



UNIVERSIDAD  
**NACIONAL**  
DE COLOMBIA

# **Validación de modelos predictivos de analítica de datos de análisis de aceite usado para la toma de decisiones de mantenimiento en un turbogenerador de vapor.**

**Juan Pablo Sierra Mejía**

Universidad Nacional de Colombia  
Facultad de Minas, Departamento de Ingeniería Mecánica  
Medellín, Colombia

2021

# **Validación de modelos predictivos de analítica de datos de análisis de aceite usado para la toma de decisiones de mantenimiento en un turbogenerador de vapor.**

**Juan Pablo Sierra Mejía**

Trabajo de investigación presentado como requisito parcial para optar al título de:

**Magister en Ingeniería Mecánica**

Director:

Ph.D. Fernando Jesús Guevara Carazas

Línea de Investigación:

Machine Learning en gestión de mantenimiento

Grupo de Investigación:

GOMAC, Gestión, operación y mantenimiento de activos

Universidad Nacional de Colombia

Facultad de Minas, Departamento de Ingeniería Mecánica

Medellín, Colombia

2021

*A mi familia, por darme las herramientas y el impulso para asumir mis retos.*

# Declaración de obra original

Yo declaro lo siguiente:

He leído el Acuerdo 035 de 2003 del Consejo Académico de la Universidad Nacional. «Reglamento sobre propiedad intelectual» y la Normatividad Nacional relacionada al respeto de los derechos de autor. Esta disertación representa mi trabajo original, excepto donde he reconocido las ideas, las palabras, o materiales de otros autores.

Cuando se han presentado ideas o palabras de otros autores en esta disertación, he realizado su respectivo reconocimiento aplicando correctamente los esquemas de citas y referencias bibliográficas en el estilo requerido.

He obtenido el permiso del autor o editor para incluir cualquier material con derechos de autor (por ejemplo, tablas, figuras, instrumentos de encuesta o grandes porciones de texto).

Por último, he sometido esta disertación a la herramienta de integridad académica, definida por la universidad.

---

Juan Pablo Sierra Mejía

01/05/2021

Fecha

# Agradecimientos

Este trabajo merece un reconocimiento especial a dos personas, sin las cuales el presente trabajo no hubiese sido posible. En primer lugar, se agradece al doctor Fernando Jesús Guevara, quien dirigió el presente trabajo de investigación y quien fue parte fundamental de mi formación académica desde su inicio. Gracias al profesor Guevara encontré en la disciplina del mantenimiento un área de estudio fascinante y con todo el potencial para aportar al desarrollo del país.

En segundo lugar, quisiera agradecer al señor Daniel Fernando Acosta, futuro ingeniero industrial, quien tuvo un aporte fundamental desde su conocimiento de la analítica de datos. El señor Acosta aportó al presente trabajo con un análisis profundo de los datos disponibles, lo que facilitó la comprensión de los mismos y la generación de una herramienta de toma de decisiones para conservar la vida disponible de activos productivos, con base en los resultados obtenidos a partir de los análisis de aceite usado.

Por último, debo agradecer a la Universidad Nacional de Colombia. Gracias a esta gloriosa institución he podido alcanzar todas las metas profesionales y personales que me he propuesto a lo largo de mi vida.

## Resumen

**Validación de modelos predictivos de analítica de datos de análisis de aceite usado para la toma de decisiones de mantenimiento en un turbogenerador de vapor.**

En el presente trabajo se desarrollan modelos descriptivos, clasificatorios y predictivos de la analítica de datos, con el fin de generar una herramienta de toma de decisiones basadas en las observaciones capturadas de diferentes pruebas realizadas al aceite usado de un turbogenerador de vapor marca Siemens de una industria papelera. Se estructura una base de datos con la información recopilada en un periodo de seis años (81 registros).; allí se cuenta con mediciones de diferentes propiedades del lubricante, por lo que se seleccionan 4 variables principales para el análisis. Las variables seleccionadas son el Número ácido total (TAN), el porcentaje de agua disuelta en el aceite, la concentración de fósforo en el aceite y la viscosidad a 40°C. Se implementan modelos de clusterización jerárquica, series de tiempo, aproximación por medias móviles y cartas de control. Por último, se presentan las conclusiones derivadas de la implementación de dichos modelos.

**Palabras clave:** Análisis de aceite usado, Machine Learning, Turbogenerador de vapor, Analítica de datos, Mantenimiento predictivo

- X Validación de modelos predictivos de analítica de datos de análisis de aceite usado para la toma de decisiones de mantenimiento en un turbogenerador de vapor.
- 

## Abstract

### **Validation of predictive models on used oil analysis data for maintenance decision making in a steam turbo generator.**

In this study, Data analytic models (descriptive, classificatory and predictive) are developed, in order to generate a decision-making tool based on observations obtained from different tests carried out on used oil of a Siemens brand steam turbogenerator from paper industry. A database is structured with information collected over a period of six years (81 records). There are measurements of different properties of lubricant, Then, 4 main variables are selected for analysis. Selected variables are Total Acid Number (TAN), percentage of water dissolved in oil, phosphorus concentration in oil and viscosity at 40 ° C. Hierarchical clustering models, time series, moving average approximation and control charts are implemented. Finally, Conclusions derived from the implementation of these models are presented.

**Keywords:** Used Oil Analysis, Machine Learning, Steam Turbogenerator, Data Analytics, Predictive Maintenance.

# Contenido

	Pág.
<b>Resumen .....</b>	<b>IX</b>
<b>Lista de figuras.....</b>	<b>XIII</b>
<b>Lista de tablas .....</b>	<b>XV</b>
<b>1. Introducción .....</b>	<b>1</b>
1.1 Justificación.....	1
1.2 Problema de Investigación .....	5
1.3 Hipótesis .....	7
1.4 Objetivos .....	8
1.4.1 Objetivo General.....	8
1.4.2 Objetivos Específicos.....	8
<b>2. Marco Teórico.....</b>	<b>9</b>
2.1 Definiciones Básicas .....	10
2.1.1 Triblogía.....	10
2.1.2 Fricción.....	10
2.1.3 Desgaste. ....	10
2.1.4 Lubricación. ....	11
2.2 Estado del Arte.....	12
2.3 Turbogeneradores de vapor .....	15
2.3.1 Sistemas de Lubricación.....	18
2.3.2 Lubricantes .....	21
2.4 Análisis de Aceite usado. ....	22
2.4.1 Técnicas de Medición de propiedades de aceite.....	24
2.4.1.1 Técnicas de medición de fósforo.....	24
2.4.1.2 Técnicas de medición del TAN.....	24
2.4.1.3 Técnicas de Medición de Viscosidad.....	25
2.4.1.4. Técnicas de Medición de porcentaje de humedad.....	25
2.5 Modelos Estadísticos y Analítica de Datos. ....	26
2.5.1 Métodos Descriptivos.....	27
2.5.1.1 Población, individuo y muestra.....	27
2.5.1.2 Medidas de tendencia central.....	27
2.4.1.3. Medidas de dispersión .....	29
2.4.1.4. Diagrama de cajas y bigotes .....	30
2.5.2 Métodos Clasificatorios.....	32
2.4.2.1. Clusterización .....	33
2.4.2.2. Cartas de control.....	33



2.5.3	Métodos Predictivos.....	35
2.5.3.1.	Validación de Modelos Predictivos.....	36
2.5.3.2.	Cálculo del error en procesos de validación.....	37
<b>3.</b>	<b>Materiales y Métodos.....</b>	<b>41</b>
3.1	Datos Recopilados.....	41
3.2	Modelos Estadísticos.....	45
3.2.1	Validación de Modelo Predictivo.....	48
<b>4.</b>	<b>Resultados y Análisis.....</b>	<b>51</b>
4.1	Análisis Descriptivo.....	51
4.1.1	Análisis de las variables seleccionadas.....	52
4.2	Análisis Clasificadorio.....	57
4.3	Análisis Predictivo.....	62
4.3.1	Validación del análisis predictivo.....	74
4.4	Análisis mediante Cartas de Control.....	79
<b>5.</b>	<b>Conclusiones y recomendaciones.....</b>	<b>83</b>
5.1	Conclusiones.....	83
5.2	Recomendaciones.....	84
<b>A.</b>	<b>Anexo: Ficha Técnica Aceite Mobil SHC 825.....</b>	<b>87</b>
	<b>Bibliografía.....</b>	<b>89</b>

## Lista de figuras

	Pág.
<b>Figura 1:</b> La función del mantenimiento en una perspectiva temporal. ....	2
<b>Figura 2:</b> Documentos publicados por año en Machine Learning .....	5
<b>Figura 3:</b> Artículos referenciados para los términos "Machine Learning" y "Maintenance" 6	6
<b>Figura 4:</b> Curvas comunes de comportamiento del desgaste por deslizamiento.....	11
<b>Figura 5:</b> Mecanismo de separación de superficies mediante lubricación.....	12
<b>Figura 6:</b> Inversión en Lubricantes en el Continente Asiático.....	13
<b>Figura 7:</b> Artículos referenciados para los términos "lubricant" y "analytics" .....	14
<b>Figura 8:</b> Turbina de Vapor.....	16
<b>Figura 9:</b> Turbina de Vapor con sobrecalentamiento .....	17
<b>Figura 10:</b> Componentes típicos del sistema de lubricación de una turbina de vapor ....	21
<b>Figura 11:</b> Tipos de Histograma .....	29
<b>Figura 12:</b> Diagrama de cajas y bigotes característico.....	30
<b>Figura 13:</b> Distribución de Datos para diferentes coeficientes de correlación .....	32
<b>Figura 14:</b> Dendograma.....	33
<b>Figura 15:</b> Técnicas de validación cruzada para modelos predictivos.....	37
<b>Figura 16:</b> Ejemplo de reporte análisis de aceite usado.....	44
<b>Figura 17:</b> Diagrama de flujo metodológico. ....	47
<b>Figura 18:</b> División del dominio para validación cruzada de orden 5.....	49
<b>Figura 19:</b> BoxPlot para el TAN .....	53
<b>Figura 20:</b> BoxPlot para P (Phosphorus (ppm)) .....	54
<b>Figura 21:</b> Boxplot para la Viscosidad a 40°C.....	55
<b>Figura 22:</b> Boxplot para el contenido de Agua disuelta.....	56
<b>Figura 23:</b> Matriz de Correlaciones.....	57
<b>Figura 24:</b> Dendograma de clusterización. ....	58
<b>Figura 25:</b> Gráfico de dispersión cruzada entre TAN y Fósforo (ppm) .....	59
<b>Figura 26:</b> Gráfico de dispersión cruzada entre TAN y viscosidad a 40°C .....	59
<b>Figura 27:</b> Gráfico de dispersión cruzada entre TAN y % de agua disuelta .....	60
<b>Figura 28:</b> Gráfico de dispersión cruzada entre TAN y reporte de status .....	60
<b>Figura 29:</b> Diagrama de dispersión cruzada entre P (ppm) y la viscosidad a 40°C. ....	61
<b>Figura 30:</b> Diagrama de dispersión cruzada entre fósforo (ppm) y % de agua disuelta..	61
<b>Figura 31:</b> Diagrama de dispersión entre viscosidad a 40°C y % de agua disuelta.....	62
<b>Figura 32:</b> Gráfica de serie temporal para el TAN.....	63
<b>Figura 33:</b> Serie temporal para el TAN mayor a 0.5.....	64
<b>Figura 34:</b> Serie temporal de la variable P (Phosphorus (ppm)) .....	65

<b>Figura 35:</b> Serie temporal para la variable P (ppm) con filtro de TAN > 0.5 .....	65
<b>Figura 36:</b> Serie temporal para la viscosidad a 40°C.....	66
<b>Figura 37:</b> Serie de Tiempo para la viscosidad a 40°C. Filtro TAN > 0.5 .....	67
<b>Figura 38:</b> Serie temporal para el porcentaje de agua disuelta.....	67
<b>Figura 39:</b> Serie Temporal para porcentaje de agua. Filtro TAN > 0.5.....	68
<b>Figura 40:</b> Comparativo series de tiempo TAN y P (ppm) .....	69
<b>Figura 41:</b> Comparativo Series de Tiempo TAN y Viscosidad a 40°C .....	69
<b>Figura 42:</b> Comparativo Series de Tiempo TAN y % de Agua Disuelta .....	69
<b>Figura 43:</b> Comparativo Series de Tiempo P (ppm) y Viscosidad a 40°C.....	70
<b>Figura 44:</b> Comparativo Series de Tiempo P (ppm) y % de agua disuelta. ....	70
<b>Figura 45:</b> Comparativo Viscosidad a 40°C y % de agua disuelta .....	70
<b>Figura 46:</b> Aproximación mediante medias móviles para el TAN.....	72
<b>Figura 47:</b> Aproximación mediante medias móviles para el Fósforo (ppm).....	73
<b>Figura 48:</b> Aproximación mediante medias móviles para Viscosidad a 40°C.....	73
<b>Figura 49:</b> Aproximación mediante medias móviles para % de agua disuelta .....	74
<b>Figura 50:</b> Validación cruzada de orden 5 para el fósforo (ppm). ....	75
<b>Figura 51:</b> Validación cruzada de orden 5 para el TAN (mg KOH/g) .....	75
<b>Figura 52:</b> Validación cruzada de orden 5 para la Viscosidad @40°C (cSt) .....	76
<b>Figura 53:</b> Validación cruzada de orden 5 para el % de Agua (Vol%) .....	76
<b>Figura 54:</b> Cartas de control para el TAN .....	79
<b>Figura 55:</b> Carta de control para la variable P (Phosphorus) .....	80
<b>Figura 56:</b> Carta de Control para la viscosidad a 40°C.....	80
<b>Figura 57:</b> Carta de control para % de agua disuelta.....	81

## Lista de tablas

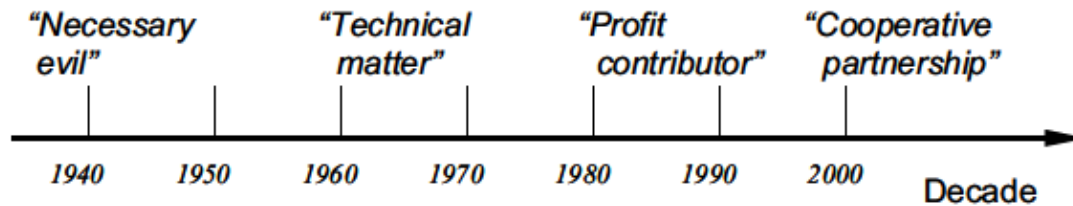
	<b>Pág.</b>
<b>Tabla 1:</b> k óptimo para modelo de medias móviles .....	72
<b>Tabla 2:</b> RMSPE obtenido en validación cruzada para el fósforo (ppm).....	77
<b>Tabla 3:</b> RMSPE obtenido en validación cruzada para el TAN (mg KOH/mg).....	77
<b>Tabla 4:</b> RMSPE obtenido en validación cruzada para la Visc @40°C (cSt) .....	77
<b>Tabla 5:</b> RMSPE obtenido en validación cruzada para el agua (%Vol) .....	78

# 1. Introducción

## 1.1 Justificación

Durante las últimas décadas, tal y como lo afirma (Deloitte Brazil, 2021) el mantenimiento de activos productivos se ha ubicado en una posición de gran importancia dentro de los organigramas de compañías en todos los sectores y tamaños que pueden existir dentro del universo macroeconómico de la sociedad actual. Lo anterior, se debe a que dichas empresas han comprendido que, al garantizar la confiabilidad y disponibilidad de sus activos, se encuentran en la capacidad de brindar a sus clientes el respaldo que ellos buscan de sus proveedores (Patiño-Rodriguez & Guevara Carazas, 2020). Por lo tanto, el mantenimiento pasa a ser un área que no solo interactúa con la producción, sino que se convierte en un área pilar de la empresa (Guevara Carazas & Martha daSouza, 2010).

