  $S_k$  tal que se compone de los valores extremos identificados en el punto anterior -concatenados en el mismo orden-. Luego, se aplica el algoritmo de la Figura 2.7 sobre la secuencia  $S_k$ . El algoritmo va considerando 4 puntos sucesivos; se determinan los rangos  $\Delta S_1 = |S_2 - S_1|$ ,  $\Delta S_2 = |S_3 - S_2|$ ,  $\Delta S_3 = |S_4 - S_3|$ . Si se cumple que  $\Delta S_2 \leq \Delta S_1$  y  $\Delta S_2 \leq \Delta S_3$ , entonces se considera dentro del contador el ciclo completo caracterizado por la amplitud  $S_3 - S_2$  y valor medio  $(S_3 + S_2)/2$ . Posteriormente, se extrae de la secuencia los puntos  $S_2$  y  $S_3$ , se actualiza los índices y se itera. En la Figura 2.8 se ilustra el proceso mencionado;  $S_1$  es el punto donde inicia la línea punteada y  $S_4$  es el punto donde llega la línea punteada. Luego la figura sombreada representa el ciclo extraído dado que se cumplen las condiciones.
3. De la secuencia resultante de iterar el punto 2, no es posible formar ciclos contenidos entre 4 puntos sucesivos como ya se ha detallado en el punto 2. Por lo tanto, cada rango formado entre picos y valles consecutivos, se cuentan como medio ciclo y se finaliza el contador.

Este algoritmo es el mismo que se detalla en el estándar de conteo de ciclos en análisis de fatiga [14], salvo que en este caso, los medios ciclos se contabilizan al finalizar la extracción de los ciclos completos identificados y extraídos de la señal.



(a) Señal cíclica-irregular original.

(b) Señal procesada

Figura 2.6: Obtención de puntos peaks para aplicar el algoritmo de rainflow counting.

## 2.5. Clustering

A continuación se presentan dos tipos de algoritmos de agrupación de datos que se utilizan dentro del desarrollo de esta tesis.



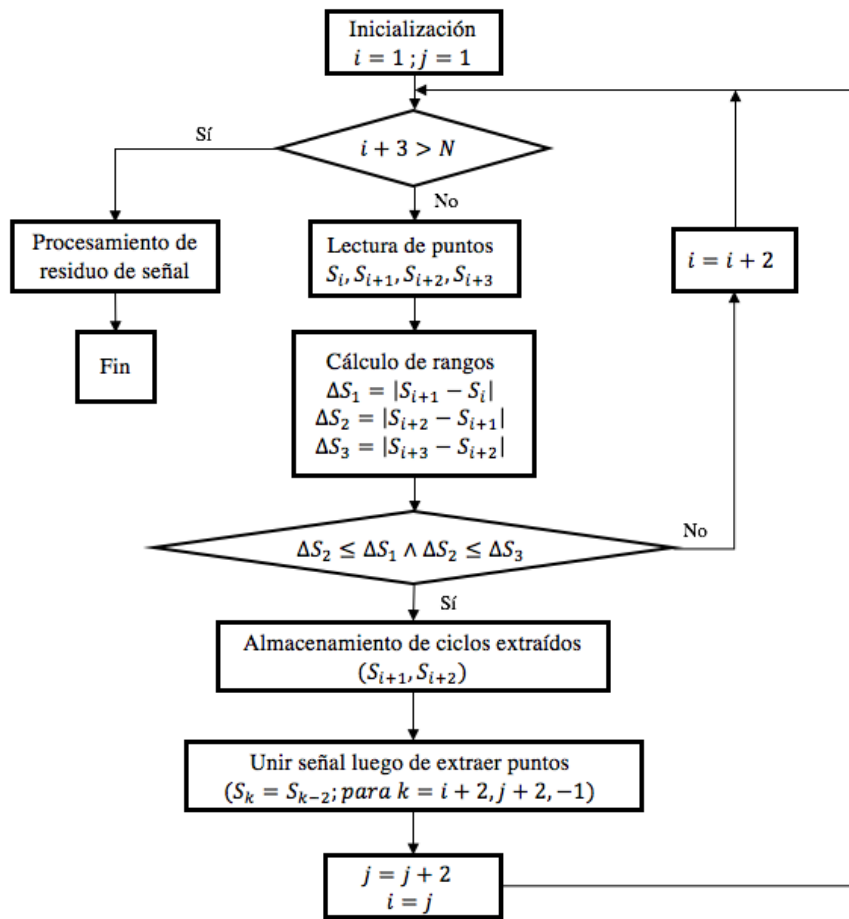


Figura 2.7: Algoritmo Rainflow Counting [1].

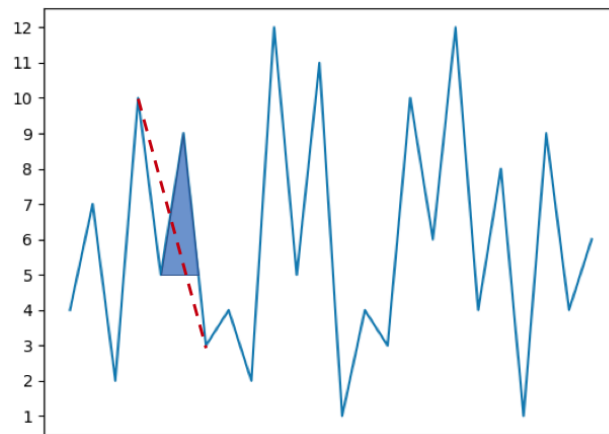


Figura 2.8: Ejemplo de extracción de ciclo para el contador [1].

### 2.5.1. DBSCAN

DBSCAN, por las siglas en inglés de Density-based spatial clustering of applications with noise, es un algoritmo de agrupación no supervisado y basado en la concentración de datos dada su distribución en el espacio a analizar. En particular, la principal idea es que los vectores de la base de datos pertenecientes a un cluster, deben tener un mínimo de vectores dentro de su vecindario definido por un parámetro radio. Para esto es necesario definir la métrica de distancia entre puntos, denotada como  $\text{dist}(p, q)$ . A continuación se formaliza los conceptos que definen las nociones de densidad y criterios para agrupar datos [6].

**Definición 2.1** ( $N_\varepsilon(p)$ ) *La vecindad dado  $\varepsilon$  de un punto  $p$ , denotada como  $N_\varepsilon(p)$  está dada por el conjunto*

$$N_\varepsilon(p) = \{q \in D \mid \text{dist}(p, q) \leq \varepsilon\}.$$

**Definición 2.2** (*Densamente alcanzable directo*) *Un punto  $p$  es densamente alcanzable directo desde un punto  $q$  con respecto a  $\varepsilon$  y a  $\text{MinPts}$  si*

1.  $p \in N_\varepsilon(q)$ .
2.  $|N_\varepsilon(q)| \geq \text{MinPts}$ .

**Definición 2.3** (*Densamente alcanzable*) *Un punto  $p$  es densamente alcanzable desde un punto  $q$  con respecto a  $\varepsilon$  y a  $\text{MinPts}$  si existe una sucesión  $p_1, \dots, p_n$  tales que  $p_1 = q$ ,  $p_n = p$  y  $p_{i+1}$  es densamente alcanzable directo desde  $p_i$ .*

**Definición 2.4** (*Densamente conectado*) *Un punto  $p$  es densamente conectado a un punto  $q$  con respecto a  $\varepsilon$  y a  $\text{MinPts}$  si existe un punto  $o$  tal que  $p$  y  $q$  son densamente alcanzables desde  $o$  con respecto a  $\varepsilon$  y a  $\text{MinPts}$ .*

Dados estos conceptos, se define qué puntos comprenden un cluster y cuales no, considerándose ruido.

**Definición 2.5** (*Cluster*) *Sea  $D$  un conjunto de puntos.  $C$  es un cluster con respecto a  $\varepsilon$  y a  $\text{MinPts}$  no vacío subconjunto de  $D$  si se cumple lo siguiente:*

1.  $\forall p, q$ : si  $p \in C$  y  $q$  es densamente alcanzable desde  $p$  con respecto a  $\varepsilon$  y  $\text{MinPts}$ , entonces  $q \in C$ .
2.  $\forall p, q \in C$ :  $p$  es densamente conectado a  $q$  con respecto a  $\varepsilon$  y  $\text{MinPts}$ .

**Definición 2.6** (*Ruido*) *Sea  $C_1, \dots, C_k$  clusters identificados de  $D$  con respecto a parámetros  $\varepsilon_i$  y a  $\text{MinPts}_i$  con  $i = 1, \dots, k$ . Se define el ruido como el conjunto de puntos de  $D$  tal que no pertenecen a ningún cluster  $C_i$ .*

Dadas estas definiciones, el algoritmo recorre punto a punto  $p$  clasificando si cumple la condición de  $\text{MinPts}$  dado  $\varepsilon$ . De ser cierto, dentro de  $N_{\text{epsilon}}(p)$  se busca todos los puntos densamente alcanzables desde  $p$ , que a su vez, cumplen la condición de  $\text{MinPts}$  dado  $\varepsilon$  y

se asigna la identificación del actual cluster. De no cumplirse las condiciones anteriormente señaladas, se asigna la identificación de ruido.

DBSCAN(SetOfPoints,  $\varepsilon$ , *MinPts*)

**Entrada:** SetOfPoints,  $\varepsilon$ , *MinPts*

**Salida:** SetOfPoints labeled

SetOfPoints *is UNCLASSIFIED*;

ClusterId  $\leftarrow$  nextId(*NOISE*);

**para**  $i \leftarrow 1$  **a** SetOfPoints.size **hacer**

    Point  $\leftarrow$  SetOfPoints.get( $i$ );

**si** Point.Ciid = *UNCLASSIFIED* **entonces**

**si** ExpandCluster(SetOfPoints, Point, ClusterId,  $\varepsilon$ , *MinPts*) **entonces** ClusterId

$\leftarrow$  nextId(ClusterId) ;

**fin**

**fin**

**Algoritmo 1:** Iniciación DBSCAN.

## 2.5.2. K-Means

Dado  $k \in \mathbb{N}$  y un conjunto de  $n$  datos tal que  $n \in \mathbb{R}^d$ , el objetivo es escoger  $k$  centroides tal que se minimiza  $\phi$ , la suma de las distancias de los  $n$  vectores al centroide más cercano [2].

Formalmente se define el problema como sigue.

$$\phi = \sum_{x \in \chi} \min_{c \in C} \|x - c\|^2 \quad (2.7)$$

Luego, se define el agrupamiento asignando a cada  $x$  el centroide  $c$  más cercano.

### Algoritmo

1. Escoger arbitrariamente  $k$  centros  $C = \{c_1, c_2, \dots, c_k\}$
2. Para cada  $i \in \{1, \dots, k\}$ , agrupar en  $C_i$  todos los  $x \in \chi$  tal que son más cercanos a  $c_i$  que a  $c_j \forall j \neq i$ .
3. Recalcular el  $c_i$  para cada  $C_i$  como el centro de masa de todos los elementos en  $C_i$ ;  
 $c_i = \frac{1}{|C_i|} \sum_{x \in C_i} x$ .
4. Repetir 2 y 3 hasta que los centros no varíen.

ExpandCluster(SetOfPoints, Point, *CIId*,  $\varepsilon$ , *MinPts*)

**Entrada:** SetOfPoints, Point, *CIId*,  $\varepsilon$ , *MinPts*

**Salida:** Boolean

```
seeds  $\leftarrow$  SetOfPoints.regionQuery(Point,  $\varepsilon$ );
si seeds.size < MinPts entonces
    SetOfPoints.changeClId(Point, NOISE);
    devolver False;
    en otro caso SetOfPoints.changeClIds(seeds, CIId);
    seeds.delete(Point);
    mientras seeds <> Empty hacer
        currentP  $\leftarrow$  seeds.first();
        result  $\leftarrow$  SetOfPoints.regionQuery(currentP,  $\varepsilon$ );
        si result.size  $\geq$  MinPts entonces
            para i  $\leftarrow$  1 a result.size hacer
                resultP  $\leftarrow$  result.get(i);
                si resultP.CIId in UNCLASSIFIED, NOISE entonces
                    si resultP.CIId = UNCLASSIFIED entonces
                        | seeds.append(resultP)
                    fin
                fin
            fin
        fin
    fin
    ;
fin
```

**Algoritmo 2:** Expansión de conjuntos DBSCAN.

# Capítulo 3

## Desarrollo

En el presente capítulo se detalla los procedimientos utilizados para conformar el esquema de diagnóstico y pronóstico de fallas, así como el desglose por sección de la metodología empleada. En ese sentido, se describe el contexto de operación del equipo minero y se definen métricas de estimación de daño sobre el componente, junto a modelos de asignación de daño por ciclos de operación del equipo en función de variables de entorno de operación, tales como la ubicación geográfica del equipo dentro de la mina. Además, se especifica el procesamiento de datos y filtrado de valores anómalos de acuerdo a consideraciones físicas del camión de extracción minero.

### 3.1. Esquema General

En la Figura 3.1 se esquematiza el proceso de pronóstico de falla propuesto. Se identifican dos módulos que engloban todo el desarrollo; el “módulo de modelamiento” y el “módulo de pronóstico”. Para el primer caso, se agrupan todos los sub-procesos necesarios que definen los criterios para la extracción de características, diagnóstico de estado de salud y pronóstico para estimar la VUR. En segundo lugar, el módulo de pronóstico representa la implementación del sistema de mantenimiento basado en la condición con los modelos ya definidos en el paso anterior. Por ende, en este módulo el flujo de información desde el equipo es procesada y evaluada para tomar decisiones sobre la operación del equipo dado el estado de salud en tiempo real.

Este trabajo se centra en gran parte en el estudio del módulo de modelamiento para caracterizar el sistema, plantear y desarrollar los supuestos que conduzcan a identificar la falla mediante métricas calculadas de datos obtenidos del equipo.

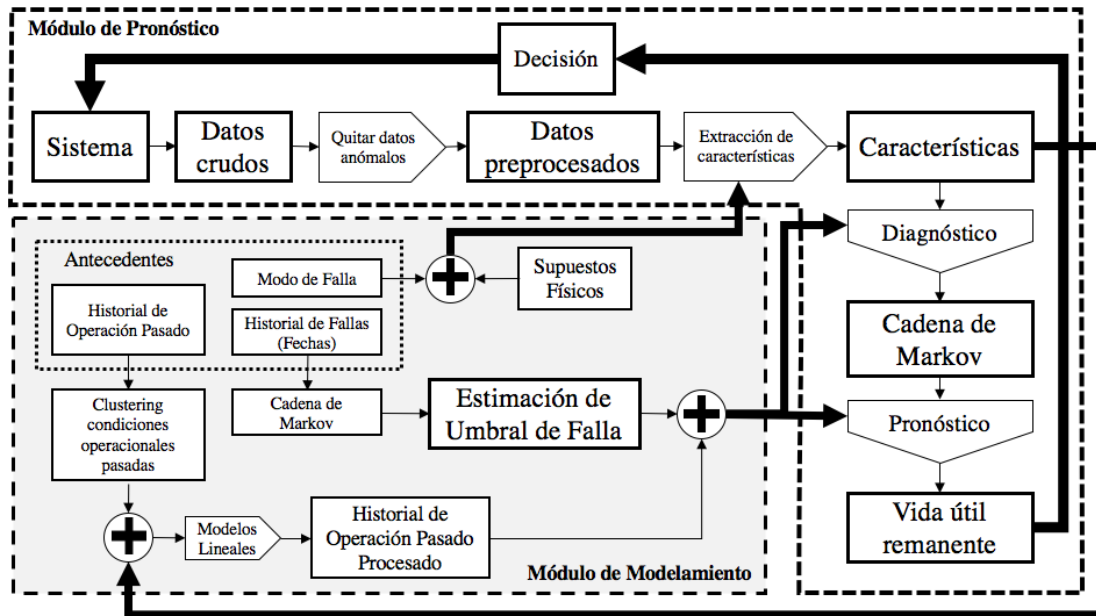


Figura 3.1: Esquema general de la metodología propuesta para el diagnóstico y pronóstico de fallas en el motor de tracción.

### 3.1.1. Módulo de modelamiento

Tal y como se ha mostrado, se reconoce una metodología basada en el esquema convencional de diagnóstico y pronóstico de fallas presentado en el capítulo anterior. Se considera en primer lugar, el análisis de antecedentes y contexto del caso, lo que da paso a la construcción de modelos y correspondiente estimación de parámetros.

Se cuenta con antecedentes de 3 tipos; el modo de falla del perno en cuestión, las fechas en que el motor de tracción fue reemplazado por ese modo de falla y, finalmente, una recopilación de datos operacionales del equipo en funcionamiento, que, como se detalla posteriormente -sección de estados de operación-, también se utiliza directamente en el proceso de extracción de características.

Dicho esto, se definen las variables físicas a monitorear así como las métricas que serán calculadas a partir de las mediciones obtenidas del sensor. Aquella métrica representa un valor proporcional al daño acumulado por el componente en estudio dado el perfil de utilización en faena. Esto se sustenta en la investigación previa del modo de falla del perno junto al estudio del estado del arte en cuanto a fatiga de materiales por daño cíclico. A considerar son las restricciones de monitoreo y datos disponibles. En particular, la definición de variables a monitorear con el fin de no sobrecargar el sistema que adquiere y almacena los datos. En esta etapa los supuestos físicos son de vital importancia, tales como efectos de la severidad operacional y relación entre variables monitoreadas con la deformación/estrés del componente. Por esta razón es fundamental el enfoque escogido, que corresponde al tipo caja negra, para modelar el componente y los efectos estructurales dada las condiciones de operación y consiguiente disminución de vida útil remanente. De esta forma, los historiales de falla desempeñan el rol que permite identificar condiciones de borde y consecuentemente

la definición de parámetros.

La segunda componente de este módulo de modelamiento, es la estimación del umbral de falla. Este se conforma a partir de la definición de daño propuesta en función de las mediciones y extracción de características, identificación de tendencias de crecimiento y reconstrucción de su evolución en el pasado para cruzar con el historial de fallas en los casos disponibles. Para aquello, se caracteriza una cadena de Markov identificando estados por el nivel de daño obtenido por ciclo (DOPC) y las transiciones de estos observando la secuencia temporal. Mediante distintas realizaciones de la cadena de Markov para la evolución del DOPC, en conjunto a las fechas de falla de los equipos, se obtiene el valor de umbral de falla.

Finalmente, se utiliza el historial de estados de operación del equipo para diagnosticar el estado de salud. Dado el contraste temporal de las dos fuentes de información que se utilizan, se diseña un método (regresión lineal por categorías) que estima el daño acumulado al momento actual en función de parámetros operacionales. Estos corresponden a: lugar de inicio del trayecto de acarreo de mineral, lugar de destino para el depósito, tiempo total empleado, equipo observado y la combinación de todas esas variables. Esto finalmente indica el valor de daño acumulado actualizado a la fecha desde el último cambio/reparación del componente en caso de que sea posible, dada la restricción de cantidad de datos en el historial por equipo. Por otra parte, este mismo modelo operaría como estimador cuando hay pérdida de información en la fuente principal de monitoreo en línea.

### **3.1.2. Módulo de pronóstico**

Con respecto al proceso en línea, se opera con todos los parámetros y reglas calculadas en el módulo de modelamiento. Una primera etapa considera desde la extracción de datos medidos por el sensor hasta el preprocesamiento como primer filtro para quitar datos anómalos. La siguiente etapa corresponde a la extracción de características de la señal medida -la cual se define en el módulo de modelamiento-. El flujo posterior diagnostica el estado de salud según el umbral de falla identificado y la estimación del daño acumulado a la fecha actual. Una vez estimado el umbral de falla en conjunto al nivel de daño actual para cada equipo monitoreado, se utiliza la misma cadena de Markov para proyectar el comportamiento del equipo y estimar cuándo -estimación de VUR- es propenso a falla mediante distintas realizaciones. En caso de estar en actual funcionamiento el equipo, la tendencia de la evolución del daño en el componente -ie. la extrapolación futura del valor acumulado de la característica extraída- permite tomar decisiones operativas con respecto al equipo. De este modo, se emplea la VUR estimada como criterio. Esto se itera a medida que se obtienen nuevos datos y los cálculos en tiempo real están acotados por lo que tardan los equipos en realizar un ciclo de carga-descarga. A su vez, los horizontes de predicción se mantienen constantes.

## **3.2. Estructura de datos adquiridos y utilizados**

La información obtenida desde el equipo se divide en dos grupos:

1. Estados de operación
2. Parámetros físicos

A continuación se lista los equipos monitoreados. Existen diferencias en la cantidad de datos obtenidos por equipo ya que se presentan intermitencias para ciertos casos. En otros, la fecha de adquisición difiere debido a la reprogramación de instancias de captura de datos en el sistema. En las Tablas 3.4 y 3.1 se detalla las diferencias :

1. 53
2. 55
3. 57
4. 59
5. 60
6. 61
7. 63
8. 64
9. 65
10. 66
11. 67
12. 69
13. 70
14. 71
15. 72
16. 73
17. 76
18. 77

A continuación se indica las características de cada uno de los grupos de información o fuentes de datos junto a las principales características del proceso de como se obtienen del sistema que almacena los datos.

### **3.2.1. Estados de operación**

Los estados de operación son variables que se adquieren y registran sólo cuando el equipo realiza determinada acción o entra en un determinado estado. Por ejemplo, cuando se está cargando material en el equipo para ser transportado o cuando el camión queda detenido en espera durante el trayecto. En este caso, el interés es conocer la fecha, hora exacta y ubicación GPS en que ocurrió dicho evento, correspondiendo a valores que se adjuntan en cada registro. Además, esta fuente de información entrega otras variables como operador de turno, interfaz de adquisición de datos y los detalles asociados para identificación en números de serie.

De todas los estados disponibles, los que se utilizan para este estudio son los siguientes:



- “Cargando”: El equipo está recibiendo material desde la pala o cargador.
- “Descargando”: El equipo libera el material acarreado en la tolva.
- “Pase”: Corresponde a la instancia en que el equipo recibe material correspondiente a un balde de la pala.
- “Toneladas”: Estimación de carga de material.

Para este estudio, se define un ciclo de operación al proceso comprendido entre la carga de material en el equipo y la descarga de dicho material. La “Estimación de carga de material” es una variable consolidada que estima por ciclo de operación el valor de carga transportado en toneladas, denominado payload.

### 3.2.2. Parámetros físicos

Los parámetros físicos corresponden a mediciones directamente de los sensores disponibles en el equipo y son adquiridos sin cumplirse condiciones previas, a menos que se especifique lo contrario. No obstante, dadas las limitaciones de las instancias de obtención de datos, denominadas sesiones de monitoreo para estas variables, es normal configurar condiciones para no saturar el sistema y evitar su eventual caída. Un caso tipo de esta situación corresponde a adquirir una determinada variable física sólo cuando sobrepasa un valor umbral predefinido. La velocidad del equipo, temperatura del motor o revoluciones por minuto de la rueda son un ejemplo de este tipo de variable.

La única variable que se extrae de esta fuente de información corresponde al torque aplicado por el motor de tracción, programando la instancia de adquisición de datos de forma tal que se guarda el dato del sensor sólo cuando el valor absoluto de esta sobrepasa los 16000  $[Nm]$ . La razón de este criterio es debido a la saturación del sistema que alberga la recolección de datos tratando de reducir lo máximo posible la carga a este mismo, como se ha explicado previamente.

El período de muestreo del sistema para el caso de los parámetros físicos es configurable. La máxima resolución corresponde a 1 segundo. Sin embargo, la información se entrega en una ventana de tiempo de 2 segundos, conteniendo el valor mínimo, máximo y promedio  $\tau_{min}$ ,  $\tau_{max}$  y  $\tau_{mean}$  respectivamente para el rango de tiempo observado. Por lo tanto, cabe mencionar que como la información se obtiene de manera discreta, para el instante de tiempo  $k$ , se tiene  $\tau_{min}(k)$ ,  $\tau_{max}(k)$  y  $\tau_{mean}(k)$ . En la Figura 3.2 se ilustra una ventana temporal para la evolución de  $\tau$  en uno de los equipos monitoreados cuando está en operación. En particular, cuando se está trasladando con carga. Para cada instante de tiempo se grafica los valores máximos, mínimo y promedio registrados.

## 3.3. Extracción de características

Uno de los enfoques utilizados para caracterizar la falla de componentes debido a fatiga cíclica es simular el comportamiento del componente realizando experimentos de laboratorio.

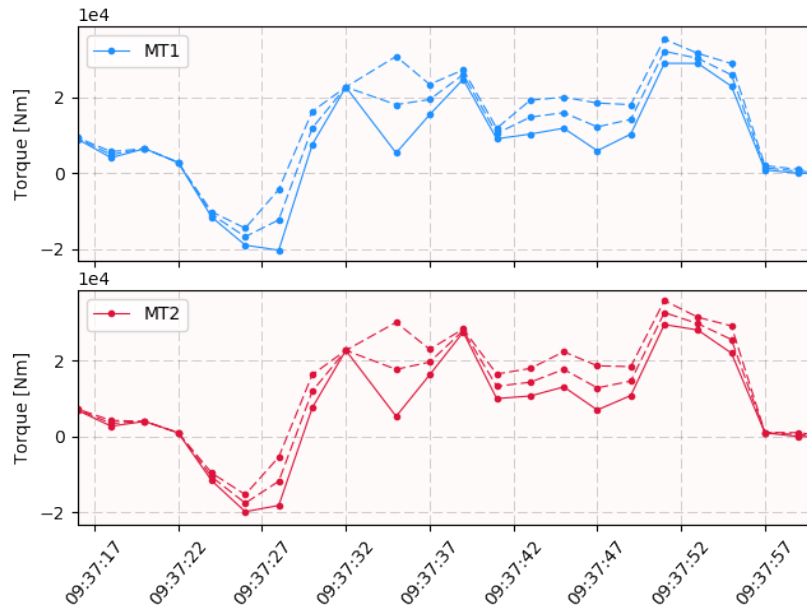


Figura 3.2: Secuencia temporal de la variable medida, torque de motor de tracción  $\tau$ .

Estos son sometidos a esfuerzos similares a los que la aplicación final le demanda. De ahí, se construyen curvas que relacionan la cantidad de ciclos a los cuales fue sometido antes de la fractura. Estas curvas son por ende las estimaciones que se utilizan para calcular la VUR como función de similitud entre la aplicación real y de laboratorio [9]. No obstante, para el caso de estudio no se cuenta con datos de laboratorio y se propone estimar la VUR en función del historial de fallas disponible con las métricas propuestas de DOPC. Para realizar dichos experimentos en laboratorio, es necesario conocer los vectores de fuerza a los que es sometido el componente en la aplicación para reproducir el comportamiento, lo que hace que sea muy costoso en tiempo y dificultoso en conocer y reproducir de manera exacta los esfuerzos a los que se somete con las máquinas de laboratorio. Por esta razón se escoge simulación mediante el modelo de cadena de Markov para caracterizar la variabilidad estadística del límite de fatiga del material, suponiendo un modelo de caja negra entre el torque de motor medido y la deformación del carrier de velocidad que contiene el componente.

De especial interés para este punto del desarrollo, es trabajar sobre el supuesto de que la variable medida directamente desde el equipo,  $\tau$ , es proporcional al estrés al cuál está siendo sometido el componente, conteniendo características cíclicas que figuran como el elemento de variación periódica presente en las ecuaciones de fatiga de materiales como la ley de Paris.

En la Figura 3.3 se muestra una ventana de tiempo en que el equipo realiza la descarga de material. Este análisis fue realizado con la adquisición de 4 parámetros previo a la decisión de monitorear sólo la variable  $\tau$ . El proceso de descarga se distingue en el parámetro payload que pasa de un valor aproximado de 340 toneladas a un valor cercano a 0 pasado el minuto 38. En el gráfico inferior de la misma figura está la señal binaria para 2 parámetros: “forward request” y “reverse request”. Estas representan la acción del operador para ir hacia adelante y hacia atrás respectivamente. Para el lapso de tiempo en que el operador activa la marcha hacia atrás, la señal de torque de ambos motores sufre variaciones significativas para pequeños lapsos de tiempo del orden de 1 segundo. En particular la señal alcanzó variaciones del orden

de 25  $[KNm]$  máximo como se puede apreciar en la Figura 3.4, que corresponde al mismo evento anterior ampliada en el tiempo. No obstante, esta situación es repetitiva para la mayoría de los ciclos observados en cada uno de los equipos.

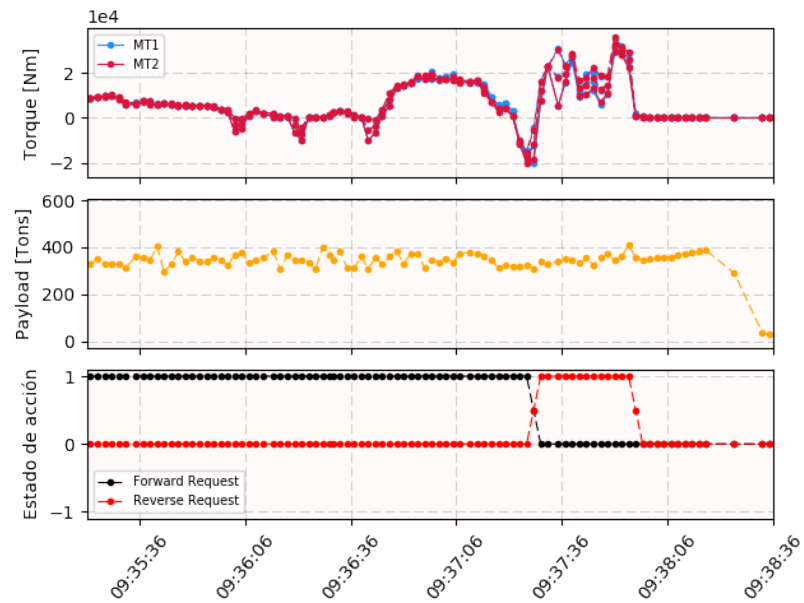


Figura 3.3: Comparación temporal de variables que indican el estado de operación del equipo para descarga de material.

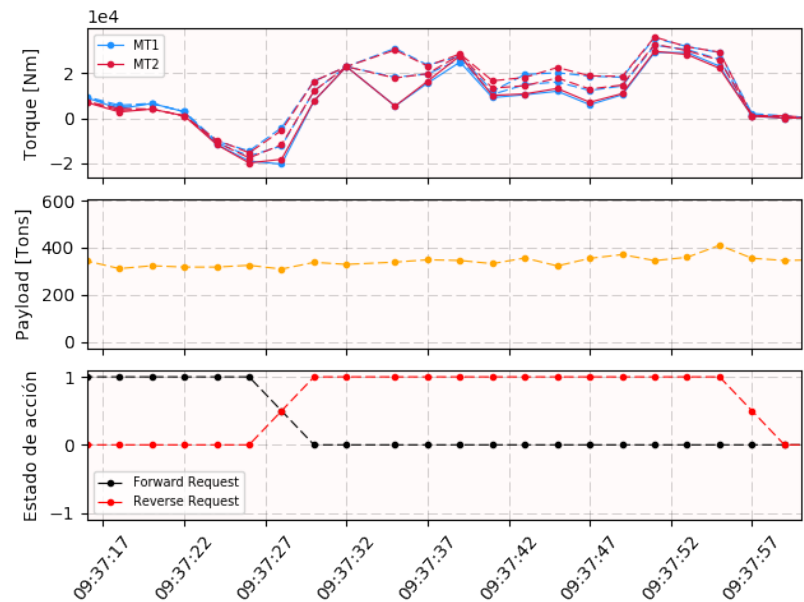


Figura 3.4: Acercamiento al instante de tiempo en que se cambia a marcha atrás.

La periodicidad del comportamiento entre ciclos se observa en la Figura 3.5. Se evidencia que el comportamiento de torque de motor de tracción aplicado alcanza grandes valores cuando el equipo va cargado. Lo que coincide con la intuición de que para mantener al sistema completo en movimiento, venciendo a las fuerzas que se oponen, requiere mayor esfuerzo al acarrear una considerable cantidad de masa en material extra. Por esta razón, la obtención

de información desde los parámetros físicos se hace efectiva bajo la condición impuesta que corresponde a equipo cargado. Esta condición se sustenta -en posterior consideración a las limitaciones del sistema de descarga de datos- bajo la premisa de que la variable monitoreada  $\tau$ , cobra mayor relevancia con el equipo cargado y que, potencialmente todas las situaciones severas que hacen que el componente entre en la zona de mayor exigencia estructural debido a los esfuerzos asociados -y por lo tanto, sea más propenso a fallar- se den en esta condición.

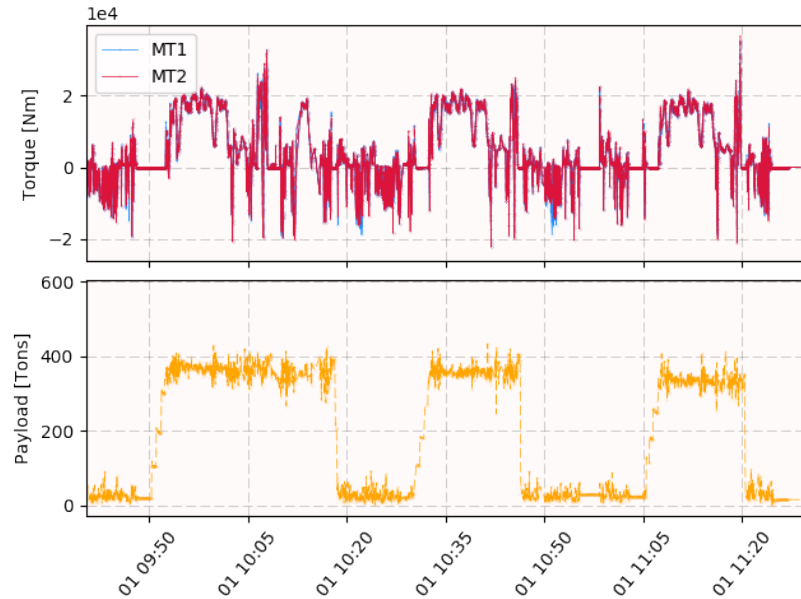


Figura 3.5: Comportamiento sistemático a través de ciclos de operación.

Por lo tanto, a continuación se detalla la conversión de los datos crudos, obtenidos desde el sensor del sistema, a las métricas propuestas para caracterizar el daño y posterior uso en los modelos para pronosticar la falla. Como se ha indicado, la variable utilizada corresponde al torque  $\tau$  aplicado por el motor de tracción. Esta se asume proporcional a los esfuerzos a los que el perno está sometido, bajo el supuesto de deformación del carrier de baja velocidad que contiene el sistema planetario de engranajes. El motivo de la obtención de este valor es la existencia de correlación con la degradación de vida útil del componente dado el estudio de contexto y antecedentes, ya que, como ha sido mencionado en el capítulo anterior, la variación de estrés a la cual el componente se somete es un buen indicador para el modelo de desgaste por fatiga cíclica de materiales. Es decir, la cuantificación de la degradación de vida útil del componente va a estar dada por el valor acumulado de la métrica extraída a lo largo de los ciclos realizados por la máquina, la cual se denota como daño obtenido por ciclo (DOPC).

Luego, la metodología contempla dividir el desarrollo de operación en varios sub-conjuntos, los cuáles están dados por el proceso de carga-descarga de mineral del equipo. Entonces, los subíndices  $k_{n-1}$  y  $k_n$  son los instantes de tiempo en que el equipo terminó de recibir la carga -comenzando el trayecto- y descarga de mineral en el botadero respectivo, en particular, para el  $n$ -ésimo ciclo de operación. La motivación de este procedimiento es sistematizar y agrupar la obtención de información utilizando los ciclos inherentes de la operación, que para una misma unidad en cuestión, comparte elementos comunes, como lo pueden ser operador de equipo, ruta que se está realizando o duración en tiempo similares dadas las localidades de punto de partida y destino. Sin embargo, estos factores eventualmente varían entre ciclos tal

y como se han definido, por lo que es importante tener en consideración para generar modelos que aprendan estas diferencias, con la intención de asociarlas a niveles de DOPC para dichas condiciones.

Antes de exponer las métricas en la extracción de características, se explica la metodología para la identificación de ciclos a los cuáles se les asigna DOPC.

### 3.3.1. DBSCAN para identificación de ciclos de operación

La identificación de los instantes  $k_i$ , para  $i = 1, \dots, n$  cuando los datos reunidos permiten observar  $n$  ciclos, se realiza mediante la información de los estados de operación, que incluye eventos “descargando” y “cargando”, entre otros, como se mencionó en la sección “Estados de operación”, para lo cual se utilizan dichos “flags” como indicadores, tal que la información más relevante, corresponde al instante de tiempo en que está en dichos estados. El procedimiento consiste en filtrar los eventos “pase” y “cargando” para identificar instantes en que el equipo está siendo cargado con material. Con los datos resultantes, se aplica el algoritmo DBSCAN de una sola dimensión, donde la variable en cuestión es el instante de tiempo de cada evento, lo cual permite distinguir entre ciclos de operación del equipo. El formato de tiempo que se emplea para utilizar DBSCAN corresponde a un número entero que indica los segundos transcurridos desde la medianoche UTC del 1 de enero de 1970, el cual es conocido como “tiempo unix” o “tiempo POSIX” ya que se utiliza principalmente en sistemas operativos unix. Luego, los parámetros para DBSCAN corresponden a los de las Ecuaciones 3.1 y 3.2 y se implementan con las librerías de aprendizaje de máquinas de scikit-learn [16].

$$\varepsilon = 420[\text{seg}] \quad (3.1)$$

$$\text{MinPts} = 1 \quad (3.2)$$

La elección de los valores para dichos parámetros está dada por la observación de los datos, considerando que los eventos de “pase” típicamente corresponden a 4 -i.e. 4 instancias de carga de material debido a la capacidad del balde de las palas- y el evento “cargando” se registra dentro de esta ventana de tiempo, los que no superan los 5 minutos en producirse. La misma metodología se emplea sobre los eventos de descarga, donde se conservan los eventos “descargando”, y “toneladas” para aplicar nuevamente DBSCAN con los mismos parámetros. Identificados cada ciclo como cluster, se procede a unir los eventos conservando la fecha del primer evento de cada cluster. El resultado de este procedimiento es la obtención de fechas que corresponden a los  $k_i$ , que definen los ciclos de operación.

Paralelamente, los valores de  $\tau$  se consiguen del “streaming” de datos.

Por lo tanto, se presentan tres métricas que se utilizan para estimar el umbral de falla: “Rango de variación mínimo-máximo”, “Longitud de arco” y “Rainflow counting”. Todas son proporcionales a la variación en la señal de torque. Esta variable se escoge debido a que se asume que la variación de estrés al cuál el componente en cuestión se somete, está dada por

esa estimación. Por consiguiente, el daño obtenido por ciclo (DOPC) va a estar caracterizado por cada una de dichas métricas y se denotarán como  $DOPC^{dmm}$  para el rango de variación mínimo-máximo,  $DOPC^{la}$  para longitud de arco y  $DOPC^{rfc}$  para rainflow counting.

Para cada uno de los cálculos, se aplican filtros en el pre-procesamiento para quitar datos anómalos erróneos por fallas en el registro del sensor. Principalmente valores atípicos que quedan fuera de los rangos físicos plausibles. Se considera un valor entre - 25000 [Nm] a 25000 [Nm] como correcto.

En la Tabla 3.1 se muestra las fechas de inicio y termino de monitoreo por equipo realizados en este estudio.

Tabla 3.1: Equipos y rangos de tiempo en el que fueron monitoreados.

Equipo	Fecha de inicio	Fecha de fin
53	2017-06-28	2017-08-15
55	2017-06-29	2017-08-15
57	2017-06-28	2017-08-08
59	2017-06-30	2017-08-11
60	2017-06-30	2017-08-15
61	2017-07-06	2017-08-14
63	2017-06-28	2017-09-19
64	2017-07-03	2017-08-08
65	2017-06-30	2017-08-12
66	2017-06-30	2017-07-13
67	2017-06-29	2017-09-19
69	2017-06-29	2017-07-12
70	2017-06-28	2017-09-19
71	2017-07-06	2017-07-13
72	2017-07-01	2017-07-10
73	2017-06-29	2017-07-12
76	2017-06-28	2017-09-19
77	2017-07-06	2017-07-13

### 3.3.2. Rango de variación mínimo - máximo

Sea  $\Delta\tau(k)$  la diferencia entre el valor máximo y mínimo medido para un instante de muestreo  $k$ , definido por la Ecuación 3.3, entendido como el rango al cuál está sometido el componente, dado cierto valor  $\tau_{mean}$ .

$$\Delta\tau(k) = \tau_{max}(k) - \tau_{min}(k) \quad (3.3)$$

Se considera el daño acumulado para una ventana de tiempo, definida por los instantes  $k_{n-1}$  y  $k_n$ , usando la Ecuación 3.4.

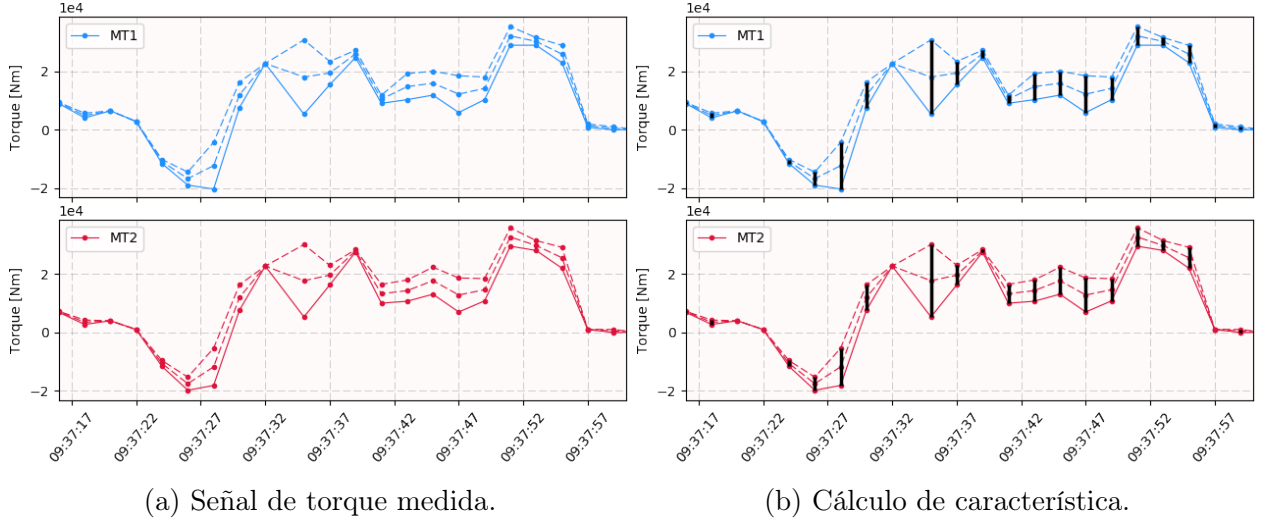


Figura 3.6: Extracción de característica; distancia mínimo-máximo.

$$DOPC_n^{dmm} = \sum_{k=k_n-1}^{k_n} \Delta\tau(k) \quad (3.4)$$

Nótese que  $\Delta\tau$  es un rango en el cual la variable tuvo alcance y no la variación definida típicamente por la sucesión temporal de valores medidos dada por  $\tau_f - \tau_i$ . De ser el último caso, la sumatoria representaría la función original, análogo a la situación de integrar una derivada, lo cual no es cierto para este caso, dado que la sucesión en el tiempo de  $\tau_{min}$  y  $\tau_{max}$  es desconocida e imposible de asumir de forma general.