Sin embargo, llegar a dicho nivel de comprensión respecto a la importancia del mantenimiento dentro del proceso productivo de una compañía, no ha sido un camino fácil y mucho menos espontáneo. De hecho (Pintelon & Parodi-Herz, 2008) hacen un análisis, sintetizado en la figura 1, en la cual se muestra que para las primeras décadas del siglo XX el mantenimiento era considerado tan solo un mal necesario y por tanto no le eran destinados los recursos requeridos, agravando en una mayor medida la causa raíz del problema.



**Figura 1:** La función del mantenimiento en una perspectiva temporal.  
**Fuente:** (Pintelon & Parodi-Herz, 2008)

La **Figura 1** es coherente con el hecho de que antes de la segunda guerra mundial, la maquinaria utilizada en procesos industriales era mucho más robusta y en general no requería sostener niveles de productividad tan altos como aquellos solicitados por la industria bélica. En la época de la posguerra las industrias de Japón y Alemania debieron reconstruirse casi en su totalidad, con una cantidad de recursos limitados; lo anterior resultó en que se construyera maquinaria más liviana, con exigencias de calidad y productividad casi iguales a aquellas requeridas durante la guerra. Esto fue lo que condujo a la generación de las primeras rutinas de mantenimiento planeado (Brown & Sondalini, n.d.)

Sin embargo, con el pasar de los años las empresas comenzaron a darse cuenta de que un mantenimiento totalmente preventivo y planeado muchas veces resultaba costoso e incluso ineficaz. El mantenimiento preventivo basado en rutinas en función del tiempo resultaba en intervenciones de mantenimiento innecesarias para equipos cuya condición operativa era normal. Por otro lado, pueden existir fallas ocultas en equipos que se manifiestan antes de que se cumpla el respectivo tiempo de inspección (Brown & Sondalini, n.d.) (Martha deSouza, 2012).

Lo anterior, permitió que durante los años 60 comenzaran a desarrollarse nuevas estrategias y filosofías de mantenimiento como el mantenimiento productivo total (TPM por sus siglas en inglés) y el mantenimiento centrado en confiabilidad (RCM), las cuales tienen como punto común la optimización de los recursos asignados a mantenimiento. Esto, sumado a los avances tecnológicos en la realización de ensayos no destructivos, ha permitido que se avance hacia un mantenimiento más proactivo, en el cual se saque el mayor provecho posible de la vida útil de los equipos con un sacrificio mínimo de la disponibilidad de planta (Gupta & Mishra, 2016).

Con el surgimiento de dichas técnicas, se ha comenzado a hablar durante las últimas décadas de mantenimiento predictivo, término que muchas compañías usan de manera desacertada puesto que la mera utilización de ensayos no destructivos no implica una estimación de la vida útil remanente del activo; únicamente permite visualizar el estado actual de componentes ocultos o de fallas subsuperficiales sin necesidad de realizar un paro de máquina que afecte la disponibilidad de la planta. Para dar al mantenimiento el carácter predictivo al cual se desea llegar, es necesario complementar estas técnicas con modelos estadísticos que permitan generar herramientas de toma de decisiones sobre cuál es el momento óptimo en el cual se debe intervenir un activo productivo. Lo anterior implica que, a partir de la existencia de datos históricos, se pueda “entrenar” un modelo que permita, tomando como base medidas que se adquieren con una determinada frecuencia, identificar si un equipo es propenso a la falla o no y en qué momento ocurrirá dicha falla, además de pronosticar las consecuencias que tendría sobre el activo determinada condición externa (Han et al., 2021).

Dentro de las principales técnicas del mantenimiento basado en condición (CBM, por sus siglas en inglés) puede resaltarse el análisis de aceite usado. Esta técnica tiene gran relevancia puesto que permite obtener información acerca del desgaste presente en los mecanismos, la contaminación del lubricante, así como la condición de este. (Zhao, 2017) hace la analogía entre el aceite y la sangre en los seres humanos. Mediante los análisis de aceite usado pueden diagnosticarse gran cantidad de problemas ocultos en los mecanismos que no necesariamente muestran síntomas evidentes como ruido o vibraciones mecánicas (Raadnui & Kleesuwana, 2005).

En este punto, resulta de gran importancia introducir conceptos como *data science* y *big data* los cuales conforman una línea de investigación de bastante auge en la actualidad. Estos conceptos surgen a partir de la revolución informática, en la cual comenzaron a surgir en el mundo gran cantidad de datos estructurados y no estructurados. Según la revista Forbes (Cave, 2017), para 2025 habrán en el mundo 163 zettabytes almacenados, lo que equivale a casi 180 billones de gigabytes. Si dicha cantidad de información se almacenase en CDs y dichos CDs se apilaran, se formaría una torre de 307.234 km, lo cual es comparable con la distancia promedio de la tierra a la luna (384.400 km).

Las principales características de la forma en la cual se generan, almacenan y procesan estos datos están sintetizadas en un concepto conocido como las 4V (Volumen, Velocidad, Variedad y Veracidad). Sin embargo, en (Elshawi, Sakr, Talia, & Trunfio, 2018) se propone una quinta V en la cual entra a jugar directamente la ciencia de los datos, “Valor”. El valor de los datos se obtiene analizándolos y obteniendo de ellos patrones escondidos, tendencias y modelos de conocimiento que permitan tomar decisiones con base a estos datos. Es en este punto donde el aprendizaje de máquina o *Machine Learning* entra a jugar un papel preponderante, pues son los diferentes modelos que cobija este concepto, aquellos que permiten encontrar dicho valor para la toma de decisiones a partir de los datos.

Puede decirse que los modelos de aprendizaje de máquinas buscan, en general, dar respuesta a dos tipos de problemas. El problema de la clasificación y el problema de la predicción, también llamada regresión. Los modelos de clasificación buscan ubicar determinado dato dentro de un grupo u otro de datos que comparten características similares. Por su lado, los problemas de regresión tratan de predecir a partir de un histórico de datos, cuál será la tendencia de estos en el futuro (Vellido, Martín-Guerrero, & Lisboa, 2012).

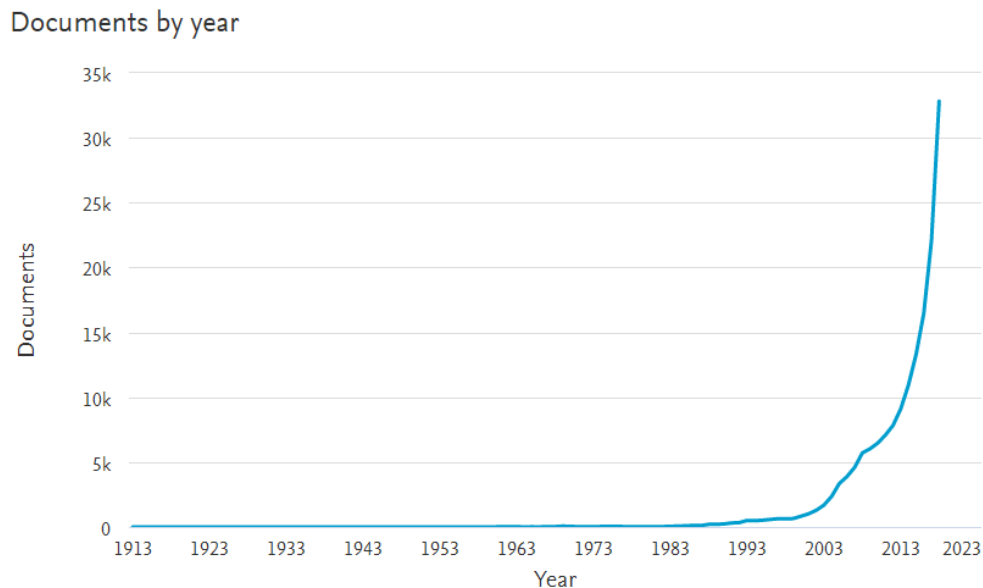
Volviendo al mantenimiento podemos ver que la solución de dichos problemas encaja perfectamente con las necesidades discutidas anteriormente. Generar modelos que permitan clasificar en qué estado se encuentra un equipo de acuerdo con características que se hallan medido anteriormente, o bien predecir el momento en el cual un equipo fallará de acuerdo con las tendencias que se generen a partir de históricos de sus variables críticas; puede marcar un gran hito en la historia de esta disciplina y ser de un grandísimo valor para compañías que deseen llevar sus políticas de mantenimiento a un nivel verdaderamente valioso (Ayvaz & Alpay, 2021).

En el presente trabajo se pretende realizar un análisis de los datos de análisis de aceite usado recolectados durante un periodo de 7 años. Se presentarán diferentes características de dichos datos y se analizarán a la luz de las tendencias o modelos de clasificación que pueden derivarse de dichos resultados. El presente trabajo busca establecer una metodología para el análisis de datos de aceite usado en función de establecer un adecuado sistema de muestreo y análisis.



## 1.2 Problema de Investigación

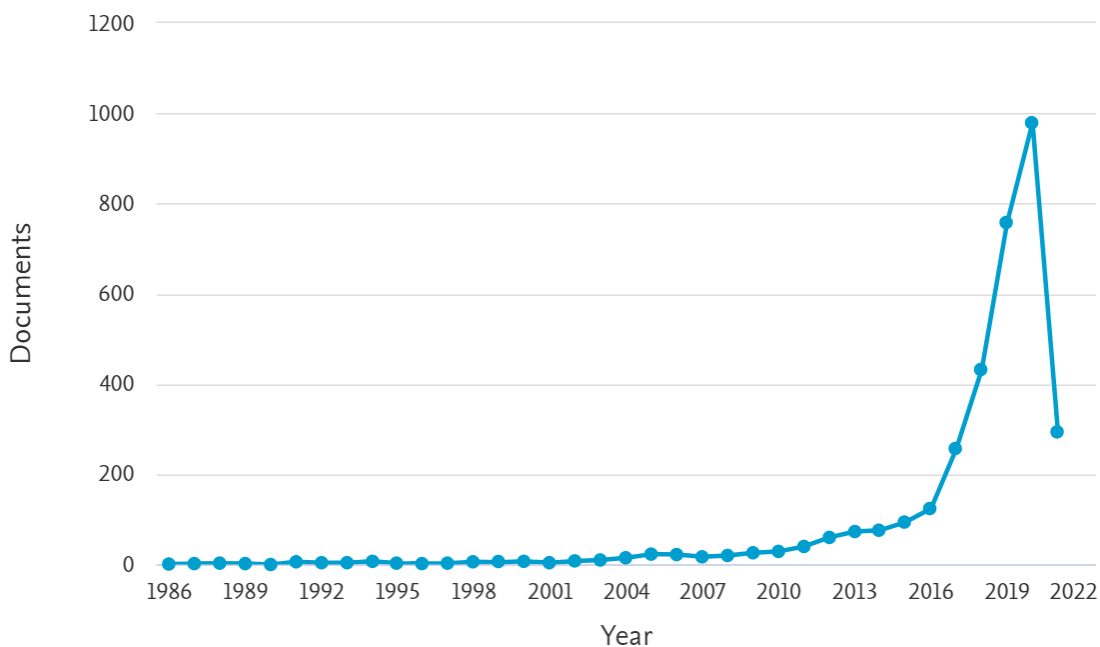
El aprendizaje de máquina o *machine learning* es una de las líneas de investigación más activas en la actualidad (Brnabic & Hess, 2021). Áreas del conocimiento como la medicina (Zhang & Sejdić, 2019)(Nicholson et al., 2019), las ciencias de la computación (Ucci, Aniello, & Baldoni, 2019), las geociencias (McCoy & Auret, 2019), el análisis numérico (Capuano & Rimoli, 2019), por mencionar solo algunas; han encontrado en las técnicas de *big data* y/o *machine learning* modelos estadísticos que les permiten realizar predicciones y tomar decisiones importantes. En la **Figura 2** puede verse como el aprendizaje de máquina se ha convertido en una tendencia de investigación durante los últimos años, esto se evidencia en la clara tendencia exponencial mostrada en la gráfica. En la actualidad, el manejo de estas herramientas es altamente valorado en el mercado laboral a nivel mundial, pues las compañías han descubierto que anualmente pagan millones de dólares por la generación y almacenamiento de datos que en muchos casos carecen de valor alguno.



**Figura 2:** Documentos publicados por año en Machine Learning  
**Fuente:** ("Scopus - Analyze search results," n.d.)

Tal y como se muestra en la **Figura 3**, el mantenimiento no ha sido ajeno a esta revolución y presenta un crecimiento exponencial con casi 1000 artículos referenciados durante el año 2021 (Scopus, 2021b). Diversos autores han comprendido que la clave principal para tener éxito en la aplicación de algoritmos de aprendizaje de máquina como método de toma de decisiones en mantenimiento industrial, son el adecuado almacenamiento y tratamiento de los datos obtenidos. Shahbazi et al. (Shahbazi et al., 2018) desarrollaron, compararon y aplicaron dos modelos de aprendizaje supervisado para hacer clasificación de órdenes de trabajo con el fin de extraer información adecuada de dichos documentos, lo cual es un ejemplo de cómo el primer paso para obtener valor de los datos con los cuales cuenta una compañía, es tener una adecuada organización de los mismos.

Documents by year



**Figura 3:** Artículos referenciados para los términos "Machine Learning" y "Maintenance"  
**Fuente:** (Scopus, 2021b)

Otra aplicación interesante que se les ha dado a estos métodos es la desarrollada por (Bousdekis, Papageorgiou, Magoutas, Apostolou, & Mentzas, 2017). Ellos desarrollaron modelos para realizar optimización de inventarios de repuestos utilizando modelos de regresión de arquitectura basada en eventos. Este tipo de modelos permiten a través de datos provenientes de la sensorica instalada en un equipo. Predecir el

momento en el cual determinado repuesto será requerido, emitiendo en el momento preciso la orden de compra del mismo. Este tipo de estrategias podrían llegar a impactar directamente los estados financieros de una compañía, ahorrando millones de dólares al año.

En este punto, puede resaltarse que las diferentes técnicas de mantenimiento basado en condición como lo son el análisis de vibraciones, el análisis por termografía, el análisis de discontinuidades por ultrasonido y el análisis de aceite usado, tienen una gran utilidad pues permiten recabar los datos del estado actual de determinado componente o mecanismo de tal manera que estos puedan ser estudiados mediante modelos predictivos o clasificatorios que ayuden a la toma de decisiones de mantenimiento y gestión de dichos activos.

El presente trabajo buscará generar un modelo que permita hacer análisis de los datos obtenidos mediante una de estas técnicas a un equipo en específico de una planta productiva. Específicamente se buscará realizar un modelo predictivo que permita la generación de tendencias, correlaciones y alertas basadas en los datos obtenidos del análisis de aceite de lubricación usado en un turbogenerador de vapor.

### **1.3 Hipótesis**

Con el fin de obtener una hipótesis que sirva como base para generar un método que busque dar respuesta a la pregunta de investigación anteriormente formulada,

*“Existe un modelo estadístico que permita predecir con un grado aceptable de certeza el estado de un equipo productivo basado en los datos obtenidos de análisis de aceite usado.”*

## 1.4 Objetivos

A continuación, se presentarán el objetivo general y los objetivos específicos del trabajo de investigación que se formula en la presente propuesta:

### 1.4.1 Objetivo General

- Generar un modelo estadístico que permita elaborar una herramienta de toma de decisiones basado en los datos obtenidos en el tiempo de los análisis de aceite usado de un turbogenerador de vapor en una planta productiva.

### 1.4.2 Objetivos Específicos

- Identificar las variables o propiedades críticas que deben ser conocidas o medidas en el aceite usado de un activo con el fin de predecir o clasificar su comportamiento.
- Establecer un esquema de organización de los datos de tal manera que la información pueda ser capturada por una herramienta computacional para la realización de predicciones.
- Determinar el modelo predictivo o clasificadorio que pueda aplicar en mejor medida con el fin de caracterizar el estado actual y futuro de un activo en función del comportamiento de las propiedades medidas mediante análisis de aceite usado.
- Validar la precisión en las predicciones realizadas por el modelo y la metodología diseñada, para su aplicación real en un activo crítico de una planta productiva.

## 2. Marco Teórico

En este capítulo se presentará un resumen de los principales elementos que deben estar en conocimiento del lector para facilitar el entendimiento del trabajo de investigación. En primer lugar, se presentarán algunos conceptos básicos que deben estudiarse a la hora de comprender este estudio. Luego, se hará un breve recorrido por el estado del arte en materia de analítica de datos aplicada al análisis de aceite usado. Se describirá brevemente la historia y funcionamiento de los turbogeneradores de vapor, recalcando la importancia de este tipo de maquinaria para la industria productiva. Se hará mención especial a la relevancia de los sistemas de lubricación en dichos equipos y la importancia que tienen en garantizar la confiabilidad de estos.

Por otro lado, se hablará del análisis de aceite usado. Allí se describirá cuáles son las principales variables que se monitorean en un programa de análisis de aceite usado haciendo énfasis en que influencia tienen dichas variables en la determinación de la idoneidad del aceite en servicio y en la vida misma de los equipos que se están monitoreando.

Por último, se hablará de los modelos estadísticos que fueron utilizados en el presente trabajo de investigación, haciendo énfasis en cómo dichos modelos pueden ser utilizados para generar herramientas de clasificación o predicción que permitan tener un mejor aprovechamiento del programa de análisis de aceite usado de un proceso productivo.

## 2.1 Definiciones Básicas

### 2.1.1 Tribología.

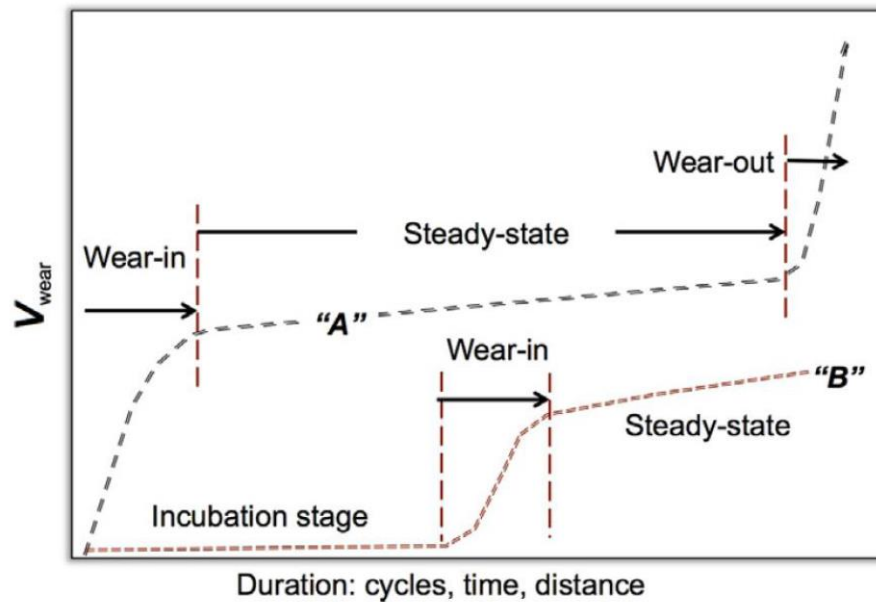
La tribología es una ciencia que estudia las superficies y las fuerzas que se desarrollan entre ellas cuando las mismas interactúan. Específicamente, estudia la fuerza de fricción, cómo se produce, cuantifica y sobre todo cómo puede intervenir el ser humano con el fin de disminuir los efectos producidos por la misma (Neale, 1973).

### 2.1.2 Fricción.

La fricción es la fuerza de resistencia tangencial al movimiento que se genera entre dos superficies en contacto. Esta fuerza es proporcional a la fuerza normal aplicada e independiente del área aparente de contacto (Bahari, 2017).

### 2.1.3 Desgaste.

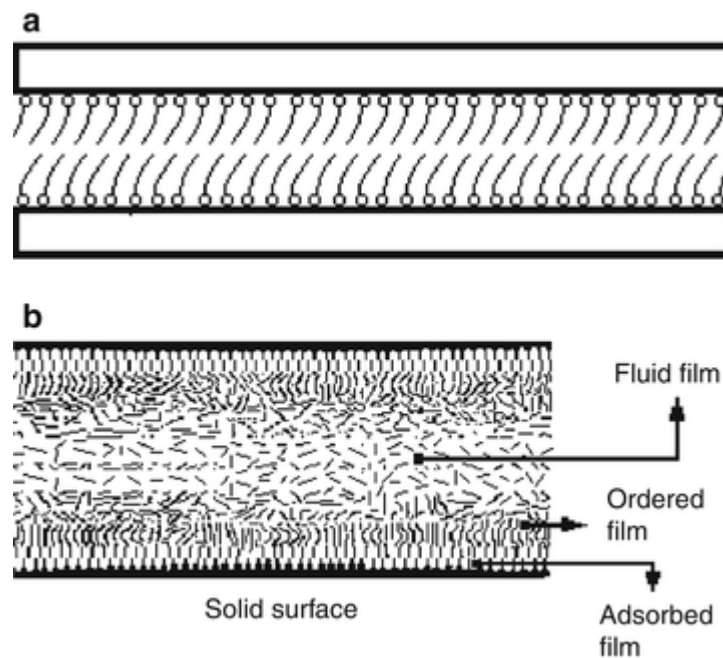
El desgaste puede definirse como la progresiva pérdida de sustancia de la superficie operativa de un cuerpo, como resultado del movimiento relativo de dicha superficie respecto a otra. La **Figura 4** muestra los típicos comportamientos de desgaste de un sistema de contacto no lineal por deslizamiento. Esta muestra que el proceso de desgaste (Curva A) puede tomar lugar en diferentes etapas. El desgaste aumenta rápidamente durante lo que se conoce como el proceso de asentamiento, desarrollando pronto un periodo de estado estable. Por último, nuevamente se incrementa cuando se comienza la etapa de fatiga superficial. En caso de que tener cargas bajas, la etapa de asentamiento puede no producirse inmediatamente, por lo tanto esta entra en una etapa de incubación del asentamiento, tal y como se aprecia en la curva B.



**Figura 4:** Curvas comunes de comportamiento del desgaste por deslizamiento  
**Fuente:** (Bahari, 2017)

#### 2.1.4 Lubricación.

La lubricación juega un papel fundamental pues representa una de las técnicas más eficaces y económicamente viables para disminuir los efectos de la fricción, disminuyendo el consumo energético generado por el fenómeno y protegiendo las superficies de sufrir desgaste, el cual es uno de los modos de falla más comunes e inevitables dentro de la vida útil de cualquier activo productivo (Albarracín, 2015). En la **Figura 5** puede apreciarse el mecanismo mediante el cual una película de lubricante separa dos superficies que se mueven una respecto a la otra, con el fin de disminuir los efectos de la fricción (Luo, 2013).

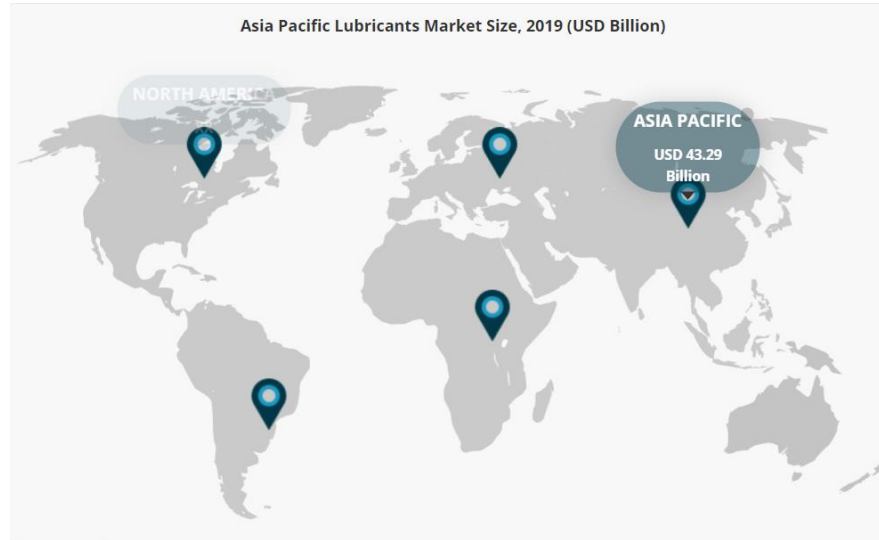


**Figura 5:** Mecanismo de separación de superficies mediante lubricación.  
**Fuente:** (Luo, 2013)

## 2.2 Estado del Arte

La lubricación es una de las actividades de mantenimiento preventivo fundamentales dentro de cualquier proceso productivo. Durante el año 2019 124 billones de dólares fueron invertidos por la industria en el rubro de lubricantes. (Fortune Business Insight, 2020). Por ejemplo, el continente asiático tuvo en el año 2019 un nivel de inversión de 43.3 billones de dólares en lubricantes, tal y como se muestra en la **Figura 6**. Esto con el objetivo de garantizar la disponibilidad de sus equipos productivos, preservando o incluso aumentando la vida disponible de los mismos.





**Figura 6:** Inversión en Lubricantes en el Continente Asiático  
**Fuente:** (Fortune Business Insight, 2020)

Dada la magnitud de dicha inversión, durante las últimas décadas se ha dado un cambio sustancial en el enfoque con el cual se administran los programas de lubricación en la industria. Se ha migrado de una filosofía netamente preventiva en la cual las actividades de lubricación (cambio de aceites, reengrase de componentes, entre otras) son basadas únicamente en una base temporal, a una visión basada en la condición de los activos y lubricantes en servicio. Es por ello, que el análisis de aceites y grasas usadas cada vez cobra mayor relevancia, permitiendo que la vida útil de los lubricantes sea llevada a su máxima capacidad.

El cambio de enfoque anteriormente descrito ha generado dentro de la industria una gran cantidad de datos que, en raras ocasiones, son analizados a consciencia por los responsables de la administración de programas de lubricación en procesos productivos. La propuesta presentada en este trabajo de investigación (con carácter de profundización) consiste en la aplicación de modelos de analítica de datos que permitan entender en una mejor medida la información que es suministrada a raíz de estos ensayos.

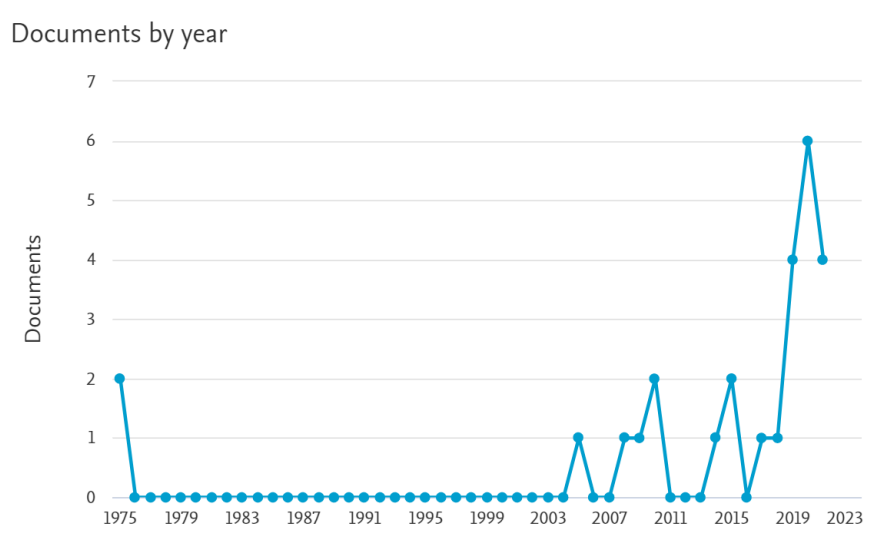
El desarrollo de este tema en la literatura no es muy amplio, algunos trabajos como el desarrollado por (Raposo, Farinha, Fonseca, & Galar, 2019) presentan modelos de predicción de la condición para aceites lubricantes de motores de combustión interna (Aceites SAE). Allí se utilizan modelos de suavizamiento exponencial, con el fin de lograr

extrapolaciones de las series temporales que permitan predecir la condición del aceite en un determinado tiempo.

Por otro lado (Vališ, Žák, & Pokora, 2015), presentan un modelo para la predicción de falla en motores Diesel, basado en la tendencia al desgaste evidenciada en los análisis de aceite usado. Allí se utilizan métodos de regresión lineal, con el fin de establecer tendencias que permitan anticiparse a la falla de componentes críticos de estos equipos.

Asimismo, (Macián, Tormos, Ruíz, & Ramírez, 2015) realizaron un análisis de la influencia de las condiciones del aceite para lubricación de motores Diesel en el consumo de combustible de una flota de buses urbanos. Allí, analizaron la correlación entre diferentes variables obtenidas del análisis de aceite usado con los datos registrados de consumo de combustible.

Tal y como puede observarse, hay poco desarrollo del tema de análisis de datos provenientes de un programa de muestreo de aceite, lo anterior puede observarse claramente en la **Figura 7**, la cual presenta los resultados de consulta obtenidos de la base de datos Scopus ante las palabras clave “*lubricant*” y “*analytics*”. Allí se observa que el máximo de artículos referenciados ocurrió en el año 2020 y fue de solo 6 artículos.