Finalmente, la característica que resume el daño acumulado considerando el rango de variación mínimo-máximo, para iterar el modelo de diagnóstico y posterior pronóstico de la futura falla, corresponde al cálculo de  $\Delta\tau$  de la Ecuación 3.4 entre los límites de los ciclos carga-descarga identificados. Así, para todos los ciclos que fueron ejecutados por el equipo en la fecha de evaluación de la Ecuación 3.4, se tiene la acumulación de daño respectiva. Finalmente, considerando todos los ciclos desde el último recambio de componente, por hipótesis la métrica converge a un límite de fatiga de material en que se produce la fractura.

### 3.3.3. Longitud de Arco

La siguiente métrica para estimar el nivel de fatiga o acumulación de daño en el material corresponde a la longitud de arco.

Dada la señal de torque, se calcula como:

$$L(k) = \sqrt{\delta\tau_{mean}^2(k) - \delta t^2(k)} \quad (3.5)$$

Donde

$$\delta\tau_{mean}(k) = \tau_{mean}(k) - \tau_{mean}(k - 1) \quad (3.6)$$

$$\delta t(k) = t(k) - t(k - 1) \quad (3.7)$$

$$DOPC_n^{la} = \sum_{k=k_n-1}^{k_n} L(k) \quad (3.8)$$

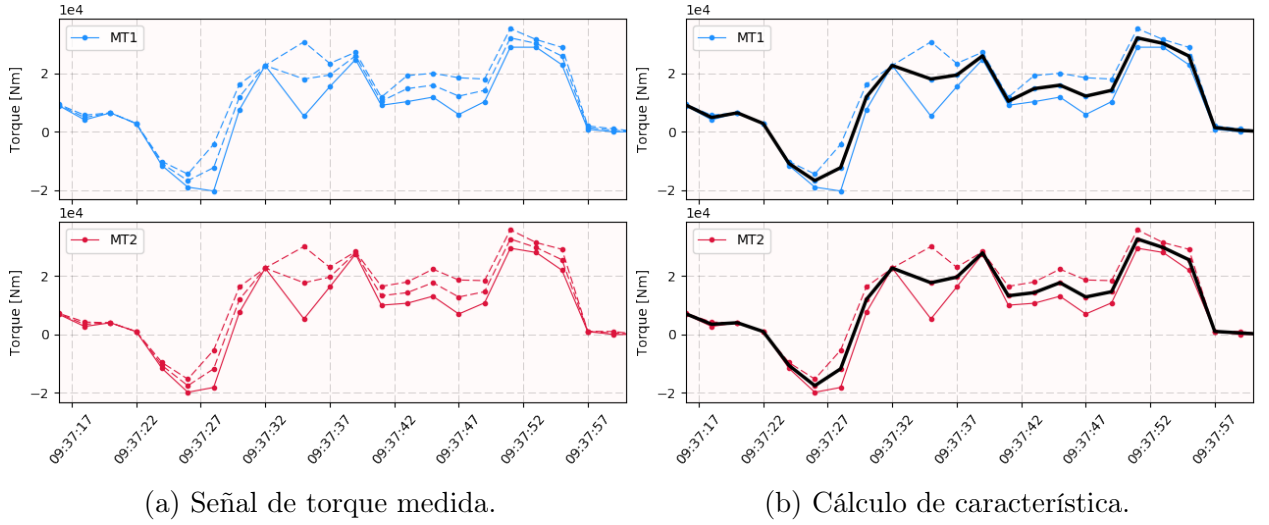


Figura 3.7: Extracción de característica; longitud de arco.

### 3.3.4. Rainflow Counting

En la Figura 3.4 se observa una muestra representativa de la variable torque de motor. Se observa el carácter cíclico-irregular de la señal. Esto es común y lo esperado en operaciones y procesos físicos que ocurren en situaciones reales - i.e. no necesariamente ideales y que se rigen por la contingencia de la aplicación-. En consecuencia, se utiliza una herramienta para estimar y extraer la cantidad de ciclos completos realizados por una señal irregular y periódica.

Dada la señal de torque, rainflow counting se calcula como:

$$DOPC_n^{rfc} = \sum_i Count_{i,n} \cdot Range_{i,n} \quad (3.9)$$

Donde  $Count_{i,n}$  es la cantidad de ciclos completos observados de rango de variación  $Range_{i,n}$  por el algoritmo de rainflow counting en el  $n$ -ésimo ciclo de operación



Nótese por tanto que, a medida que la señal sufre variaciones de gran magnitud, el contador los considera ponderando más en comparación a una variación menor. Como resultado, la acumulación de daño calculada aumenta a una tasa mayor en el tiempo y por ende, la VUR estimada disminuye. Precisamente este comportamiento de la señal de torque se ve reflejada en operaciones calificadas de mayor severidad. Este hecho es consistente con la intuición y teoría de fatiga de materiales presentada previamente en la sección de contexto teórico.

En las Figuras 3.8 - 3.10 se muestra los histogramas para los DOPC de cada caso (variación mínimo-máximo, longitud de arco y rainflow counting), observados tras la extracción de características.

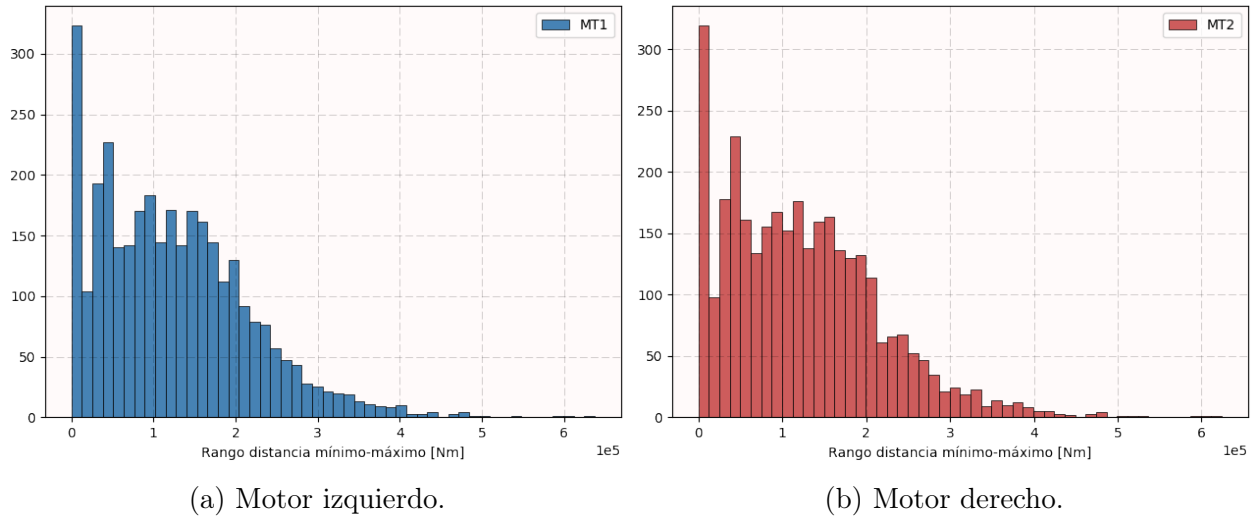


Figura 3.8: Histograma de DOPC para la métrica distancia mínimo-máximo.

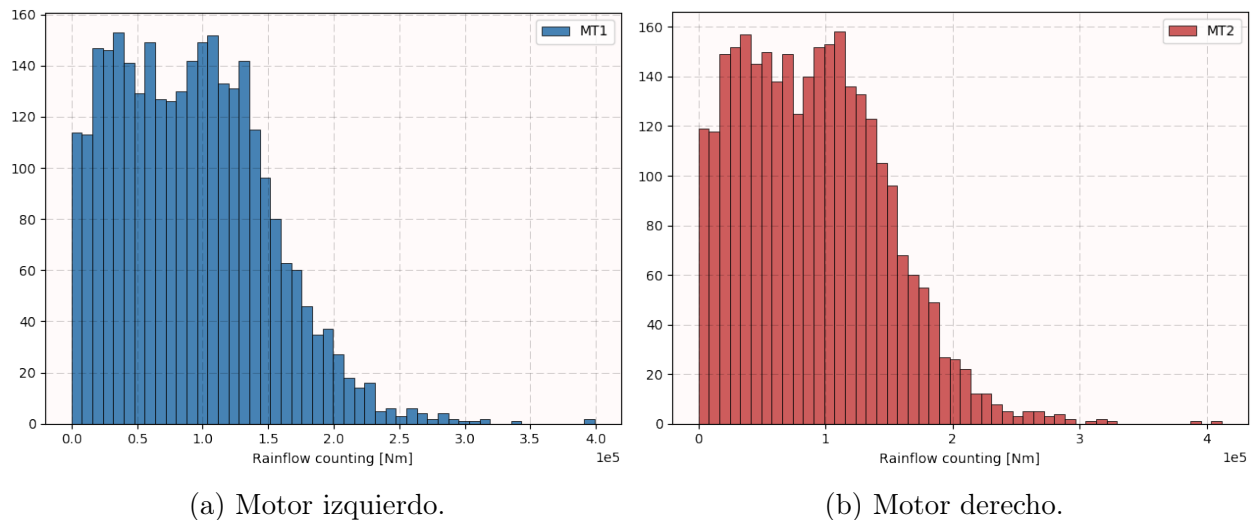


Figura 3.9: Histograma de DOPC para la métrica rainflow counting.

Los valores característicos de cada métrica se indican en las Tablas 3.2 y 3.3. Nótese que al presentarse diferentes valores, para cada métrica se identifica un valor de umbral de falla. Luego el pronóstico de falla se realiza sin mezclar estas métricas. Se compara y considera la que entregue una mejor predicción considerando el tiempo de falla real versus el predicho.

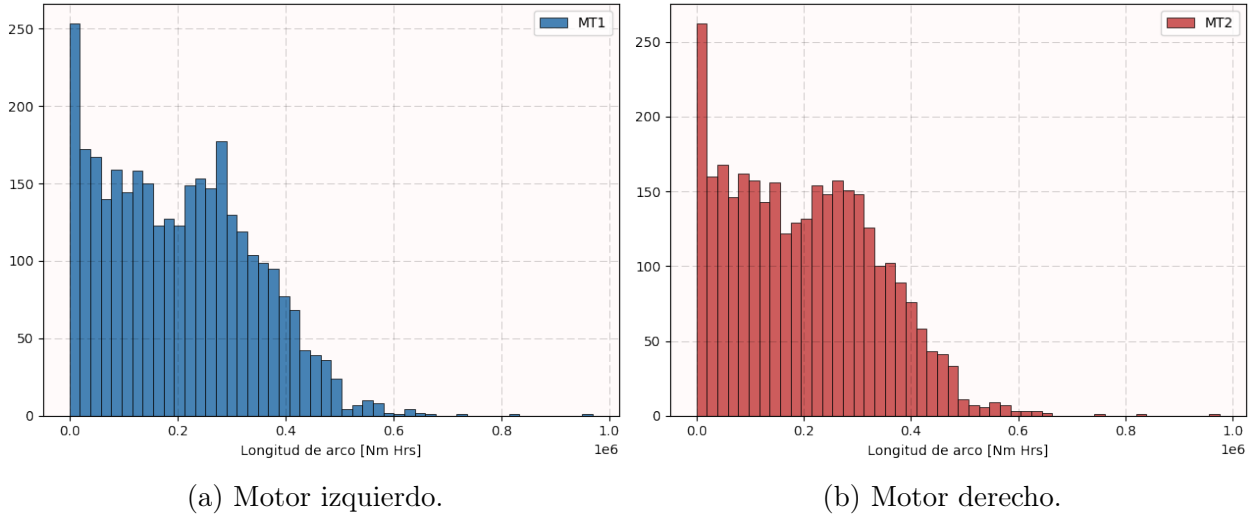


Figura 3.10: Histograma de DOPC para la métrica longitud de arco.

Tabla 3.2: Promedio de todos los equipos monitoreados por cada métrica.

	MT1 Promedio	MT2 Promedio
Distancia mínimo-máximo [Nm]	1.25e+5	1.25e+5
Rainflow counting [Nm]	9.14e+4	9.07e+4
Longitud de arco [Nm Hrs]	2.03e+5	2.03e+5

Tabla 3.3: Mediana de todos los equipos monitoreados por cada métrica.

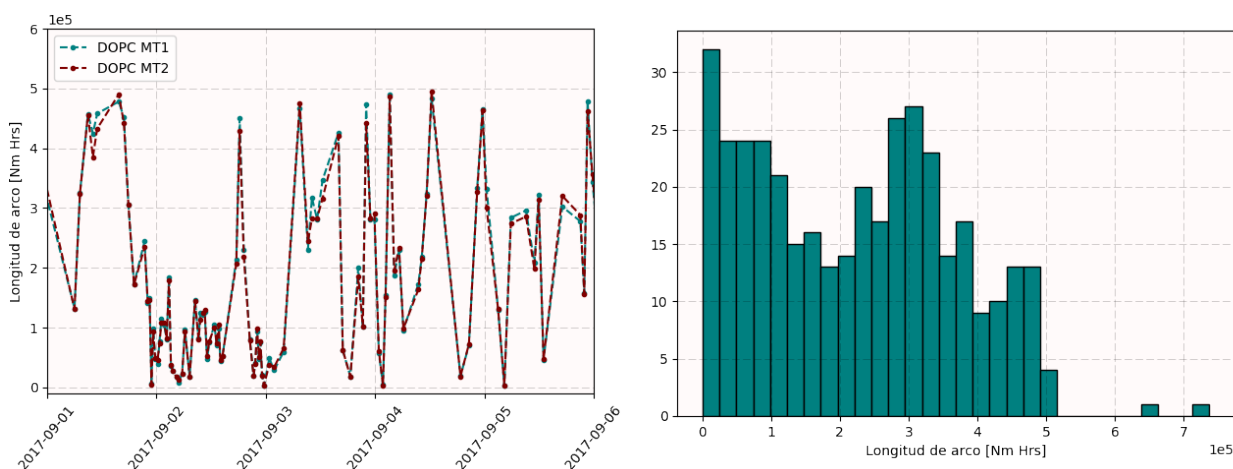
	MT1 Mediana	MT2 Mediana
Distancia mínimo-máximo [Nm]	1.14e+5	1.14e+5
Rainflow counting [Nm]	8.81e+4	8.72e+4
Longitud de arco [Nm Hrs]	1.96e+5	1.96e+5

### 3.4. Umbral de Falla

Sea el daño acumulado total (DAT) la suma de todos los DOPC dada la utilización del equipo, considerando una cantidad fija de ciclos de transporte completos ya realizados, se define el umbral de falla (UDF) como el valor de DAT para el cual el componente falla. El UDF se considera, por lo tanto, una variable aleatoria, ya que se ajusta al historial de fallas disponible y se obtiene como el promedio de los DAT en que cada caso de dicho historial de fallas. Cabe mencionar que la condición inicial, definida cuando el DAT es igual a cero, se tiene cuando el camión comienza a operar con un perno nuevo de fábrica, es decir, cuando se hace la instalación de un motor nuevo generalmente. Como ya se ha expresado, el fundamento para basar el diagnóstico y pronóstico de falla en este valor está dado por la fisura del componente por fatiga cíclica, además de considerar que el desgaste se acumula hasta llegar a un valor crítico que, se supone, obedece a una distribución con variabilidad aceptable para caracterizar un valor considerando todos los casos de falla juntos. Esto ya que se asume que el componente está constituido por el mismo material y fabricado de la misma forma. El concepto de límite de fatiga también ha sido explorado en el campo de la termodinámica, definiendo el concepto

de daño entrópico, donde se señala y se corrobora con experimentos, que la fractura de los materiales ocurre cuando su estructura se acerca a un límite de daño, debido a la evolución de los estados moleculares que constituyen el componente [13].

En las Figuras 3.11 - 3.13 se grafica la secuencia temporal para el DOPC en un equipo para cada una de las métricas definidas. De ahí se puede observar que sistemáticamente el valor del DOPC oscila entre valores altos, bajos y medios sistemáticamente. Esta identificación de valores se asocia a estados de operación en que el equipo traslada material en diferentes grados de severidad. Es decir, los valores de DOPC altos tienen mayor probabilidad de estar asociados a ciclos de mayor duración, trayectos más severos en cuanto al nivel de inclinación de los caminos, de distancias mayores, sobrecarga de material transportado o incluso por factores humanos en la operación del equipo. En el caso de valores bajos corresponde a condiciones opuestas y así, una combinación de estos factores para el caso medio. Este supuesto se utiliza para definir un modelo de cadena de Markov para reproducir el comportamiento de los equipos en operación.



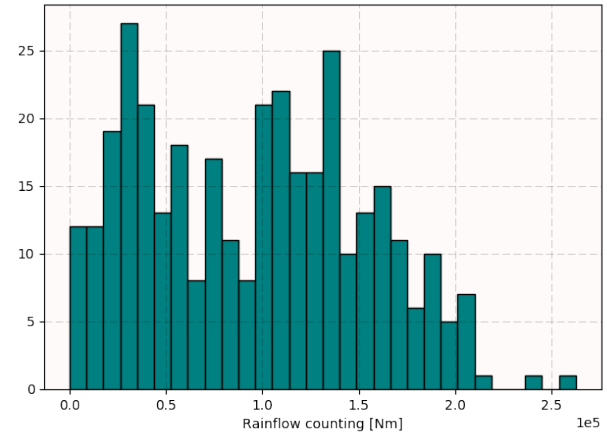
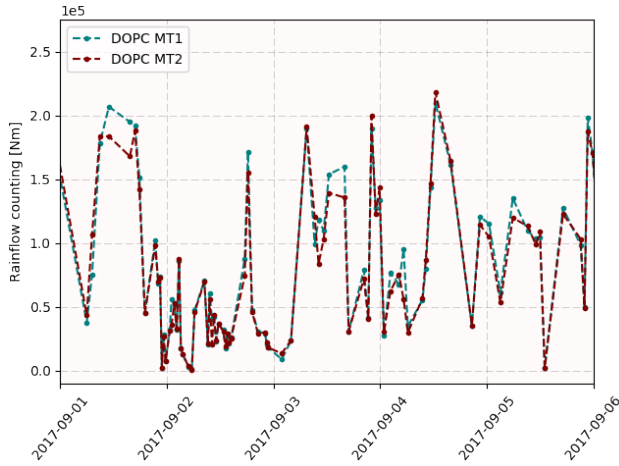
(a) Secuencia temporal de los DOPC de longitud de arco para el equipo 876. (b) Histograma de todos los DOPC por los ciclos observados del equipo 876.

Figura 3.11: Resumen longitud de arco para la aplicación de un equipo.

En la Tabla 3.4 se indica la cantidad de ciclos identificados a los cuales se pudo asociar DOPC. Esta diferencia en cantidad de ciclos observados responde a la diferencia en duración de ciclos entre equipos, asignación variable para realización de ciclos y mala obtención de datos de sensores en los parámetros físicos.

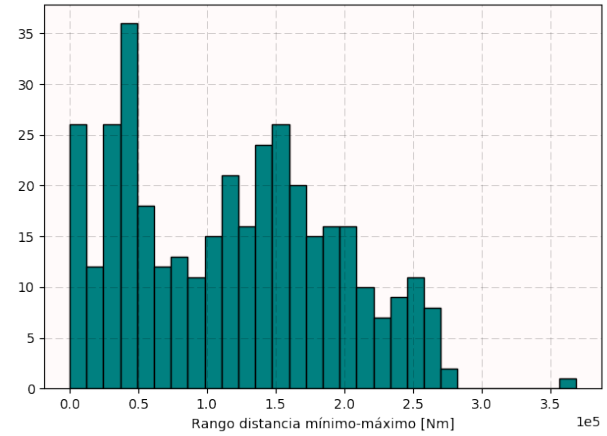
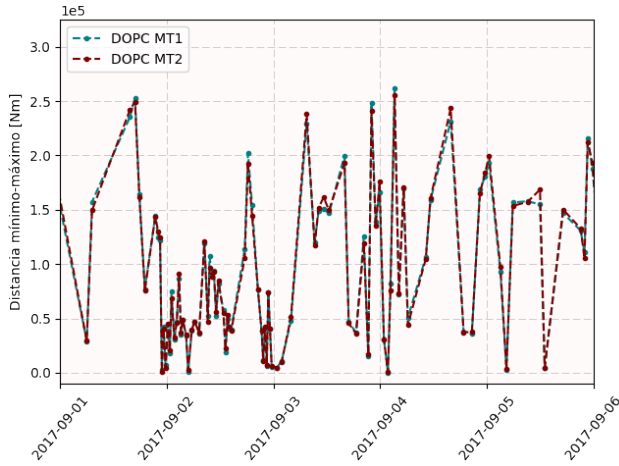
### 3.4.1. Cadena de Markov

La metodología para estimar el UDF es la siguiente; se utiliza una muestra de la secuencia de DOPC por equipo y se caracteriza la acumulación de daño mediante un modelo de cadena de Markov, donde se definen 3 estados identificados por el valor alto, medio y bajo de DOPC en el motor de tracción. Luego para cada uno de los tres estados se construye la distribución de probabilidad de manera discreta para los DOPC observados dentro de cada categoría. Una vez que ya se ha caracterizado el modelo de cadena de Markov que simula el comportamiento



(a) Secuencia temporal de los DOPC para el equipo 876. (b) Histograma de todos los DOPC por los ciclos observados del equipo 876.