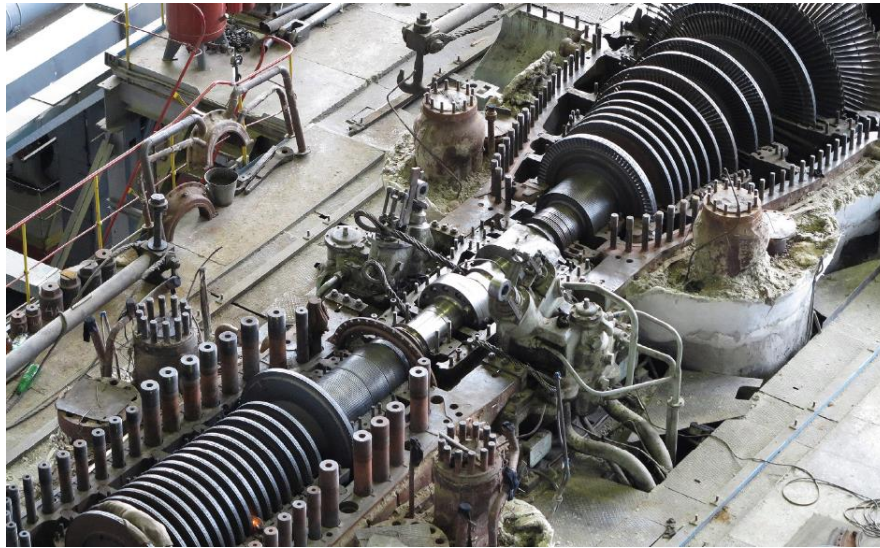
**Figura 7:** Artículos referenciados para los términos "lubricant" y "analytics"  
**Fuente:** (Scopus, 2021a)

Más aun, no se encontró ningún trabajo donde el asunto sea abordado para aceites de tipo industrial (ISO). Diferentes productores de lubricantes como Mobil, Shell, Petrobras prestan el servicio de análisis de aceite usado para sus productos. En esta misma medida, en este momento dichas compañías trabajan arduamente en construir modelos que permitan recopilar los datos obtenidos y ofrecer a sus diferentes clientes una solución con un valor agregado al problema anteriormente expuesto.

## **2.3 Turbogeneradores de vapor**

Una turbina de vapor es un dispositivo mecánico que se encarga de extraer energía térmica del vapor presurizado y convertirlo en energía mecánica rotacional. Su manifestación moderna fue inventada en 1884 por Sir Charles Parsons (Maxell, 1996).

El principal uso que se le da a la energía rotacional convertida en una turbina de vapor es generar energía eléctrica a través de un generador eléctrico. De allí que este tipo de mecanismos acoplados se conozcan como turbogeneradores de vapor. Se dice que aproximadamente el 80% de la energía eléctrica generada en el mundo, proviene de turbogeneradores de vapor, esto se debe principalmente a que son equipos de alta eficiencia termodinámica que permiten generar energía eléctrica a bajo costo y con impactos ambientales controlables en comparación a otras alternativas (Joseph Omosanya, Titilayo Akinlabi, & Olusegun Okeniyi, 2019).



**Figura 8:** Turbina de Vapor  
**Fuente:** (Phillips 66, 2019)

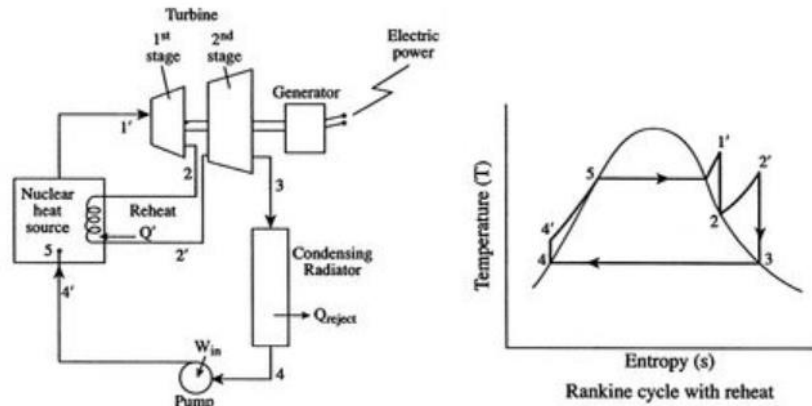
En la **Figura 8** se muestra la configuración física de un turbogenerador de eje horizontal. Los turbogeneradores de vapor son construidos en una gran variedad de tamaños y configuraciones, las cuales van desde equipos pequeños de 1 KW de potencia, los cuales son utilizados para mover bombas de agua hasta turbinas con capacidad de generar hasta 1500 MW de potencia eléctrica. De acuerdo con esto, las turbinas de vapor pueden clasificarse de diversas maneras de acuerdo con las condiciones de suministro y descarga de vapor o según la configuración de sus ejes y carcazas (Maxell, 1996).

Según las condiciones de suministro y descarga de vapor, las turbinas de vapor pueden clasificarse en turbinas condensadoras, no condensadoras, de sobrecalentamiento, de extracción y de inducción.

Las turbinas no condensadoras o de contrapresión son aquellas cuya presión de descarga es controlada por una válvula reguladora de presión, lo anterior con el fin de entregar vapor a otras fases de un proceso. Son comúnmente usadas en refinerías, la industria de pulpa y papel y unidades desalinizadoras. En este tipo de turbinas no se aprovecha el 100% de la entalpia disponible en el vapor, sino que se busca reducir la presión del mismo para su utilización en otros procesos (Maxell, 1996).

Por otro lado, las turbinas condensadoras son aquellas utilizadas principalmente en plantas de generación eléctrica. El vapor de la descarga se entrega con una calidad cercana al 90% con el fin de ser condensado y retornado al ciclo de generación (Maxell, 1996).

En las turbinas de sobrecalentamiento el flujo de vapor sale de una etapa de alta presión en la turbina y es retornado nuevamente a la caldera de potencia donde se entrega energía térmica adicional al vapor, tal y como logra apreciarse en la **Figura 9**. Luego el vapor reingresa a la turbina en una sección de presión intermedia donde continua su expansión (Maxell, 1996).



**Figura 9:** Turbina de Vapor con sobrecalentamiento  
Fuente: (Spakovszky, 2007)

Las turbinas de extracción se utilizan en casi todas las aplicaciones y son aquellas en las que se extrae vapor a diferentes presiones en etapas diferentes. Este vapor se utiliza en diferentes procesos industriales (Maxell, 1996).

Además de la clasificación de acuerdo con la condición del vapor de descarga de la turbina, también puede hacerse una clasificación de acuerdo con la disposición de los ejes de esta. Así, existen turbinas de eje simple, en las cuales una sola turbina está acoplada a un solo generador, turbinas de ejes en tándem donde dos o más turbinas se acoplan a un solo generador por medio de una caja de engranajes. Por último, se encuentran las turbinas de ejes cruzados en las cuales dos o más ejes se acoplan a dos o más generadores mediante cajas reductoras, los generadores usualmente operan a

velocidades distintas de acuerdo con su número de polos magnéticos (Reddy, Ahmed, Kumar, Reddy, & Bharathi, 2014).

A pesar de los diferentes tipos de turbogeneradores de vapor que pueden encontrarse en la industria, todos ellos comparten el mismo principio de funcionamiento. Una turbina de vapor tiene en su interior varios conjuntos de aspas o alabes como se les conoce regularmente. Un conjunto de alabes estacionarios esta unido a la carcasa de la turbina  $\zeta$ , mientras otro conjunto esta unido a un eje rotatorio, soportado usualmente sobre cojinetes lisos que le permiten girar. El flujo de vapor que ingresa a la turbina sufre una caída de presión, la cual se convierte en energía cinética que es entregada al eje a través de los alabes (Reddy et al., 2014).

Generalmente las velocidades de rotación de una turbina de vapor son de aproximadamente 3000 RPM, es por ello por lo que este tipo de equipos tienen una sensibilidad bastante elevada a condiciones vibratorias. Por tal motivo, las tolerancias que se presentan en los apoyos o cojinetes lisos de la turbina son bastante cerradas, evitando la aparición de holguras. Es por ello por lo que la lubricación de estos cojinetes es un aspecto crítico cuando se trata de garantizar la confiabilidad del equipo, evitando posibles fallas catastróficas (Banaszkiewicz, 2014).

A continuación, se presentará un breve resumen de la importancia de los sistemas de lubricación en turbinas de vapor, así como las principales características que debe tener un lubricante para garantizar que las condiciones de fricción en el equipo sean las adecuadas para un funcionamiento adecuado.

### **2.3.1 Sistemas de Lubricación**

En vista de la criticidad que los sistemas de lubricación representan para el adecuado funcionamiento de los turbogeneradores de vapor, es importante mencionar que estos por lo general son sistemas redundantes, que buscan que las condiciones de lubricación sean mantenidas en la mayoría de escenarios operacionales posibles, con el

fin de que una falla catastrófica en el sistema de suministro de energía eléctrica no tenga una incidencia en el funcionamiento del sistema de lubricación.

Aunque el sistema de lubricación es considerado un componente periférico a las turbinas de vapor, es un componente completamente indispensable para el funcionamiento de la turbina. Por lo general los sistemas de lubricación cuentan con una bomba principal, que es la encargada de llevar el aceite desde el depósito de lubricación hacia los cojinetes, esta bomba puede ser accionada por medio de un motor eléctrico o por medio de una pequeña turbina de vapor alimentada desde una de las descargas de vapor del turbogenerador. Además, estos sistemas deben contar con una bomba auxiliar que sirva de stand-by en caso de que se presente alguna falla en la bomba principal (Exxon Mobil, 2009).

Por otro lado, el sistema debe contar con una bomba de emergencia, la cual usualmente es movida por un motor eléctrico de corriente directa alimentado desde un banco de baterías. Esta bomba debe tener la capacidad suficiente para sostener las condiciones mínimas de lubricación del turbogenerador ante un eventual corte total de la energía eléctrica que deje por fuera la bomba de lubricación principal y la bomba auxiliar (Lazovic & Marinkovic, 2015).

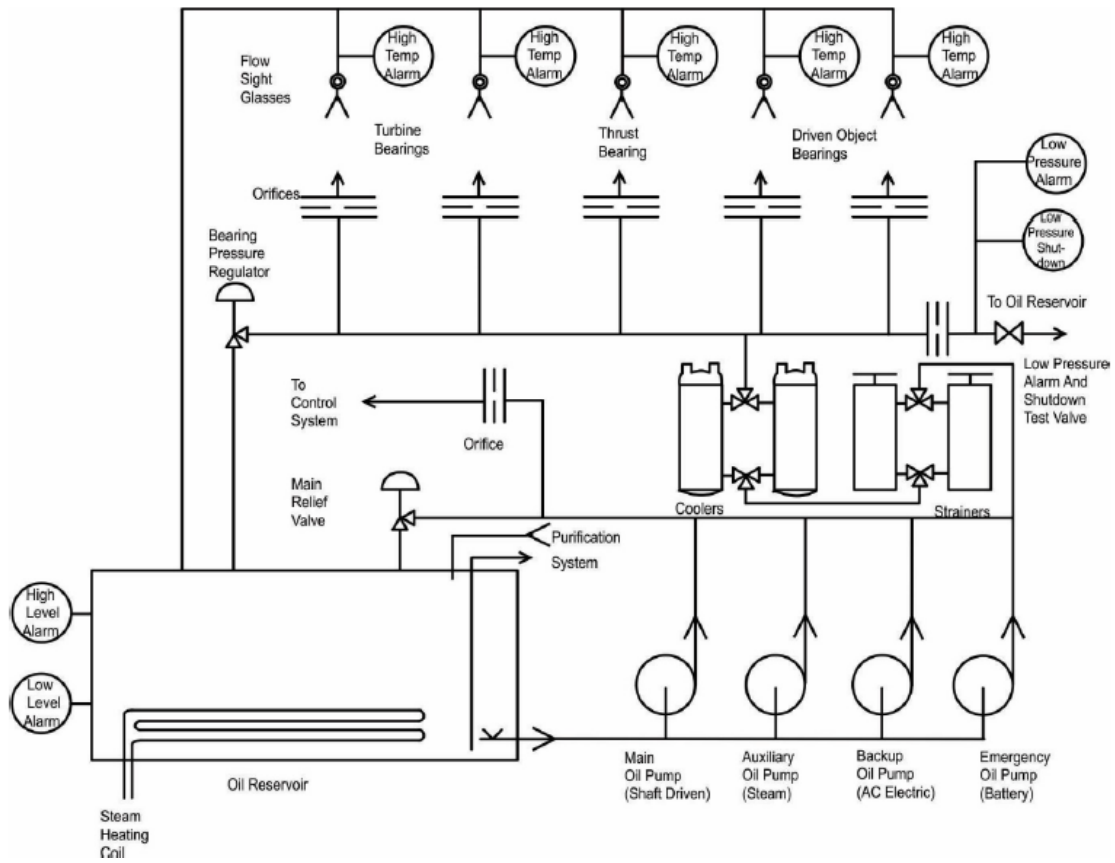
Dependiendo de la configuración de la turbina, también puede encontrarse dentro del sistema de lubricación una bomba adicional llamada bomba de levantamiento. La función principal de la bomba de levantamiento es generar una película de lubricante que separe las superficies móviles y fijas de los cojinetes de la turbina, de tal manera que la misma pueda tener un arranque seguro (Lazovic & Marinkovic, 2015).

Otros de los componentes que tienen gran relevancia dentro del sistema de lubricación de una turbina de vapor son los sistemas de filtración. Tal y como se mencionó anteriormente, una de las características principales de los cojinetes lisos que sirven de apoyo al eje de una turbina de vapor es el tener tolerancias estrechas que evitan la aparición de holguras que induzcan vibraciones mecánicas. Es por ello por lo que la limpieza del aceite se vuelve un factor crucial a la hora de garantizar la integridad de un turbogenerador de vapor. Partículas con tamaños superiores al espesor de la película de lubricante, pueden generar desgaste abrasivo en los cojinetes (Hausermann, 1961).

Dadas las condiciones anteriormente mencionadas, los sistemas de lubricación de turbinas de vapor, incluyen diferentes sistemas de filtración de alta eficiencia, los cuales buscan garantizar que el código ISO 4406 del lubricante se encuentre dentro de los parámetros recomendados por el fabricante del equipo. Dentro de los diferentes sistemas de filtración que pueden mencionarse se encuentran, los sistemas de filtración en la descarga de las bombas, los cuales son los encargados de proteger como última barrera, los cojinetes lisos del turbogenerador. Además, vale la pena mencionar los sistemas de filtración de retorno, los cuales se encargan de retirar las diferentes partículas que pudieron haberse generado producto del desgaste normal de los componentes. Por último, debe mencionarse el sistema de filtración offline, el cual es una recirculación de aceite al depósito que busca mantener el nivel de limpieza requerido por el fabricante en el aceite de la turbina (Hausermann, 1961).

También, es importante mencionar como componente importante de los sistemas de lubricación en un turbogenerador de vapor, el sistema de refrigeración de aceite. Una de las principales funciones de cualquier lubricante es el de refrigerar; el aceite recibe el calor generado por fricción entre las superficies, evitando un aumento de temperatura de estas. Es por ello, que los sistemas de lubricación incluyen intercambiadores de calor los cuales buscan reducir la temperatura del aceite con el fin de que este pueda conservar su viscosidad de operación (AXA Risk Consulting, 2020). Los elementos anteriormente descritos pueden apreciarse en la **Figura 10**.





**Figura 10:** Componentes típicos del sistema de lubricación de una turbina de vapor  
**Fuente:** (AXA Risk Consulting, 2020)

### 2.3.2 Lubricantes

Además, es importante entender la importancia de las características principales que posee un aceite de lubricación utilizado en turbogeneradores de vapor. Tal y como se comentó en la sección 2.3.1, las velocidades de operación de una turbina de vapor son en general velocidades altas por encima de las 3000 RPM. En general, la viscosidad de operación de un lubricante es inversamente proporcional a la velocidad de operación, razón por la cual en general los aceites utilizados en turbinas son de baja viscosidad. Fabricantes como Solar o Siemens, comúnmente recomiendan aceites de un grado ISO VG 32 a ISO VG 46, los cuales garantizan la formación de una película de lubricante adecuada, evitando un posible sobrecalentamiento del equipo por exceso de fricción fluida (Sander, 2012).

Las principales funciones que debe cumplir un aceite para turbina son las siguientes:

- Reducir la fricción en los cojinetes de la turbina.
- Prevenir la corrosión y oxidación de los componentes.
- Refrigerar los cojinetes y la turbina.
- Reducir el desgaste.
- Remover partículas

Asimismo, uno de los principales factores que puede afectar el adecuado desempeño de un aceite para turbinas es la presencia de agua producto de fugas de vapor en el sistema de sellado de los cojinetes. La presencia de agua en el aceite puede llevar a la acidificación y formación de barnices en el mismo. Es por ello que los lubricantes para turbinas más avanzados incluyen agentes químicos que previenen o reducen la formación de ácidos, oxidación y corrosión. Además, poseen aditivos demulsificantes que ayudan a una separación más efectiva entre el agua y el aceite (Coleman, 1981).

## **2.4 Análisis de Aceite usado.**

De la mano con el análisis de vibraciones mecánicas, el análisis de termografía y el análisis de ultrasonido, el análisis de aceite usado es una de las principales técnicas de mantenimiento basado en condición (Ahmad & Kamaruddin, 2012). El mantenimiento basado en condición es una estrategia en la cual se monitorean condiciones de operación de determinado equipo con el fin de detectar fallas incipientes antes de que estas ocurran (Syan, Ramsoobag, Mahabir, & Rajnauth, 2020). Asimismo, pueden programarse intervenciones en el tiempo correcto para evitar fallas catastróficas que generen indisponibilidad de los activos de una compañía.

De acuerdo con lo anterior, el análisis de aceite usado logra no solo el objetivo de detectar un envejecimiento del lubricante sino ayuda en la detección y prevención de fallas en los activos. Según (Troyer & Fitch, 2004), siempre que se aplique una estrategia de mantenimiento proactivo, son necesarios tres pasos para asegurar que los beneficios de la misma se conseguirán.

El primer paso es fijar un objetivo o estándar, asociado con cada causa de falla. El segundo paso es mantener el control de las condiciones dentro de dicho estándar. El tercer paso es crear un círculo de retroalimentación implementando acciones correctivas una vez se traspasan los límites establecidos en el primero de los pasos (Troyer & Fitch, 2004).

En términos de análisis de aceite usado, dichos pasos deben cumplirse a cabalidad con el fin de lograr el objetivo principal de tener un programa análisis de aceite usado. En primer lugar, se deben establecer los límites y alarmas en cada una de las diferentes variables que se pretenden muestrear, posteriormente se deben establecer estrategias preventivas para que dichas variables se mantengan dentro de los rangos establecidos. En temas de análisis de aceite estas estrategias corresponden a las actividades de control de contaminación, buenas prácticas de lubricación de maquinaria y una adecuada frecuencia de reposición preventiva de lubricante. En tercer lugar, existen estrategias de retroalimentación ante posibles desviaciones de los parámetros. Dichas actividades incluyen filtración de lubricante, deshidratación y diálisis de aceite (Troyer & Fitch, 2004).

El análisis de aceite es un extenso campo que comprende cientos de pruebas diferentes que valoran una o más de las propiedades de un lubricante o una máquina. El objetivo del análisis de aceite usado es evaluar la condición de las máquinas que son lubricadas. El lubricante sirve como medio de información generada en el equipo en forma de contaminación o partículas de desgaste (Troyer & Fitch, 2004).

Hay tres categorías importantes del análisis de aceite:

- Análisis de propiedades del fluido: Aquí se evalúan propiedades físicas, químicas y de los aditivos del aceite.
- Análisis de contaminación: Los contaminantes son elementos diferentes al aceite que se introducen al sistema desde el medio ambiente o son generadas internamente. En general, los contaminantes en el aceite afectan negativamente no solo la salud del aceite, sino también la vida útil de los activos lubricados. En general los contaminantes pueden ingresar a un mecanismo a través de ventilas y respiradores, sellos de ejes ineficientes o dañados, aceite nuevo o filtros defectuosos o saturados.

- Análisis de partículas de desgaste: Cuando los componentes se desgastan por su operación generan partículas que quedan disueltas en el aceite. La identificación de estas partículas y su composición facilita la identificación de fallas y brinda una herramienta para la toma de decisiones proactivas de mantenimiento.

## 2.4.1 Técnicas de Medición de propiedades de aceite.

En el presente trabajo se analizan principalmente cuatro propiedades fundamentales en el aceite las cuales son la concentración de fósforo, el TAN, la viscosidad y el porcentaje de humedad. En vista de que los métodos de medición de dichas variables son relevantes para el presente estudio, a continuación se presenta una reseña de los principales métodos utilizados para la medición de las mismas.

### 2.4.1.1 Técnicas de medición de fósforo

El fósforo es un elemento presente principalmente en los compuestos organometálicos encargados de controlar la corrosión en piezas metálicas lubricadas. La medición de este elemento se realiza a través del ensayo normalizado ASTM D4951, también conocido como *Inductively Coupled Plasma Atomic Emission Spectrometry* (Shimadzu, 2003).

Este método consiste en la ionización de los diferentes elementos presentes en el aceite exponiéndolos a plasma de alta temperatura. Esta ionización hace que los electrones de cada átomo suban a niveles de energía superiores. Al retornar a su estado energético natural estos emiten luz a una longitud de onda propia de cada elemento. Al medir la intensidad de la luz emitida en la longitud de onda propia del fósforo puede correlacionarse la concentración del mismo en la muestra. Este tipo de técnicas permiten diferenciar concentraciones verdaderamente pequeñas (en la escala de partes por millón) de determinado elemento en una muestra de aceite. (Michalke & Nischwitz, 2017).

### 2.4.1.2 Técnicas de medición del TAN.

El TAN o número ácido total es una medida de la concentración de ácidos presentes en el aceite. Esta variable puede verse afectada por la presencia de ácidos en el paquete

de aditivos del lubricante, contaminación con ácidos u oxidación del producto. Esta medida es de gran importancia en aplicaciones de maquinaria industrial. Para motores de combustión interna, es más relevante el uso del TBN o Número básico total (Scientific Spectro, 2000).

Pese a que existen varias técnicas para la medición del TAN, la más relevante de ellas es la Titración potenciométrica, la cual está estandarizada por la prueba ASTM D664. Esta prueba consiste en disolver la muestra en tolueno, isopropanol y agua. Un electrodo de referencia es puesto en la mezcla y es conectado a un potenciómetro. El punto final de titración se alcanza cuando hay un pico bien definido de voltaje en el potenciómetro. La medida de voltaje en dicho punto puede correlacionarse directamente con el TAN (Scientific Spectro, 2000).

#### *2.4.1.3 Técnicas de Medición de Viscosidad.*

La viscosidad es una variable que hace referencia a la capacidad de un fluido para recorrer determinada distancia, producto de una presión determinada y en un tiempo determinado. Existen gran cantidad de técnicas para la medición de la viscosidad; la selección de la técnica más adecuada depende del tipo de fluido que se está testeando. Para aceites lubricantes, la técnica más utilizada es el viscosímetro Cannon-Fenske o viscosímetro de vasos capilares (Sibata, 2020).

Esta técnica consiste en medir el tiempo que tarda en fluir determinada cantidad de muestra en un tubo capilar. Es importante resaltar que la viscosidad tiene una relación inversa con la temperatura. Para aceites lubricantes industriales las temperaturas de referencia más utilizadas son 40°C y 100°C (Sibata, 2020). En el presente trabajo se analiza únicamente la viscosidad a 40°C.

#### *2.4.1.4 Técnicas de Medición de porcentaje de humedad.*

El agua es uno de los contaminantes más comunes y perjudiciales que puede existir en el aceite pues puede producir alteraciones en la viscosidad del aceite, pérdida de los principales aditivos, formación de ácidos, lodos y barnices. Por lo tanto, la medición de humedad en el aceite es una parte fundamental de cualquier programa de análisis de lubricante (Spectro Scientific, 2015).

El método más ampliamente utilizado para medición de la concentración de agua en el aceite es la de titración colorimétrica de Karl-Fisher. La titración es una técnica por medio de la cual una solución de concentración conocida es utilizada para hallar la concentración de una solución desconocida. Típicamente, el tritante (la solución conocida) se añade de manera controlada a un recipiente con una cantidad determinada de la muestra a analizar, hasta que se complete una reacción. En la prueba de Karl-Fisher, se añade yodo a la muestra hasta que esta reaccione completamente. La cantidad de yodo añadida es equivalente a la cantidad de agua presente en el aceite. Entonces, midiendo la cantidad de yodo que debe ser añadido a la muestra, puede determinarse de manera muy precisa la concentración de agua en el mismo (Spectro Scientific, 2015).

En este punto es importante aclarar que los valores límites de humedad en el aceite son de alrededor de 200 ppm, lo que equivale a tan solo un 0.02% de porcentaje en volumen. El lector podría preguntarse si una concentración en escala tan baja si es fácilmente detectable en una muestra de 100 ml, como las que regularmente se toma. Sin embargo métodos como el de Karl Fisher tienen la precisión necesaria para identificar moléculas de agua en el aceite aun en proporciones tan pequeñas.

## **2.5 Modelos Estadísticos y Analítica de Datos.**

Con el fin de sacar el adecuado provecho y adquirir herramientas que permitan hacer de la toma de decisiones de mantenimiento un proceso estructurado y basado en la realidad, debe adquirirse un conocimiento profundo en los diferentes métodos estadísticos que permiten caracterizar los datos y generar modelos de clasificación y predicción.

En el presente trabajo, se utilizarán tres familias principales de herramientas estadísticas básicas que permitirán tener una mejor comprensión de los datos recopilados y analizar desde diferentes puntos de vista la información suministrada por esto. Estas familias son las herramientas descriptivas, las herramientas de clasificación y las herramientas de predicción. A continuación, se realizará un pequeño resumen de cada una

de estas familias y se describirán los métodos estadísticos empleados para lograr los objetivos del presente trabajo.

## 2.5.1 Métodos Descriptivos

Según (Faraldo, 2013) la estadística descriptiva es un conjunto de técnicas numéricas y gráficas para describir y analizar un grupo de datos, sin extraer conclusiones (inferencias) sobre la población a la que pertenecen. Lo anterior indica que los métodos descriptivos buscan construir parámetros que representen o logren sintetizar información de una muestra de datos sin realizar una extrapolación del comportamiento general de la población.

Con el fin de entender correctamente el concepto de estadística descriptiva, es importante definir los conceptos de población, individuo y muestra.

### 2.5.1.1 Población, individuo y muestra

Una población es un colectivo de individuos sobre los cuales se quiere extraer alguna conclusión. Lo anterior implica que un individuo puede definirse como cada uno de los elementos que conforman una población. De esta manera, una muestra es un subconjunto aleatorio de la población sobre la cual se requiere realizar el análisis.

De acuerdo con lo anterior, la estadística descriptiva busca sintetizar la información proporcionada desde la muestra, sin que ello implique la formulación de una conclusión que afecte la población. Para esto, existen ciertos parámetros que llamaremos descriptores que permiten representar el comportamiento de los datos. Los descriptores más ampliamente utilizados son las medidas de tendencia central y de dispersión.

### 2.5.1.2 Medidas de tendencia central

Según (Barrios, 2015) las medidas de tendencia central son medidas que indican hacia donde tienden a concentrarse los valores contenidos en una muestra. Estas medidas son valores típicos o representativos de la muestra. Las principales medidas de tendencia central son la media, la moda y la mediana.

La media, también llamada esperanza o valor esperado que resulta de sumar todos los datos de una muestra y dividirla entre el tamaño de la muestra. Es una medida representativa de todos los datos ampliamente utilizada en las ciencias económicas y sociales por su versatilidad y fácil comprensión. En la **Ecuación 1** se muestra el desarrollo matemático para el cálculo de la media en un conjunto de datos. Allí el término del numerador representa la suma de todos los datos, mientras que el denominador representa el total de datos de la muestra.

**Ecuación 1:** Cálculo de la media en un conjunto de datos discretos

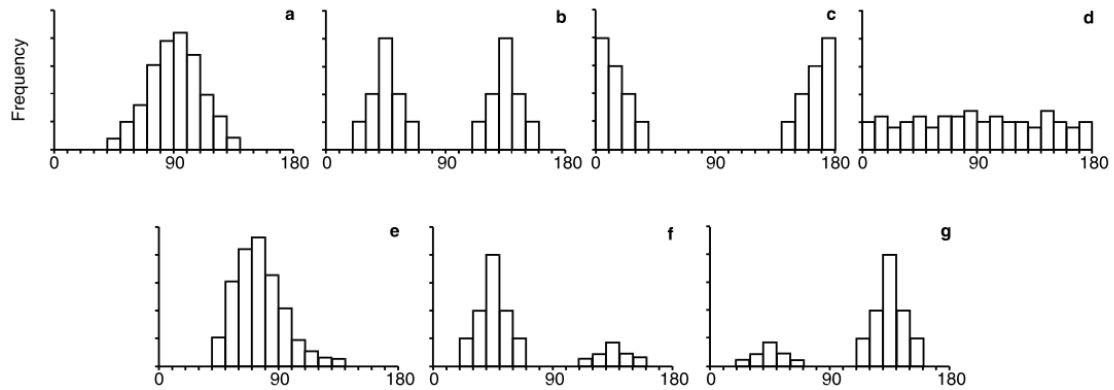
$$\bar{X} = \frac{\sum_{i=1}^n x_i}{N}$$

Por otro lado, la mediana es una medida de tendencia central cuyo valor se encuentra a la mitad de una serie de datos ordenados. Por encima de la mediana se encuentran el 50% de los datos con mayor valor de la serie y por debajo el 50% de los datos de menor valor (Barrios, 2015). También se le conoce como el segundo cuartil.

Por último, la moda se define como la medida de tendencia central cuyo el valor es el más común de los datos de una muestra. En otras palabras, la moda corresponde al dato que más se repite dentro de una muestra.

La relación entre las diferentes medidas de tendencia central permite identificar la simetría de la muestra. De acuerdo a esto las distribuciones de frecuencias o histogramas de una muestra pueden ser simétricas unimodales, simétricas multimodales, asimétricas con sesgo positivo y asimétricas con sesgo negativo (Hejnowicz, Burian, Dobrowolska, & Kolano, 2006). La **Figura 11** muestra los diferentes tipos de histogramas que pueden presentarse según la distribución de los datos.





**Figura 11:** Tipos de Histograma  
**Fuente:** (Hejnowicz et al., 2006)

En la **Figura 11**, el gráfico **a** representa un histograma con un máximo simple en la orientación transversal. El gráfico **b** representa un histograma con dos máximos simétricos. El gráfico **c** representa un histograma con dos máximos perfectamente asimétricos. El gráfico **d** representa un histograma con distribución uniforme. El gráfico **e** representa un histograma con un solo máximo asimétrico. Los gráficos **f y g** representan histogramas con múltiples máximos simétricos de diferentes magnitudes.