Figura 3.12: Resumen rainflow counting para la aplicación de un equipo.



(a) Secuencia temporal de los DOPC para el equipo 876. (b) Histograma de todos los DOPC por los ciclos observados del equipo 876.

Figura 3.13: Resumen distancia mínimo-máximo para la aplicación de un equipo.

del equipo en el tiempo, se simula la evolución del DOPC, para cada uno de los equipos, haciendo distintas realizaciones. Este método de cálculo está basado en la caracterización de demanda energética en baterías de Ion-Litio, donde se definen por discretización estados de consumo y se calcula la probabilidad de transición entre dichos estados [17]. Finalmente se ocupa el historial de fallas para el modo de falla estudiado y se realiza un cruce de información entre las horas efectivas de funcionamiento antes de que la falla ocurriese, con la simulación de la evolución del DOPC para dicho equipo. De esta forma se estima el UDF como el promedio del valor al cual todas las realizaciones llegaron en tal cantidad de horas de funcionamiento antes de fallar.

A continuación se detalla los pasos para definir el modelo de cadena de Markov según las componentes explicitadas en la sección de antecedentes.

Tabla 3.4: Equipos y la cantidad de ciclos identificados y considerados dejando de lado datos anómalos.

Equipo	Ciclos observados
53	455
55	162
57	236
59	63
60	66
61	39
63	695
64	57
65	128
66	33
67	407
69	34
70	309
71	25
72	46
73	46
76	378
77	38

### Determinación de estados

Para cada equipo existen diferentes rangos de DOPC, para los cuales se distinguen tres estados: daño alto, medio y bajo. Esto debido a que pueden estar sujetos a ciclos diferentes de manera sostenida en el tiempo. Entonces, se emplea el algoritmo K-means para identificar 3 estados por cada equipo, donde el espacio de características es de una dimensión y corresponde a los DOPC observados por equipo.

En las Tablas 3.5 - 3.7 se muestran los resultados con el valor medio de los DOPC por estado identificado.

### Determinación de probabilidad de transición entre estados

Una vez definidas las etiquetas de cada DOPC según los 3 estados identificados en el punto anterior con K-means, se construye con un enfoque frecuentista la probabilidad de transición de un estado a otro; para cada ciclo, se observa el estado actual y el estado siguiente. Entonces para cada estado se resume en una matriz todos los siguientes estados observados en la ventana de tiempo y la distribución en cantidad de estos define la probabilidad de transición para ese estado.

Tabla 3.5: Promedio de  $DOPC_{dmm}$  en cada uno de los estados determinados mediante K-means. Valores en  $[Nm]$ .

Equipo	Estado 1	Estado 2	Estado 3
53	3.75e+4	1.42e+5	2.51e+5
55	3.90e+4	1.71e+5	3.27e+5
57	3.08e+4	1.09e+5	1.96e+5
59	5.84e+4	2.02e+5	3.42e+5
60	6.10e+4	2.75e+5	5.38e+5
61	3.71e+4	2.01e+5	4.38e+5
63	3.45e+4	1.22e+5	2.09e+5
64	2.75e+4	1.32e+5	3.00e+5
65	6.42e+4	2.28e+5	4.04e+5
66	5.31e+4	2.20e+5	3.49e+5
67	3.32e+4	1.04e+5	1.79e+5
69	8.17e+4	2.39e+5	3.54e+5
70	2.79e+4	9.42e+4	1.64e+5
71	2.37e+4	8.98e+4	1.44e+5
72	2.73e+4	1.18e+5	2.49e+5
73	4.70e+4	1.76e+5	2.84e+5
76	4.01e+4	1.39e+5	2.24e+5
77	2.29e+4	1.18e+5	2.20e+5

Tabla 3.6: Promedio de  $DOPC_{rfc}$  en cada uno de los estados determinados mediante K-means. Valores en  $[Nm]$ .

Equipo	Estado 1	Estado 2	Estado 3
53	3.24e+4	9.98e+4	1.60e+5
55	2.83e+4	9.38e+4	1.72e+5
57	2.08e+4	7.26e+4	1.39e+5
59	4.61e+4	1.11e+5	2.13e+5
60	2.41e+4	1.09e+5	2.18e+5
61	5.63e+4	2.21e+5	9.92e+5
63	3.95e+4	1.10e+5	1.76e+5
64	4.78e+4	1.25e+5	1.83e+5
65	4.89e+4	1.21e+5	2.27e+5
66	4.23e+4	1.35e+5	2.23e+5
67	3.17e+4	8.31e+4	1.42e+5
69	4.14e+4	1.02e+5	1.76e+5
70	2.53e+4	6.81e+4	1.21e+5
71	1.09e+4	4.63e+4	9.31e+4
72	3.16e+4	9.13e+4	1.44e+5
73	2.01e+4	8.13e+4	1.39e+5
76	3.48e+4	1.10e+5	1.75e+5
77	3.10e+4	7.54e+4	1.34e+5

Tabla 3.7: Promedio de  $DOPC_{la}$  en cada uno de los estados determinados mediante K-means. Valores en  $[Nm \cdot Hrs]$ .

Equipo	Estado 1	Estado 2	Estado 3
53	6.65e+4	2.28e+5	3.63e+5
55	5.43e+4	2.29e+5	4.21e+5
57	3.63e+4	1.48e+5	2.98e+5
59	4.78e+4	2.20e+5	4.26e+5
60	4.39e+4	2.25e+5	4.44e+5
61	3.78e+4	2.37e+5	5.68e+5
63	7.48e+4	2.49e+5	4.11e+5
64	8.75e+4	2.65e+5	3.93e+5
65	8.06e+4	2.68e+5	4.71e+5
66	7.70e+4	2.80e+5	4.67e+5
67	5.88e+4	1.93e+5	3.42e+5
69	8.82e+4	2.20e+5	3.55e+5
70	5.61e+4	1.78e+5	3.04e+5
71	2.69e+4	9.97e+4	1.63e+5
72	5.71e+4	1.90e+5	3.00e+5
73	3.90e+4	1.93e+5	2.99e+5
76	7.63e+4	2.74e+5	4.32e+5
77	9.07e+4	1.83e+5	2.85e+5

### Probabilidad de DOPC dado un estado y número de posibles observaciones

Las cantidad de posibles observaciones se considera como la cardinalidad del conjunto de DOPC observados, no repetidos dado un estado.

Para definir un valor de DOPC dado un estado, se considera la distribución resultante de cada estado. Por simplicidad, no se parametriza una distribución acorde a los datos, si no que para la asignación en las realizaciones, dado un estado, se ordena los valores de forma creciente y se obtiene un valor generando una posición de esta lista de forma aleatoria con distribución uniforme.

### Probabilidad para el estado inicial

Se considera uniforme para los 3 estados bajo el supuesto que en un instante de tiempo cualquiera en el pasado o futuro, el equipo puede estar en cualquiera de estos 3 estados y la evolución del DAT no presenta gran variación, una vez que el modelo de cadena de Markov construye las realizaciones a partir de aquel estado.

### 3.4.2. Simulaciones

A continuación se presenta los resultados de simulación en las Figuras 3.14, 3.15 y 3.16 con los modelos de cadena de Markov construidos. Para cada métrica se identifica un umbral de falla como ya se ha mencionado. El aspecto más importante es que la acumulación de daño difiere entre equipos y más aún, entre métricas. Esto indica que existe una métrica que mejor se ajusta al historial de fallas y al comportamiento de fractura por fatiga cíclica.

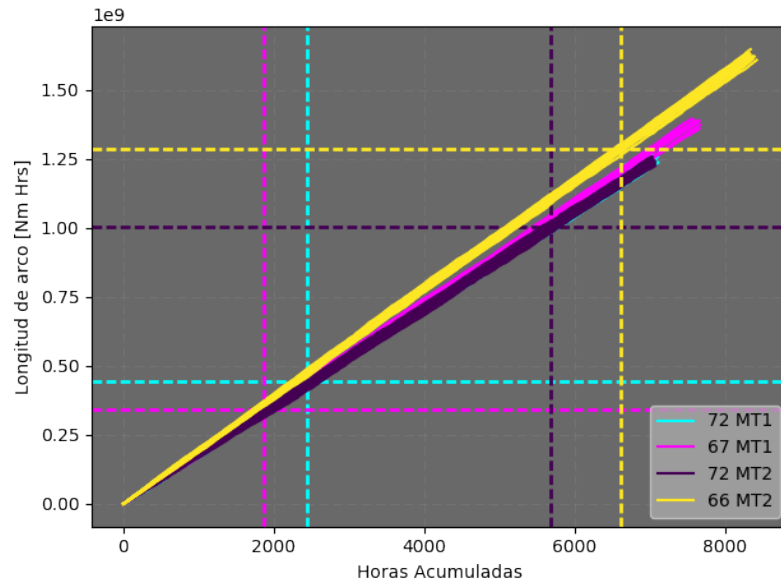


Figura 3.14: Estimación de umbral de falla para métrica longitud de arco.

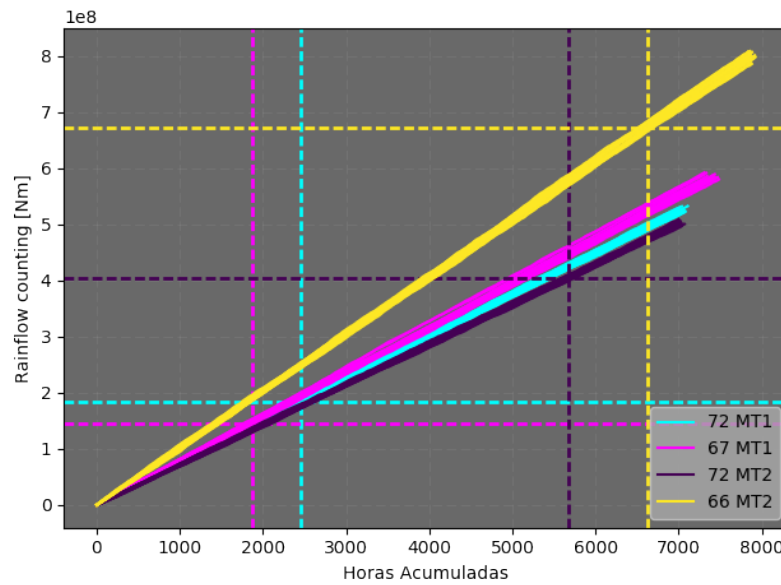


Figura 3.15: Estimación de umbral de falla para métrica rainflow counting.



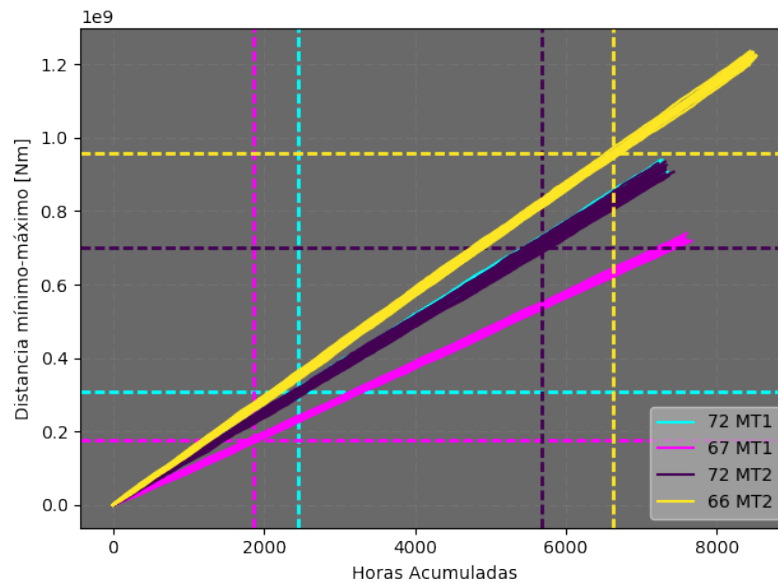


Figura 3.16: Estimación de umbral de falla para métrica distancia mínimo-máximo.

### 3.5. Diagnóstico de Falla

El planteamiento para diagnosticar el estado de salud del motor de tracción consiste en actualizar el DAT considerado desde la instalación del componente nuevo y compararlo con el UDF. Luego, si aún opera con un DAT inferior al UDF se considera en condiciones para seguir operando. Por lo tanto, esto implica conocer la historia de los equipos y contar con los datos para el cálculo de los DOPC y en consecuencia, el DAT a la fecha de evaluación. Esta información no está disponible dado que el inicio de monitoreo de parámetros físicos no coincide con ninguna fecha de instalación de motor nuevo.

La particularidad de los datos de estados de operación es que se registran históricamente y es posible acceder a ellos para un tiempo en el pasado de aproximadamente un año desde la fecha actual. Esto motiva el estudio y aproximación del valor de DOPC en función de variables de estados de operación, lo que permite explicar y relacionar el daño en cuanto a condiciones de entorno. Esto implica poder estimar el valor de daño acumulado hasta la fecha actual, desde que el componente se reemplazó, si se cumple que la fecha de reemplazo está dentro del rango temporal de estados de operación observados. Por otra parte, muchas veces hay pérdida de datos en los parámetros físicos -no así en los de estados de operación-, principalmente debido a fallas de sensor, de transmisión de datos o de almacenamiento. Esta pérdida de datos se pretende suplir con los modelos de aproximación de daño en cuestión.

Para aquello, se definen las variables de estado de operación que explicarían el DOPC y se categorizan. El objetivo es ajustar un modelo lineal en el subconjunto de datos filtrados para cada categoría identificada. Dichas variables -entendidas para un mismo ciclo- son:

- GPS de zona de carga de material y GPS de botadero
- Tiempo empleado en realizar el ciclo

- Equipo

Para cada uno de los 6 modelos creados a partir de las categorías, se estima el daño en función de:

1. Tiempo empleado
2. Equipo
3. Zona carga-descarga
4. Equipo y zona carga-descarga
5. Equipo y tiempo empleado
6. Equipo, tiempo empleado y zona carga-descarga

### 3.5.1. Definición de categorías para los datos medidos (clusters)

La motivación de agrupar los datos medidos según las 6 categorías anteriores se debe al supuesto de acumulación de daño en función de las condiciones de operación. Intuitivamente, el hecho de que el equipo se enfrente, por ejemplo, a diferentes rampas dada una condición de carga en comparación a otra de menor grado de inclinación, duración en tiempo o curvas presentes, motiva la idea de que la condición de operación está determinada completamente por estos factores, y por ende, el componente estará sometido a diferentes esfuerzos en función de aquello. Es por esto que, se examinan estos factores de entorno, para estimar cuál es el daño acumulado obtenido para cada una de dichas condiciones.

En las Figuras 3.17-3.19 se compara las tres métricas para el daño obtenido por ciclo (DOPC) en dos equipos. La hipótesis es que la agrupación de puntos por categoría evidenciarían una relación lineal dependiente de la carga transportada y el tiempo empleado en cada ciclo.

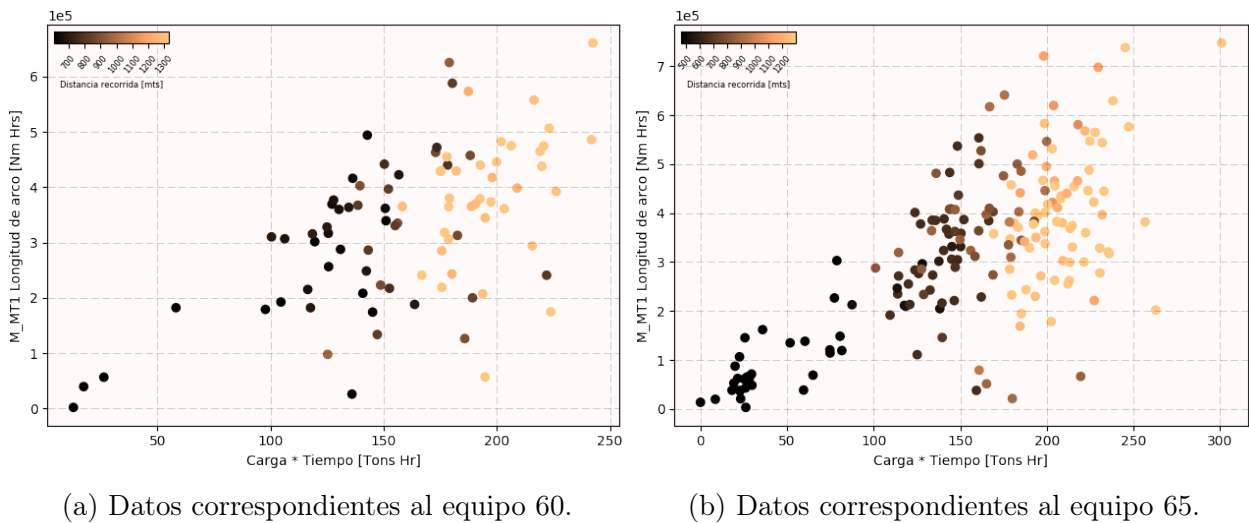
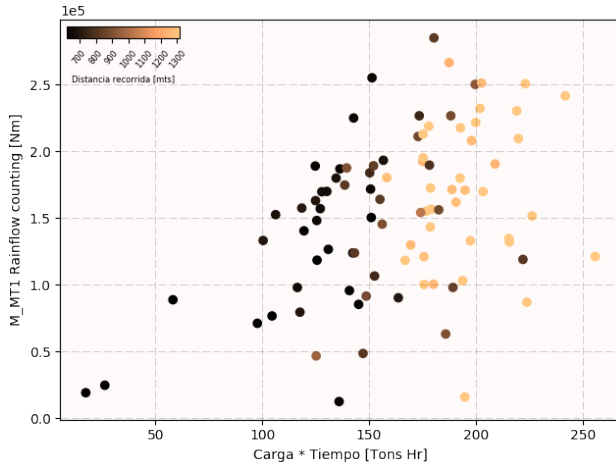
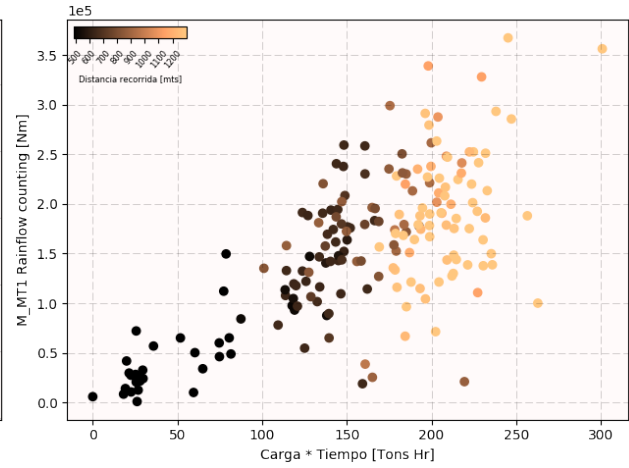


Figura 3.17: Relación de longitud de arco y carga-tiempo empleado por ciclo.

Luego, utilizando los datos de “estados de operación”, se identifican dichas condiciones de operación por ciclo. Dicho valor se aproxima con una regresión lineal para cada división de

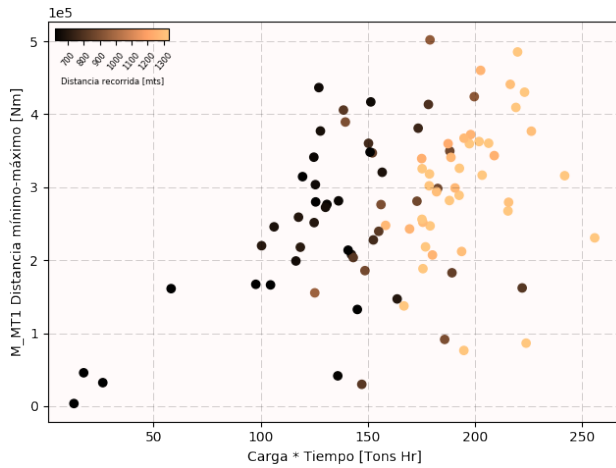


(a) Datos correspondientes al equipo 60.