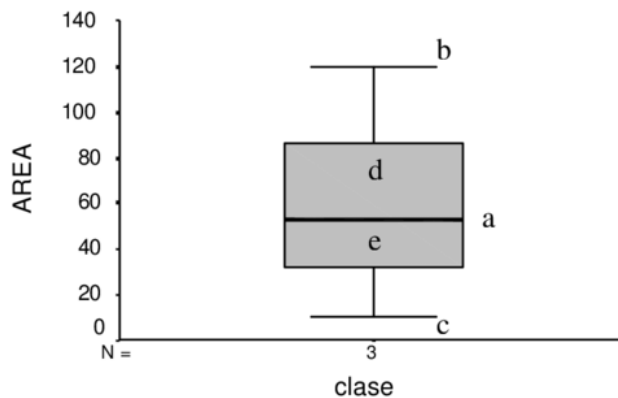
Dependiendo de las características de dicha distribución de frecuencias, puede determinarse cuál de las medidas de tendencia central resulta más representativa del conjunto de datos. En general si una distribución de frecuencias es perfectamente simétrica y unimodal, todas sus medidas de tendencia central son iguales. Cuando la distribución es unimodal y asimétrica es recomendable utilizar la mediana como la mejor medida de tendencia central, lo anterior pues la asimetría de los datos desplaza la media hacia la zona de mayor densidad de datos y por tanto divide la muestra en dos grupos desiguales donde se dejan de representar la mayoría de los datos, los cuales tienen una menor frecuencia. Por último, cuando existen dos o más modas, dichas modas son las medidas de tendencia central que representan en mejor medida el conjunto de datos.

#### 2.4.1.3. Medidas de dispersión

Por otro lado, es importante mencionar las medidas de dispersión, las cuales permiten conocer que tanto varían o se distribuyen los datos de una muestra. Valores grandes en las medidas de dispersión indican gran variación de los datos, mientras que valores pequeños indican uniformidad en la muestra (todos los datos son muy cercanos a las medidas de tendencia central). Las medidas de dispersión más conocidas son el rango (R), la varianza ( $\sigma^2$ ), la desviación estándar ( $\sigma$ ), el coeficiente de variación (CV), el índice de asimetría ( $\gamma_1$ ) y el índice de curtosis (Kurt).

#### 2.4.1.4. Diagrama de cajas y bigotes

Una herramienta de gran utilidad a la hora realizar análisis descriptivos a una muestra de datos es el gráfico de caja o “box-plot”, el cual según (Palladino, 2011) es una herramienta gráfica que permite resaltar aspectos de la distribución de observaciones dentro de una o más series de datos cuantitativos. La **Figura 12** muestra un diagrama clásico de cajas y bigotes.



**Figura 12:** Diagrama de cajas y bigotes característico.  
**Fuente:** (Jiménez Rodríguez & Arias Aguilar, 2004)

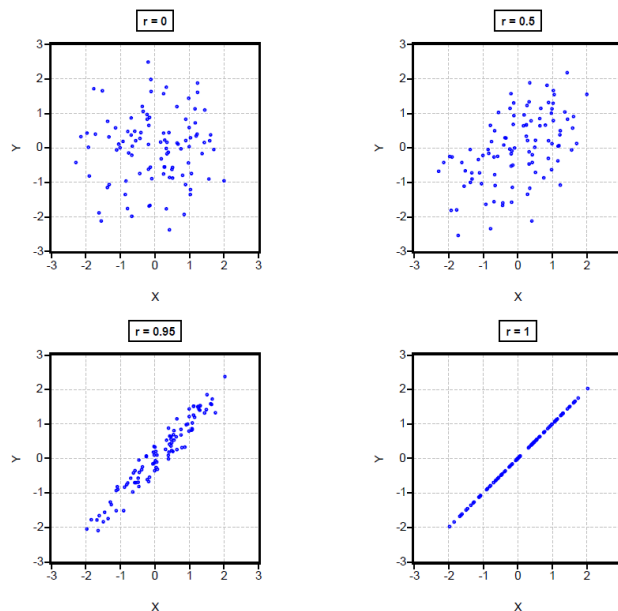
El gráfico de caja utiliza una sola escala correspondiente a la de la variable de los datos que se están analizando. La caja es un rectángulo que abarca el recorrido entre el primer (marcado con la letra e en la **Figura 12**) y tercer cuartil (marcado con la letra d en la **Figura 12**). Es decir, el rango en el cual se abarca el 50% de las observaciones

centrales. En el centro de dicha caja se traza con una línea el valor de la mediana en la escala de los datos (marcada con la letra a en la **Figura 12**).

Por fuera de la caja se dibujan líneas verticales o bigotes que representan los datos que se encuentran por fuera del rango de los datos centrales. Los límites de los bigotes son los valores máximo y mínimo de en el rango de las observaciones (marcados en la **Figura 12** con las letras b y c respectivamente).

El último de los métodos descriptivos que se utilizan en el presente trabajo es la matriz de correlaciones. La matriz de correlaciones es una matriz cuadrada y simétrica que tiene unos en su diagonal y fuera de ella los coeficientes de correlación entre las variables. Según (Lahura, 2003) el coeficiente de correlación es un instrumento estadístico que permite establecer la fuerza y dirección de una relación lineal estadística entre dos variables. El coeficiente de correlación se caracteriza por tomar valores cerrados en el intervalo entre -1 y 1.

El signo del coeficiente de correlación indica la dirección de la relación lineal, así valores positivos indican una relación directa entre dos variables y valores negativos indican una relación inversa. Por otro lado, el valor absoluto del coeficiente de correlación indica la fuerza de la relación lineal. Un coeficiente de correlación muy cercano a uno en su valor absoluto indica una fuerte relación lineal entre las dos variables, mientras que si es cercano a cero indica que la relación es muy débil. Para ilustrar lo anterior la **Figura 13** muestra el grado de asociación lineal con diversos valores del coeficiente de correlación lineal.



**Figura 13:** Distribución de Datos para diferentes coeficientes de correlación  
**Fuente:** (Lahura, 2003)

## 2.5.2 Métodos Clasificatorios.

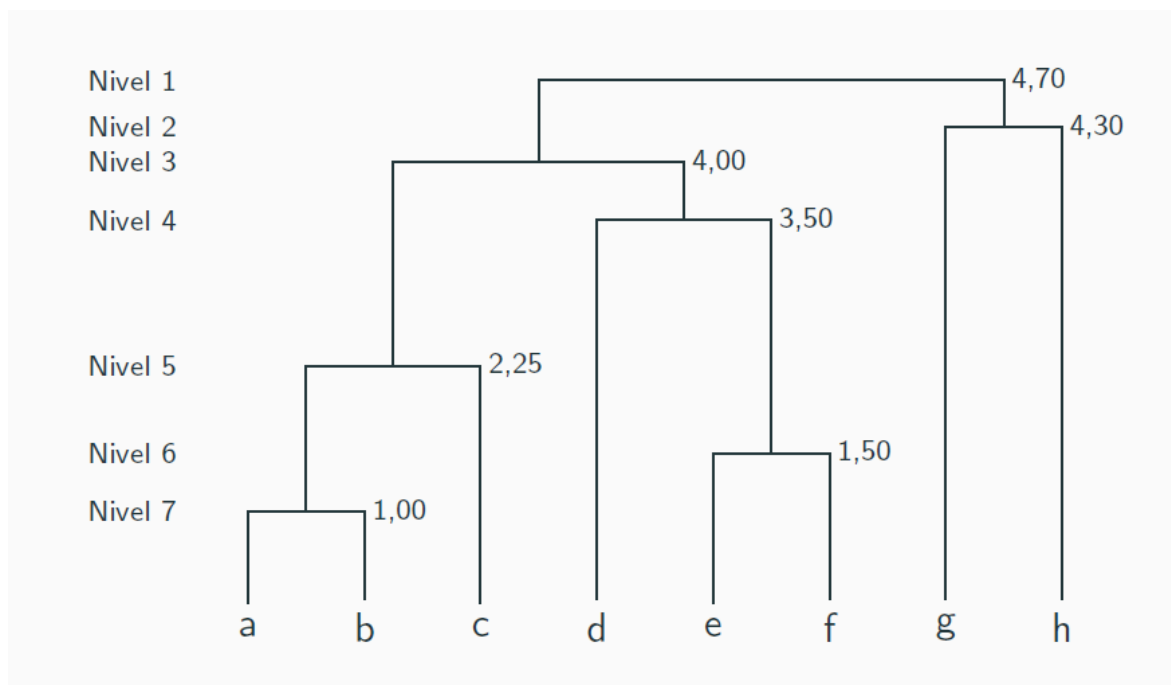
Según (Chiu & Tavella, 2008) la clasificación es una tarea que consiste en asignar objetos a una o más categorías predeterminadas. Para ello, se tiene como datos de entrada una colección de datos. Cada dato puede expresarse como una dupla o vector (x,y) donde x es el dato, atributo o propiedad de interés y la y es la categoría asignada a dicho atributo. Es importante mencionar que a diferencia de los modelos predictivos o de regresión, las categorías y son variables discretas y conocidas.

El objetivo de un modelo clasificatorio es poder asignar una categoría y a los atributos x que se ingresen al modelo. Para ello, se utilizan atributos de entrenamiento, para los cuales ya se conoce previamente la categoría a la cual pertenecen. Una vez se establecen patrones y características comunes de los datos relacionados a cada atributo, puede asignarse una categoría determinada.

Para efectos del presente trabajo se analizarán dos modelos clasificatorios los cuales son la clusterización jerárquica y el modelo de cartas de control.

### 2.4.2.1. Clusterización

El proceso de clusterización o *clustering* es un método de clasificación en el cual se construye una estructura en la que los elementos se agrupan en subconjuntos cada vez mayores hasta que todos pertenecen al mismo conjunto (Exposito, 2020). El clustering jerárquico suele representarse a través de un dendograma (**Figura 14**) que muestra en que orden se han unido los clusters y cuál es el grado de proximidad entre ellos una vez se unen.



**Figura 14:** Dendograma.  
Fuente: (Exposito, 2020)

### 2.4.2.2. Cartas de control

Por su parte, las cartas de control son un método muy usado para el control de calidad de procesos, en el cual pueden establecerse límites de aceptabilidad basados en el histórico de datos que se tiene de determinada variable. Según (Montoya-Restrepo &

Correa-Morales, 2009) la variabilidad en la calidad de un producto puede ser asignada a causas aleatorias o a causas especiales o asignables sobre las que se puede tener algún grado de control. Una de las principales ventajas de las cartas de control es su capacidad de distinguir las causas aleatorias de las causas asignables.

Las cartas de control se clasifican de acuerdo con el tipo de la característica de la calidad monitoreada. Pueden encontrarse cartas de control por variables y cartas de control por atributos. En las cartas de control por variables se lleva el registro real de una característica de la calidad, son más sensibles y alertan rápidamente sobre problemas de calidad. En las cartas de control por atributos, no hay medición continua sobre una característica de la calidad y solo se desea identificar si el producto se adapta al estándar establecido. (Montoya-Restrepo & Correa-Morales, 2009)

Particularmente, una carta de individuales analiza variables continuas que se emplea en procesos en los que se dificulta la toma de muestras masivas. Este diagrama implica el registro de cada medición particular de la variable se registra en la carta.

Esta carta de control está compuesta por diferentes medidas como los límites de control, el centro y la desviación. El centro corresponde con la estimación de la media  $\bar{X}$ , cuya formulación matemática se aprecia en la **Ecuación 1** que a su vez, corresponde con el promedio de las observaciones (Gutierrez Pulido & De la Vara Salazar, 2009).

La desviación  $\sigma_x$  se calcula según la **Ecuación 2**:

**Ecuación 2:** Cálculo de la Desviación

$$\sigma_x = \frac{\bar{R}}{d_2} = \frac{\bar{R}}{1.128}$$

Donde  $\bar{R}$  corresponde a la media de los rangos móviles de orden 2 y  $d_2 = 1.128$  es una constante.

Los límites de control, aquellos valores que representan los límites “habituales” de variabilidad del sistema, se calculan a partir de la **Ecuación 3**:

**Ecuación 3:** Límites de control

$$\bar{X} \pm 3\sigma$$

### 2.5.3 Métodos Predictivos.

Los métodos predictivos o de regresión son un área de la minería de datos que consiste en la extracción de información existente en los datos y su utilización para predecir tendencias y patrones de comportamiento, pudiendo aplicarse sobre cualquier evento desconocido ya sea en el pasado, en el presente o en el futuro. Para llevar a cabo análisis predictivos es indispensable contar con una cantidad considerable de datos, tanto actuales como pasados. Los datos permiten establecer patrones de comportamiento y así permitir el reconocimiento de patrones (Espino Timón & Martínez Fontes, 2017).

Los modelos de regresión son fundamentales cuando se habla de analítica predictiva. Estos modelos consisten en la deducción de una ecuación matemática que represente las interacciones entre las diferentes variables que se están analizando. Otro de los principales modelos que se utilizan en analítica predictiva es el análisis de series temporales.

Una serie temporal o serie de tiempo es un conjunto de observaciones tomadas periódicamente en un lapso. Usualmente las observaciones están sobre todo el intervalo y pueden ser adquiridas aleatoriamente o con una frecuencia de muestreo fija (Pourahmadi, 2002). En el presente trabajo para el análisis de series temporales se utiliza un modelo de medias móviles.

La media o promedio móvil para cierto tiempo  $t$ , es la media aritmética de las  $k$  observaciones más recientes, a través de la cual se pronostica el valor de la serie en el tiempo  $t + 1$  (Hanke & Wichern, 2010). Este promedio se calcula a través de la **Ecuación 4:**

**Ecuación 4:** Pronóstico a través de Medias Móviles

$$\hat{Y}_{t+1} = \frac{Y_t + Y_{t-1} + \dots + Y_{t-k+1}}{k}$$

Donde

$\hat{Y}_{t+1}$  = valor pronosticado para el período  $t + 1$ .

$Y_t$  = valor real de la serie en el período  $t$ .

$k$  = número de términos en el promedio móvil.

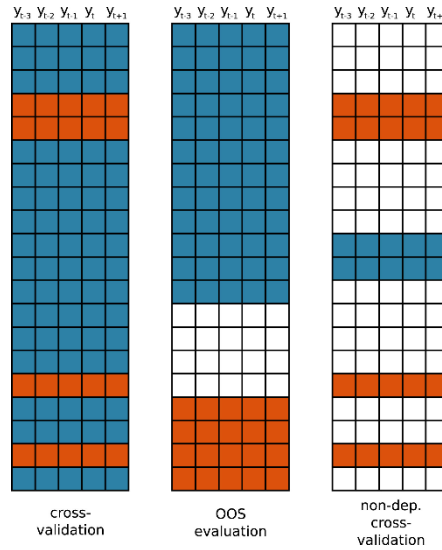
### 2.5.3.1. Validación de Modelos Predictivos.

La validación de modelos estadísticos y de analítica de datos que permiten predecir el comportamiento de un conjunto de datos en periodos de tiempo futuros es un proceso de importancia equivalente a la formulación misma del modelo. Este proceso permite establecer la confiabilidad de las predicciones a través del cálculo del error del modelo. Por otro lado, la validación de un modelo mediante la estimación del error se convierte en una herramienta indispensable para generar comparación entre varios modelos y así seleccionar aquel que presenta los mejores resultados en el contexto de los datos.(Bergmeir & Benítez, 2012)

La validación cruzada es uno de los métodos más importantes para la validación de modelos predictivos y clasificatorios. Uno de los tipos más conocidos de validación cruzada es la validación  $k$ -cruzada (*k-fold cross validation*), la cual consiste en subdividir el dominio en  $k$  subconjuntos. Luego, se realizan  $k$  validaciones del modelo donde cada validación consiste en apartar uno de los subconjuntos generados y utilizar los datos restantes como datos de entrenamiento del modelo. Luego se predicen o clasifican los datos que fueron separados del dominio y se comparan con los valores reales. Por último se promedian los errores obtenidos en cada validación con el fin de obtener un error total del modelo (Bergmeir, Hyndman, & Koo, 2018). En la **Figura 15**, pueden apreciarse diferentes técnicas para la validación cruzada de modelos predictivos, allí las filas azules hacen referencia a datos de entrenamiento, las filas color naranja hacen referencia a los



datos de validación y los datos blancos son datos omitidos por consideraciones de dependencia.



**Figura 15:** Técnicas de validación cruzada para modelos predictivos  
**Fuente:** (Bergmeir et al., 2018)

La técnica de validación cruzada es especialmente efectiva para estimar el error dentro de un modelo cuando este cuenta con pocas datos y no hay estacionalidad en los mismos (Jiang & Wang, 2017). Dicha técnica será utilizada con el fin de validar los modelos predictivos propuestos en el presente trabajo.

2.5.3.2. *Cálculo del error en procesos de validación.*

El propósito principal de la medición y cálculo del error en un modelo es obtener una distribución clara del mismo, llevando a conocer la confiabilidad del modelo. Normalmente el error siempre se calcula a través de una función de pérdida. El error absoluto de un modelo será igual al promedio de la función error en todos los intervalos de validación. Para el cálculo del error consideremos a  $y_t$  los valores observados en un tiempo  $t$ .  $\hat{y}_t$  corresponde al valor pronosticado mediante el modelo para  $y_t$ . La función de pérdida del error se computa entonces como  $E_t = y_t - \hat{y}_t$  (Bergmeir & Benítez, 2012).

Puesto que la predicción generada para determinado evento puede ser mayor o menor que el dato real y el error debe ser siempre positivo, debe considerarse el valor

absoluto de la función de pérdida o bien el cuadrado de esta. A estas modificaciones de la función de pérdida se les llama respectivamente error absoluto o error cuadrático. La **Ecuación 5** y la **Ecuación 6** muestran el cálculo de dichas expresiones para el error.

**Ecuación 5:** Error Absoluto

$$AE_t = |y_t - \hat{y}_t|$$

**Ecuación 6:** Error Cuadrático

$$SE_t = (y_t - \hat{y}_t)^2$$

Existen entonces varias formas de promediar el error obtenido en varios puntos de validación, con el fin de evaluar la exactitud de un modelo predictivo. Las más representativas son el error medio absoluto (MAE, por sus siglas en ingles) y la raíz del error cuadrático medio (RMSE, por sus siglas en ingles). La **Ecuación 7** y **Ecuación 8** la presentan el cálculo de dichas expresiones para la evaluación de modelos estadísticos.

**Ecuación 7:** Error absoluto medio

$$MAE_t = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|$$

**Ecuación 8:** Raíz del error cuadrático medio

$$RMSE_t = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}$$

Es importante resaltar que dichas mediciones del error son dependientes de la escala en la cual se encuentre la variable  $y$ . Con el fin de eliminar la dependencia de la escala, el error puede ser dividido entre el valor de referencia, definiendo así el porcentaje de error. Puede entonces definirse el porcentaje de error medio absoluto (MAPE) y la raíz media del porcentaje de error cuadrático (RMSPE), los cuales se definen de acuerdo con la **Ecuación 9** y la **Ecuación 10** respectivamente:

**Ecuación 9:** Porcentaje de error medio absoluto

$$MAPE_t = \frac{1}{n} \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{y_t}$$

**Ecuación 10:** Raíz Media del Porcentaje de error cuadrático

$$RMSPE_t = \sqrt{\frac{1}{n} \sum_{t=1}^n \left( \frac{y_t - \hat{y}_t}{y_t} \right)^2}$$

El principal problema de las anteriores definiciones es que estas tienden a infinito cuando hay valores de referencia cercanos a cero (datos nulos). Por tal motivo dichos porcentajes de error no deben utilizarse para la validación de modelos que incluyan este tipo de datos.

Para la validación del modelo presentado en el presente trabajo se utilizará el RMSPE como parámetro de evaluación de las predicciones realizadas mediante la técnica de validación cruzada.

## 3. Materiales y Métodos

En el presente capítulo se presentará la metodología utilizada con el fin de identificar y utilizar los modelos estadísticos reseñados en el marco teórico. En primer lugar, se describirá el proceso de adquisición de datos, junto con la descripción de la base de datos recopilada. Luego, se procederá a describir el orden y el motivo por el cual se emplea cada modelo estadístico y su utilidad para analizar los datos y generar modelos de toma de decisiones de mantenimiento.

### 3.1 Datos Recopilados

Los datos recopilados y que sirven como insumo principal al presente trabajo de investigación provienen del programa de análisis de aceite usado de un turbogenerador de vapor Siemens de 18 MW de potencia. En este equipo se toman muestras de aceite con frecuencia mensual y son enviadas a un laboratorio certificado en análisis de aceite usado.

Para la selección del este Turbogenerador como el activo crítico sobre el cual se centrará el análisis del presente trabajo, se tomaron en cuenta las siguientes consideraciones:

- De este activo depende aproximadamente un 30% del suministro eléctrico del complejo industrial donde se ubica. Por lo tanto, es un factor importante en la estructura de costos del negocio.
- Este turbogenerador permite la reducción de presión en el vapor generado por una caldera de recuperación de químicos, desde 600psig a 80 psig, la cual es la presión necesaria para el proceso de producción de papel.
- La disponibilidad requerida para este activo es alta pues únicamente se hace su mantenimiento programado en intervalos de 18 meses, la operación del

mismo es 24/7, razón por la cual el monitoreo por condición del mismo (allí se incluye el análisis de aceite) es crítico.

Es importante resaltar que el aceite de lubricación del turbogenerador de vapor en estudio es de la referencia Mobil SHC 825. Este es un lubricante sintético de alto desempeño con un grado ISO VG 46 y un índice de viscosidad de 145 (Exxon Mobil, 2020). En el anexo A se incluye la ficha técnica de este lubricante. Los datos analizados en el presente trabajo corresponden a muestras de aceite tomadas directamente del depósito de lubricación de la turbina, antes de pasar por cualquier sistema de filtración.

Los datos recopilados comprenden los resultados de análisis de aceite usado desde el año 2016. Lo anterior equivale a 82 registros de datos y 67 variables, cada una de las cuales contiene diferentes propiedades que fueron medidas en el aceite. La **Figura 16** presenta un ejemplo de reporte de análisis de aceite usado. Es importante resaltar que en cada reporte se muestra una tendencia de los resultados obtenidos en los último 5 análisis. Tal y como se muestra en la imagen, el reporte incluye una tendencia para las propiedades principales agrupadas en cinco grupos los cuales son información de la muestra, características del lubricante, partículas de desgaste, partículas contaminantes y concentración de aditivos. Los datos fuente de este análisis fueron extraídos de los reportes de este tipo en un periodo de 81 meses.

Dentro del grupo información de la muestra, se recopilan datos propios de la muestra y no necesariamente relacionados con la condición actual del lubricante. Entre estos datos se encuentran la referencia y edad del aceite, la fecha de la toma de la muestra, el volumen de reposición de aceite en cada periodo, y si dentro del periodo en mención se realizó algún tipo de cambio en el aceite o en los sistemas de filtración. Estos datos sirven para dar un contexto al analista y explicar ciertos cambios en las tendencias de cada una de las variables.

El segundo grupo, hace referencia a las características físico químicas del lubricante. Allí se encuentran parámetros como la viscosidad a 40°C, el índice PQ, el número ácido total, la humedad disuelta y la cantidad de partículas según la norma ISO 4406.

Al tercer grupo corresponden los parámetros que son indicadores de desgaste de alguno de los componentes del mecanismo lubricado. No están relacionados directamente con el estado del aceite, sino que corresponden a partículas contaminantes provenientes de algún componente. Dependiendo de los elementos que sean encontrados en esta categoría, puede diagnosticarse falla en componentes específicos fabricados a partir de materiales que contienen dichos elementos.

Al cuarto grupo pertenecen parámetros relacionados con partículas contaminantes que provienen del ambiente en el cual trabaja el sistema de lubricación. Un aumento en dichas partículas puede corresponder a una posible saturación o ineficiencia en los filtros respiradores del sistema hidráulico, o a contaminación con otros fluidos como refrigerantes o combustibles.

Por último, se encuentra la concentración de aditivos. Una disminución en el contenido de estos elementos significa una pérdida en la concentración de aditivos del lubricante como lo son los aditivos antidesgaste, demulsificantes, detergentes, anticorrosivos, entre otros. La degradación de aditivos se genera por elevado tiempo de servicio del lubricante o por procesos agresivos de recuperación de aceite.

44 Validación de modelos predictivos de analítica de datos de análisis de aceite usado para la toma de decisiones de mantenimiento en un turbogenerador de vapor

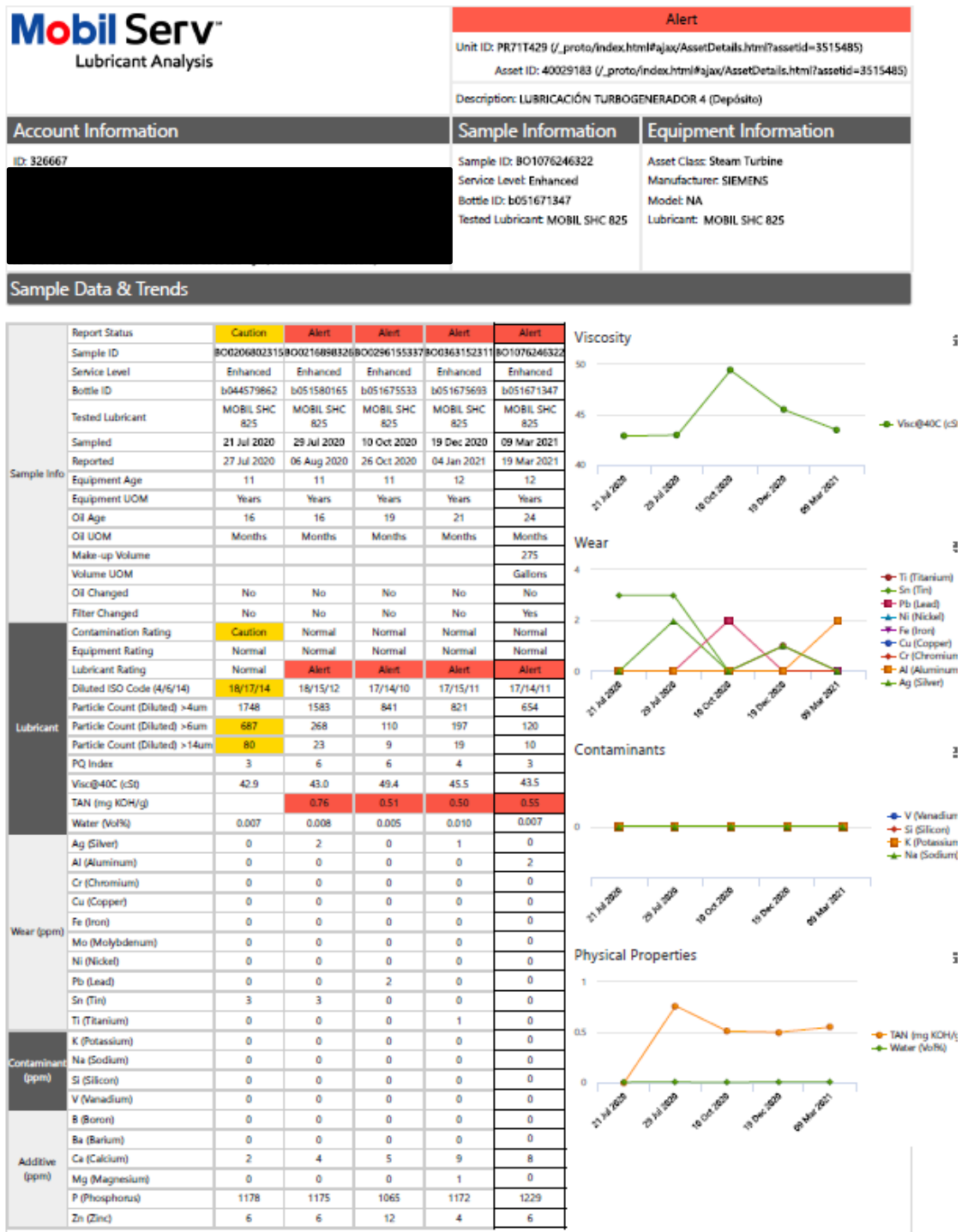


Figura 16: Ejemplo de reporte análisis de aceite usado.  
Fuente: Base de datos anónima

A partir de dichos reportes se construyó la base de datos anteriormente mencionada, con el fin de llevar a cabo los modelos estadísticos que permitieran tener un mejor análisis de la condición actual del lubricante del turbogenerador de vapor.

## 3.2 Modelos Estadísticos

Una vez se contó con la base de datos estructurada, se procedió a realizar el análisis estadístico de estos. El análisis realizado sobre los datos recopilados puede dividirse en cuatro etapas metodológicas. La primera etapa consistió en realizar una caracterización de la base de datos, verificando la presencia de datos nulos o faltantes, se seleccionaron para el análisis aquellas variables representativas que no contaban con datos faltantes. Sobre dichas variables, se realizó un análisis descriptivo, caracterizando los datos mediante diagramas de cajas y bigotes (boxplot), con el fin de identificar las diferentes medidas de tendencia central que representan los datos. Posteriormente, se construyeron las diferentes matrices de correlación entre las variables seleccionadas. Lo anterior buscando encontrar posibles relaciones entre las variables analizadas.

La segunda etapa del análisis consistió en aplicar estrategias clasificatorias a los datos con el fin de generar clusters o grupos de variables relacionadas entre sí. Esto con el fin de analizar los diferentes patrones que pueden derivarse de los datos y adquirir características comunes entre ellos que permitan diferenciar los estados de falla o alerta de las condiciones normales en el aceite. Además, la clusterización permite identificar cuáles variables tienen relación directa con los estados anteriormente mencionados y por tanto deben ser analizadas con especial cuidado.