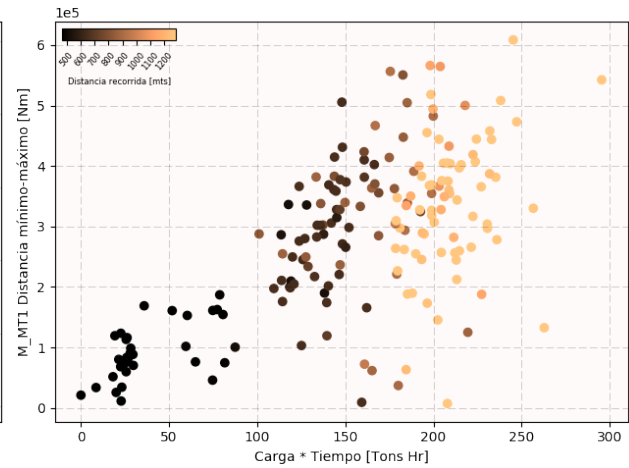


(b) Datos correspondientes al equipo 65.

Figura 3.18: Relación de rainflow counting y carga-tiempo empleado por ciclo.



(a) Datos correspondientes al equipo 60.



(b) Datos correspondientes al equipo 65.

Figura 3.19: Relación de distancia mínimo-máximo y carga-tiempo empleado por ciclo.

categoría o subconjunto de categorías que determinan un estado de operación por entorno.

Para realizar el ajuste de modelos lineales por categoría, es necesario clasificar los datos que se utilizan para el entrenamiento y utilizar sólo los que quedan dentro de cada categoría. Para esto se ocupan diferentes métodos dependiendo de la variable a clasificar. A continuación se detallan las agrupaciones.

## Variable de GPS

Esta variable se clasifica bajo el supuesto de que el camino que emplean los equipos para llevar el material entre zonas similares, para cada realización de los ciclos, debiesen ser parecidos si es que no son los mismos. Esto implica que la operación, y por ende, el desarrollo de la variable  $\tau$ , también debiese ser parecida. Se cuenta con un conjunto de locaciones GPS

clasificados como zonas de carguío o botaderos. Luego, se utiliza K-means para clasificar a qué zonas pertenece cualquier punto de la faena, contando a simple vista cuántos lugares o aglomeraciones se pueden distinguir de entre todos los datos expuestos, como se aprecia en la Figura 3.20. El criterio para utilizar K-means es la característica de predefinir la cantidad de zonas que se quiere identificar, lo cual es útil para no separar los datos en muchos grupos, ya que se busca hacer una relación lineal representativa entre el DOPC y la carga-tiempo por ciclos.

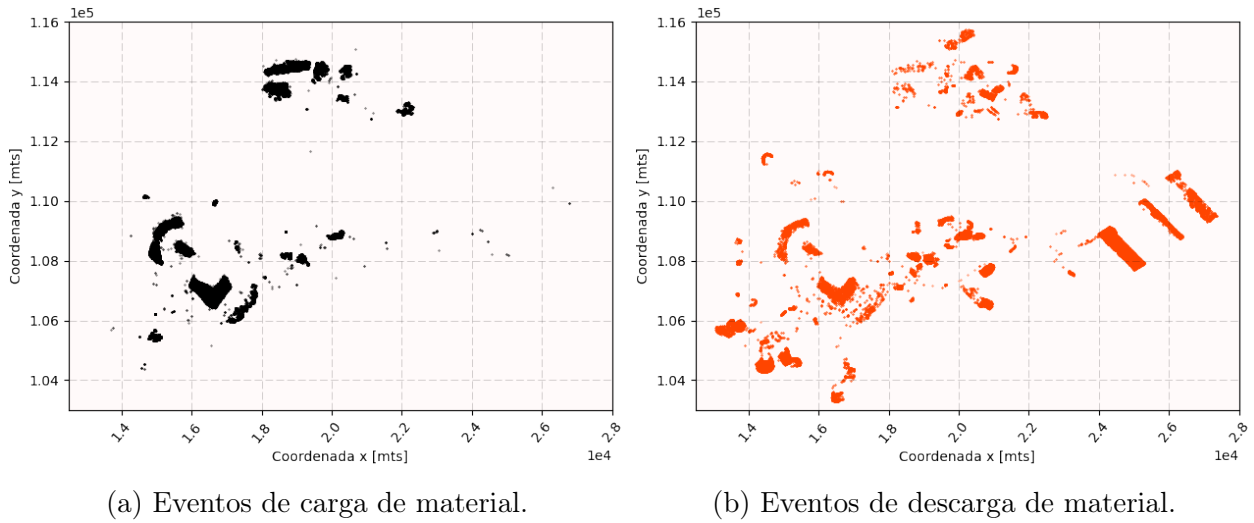


Figura 3.20: Distribución espacial de zonas de carga y descarga.

Luego de aplicar K-means, se obtiene una clasificación de zonas como se muestra en la Figura 3.21. Para las zonas de carga, el parámetro predefinido para k-means es 9 zonas, mientras que para las zonas de descarga es 15 zonas.

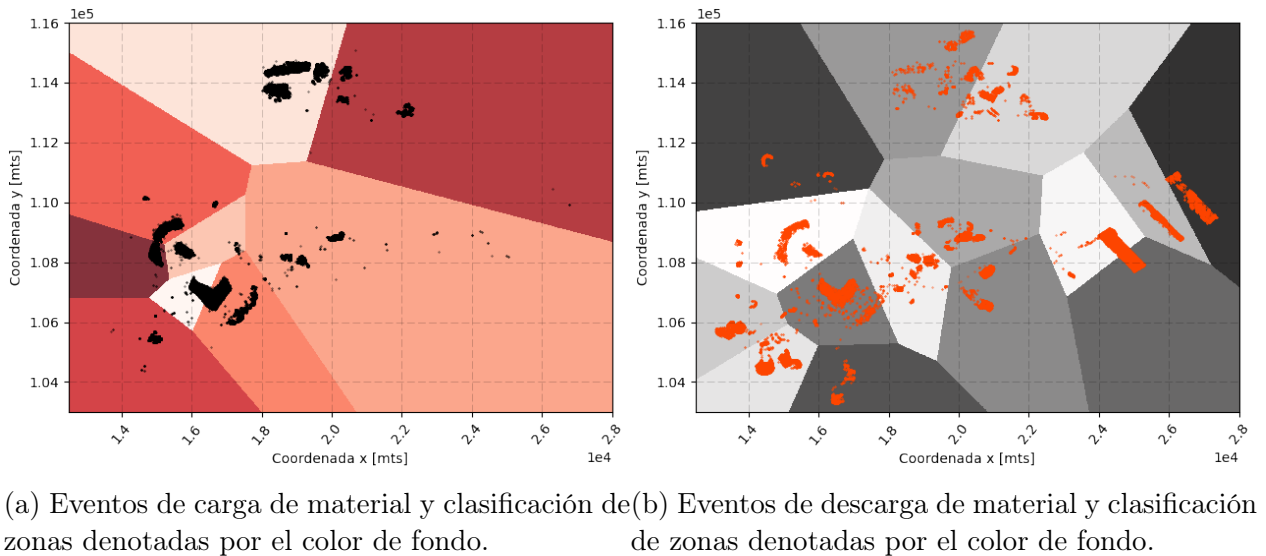


Figura 3.21: Agrupación y clasificación por conjunto de zonas de carga y descarga.

## Tiempo empleado en realizar el ciclo

El tiempo empleado en realizar el ciclo se categoriza bajo la premisa de que aproxima la ruta transitada por el equipo suponiendo que las realizaciones de trayectorias sistemáticas tienden a comportarse de forma similar en velocidad y por lo tanto debiesen tardar aproximadamente lo mismo. Por esta razón la importancia del tiempo empleado, a costa de que no se tiene la localización GPS en todo momento, se usa como información acerca del camino recorrido dado un punto de inicio y punto de fin. En la Figura 3.22 se muestra los tiempos empleados por ciclo de todos los equipos.

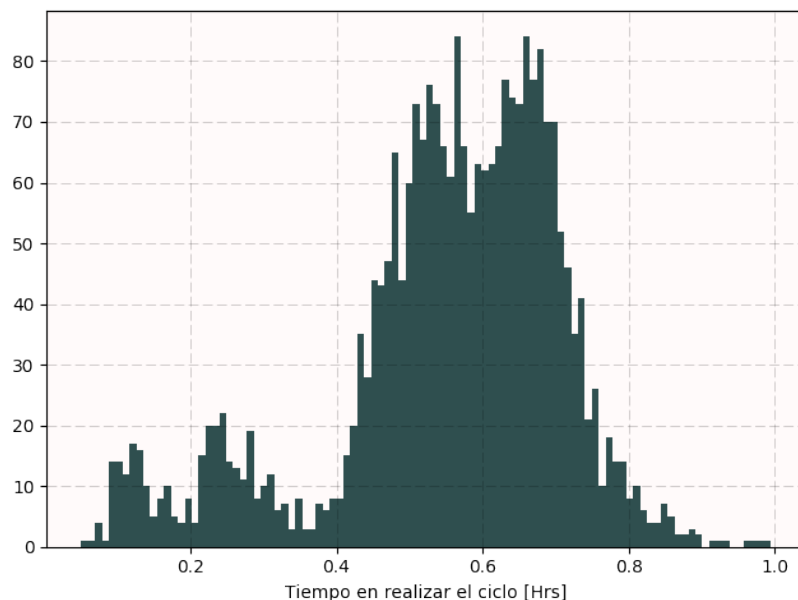


Figura 3.22: Tiempo empleado por ciclo de todos los equipos monitoreados.

## Equipo

La diferenciación por equipos se realiza bajo el supuesto de que las asignaciones de zona de carga-botadero dentro de ventanas de tiempo son sistemáticas y no varían de forma extrema. Es decir, el mismo equipo transita por la misma ruta que tiene asignada por un tiempo considerable. Por lo tanto, entre equipos se harían evidentes las diferencias de daño acumulado. Por otro lado, es una posibilidad que los sensores entre equipos discrepen por algún mal funcionamiento en alguno en particular. Luego estas mediciones son más fáciles de aislar con respecto a todos los equipos medidos.

### 3.5.2. Definición de modelos

El objetivo es encontrar la correlación entre el valor de DOPC de cada métrica y el set de variables operacionales que definen aquel ciclo.

Para cada una de las categorías y las combinaciones de estas, se realiza una regresión lineal entre el valor DOPC observado (calculado en la extracción de características) y una métrica dada por la multiplicación de la carga acarreada y el tiempo empleado en completar el ciclo. Este valor está motivado por la intuición de que la combinación de ambas variables hacen que la operación aumente en severidad de manera proporcional a su producto.

Así, los modelos que se construyen son los siguientes:

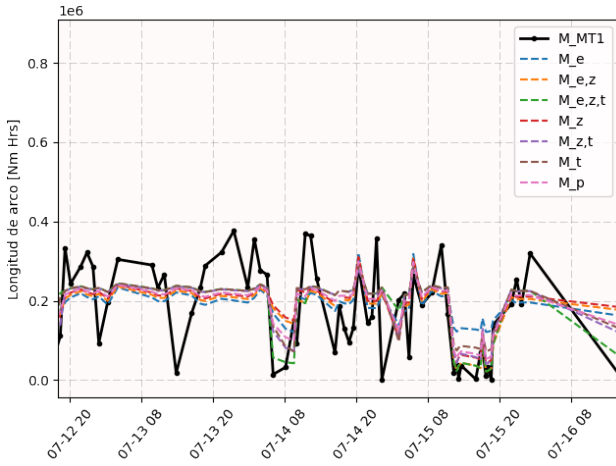
1.  $M_t$ : modelo que filtra por agrupación de tiempo empleado por ciclo.
2.  $M_e$ : modelo que filtra por equipo.
3.  $M_{e,z}$ : modelo que filtra por equipo y por zonas de carga-descarga.
4.  $M_{e,z,t}$ : modelo que filtra por equipo, zonas de carga-descarga y agrupación de tiempo empleado por ciclo.
5.  $M_z$ : modelo que filtra por zonas de carga-descarga.
6.  $M_{z,t}$ : modelo que filtra por zonas de carga-descarga y agrupación de tiempo empleado por ciclo.
7.  $M_p$ : modelo que promedia todos los modelos anteriores.

### 3.5.3. Resultados de los modelos ajustados

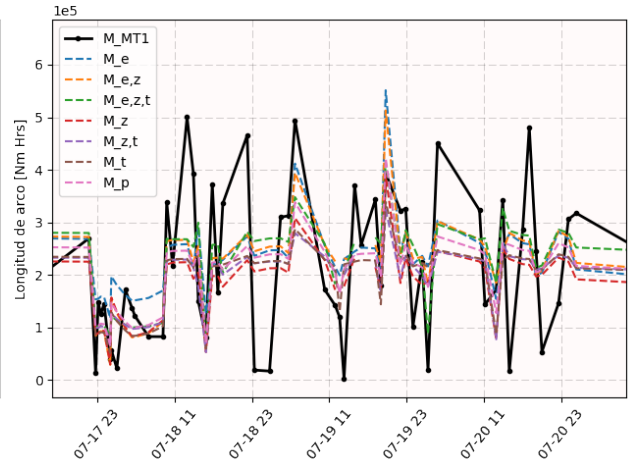
En esta sección se presentan los resultados de los modelos que aproximan el DOPC dadas las categorías de operación. En las Figuras 3.23, 3.24 y 3.26 se puede apreciar cómo aproximan este valor por cada métrica. La observación que se hace al respecto es que los modelos sólo se ajustan al promedio de los DOPC observados, mientras que a mayor diferenciación por categoría, son capaces de alcanzar valores más extremos. Este es el caso de los modelos obtenidos por filtrar equipo, zona de carga-descarga y agrupación de tiempo empleado por ciclo. Sin embargo, a medida que se profundiza en más categorías, los datos observados disminuyen por cada uno de los modelos y se aproxima una curva con pocos datos. Como consecuencia de esto, llega a ser poco representativo e incluso se puede estar modelando valores extremos y/o ruidosos, a pesar de los filtros para quitar datos anómalos. Este es el caso eventual por fallas en los “datos de operación”, las cuales en los peores casos indican fechas desfasadas o incluso dejaron de transmitir, para lo cual los ciclos quedan mal definidos y se calcula un daño entre límites erróneos.

La importancia de los valores de DOPC aproximados por los modelos es que su valor acumulado aproxima de mejor manera el DAT del componente. La consecuencia directa es que con esto se aproxima el valor actual de DAT para realizar el pronóstico de falla, aún cuando no se ha monitoreado la variable torque desde el inicio cuando se realizó el último recambio de motor. En las Figuras 3.26, 3.27 y 3.28 se muestra la aproximación del DAT real -en línea negra continua- y cada uno de los modelos. Nótese que para el equipo 53, el modelo  $M_{e,z,t}$ , el cual es el que profundiza en mayor detalle las categorías presenta ciclos en que no hubo un modelo que lo aproximase. Esto se debe a que esos ciclos presentaban un solo dato para representar la categoría.

La estimación de daño dada la información operativa del equipo en el pasado acumula su-

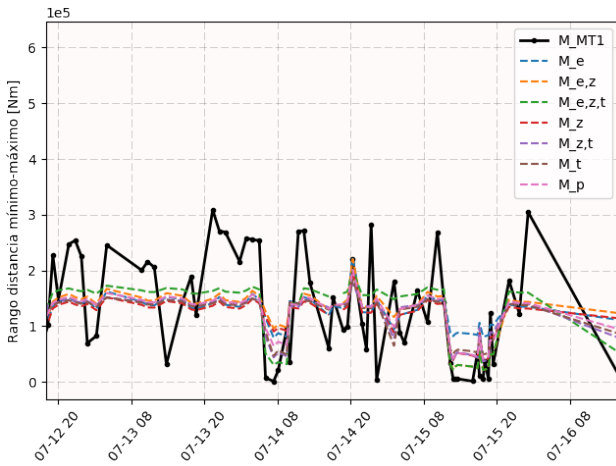


(a) Equipo 53.

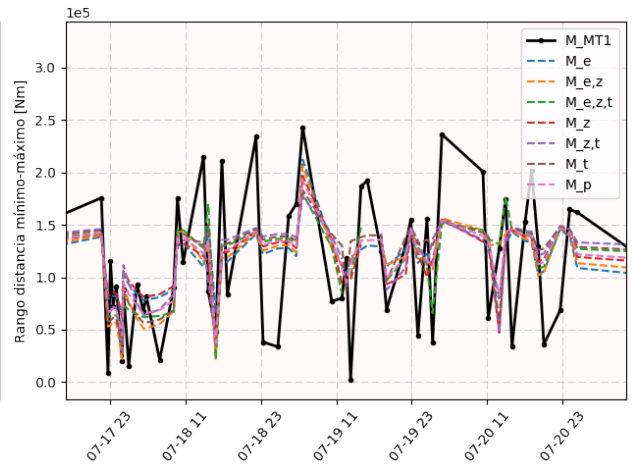


(b) Equipo 63.

Figura 3.23: Aproximación de modelos lineales por categorías del DOPC.



(a) Equipo 53.

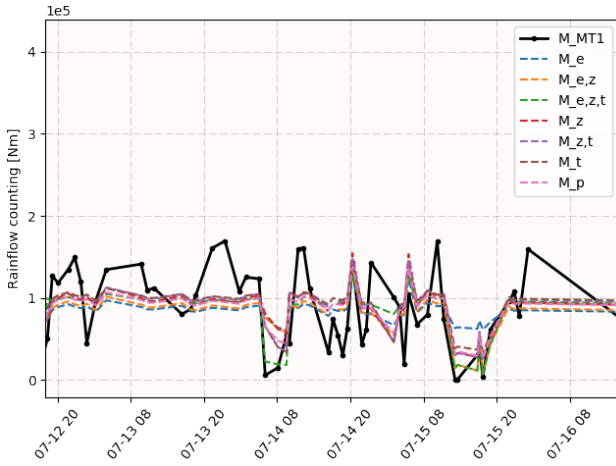


(b) Equipo 63.

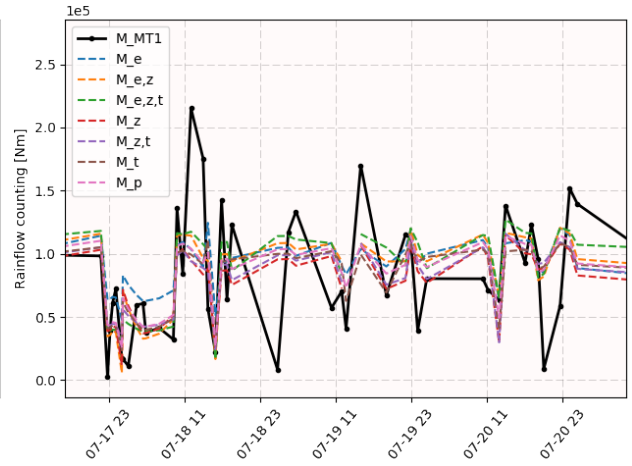
Figura 3.24: Aproximación de modelos lineales por categorías del DOPC.

poniendo daño promedio. Esto porque los modelos obtenidos mediante regresión lineal no consideran los valores peaks de las métricas predefinidas. No obstante aquello, acumulan con la misma pendiente en el tiempo en comparación a los valores medidos. Este resultado es de gran significancia para el pronóstico de falla dados los supuestos de fatiga de material acorde al desgaste cíclico y el modelo de caja negra que relaciona la fatiga con las características calculadas.

En las Tablas 3.8, 3.9 y 3.10 se indica el error cuadrático medio normalizado por el producto del valor promedio de DOPC, multiplicado por la cantidad de ciclos observados para cada una de las métricas y modelos obtenidos, con el propósito de observar las desviaciones por ciclo de los modelos y contrastar entre equipos independientemente de la cantidad de ciclos observados.

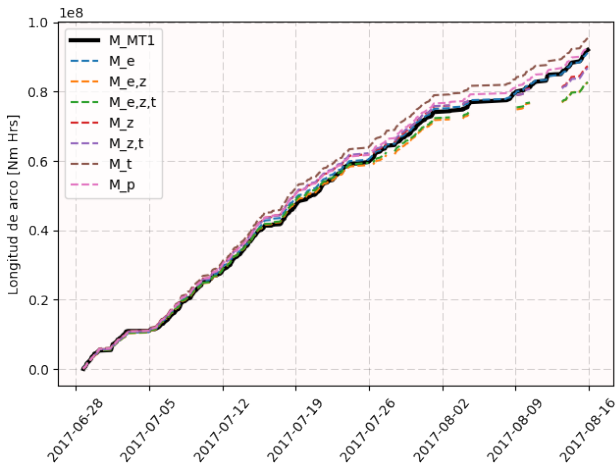


(a) Equipo 53.

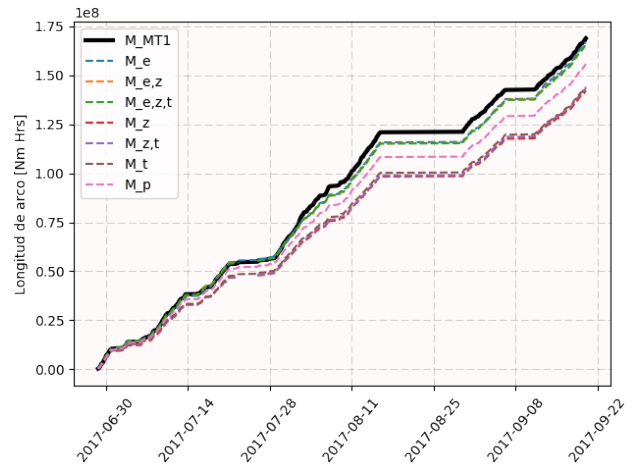


(b) Equipo 63.

Figura 3.25: Aproximación de modelos lineales por categorías del DOPC.



(a) Equipo 53.



(b) Equipo 63.

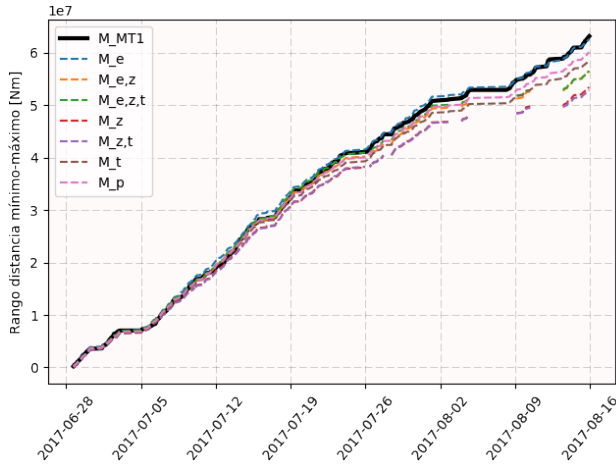
Figura 3.26: Aproximación del DAT de métrica longitud de arco por los modelos lineales por categoría.

### 3.6. Pronóstico de Falla

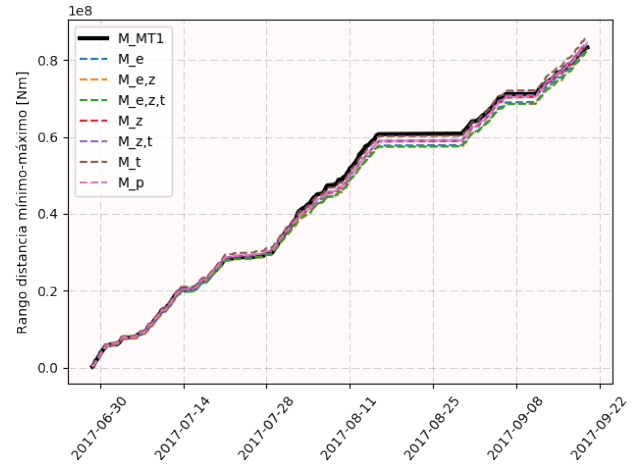
Obtenidas las estimaciones de DAT de los equipos monitoreados, se proyecta a tiempo futuro el daño que recibirá y acumulará el componente con el fin de pronosticar la VUR. La herramienta utilizada en este trabajo es una cadena de Markov.

Las operaciones dentro de la mina en relación al desempeño de los equipos se suponen sistemáticas, dado que los caminos permanecen relativamente constantes en relación a las inclinaciones por zona y/o curvas si se observa para un período de tiempo del orden de un año. Los operadores/as tienen formas de manejo características entre sí, la asignación de equipos y cantidad de estos para ciertas zonas de carga-descarga depende a priori de la disposición de equipos dados por un horario predefinido y la cantidad de mineral que se traslada debiese ser un estándar -que varía entre faenas mineras y no necesariamente ser el nominal del equipo-.



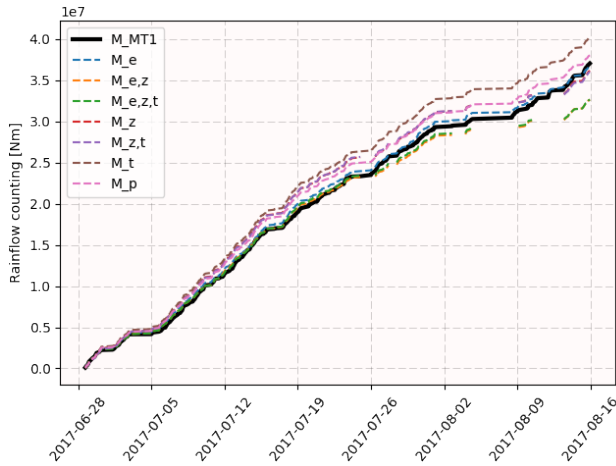


(a) Equipo 53.

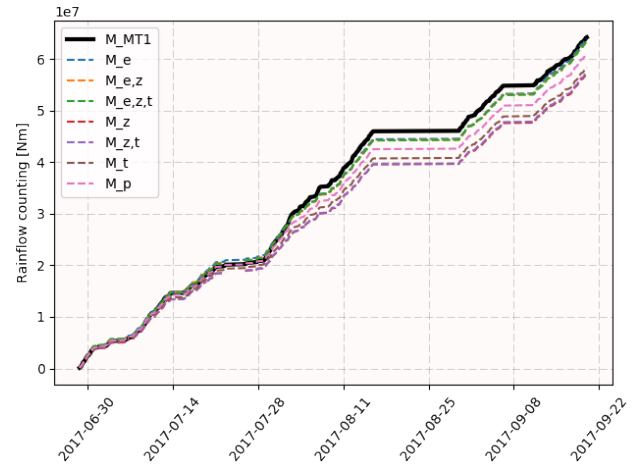


(b) Equipo 63.

Figura 3.27: Aproximación del DAT de métrica longitud de arco por los modelos lineales por categoría.



(a) Equipo 53.



(b) Equipo 63.

Figura 3.28: Aproximación del DAT de métrica longitud de arco por los modelos lineales por categoría.

Este supuesto motiva la idea de proyectar dichos comportamientos observados hacia el futuro, ya que no hay razón aparente para que las condiciones cambien de manera drástica.

Por lo tanto, el pronóstico de falla y la estimación de DAT hacia el futuro del componente, se realiza extrapolando los valores observados bajo la premisa de que el equipo se comporta de manera sistemática a lo largo del tiempo. Para esto, se utiliza el modelo de cadena de Markov identificado para estimar el umbral de falla y, posteriormente utilizar dicho valor de umbral para compararlo con el nivel de DAT de cada realización en el tiempo y de esta forma definir la fecha de falla más probable [10]. Ya conocidas las probabilidades de transición de estados para cada equipo, se simulan realizaciones para la evolución del daño. Luego, como ya ha sido estimado el umbral de falla para cada métrica, la zona de peligro se estima como un intervalo que contiene el valor el valor de UDF dado el DAT del instante actual.

Tabla 3.8: Error cuadrático medio normalizado entre el DAT y la aproximación de modelos lineales por categoría para métrica distancia mínimo-máximo.

Equipo	$M_p$	$M_t$	$M_e$	$M_{e,z}$	$M_{e,z,t}$	$M_z$	$M_{z,t}$
53	0.38	0.75	0.45	0.68	0.66	1.41	1.46
55	0.77	0.67	1.21	0.65	0.65	0.81	0.74
57	1.75	3.05	0.62	0.45	0.41	2.79	2.96
59	0.32	0.62	0.17	0.87	0.87	1.22	1.16
60	0.53	0.98	0.17	0.11	0.13	0.97	0.94
61	0.11	0.11	0.16	0.14	0.10	0.11	0.12
63	0.42	0.42	0.76	0.87	0.85	0.48	0.41
64	0.41	0.88	0.14	0.12	0.13	0.91	0.81
65	0.90	2.20	0.52	0.54	0.50	2.15	2.21
66	0.17	0.30	0.09	0.10	0.11	0.36	0.37
67	1.26	2.58	0.16	0.14	0.17	2.46	2.32
69	0.33	0.66	0.07	0.07	0.06	0.65	0.66
70	1.09	2.19	0.15	0.11	0.11	2.18	2.09
71	0.10	0.12	0.05	0.06	0.07	0.16	0.13
72	0.08	0.11	0.11	0.08	0.07	0.08	0.06
73	0.17	0.37	0.12	0.11	0.11	0.40	0.36
76	0.27	0.47	0.22	0.16	0.15	0.21	0.23
77	0.10	0.18	0.09	0.09	0.09	0.18	0.19

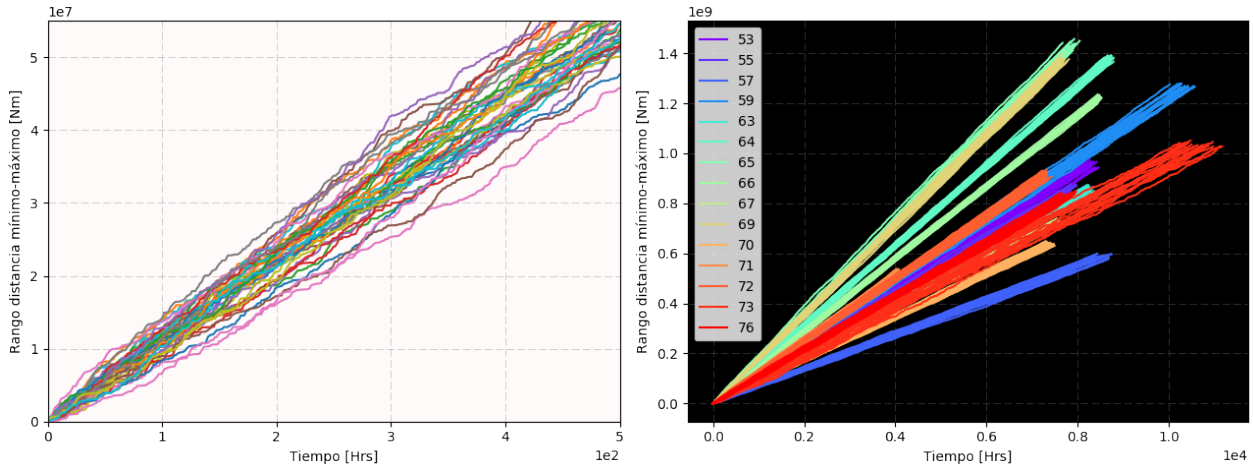
Tabla 3.9: Error cuadrático medio normalizado entre el DAT y la aproximación de modelos lineales por categoría para métrica rainflow counting.

Equipo	$M_p$	$M_t$	$M_e$	$M_{e,z}$	$M_{e,z,t}$	$M_z$	$M_{z,t}$
53	0.91	1.61	0.42	0.66	0.65	1.00	1.00
55	0.29	0.21	0.50	0.27	0.24	0.30	0.26
57	1.42	2.31	0.59	0.53	0.55	2.18	2.24
59	0.21	0.42	0.15	0.78	0.78	1.05	0.99
60	0.18	0.38	0.08	0.08	0.08	0.36	0.31
61	0.51	1.06	0.63	0.25	0.08	0.95	0.96
63	1.36	2.22	0.54	0.59	0.60	2.77	2.70
64	0.14	0.25	0.08	0.05	0.05	0.35	0.24
65	0.45	1.03	0.46	0.45	0.46	1.05	1.05
66	0.07	0.13	0.06	0.05	0.04	0.16	0.18
67	0.70	1.44	0.14	0.12	0.14	1.37	1.28
69	0.09	0.12	0.09	0.07	0.07	0.09	0.10
70	0.76	1.59	0.20	0.10	0.10	1.54	1.51
71	0.11	0.16	0.02	0.01	0.01	0.22	0.15
72	0.12	0.26	0.08	0.08	0.07	0.24	0.22
73	0.04	0.03	0.04	0.07	0.06	0.04	0.04
76	0.18	0.42	0.32	0.20	0.18	0.63	0.55
77	0.16	0.29	0.04	0.04	0.04	0.32	0.31

Tabla 3.10: Error cuadrático medio normalizado entre el DAT y la aproximación de modelos lineales por categoría para métrica longitud de arco.

Equipo	$M_p$	$M_t$	$M_e$	$M_{e,z}$	$M_{e,z,t}$	$M_z$	$M_{z,t}$
53	0.63	1.04	0.36	0.64	0.62	0.61	0.63
55	0.84	1.09	0.83	0.46	0.44	1.14	1.07
57	1.72	2.88	0.71	0.64	0.61	2.63	2.78
59	0.12	0.18	0.17	0.69	0.69	0.69	0.63
60	0.16	0.25	0.09	0.07	0.08	0.29	0.25
61	0.09	0.21	0.10	0.12	0.08	0.17	0.17
63	2.29	4.18	0.63	0.73	0.74	4.59	4.42
64	0.10	0.15	0.07	0.04	0.05	0.24	0.18
65	0.28	0.43	0.42	0.44	0.44	0.42	0.44
66	0.05	0.06	0.06	0.06	0.07	0.08	0.09
67	0.45	0.75	0.16	0.18	0.18	0.90	0.69
69	0.06	0.08	0.05	0.04	0.05	0.07	0.10
70	0.47	1.12	0.48	0.33	0.31	1.16	1.05
71	0.14	0.24	0.02	0.02	0.02	0.26	0.21
72	0.09	0.19	0.08	0.06	0.05	0.14	0.15
73	0.09	0.18	0.05	0.06	0.05	0.13	0.15
76	0.58	1.29	0.41	0.32	0.32	1.44	1.40
77	0.14	0.28	0.03	0.02	0.03	0.30	0.27

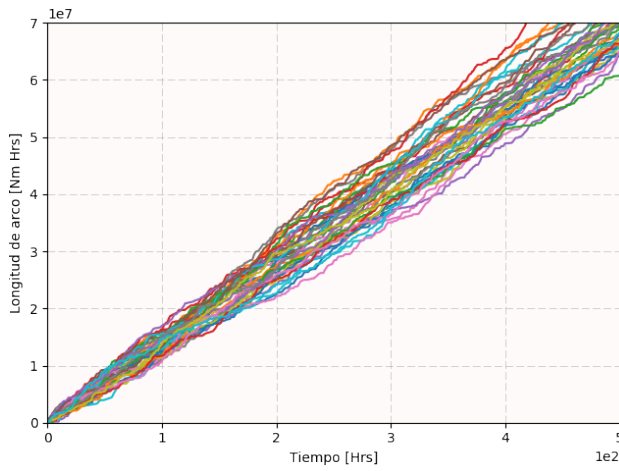
Luego, en las Figuras 3.29, 3.30 y 3.31 se puede observar 20 realizaciones para cada uno de los equipos monitoreados. Nótese la diferencia en la tasa de acumulación de daño. Esto refleja los distintos comportamientos dadas las operaciones a las cuales son expuestos los equipos entre sí.



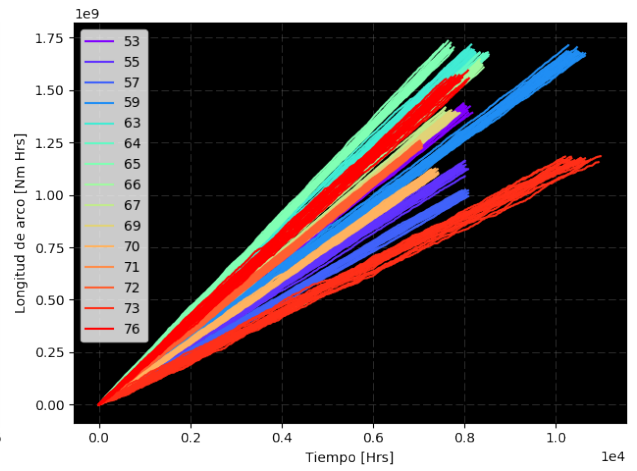
(a) 20 realizaciones para un equipo.

(b) Realizaciones para cada equipo.

Figura 3.29: Simulación de comportamiento para el crecimiento de DAT por distancia mínimo-máximo.

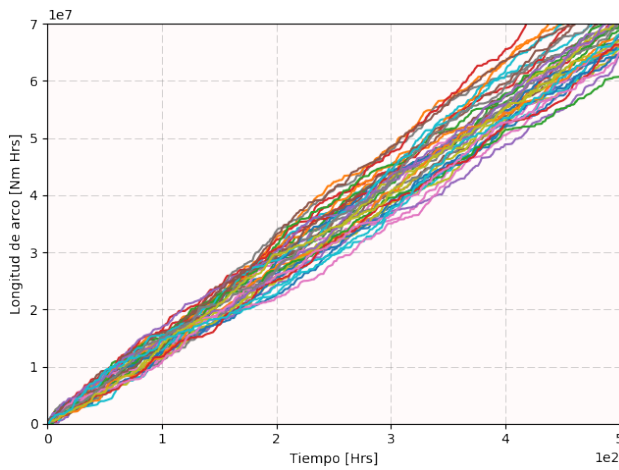


(a) 20 realizaciones para un equipo.

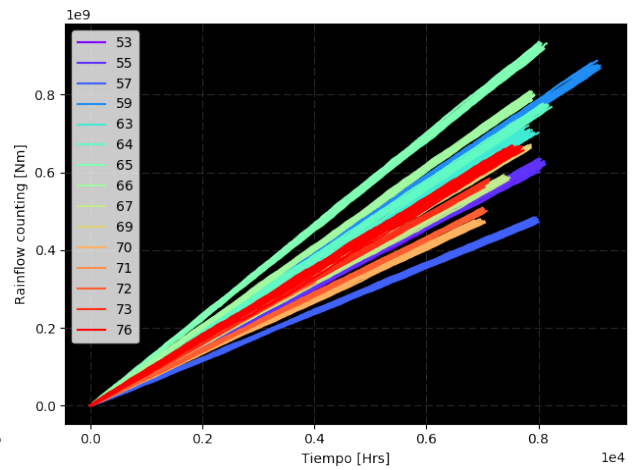


(b) Realizaciones para cada equipo.

Figura 3.30: Simulación de comportamiento para el crecimiento de DAT por longitud de arco.



(a) 20 realizaciones para un equipo.



(b) Realizaciones para cada equipo.

Figura 3.31: Simulación de comportamiento para el crecimiento de DAT por rainflow counting.

# Capítulo 4

## Resultados

En el siguiente capítulo se entrega los resultados de pronóstico de falla realizados sobre los equipos. Los casos expuestos corresponden a los que fue posible observar una realización de cambio de motor dentro de la ventana observable de historia recopilada con los datos de eventos. De esta forma se es consistente con la condición de daño inicial igual a cero. En la Tabla 4.1 se detalla dichas fechas.

Los UDF calculados a partir de las simulaciones de cadena de markov e historial de fallas se presentan en la Tabla 4.2.

Tabla 4.1: Fecha del último recambio de motor para la ventana de tiempo observada con los datos de eventos recolectados.

Equipo	Motor	Fecha recambio
55	MT2	2017-02-22
57	MT1	2017-12-12
65	MT2	2016-11-17
71	MT1	2016-12-25
72	MT2	2016-12-04
77	MT2	2017-07-24

Tabla 4.2: Valores estimados de umbral de falla.

Métrica	MT1	MT2	Promedio
Rango distancia mínimo-máximo [Nm]	2.43e+08	8.30e+08	5.37e+08
Rainflow counting [Nm]	1.64e+08	5.37e+08	3.51e+08
Longitud de arco [Nm Hrs]	3.92e+08	1.14e+09	7.68e+08

En la Tabla 3.1 se detalla la última fecha en que se obtuvo datos por cada equipo. Posterior a esa fecha, se realiza la extensión y pronóstico de comportamiento futuro en relación al DAT que el equipo obtendría. De los casos expuestos en la Tabla 4.1 sólo se obtuvo falla del equipo 55, caso que es especificado en la Tabla 4.3. En las Figuras 4.1, 4.2 y 4.3 se esquematiza las proyecciones realizadas para las métricas longitud de arco, distancia mínimo-máximo y

rainflow counting. Para cada una de estas figuras, se tiene en el gráfico de la izquierda en línea azul, la ubicación de fecha de falla que se especifica en la Tabla 4.3. Dentro de las mismas, a la izquierda en línea vertical roja punteada está la fecha de recambio del motor. Por lo tanto, La estimación futura de falla se realiza con el UDF calculado para cada métrica a partir de dicha fecha como condición inicial de daño igual a cero -línea horizontal roja inferior-. De ahí, el UDF está dado por la línea horizontal roja superior. Por lo tanto, el DAT para cada caso se representa con la curva creciente a la izquierda de la línea negra vertical punteada, que es la que indica la última fecha de obtención de datos. La proyección desde esa fecha se realiza con 20 realizaciones del modelo de cadena de Markov, las cuales acumulan DOPC al DAT hasta que eventualmente alcanzan el UDF. La banda roja representa la zona de peligro calculado como el intervalo desde el 90 % hasta el 110 % del UDF. Finalmente, cada realización de la cadena de Markov supera el UDF en una fecha que se agrega al histograma de la derecha de cada figura. Dado que el rango de las 20 realizaciones no supera una semana en rango de tiempo, se decide obtener la fecha más representativa como el promedio de todas las fechas obtenidas por realización. Estas fechas de predicción de falla por cada una de las métricas se especifican en la Tabla 4.4.

Tabla 4.3: Registro de falla posterior a la fecha de fin de obtención de datos. Este caso se utiliza como prueba de las proyecciones y comparativa entre las diferentes métricas planteadas.

Equipo	Fecha de Falla
55	2017-10-20

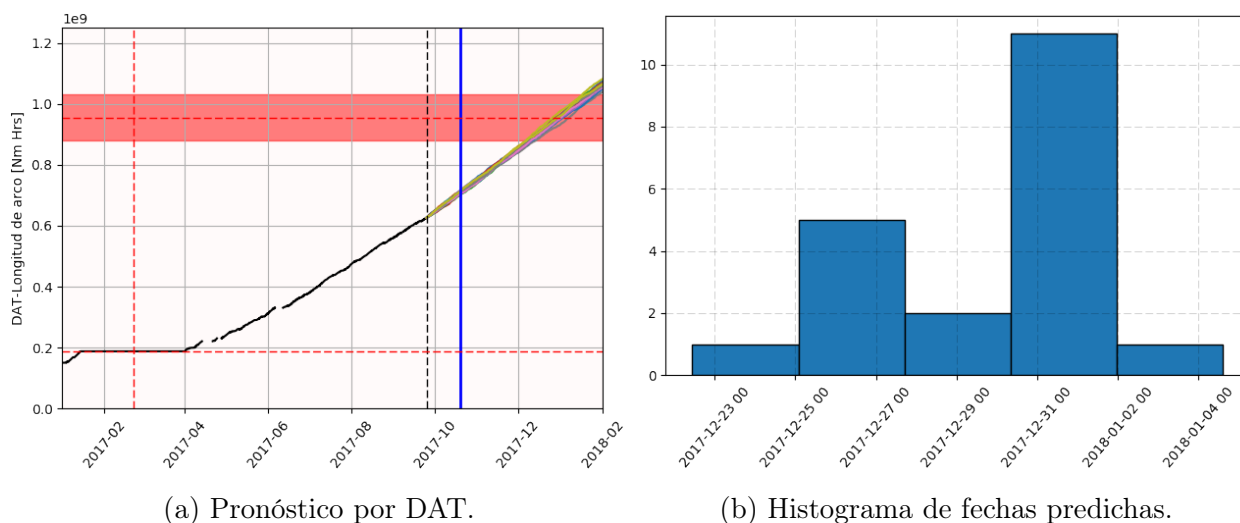
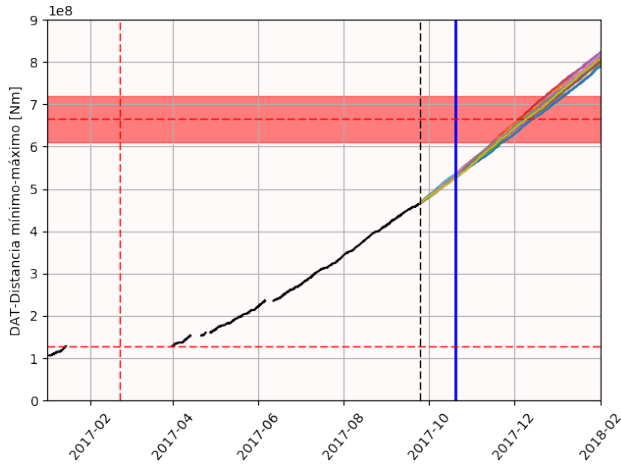


Figura 4.1: Pronóstico de falla utilizando métrica longitud de arco.

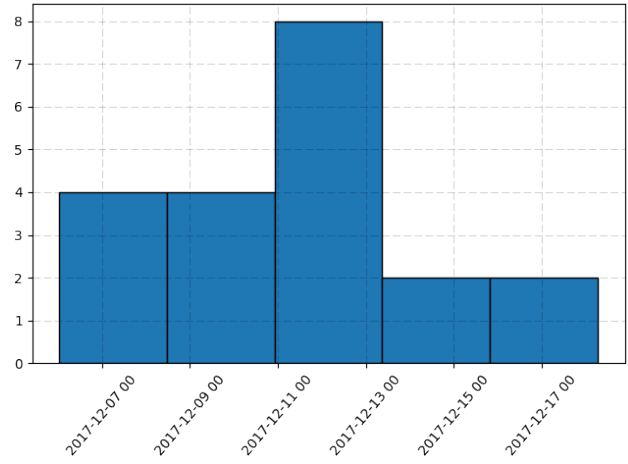
Tabla 4.4: Fechas de falla predichas.

Equipo	Métrica	Predicción de fecha de falla
55	Longitud de arco	2017-12-29
55	Distancia mínimo-máximo	2017-12-11
55	Rainflow counting	2017-11-24

De estos resultados se observa que la métrica más cercana a predecir la fecha real de falla es la de rainflow counting, sin embargo, predice la falla un mes posterior a la fecha verdadera. De

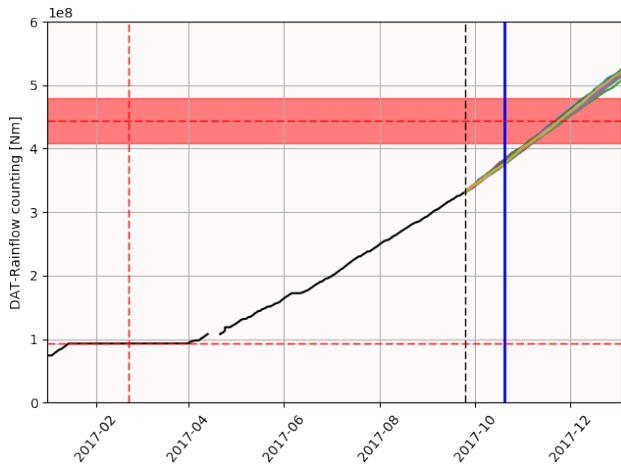


(a) Pronóstico por DAT.

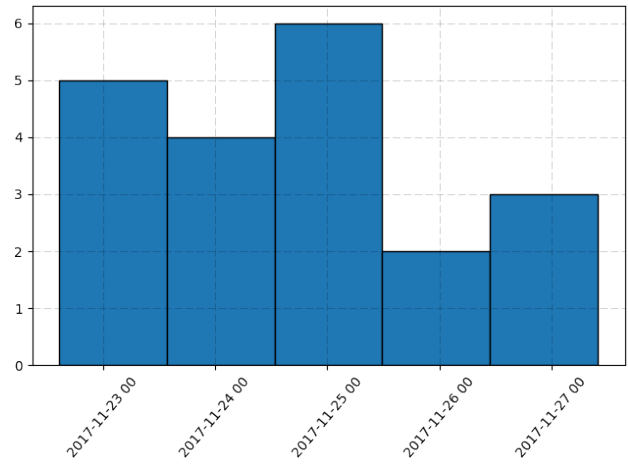


(b) Histograma de fechas predichas.

Figura 4.2: Pronóstico de falla utilizando métrica distancia mínimo-máximo.



(a) Pronóstico por DAT.



(b) Histograma de fechas predichas.

Figura 4.3: Pronóstico de falla utilizando métrica rainflow counting.

estos resultados se propone una métrica obtenida a partir de rainflow counting para mejorar los resultados. Esta corresponde al mismo contador, sólo que ponderando cada ciclo de la señal de torque por el valor medio de variación del mismo ciclo completo, la que se expresa matemáticamente en la Ecuación 4.1. Es decir, ahora los ciclos que realizaron una variación completa de  $y$  [Nm] son más severos si se realizaron en torno a  $x_1$  [Nm] que a  $x_2$  [Nm], si y sólo si  $|x_1| > |x_2|$ .

#### 4.0.1. Mean rainflow counting

Dada la señal de torque, rainflow counting se calcula como:

$$DOPC_n^{mrfc} = \sum_i Count_{i,n} \cdot Range_{i,n} \cdot Mean_{i,n} \quad (4.1)$$

Donde  $Count_{i,n}$  es la cantidad de ciclos completos observados de rango de variación  $Range_{i,n}$  por el algoritmo de rainflow counting en el  $n$ -ésimo ciclo de operación y valor promedio de ciclo  $Mean_{i,n}$ .

A continuación se muestra los resultados de esta métrica motivada por la mejora a los resultados de rainflow counting. En la Figura 4.4 se muestra los histogramas de los valores de métrica considerando todos los equipos monitoreados. Nótese que aumenta el orden de magnitud, debido a que se multiplica por el valor de torque nuevamente como se ve en la Ecuación 4.1.

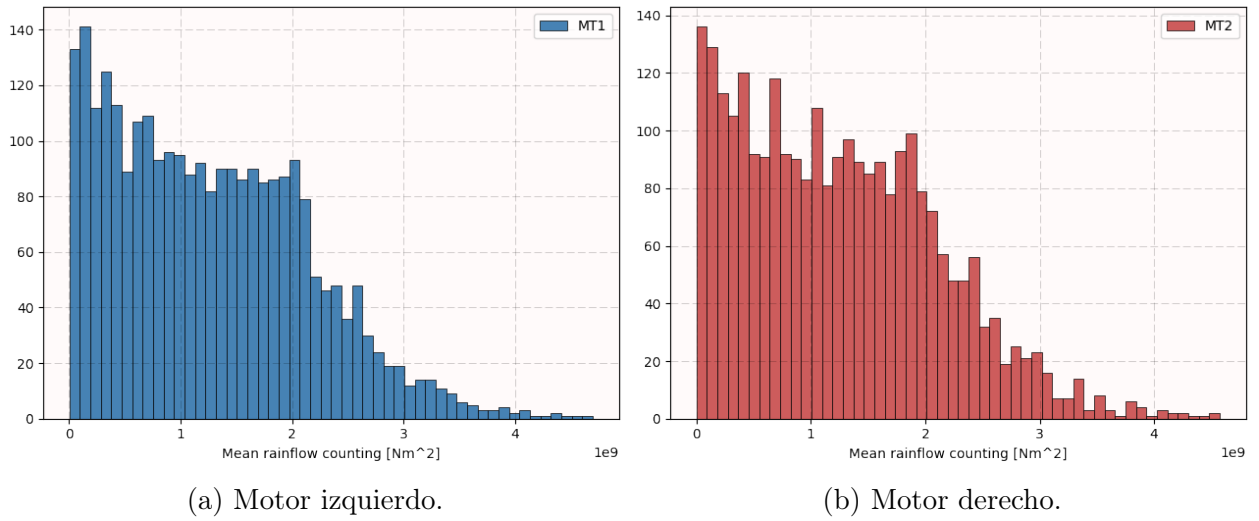


Figura 4.4: Histograma de  $DOPC_{mrfc}$  por motor de tracción.

En las Tablas 4.5 y 4.6 se indica los valores promedio y mediana de la característica. Los valores promedios se utilizan posteriormente para normalizar el error cuadrático medio entre el  $DAT_{mrfc}$  medido en contraste a los modelos que estiman los valores de  $DOPC_{mrfc}$ .

Tabla 4.5: Valor promedio de mean rainflow counting por motor de todos los equipos.

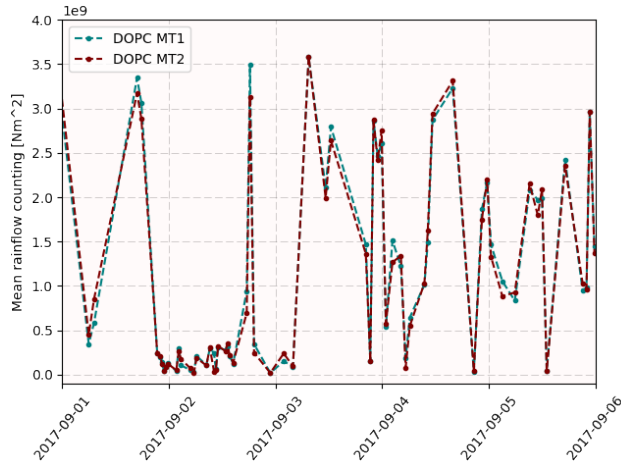
	MT1 Promedio	MT2 Promedio
Mean rainflow counting $[Nm]^2$	1.26e+9	1.25e+9

Tabla 4.6: Mediana de mean rainflow counting por motor de todos los equipos.

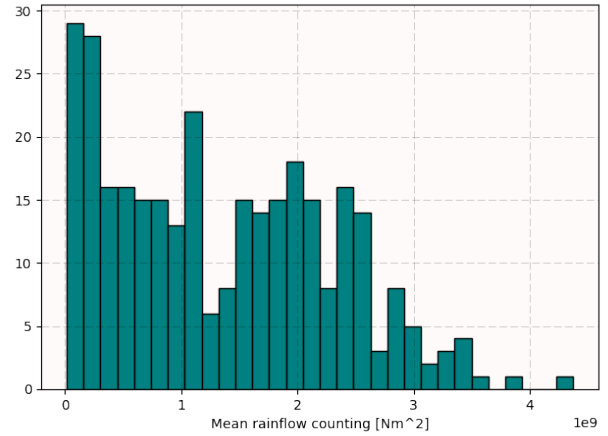
	MT1 Mediana	MT2 Mediana
Mean rainflow counting $[Nm]^2$	1.16e+9	1.17e+9

La Figura 4.5 esquematiza la secuencia temporal de un equipo junto al histograma de valores de  $DOPC_{mrfc}$  obtenidos para un mismo equipo. Se observa que los estados presentan mayor distancia en comparación a las tres métricas calculadas anteriormente, lo cual facilita la separación e identificación de estados.





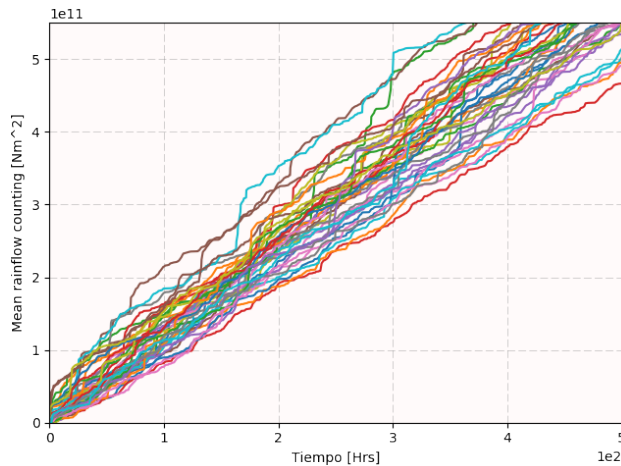
(a) Secuencia temporal de DOPC.



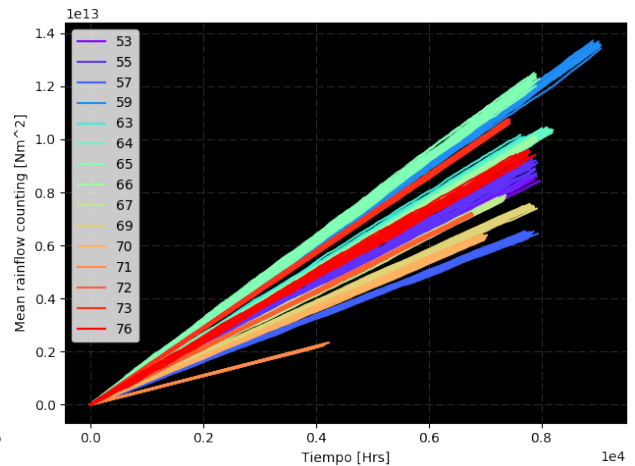
(b) Histograma de DOPC.

Figura 4.5: Resumen mean rainflow counting para la aplicación de equipo 76.

Los valores promedio por estado considerando cada uno de los equipos son los que se muestran en la Tabla 4.7. Nótese que hay algunos casos en que el orden de magnitud de los estados más altos de daño supera al resto. No se consideran datos anómalos debido a que los cálculos se obtienen después de la etapa de preprocesamiento y exclusión de datos anómalos, los que filtran y dejan los valores medidos de torque que son físicamente posibles. En los casos de daño muy alto, el modelo de cadena de markov identifica la transición a aquellos estados muy poco probables si es que son una pequeña muestra de la población. Esto significa que la mayor parte del tiempo las realizaciones caen en los estados 1 y 2 la mayor parte del tiempo.



(a) 20 realizaciones para un equipo.



(b) Realizaciones para cada equipo.

Figura 4.6: Simulación de comportamiento para el crecimiento de DAT por mean rainflow counting.

La Tabla 4.8 muestra por equipo los errores cuadrático medio entre el  $DAT_{mrfc}$  medido y los modelos obtenidos por segregación de categorías. En el diagnóstico de falla para cada equipo utiliza el modelo que aproxima con menor error cuadrático medio las observaciones de  $DAT_{mrfc}$ .

Tabla 4.7: Promedio de  $DOPC_{mrfc}$  en cada uno de los estados determinados mediante K-means. Valores en  $[Nm]^2$ .

Equipo	Estado 1	Estado 2	Estado 3
53	4.24e+8	1.31e+9	2.16e+9
55	5.58e+8	2.37e+9	1.02e+10
57	3.10e+8	1.10e+9	2.16e+9
59	7.68e+8	2.16e+9	3.28e+9
60	5.81e+8	2.08e+9	3.61e+9
61	1.32e+9	1.30e+10	2.04e+10
63	4.57e+8	1.53e+9	2.48e+9
64	5.64e+8	1.75e+9	2.97e+9
65	8.51e+8	2.15e+9	3.41e+9
66	5.41e+8	1.87e+9	2.99e+9
67	4.08e+8	1.18e+9	1.92e+9
69	1.53e+8	8.95e+8	2.14e+9
70	3.05e+8	1.02e+9	1.75e+9
71	1.05e+8	4.62e+8	8.62e+8
72	4.39e+8	1.13e+9	2.02e+9
73	5.20e+8	1.70e+9	2.56e+9
76	4.88e+8	1.70e+9	2.73e+9
77	2.89e+8	8.14e+8	1.66e+9

Tabla 4.8: Error cuadrático medio normalizado entre el DAT y la aproximación de modelos lineales por categoría para métrica mean rainflow counting.

Equipo	$M_p$	$M_t$	$M_e$	$M_{e,z}$	$M_{e,z,t}$	$M_z$	$M_{z,t}$
53	1.42	2.43	0.64	0.63	0.64	1.75	1.81
55	0.54	0.48	0.63	0.35	0.34	0.77	0.67
57	1.68	2.70	0.67	0.62	0.62	2.58	2.67
59	0.22	0.48	0.24	0.77	0.77	1.04	1.00
60	0.14	0.44	0.08	0.10	0.09	0.31	0.28
61	4.01	8.15	1.57	1.28	1.19	6.63	6.49
63	1.11	1.44	0.62	0.48	0.47	2.70	2.60
64	0.07	0.07	0.07	0.08	0.07	0.15	0.20
65	0.34	0.69	0.50	0.53	0.48	0.64	0.68
66	0.06	0.07	0.08	0.15	0.15	0.08	0.07
67	1.31	2.77	0.35	0.23	0.24	2.25	2.17
69	0.08	0.15	0.13	0.06	0.08	0.13	0.14
70	1.27	2.67	0.14	0.15	0.17	2.37	2.29
71	0.24	0.32	0.01	0.01	0.01	0.60	0.49
72	0.20	0.30	0.07	0.09	0.09	0.46	0.48
73	0.12	0.18	0.07	0.07	0.07	0.24	0.24
76	0.41	0.52	0.47	0.51	0.52	0.43	0.45
77	0.25	0.42	0.08	0.08	0.06	0.46	0.45

Tal y como se calculó en las métricas originales, se estima el UDF como el promedio del valor al cual todas las realizaciones llegaron en tal cantidad de horas de funcionamiento antes de fallar y se utiliza el promedio de los dos motores como umbral de falla. En la Figura 4.7 se muestra las simulaciones de los cuatro casos de falla con los que se obtuvo el umbral de falla.

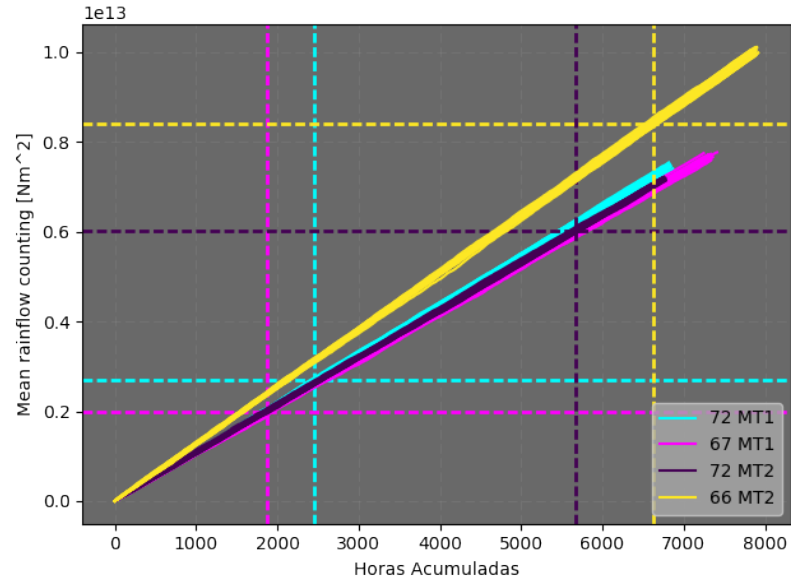
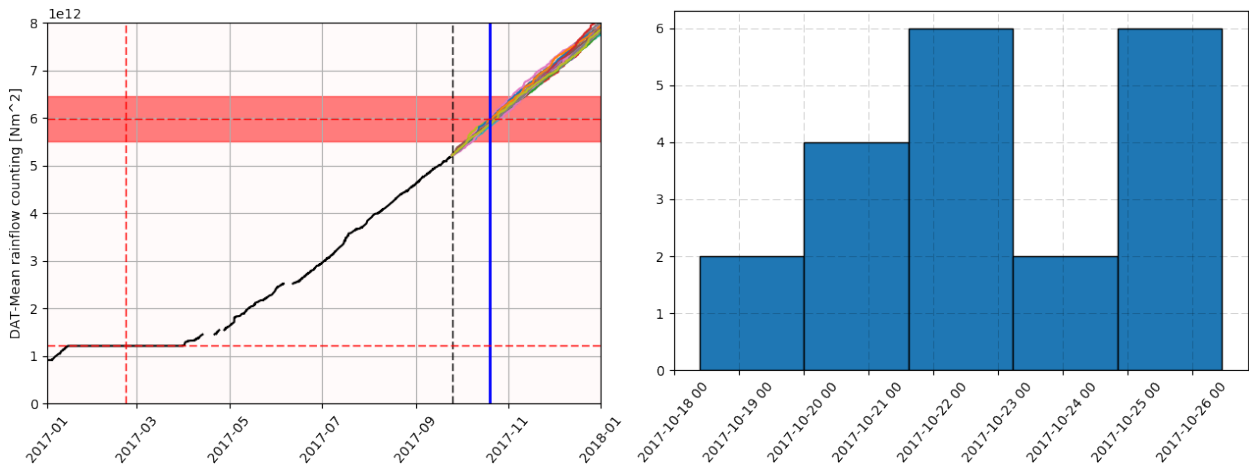


Figura 4.7: Estimación de umbral de falla para métrica mean rainflow counting.

Tabla 4.9: Valores de umbral de falla estimados usando mean rainflow counting.

Métrica	MT1	MT2	Promedio
Mean rainflow counting [ $Nm^2$ ]	$2.33e+12$	$7.19e+12$	$4.77e+12$

Finalmente el resultado de pronóstico indica una fecha predictiva de falla el 22 de octubre, la cual difiere en 2 días posteriores de la falla real.



(a) Pronóstico por DAT.

(b) Histogramas de fechas predichas.

Figura 4.8: Pronóstico de falla utilizando mean rainflow counting.

Tabla 4.10: fecha de falla predicha mediante mean rainflow counting.

Equipo	Métrica	Predicción de fecha de Falla
55	Mean rainflow counting	2017-10-22

En las Figuras 4.10 - 4.13 se muestra la predicción de evolución de DAT para el resto de los equipos de la Tabla 4.1. Para estos casos no existe falla hasta la fecha, por lo que caen en la categoría de falsos positivos. Cabe destacar que no se tiene certeza de la condición operacional desde la última fecha de medición de cada equipo, por lo que las curvas pueden trasladarse en el tiempo dependiendo si efectivamente operaron; esta metodología asume jornadas de trabajo continuas para cada equipo, por lo que para el módulo online se debe recalculan las predicciones a medida que se obtiene datos operacionales reales del equipo y se actualiza el DAT.

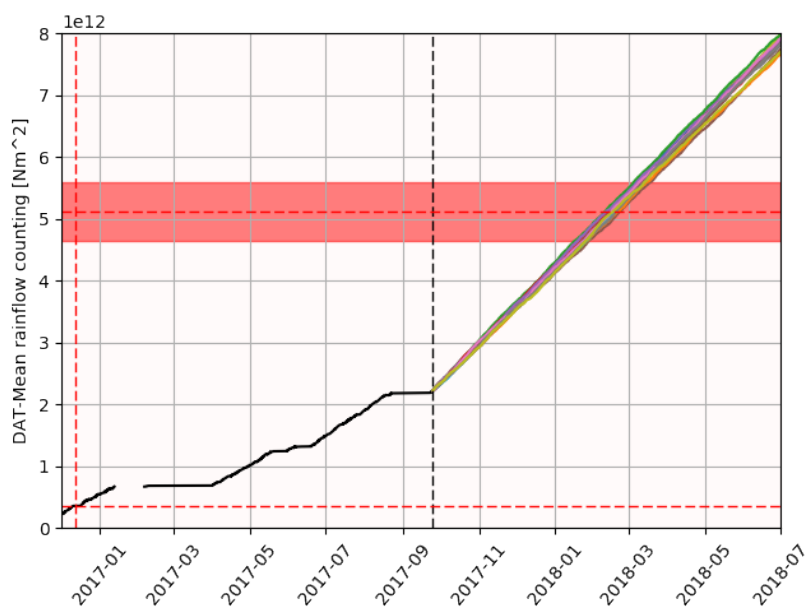


Figura 4.9: MT1 de equipo 57.

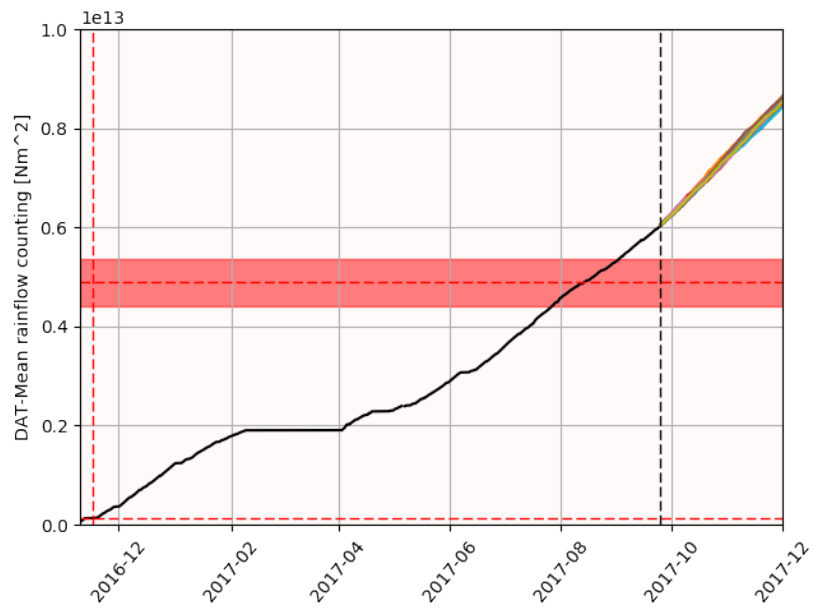


Figura 4.10: MT2 de equipo 65.

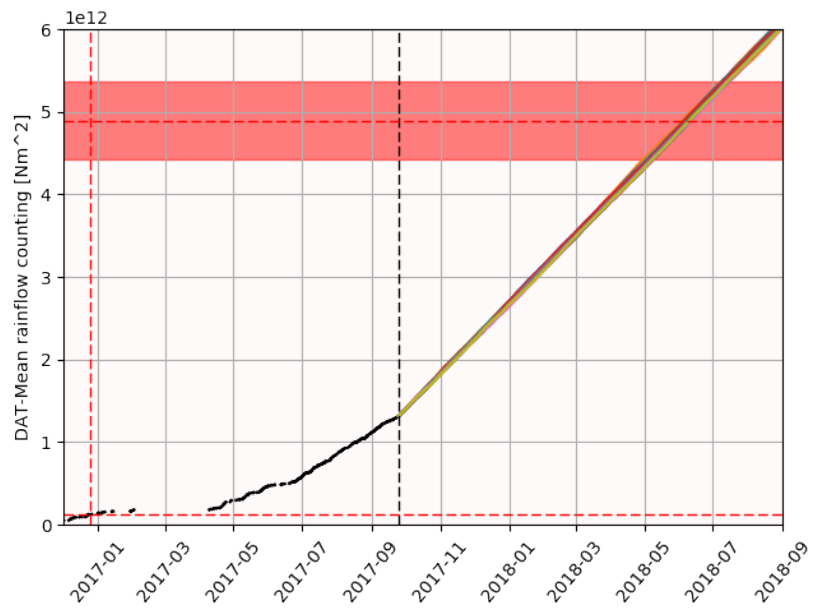


Figura 4.11: MT1 de equipo 71.

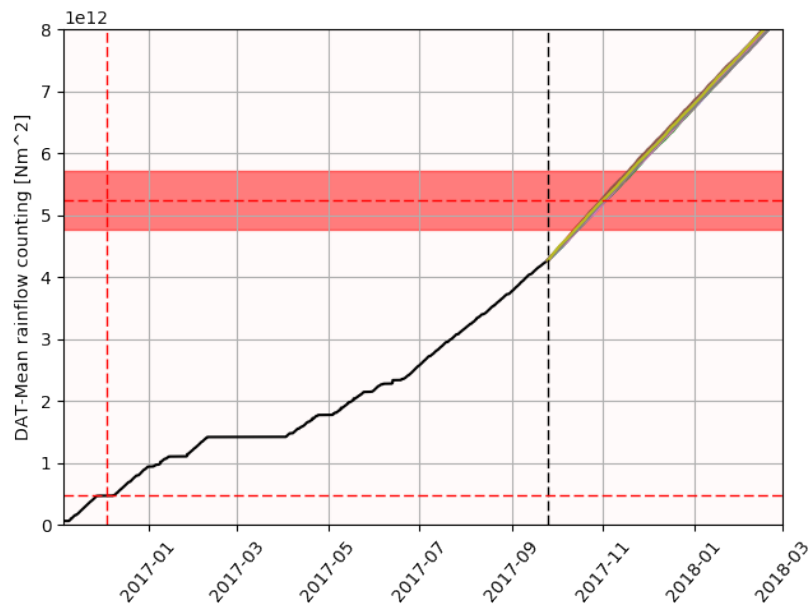


Figura 4.12: MT2 de equipo 72.

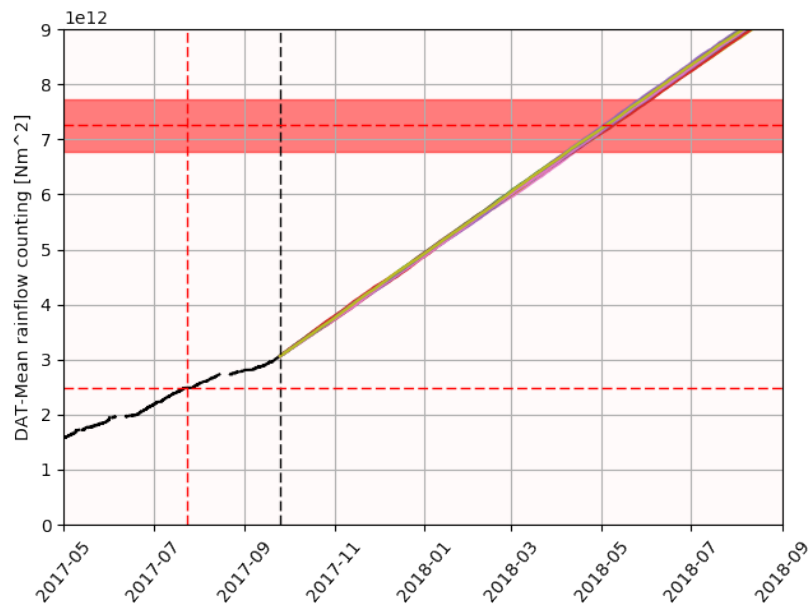


Figura 4.13: MT2 de equipo 77.

# Capítulo 5

## Conclusión

En el contexto de fallas prematuras considerando el tiempo entre fallas nominal del sistema estudiado, se introdujo una metodología que estima la vida útil remanente para el motor de tracción de equipos mineros Komatsu 960E, la cual se vale de datos operacionales obtenidos del equipo para realizar mantenimiento basado en la condición.

Estudios previos a la realización de este trabajo indican que la falla se debe a la fatiga cíclica del componente, el cual es el perno pasante del carrier de baja velocidad del sistema planetario, resaltando dos componentes principales, carga fija y variable sobre el componente. De ahí, este estudio se enfocó en la tensión variable que ocasionaba que el componente entrase en la zona de peligro e inminente falla para, posteriormente incluir un factor de carga fija utilizando mean rainflow counting y mejorar los resultados. La variable monitoreada torque de motor, se supuso proporcional a dicho esfuerzo alterno, considerando que en cada ciclo de operación del equipo, se acumula daño en el componente.

De vital importancia es la hipótesis que relaciona el torque de motor con la deformación a la que se somete el componente, dado que no es posible instalar sensores que midan dicho valor directamente en el corto plazo, principalmente debido a la severidad de operación. Es decir, se pensó un modelado del tipo caja negra y a partir de de ahí, se consideró el valor de daño acumulado.

Para este fin, se definieron métricas de daño acumulado en el componente a medida de la utilización del equipo, dados por el valor acumulado del rango de variación de  $\tau$ , longitud de arco de la señal de torque y rainflow counting para los ciclos definidos por la carga y descarga de material. Luego, estas métricas se acumularon a través del tiempo en búsqueda del valor característico de falla común a todos los especímenes por hipótesis.

Dada la información con la que se contaba, se realizó estimaciones que completaron los datos faltantes utilizando supuestos de comportamientos sistemáticos en operaciones, tomando en cuenta factores tales como zona geográfica, tiempo empleado y equipo en cuestión.

La comparación de las 4 métricas usadas indica que el crecimiento de daño relativo entre equipos no es el mismo entre cada caso, como se observa en las Figuras 3.14, 3.15, 3.16

y 4.7. Para visualizar esta idea, nótese que como los equipos corresponden al mismo color entre gráficos, la posición de las realizaciones y crecimiento del mismo color varían entre métricas. Este resultado es llamativo considerando que el daño real sobre los componentes en un instante de tiempo sólo puede ser una combinación posible, a diferencia de lo que muestran las estimaciones de daño propuestas entre sí. De este resultado se concluye que el historial de fallas y el caso de prueba del equipo 55 es trascendental para determinar la métrica que mejor describe el fenómeno y valida la hipótesis planteada de fractura al alcanzar un nivel de daño característico, UDF. La importancia de la última característica calculada, correspondiente a mean rainflow counting mejora la solución para el pronóstico de fallas en el equipo 55 ya que considera el valor en torno al cual los ciclos completos se produjeron. Este resultado es coherente con la intuición y antecedentes ya que al tener un valor de torque tan alto, se genera una mayor deformación del carrier de velocidad. Al realizarse una variación de aquel valor, esta deformación correspondería a una de mayor magnitud.

Para los casos que se asumen falsos positivos, que corresponde a los equipos 57, 65, 71, 72 y 77, hay 3 principales motivos para los que se piensa no se pronosticó de manera correcta. El posible efecto de cierre de grietas, provocando que la fractura se postergue en el tiempo e incluso, que nunca ocurra. Además, hay una variable que se asumió constante pero eventualmente varía y corresponde al “pre torque de instalación”. Este valor justamente es el valor con que se fija el perno en el carrier del sistema. Luego para valores altos de pre torque, el componente es más propenso a fallar antes dado un perfil de carga variable que lo afecte. El segundo concepto se relaciona con los filtros de información debido a la saturación del sistema que gestiona las sesiones de monitoreo de variables; como se mencionó, la condición para almacenar los datos consideraba cuando el torque de motor excedía en valor absoluto los 16000  $[Nm]$ , lo cual en la práctica significa equipo cargado. Puede darse el caso en que cuando el equipo se traslada descargado, las variaciones de torque en torno a cero influyan en el calculo del UDF. Por otro lado, no se pudo seguir monitoreando la operación de los equipos luego de la fecha indicada en línea punteada vertical negra de las Figuras 4.9, 4.10, 4.11, 4.12 y 4.13, por lo que se desconoce si realmente son falsos positivos. Finalmente, los modelos de estimación de DAT aproximaban el daño acumulado, sin embargo, la medición directa de torque de motor es la que finalmente indica el verdadero desgaste en el componente acorde a las definiciones de esta tesis. Como ya se indicó, el monitoreo de variables en ningún caso se realizó desde el recambio de algún componente, por lo que no se tuvo la medición real de DAT para los casos estudiados.

En consideración a metodologías desarrolladas que difieren de la aquí planteada para pronosticar la falla del componente se destaca la medición de nivel de concentración de material particulado en el aceite de motor. Este método presenta la problemática de que las mediciones se hacen con un período de muestreo muy extenso, del orden de las 200 horas de operación, lo cual al realizarse la fractura súbita no se detectaba de manera oportuna. Como trabajo futuro, la aplicación conjunta del método de monitoreo de aceite con el de acumulación de daño se proponen para disminuir el período de muestreo de niveles de aceite.

El uso de herramientas de agrupación en este trabajo, DBSCAN y K-means, son trascendentales para el procesamiento de datos y obtención de métricas. En primer lugar porque facilitan el procesamiento automático de datos, más aún, tomando en cuenta que se avanza hacia una mejor obtención de grandes cantidades desde los equipos, integrándolos en lagos



de datos que almacenan un vasto historial de operación. Y en segundo lugar porque integran el conocimiento de la operación mediante los parámetros que ocupan para agrupar. Para el caso de DBSCAN,  $\varepsilon$  se define con los tiempos aproximados que demoran los equipos en realizar los ciclos de operación y K-means utiliza la cantidad predefinida de zonas geográficas de carga-descarga, la cantidad de estados en que se estima separar el nivel de daño entre ciclos de operación o la cantidad de grupos en que se deseaba separar los ciclos según tiempo de realización. Todo esto tras estudiar el comportamiento operacional de los equipos.

Finalmente el trabajo realizado para determinar variables operacionales como carga transportada, tiempos de ciclo y zonas de carga-descarga para sistematizar y relacionar el daño obtenido por ciclo es determinante para futuros trabajos, con el fin de caracterizar el desgaste de componentes en el equipo de extracción de material. El caso ideal es extender la agrupación de zonas de carga-descarga con rutas de GPS continuas. De esta forma no existe incertidumbre en la trayectoria realizada por el equipo a diferencia de sólo conocer los puntos de inicio y destino. En consecuencia, el modelamiento de perfil de uso entre rutas tendería a evidenciar sólo diferencias en la acción de operadores, bajo el supuesto de mismas distancias recorridas, pendientes recorridas y tiempos utilizados para completar los ciclos de carga-descarga.

# Capítulo 6

## Bibliografía

- [1] C Amzallag, JP Gerey, JL Robert, and J Bahuaud. Standardization of the rainflow counting method for fatigue analysis. *International journal of fatigue*, 16(4):287–293, 1994.
- [2] David Arthur and Sergei Vassilvitskii. k-means++: The advantages of careful seeding. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 1027–1035. Society for Industrial and Applied Mathematics, 2007.
- [3] Leo H Chiang, Evan L Russell, and Richard D Braatz. *Fault detection and diagnosis in industrial systems*. Springer Science & Business Media, 2000.
- [4] Wai-Ki Ching and Michael K Ng. Markov chains. *Models, algorithms and applications*, 2006.
- [5] Stephen D Downing and DF Socie. Simple rainflow counting algorithms. *International journal of fatigue*, 4(1):31–40, 1982.
- [6] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, pages 226–231, 1996.
- [7] NE Frost, KJ Marsh, and LP Pook. *Metal fatigue*, clarendon, 1974.
- [8] G Glinka and JCP Kam. Rainflow counting algorithm for very long stress histories. *International Journal of Fatigue*, 9(4):223–228, 1987.
- [9] Blaise Kevin Guepie and Stephane Lecoecue. Similarity-based residual useful life prediction for partially unknown cycle varying degradation. In *Prognostics and Health Management (PHM), 2015 IEEE Conference on*, pages 1–7. IEEE, 2015.
- [10] Ya-Wei Hu, Hong-Chao Zhang, Shu-Jie Liu, and Hui-Tian Lu. Sequential monte carlo method toward online rul assessment with applications. *Chinese Journal of Mechanical Engineering*, 31(1):5, 2018.

- [11] Rolf Isermann and Peter Balle. Trends in the application of model-based fault detection and diagnosis of technical processes. *Control engineering practice*, 5(5):709–719, 1997.
- [12] Haël Mughrabi. Microstructural mechanisms of cyclic deformation, fatigue crack initiation and early crack growth. *Phil. Trans. R. Soc. A*, 373(2038):20140132, 2015.
- [13] M Naderi, M Amiri, and MM Khonsari. On the thermodynamic entropy of fatigue fracture. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 466, pages 423–438. The Royal Society, 2010.
- [14] ASTM Committee E-8 on Fatigue and Fracture. *Standard practices for cycle counting in fatigue analysis*. ASTM International, 2005.
- [15] Romano Patrick-Aldaco. *A model based framework for fault diagnosis and prognosis of dynamical systems with an application to helicopter transmissions*. Georgia Institute of Technology, 2007.
- [16] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- [17] Daniel A Pola, Hugo F Navarrete, Marcos E Orchard, Ricardo S Rabié, Matías A Cerda, Benjamín E Olivares, Jorge F Silva, Pablo A Espinoza, and Aramis Pérez. Particle-filtering-based discharge time prognosis for lithium-ion batteries with a statistical characterization of use profiles. *IEEE Transactions on Reliability*, 64(2):710–720, 2015.
- [18] L. Rabiner and B. Juang. An introduction to hidden markov models. *IEEE ASSP Magazine*, 3(1):4–16, Jan 1986.
- [19] Lawrence R Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989.
- [20] Jaap Schijve. *Fatigue of structures and materials*. Springer Science & Business Media, 2001.
- [21] Jaap Schijve. Fatigue of structures and materials in the 20th century and the state of the art. *International Journal of fatigue*, 25(8):679–702, 2003.
- [22] Raghavan Srinivasan and M Ashraf Imam. Role of dispersoids on the fatigue behavior of aluminum alloys: A review. In *Fatigue of Materials III*, pages 11–22. Springer, 2016.
- [23] George J Vachtsevanos, Frank Lewis, Andrew Hess, and Biqing Wu. *Intelligent fault diagnosis and prognosis for engineering systems*. Wiley Online Library, 2006.
- [24] XG Wang, HR Ran, C Jiang, and QH Fang. An energy dissipation-based fatigue crack growth model. *International Journal of Fatigue*, 2018.