La tercera etapa metodológica que se implementó en el presente trabajo de investigación, fue la etapa predictiva. Para ello, se emplearon dos estrategias, la primera de ellas fue realizar el análisis de series temporales de los datos existentes en cada una de las variables que fueron previamente preseleccionadas en la primera etapa. Lo anterior, con el fin de analizar gráficamente la existencia de comportamientos que puedan sugerir modelos predictivos de series de tiempo a los cuales se ajusten las variables. Dichas series temporales fueron a su vez contrastadas con los registros de adiciones

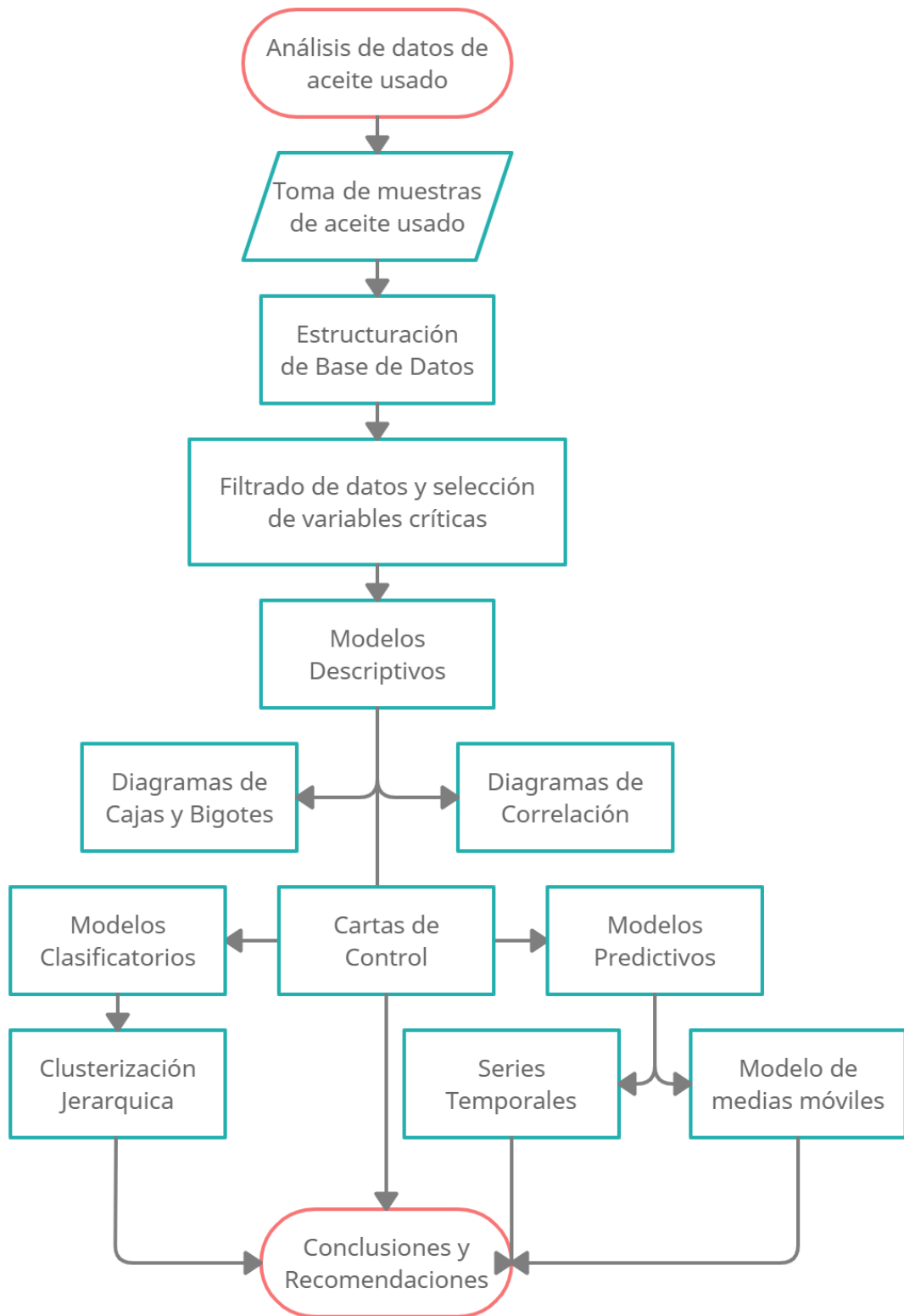


parciales o cambios totales de aceite realizados al sistema con el fin de analizar posibles variaciones en las propiedades del lubricante relacionadas con la renovación de este.

Luego se procedió a utilizar el modelo predictivo sugerido por (Hanke & Wichern, 2010) para el tratamiento de este tipo de datos. Este modelo llamado modelo de medias móviles es utilizado para predecir el comportamiento de una variable según el promedio de los últimos datos consignados en la serie de tiempo. Luego de aplicar dicho modelo, se comparó con la serie temporal, con el fin de identificar el ajuste real de las predicciones realizadas estadísticamente, junto con los datos reales obtenidos como resultado del programa de análisis de aceite usado.

Por último, la cuarta etapa metodológica consistió en la implementación del modelo de cartas de control, el cual posee algunas de las ventajas de los modelos clasificatorios y algunas de las ventajas de los modelos predictivos. Tal y como se mencionó anteriormente las cartas de control son utilizadas en calidad para generar alarmas que permita identificar cuando un dato está fuera de los intervalos normales para los parámetros analizados. Lo anterior, permite crear límites de alarma reales basados en el comportamiento de los datos.

Los pasos metodológicos descritos anteriormente logran apreciarse en la **Figura 17**. Allí puede observarse de manera gráfica los diferentes modelos que fueron utilizados y como ellos se relacionan en sus diferentes niveles con el fin de construir las diferentes conclusiones y recomendaciones de este estudio presentadas en la Sección 5.



**Figura 17:** Diagrama de flujo metodológico.  
**Fuente:** Elaboración Propia

Es importante resaltar en este punto que el software utilizado para la aplicación de los modelos antes mencionados fue el software R Studio, el cual es un software diseñado como banco de trabajo del lenguaje R. Estos son lenguajes de programación de código libre para cómputos estadísticos y generación de gráficos por lo que se configura como de los lenguajes más utilizados por los científicos de datos en la actualidad("R: The R Project for Statistical Computing," n.d.).

### 3.2.1 Validación de Modelo Predictivo.

Con el fin de realizar la validación del modelo predictivo propuesto en el presente trabajo, el cual es un modelo de medias móviles, se implementó la técnica validación cruzada como la propuesta en la sección 3.2.5.3.1. Para ello se tuvo en cuenta lo propuesto por (Bergmeir et al., 2018), quienes sugieren el uso de un  $k=5$  (5-fold cross validation) para hacer validación de modelos predictivos con un número de datos restringidos.

Posteriormente se procedió a dividir el dominio en 5 grupos de manera aleatoria y a realizar el cálculo del porcentaje de error cuadrático medio (RSMPE) para los datos de cada grupo tomándolos como datos de validación en 5 etapas diferentes.

Puesto que el modelo de medias móviles exige que exista una cantidad de datos previos para la generación de la predicción, los primeros 6 datos de la muestra fueron apartados en un grupo aparte. En cada etapa de validación este primer conjunto de datos fueron tomados como datos de entrenamiento del modelo.

La **Figura 18** muestra mediante un mapa de colores la configuración de dichos grupos. En ella, cada celda representa un dato y el color de dicha celda indica el grupo que le fue asignado. Cada etapa de validación consiste en elegir uno de los grupos como grupo de validación y dejar los demás grupos como grupos de entrenamiento del modelo. Se realizan predicciones utilizando el modelo seleccionado para los datos que fueron seleccionados como de validación en dicha etapa y se procede a hacer el cálculo del RMSPE en dichas entradas.



**Figura 18:** División del dominio para validación cruzada de orden 5  
**Fuente:** Elaboración Propia.

Una vez realizadas las 5 etapas de validación para cada variable, se procede a promediar mediante media aritmética los porcentajes de error encontrados, obteniendo el error total del modelo. Con el cual podría validarse la confiabilidad del modelo propuesto para hacer predicciones en cada una de las variables consideradas en este estudio.

## 4. Resultados y Análisis

A continuación, se procede a presentar los resultados obtenidos según la descripción metodológica anteriormente descrita. También se presentará el análisis de los resultados, dicho análisis permitirá plantear las conclusiones del trabajo de investigación, las cuales corresponden a los objetivos planteados para la confirmación de la hipótesis de investigación.

### 4.1 Análisis Descriptivo

Tal y como se mencionó anteriormente en la sección 3.1, la base de datos utilizada cuenta con 82 registros y 67 variables diferentes. Se procedió en primer lugar a realizar un análisis de los datos faltantes para cada una de las variables. En total, hay 67 registros con datos faltantes correspondientes al 81.7% de la muestra. Por tal razón, para el análisis estadístico, se seleccionaron 4 variables como las de mayor relevancia: TAN (mg KOH/g), P (Phosphorus), Visc@40c (cSt) y Water (Vol%). Dichas variables no presentan faltantes.

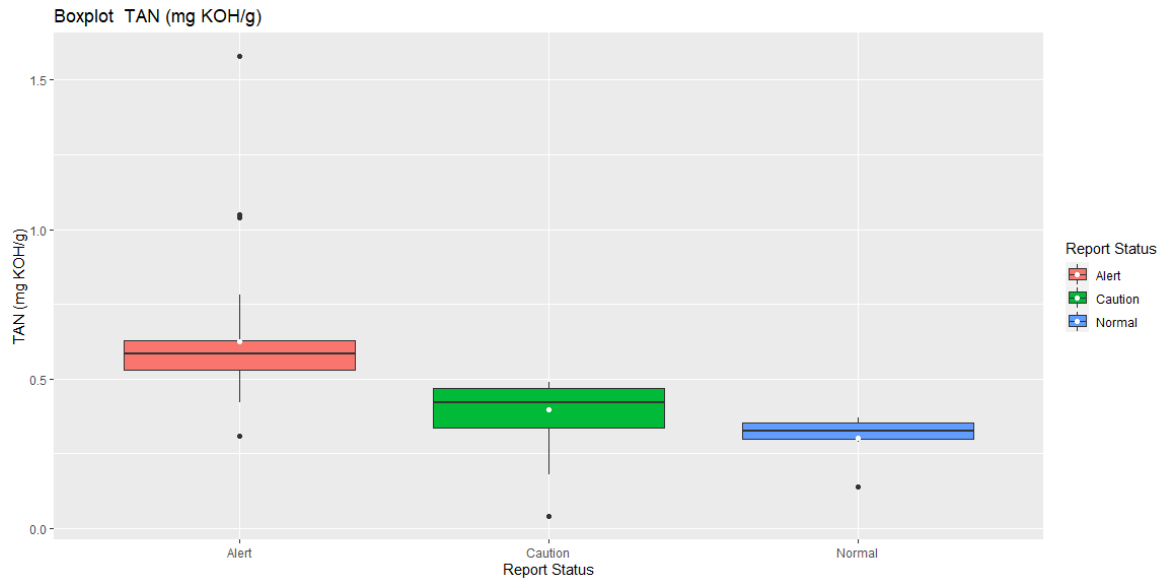
Pese a que otras variables como el código de limpieza (código ISO), el índice PQ y el contenido de partículas de desgaste también podrían contener información valiosa para entender la condición actual del equipo. Es importante tener en cuenta que el objetivo del presente trabajo es formular modelos estadísticos que permitan generar predicciones y clasificaciones sobre los datos futuros. Para tal fin es muy importante contar con la mayor cantidad de datos de entrenamiento posible. Los modelos propuestos en este trabajo pueden aplicarse a otras variables una vez sean formulados. Es por ello que como recomendación de este trabajo queda la aplicación de estos modelos a otras variables críticas obtenidas en los análisis de aceite de este activo.

### 4.1.1 Análisis de las variables seleccionadas

El análisis de las variables seleccionadas parte de los datos categorizados en función de la variable "*Report Status*". De la muestra, 38 registros aparecen marcados como *Alert*, 38 como *Caution* y 6 como *Normal*. El estado denominado Normal es aquel en el cual las condiciones del lubricante están dentro de los rangos de operación recomendados por el fabricante y por lo tanto permiten una operación confiable del equipo. Por otro lado, el estado denominado *Caution* representa condiciones en las cuales las propiedades medidas mediante los análisis de aceite están fuera de los rangos recomendados y se recomienda tomar acciones para el restablecimiento de las condiciones del lubricante. Finalmente, en los registros marcados como *Alert* las propiedades han excedido considerablemente los límites permisibles y por lo tanto no puede garantizarse una operación confiable del equipo bajo estas condiciones.

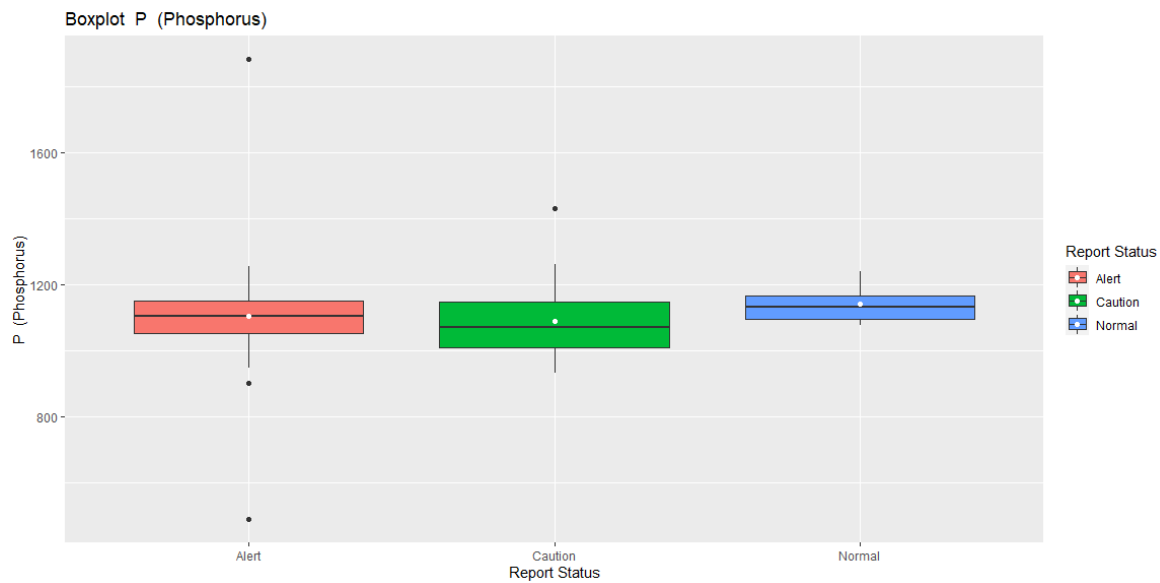
Se elaboraron diagramas de cajas (boxplot) para observar la distribución de los registros en cada una de las variables "relevantes" según el estado reportado en "*Report Status*". En los diagramas que serán presentados en las **Figura 19 a la Figura 22**, las cajas de color azul representan los estados *Normal*, las cajas color verde representan los estados de *Caution* y las cajas color rojo representan los estados de *Alert*.

El diagrama de cajas y bigotes para el TAN (mg KOH/g) se muestra en la **Figura 19**. En él se observa una diferencia marcada entre el estado de alerta y las otras dos categorías (las cajas no se superponen) y, tanto sus medias como medianas parecen diferir considerablemente. Esto puede sugerir que la variable en cuestión puede estar relacionada con el estado de alerta. Sin embargo, no se evidencia diferenciación entre el estado "*Caution*" y "*Normal*" pues las cajas de estas dos categorías se superponen entre sí. De lo anterior, puede inferirse que la variable TAN es determinante para la aparición de registros tipo *Alert*, razón por la cual debe ponerse especial atención a las condiciones operacionales del equipo que puedan favorecer procesos de acidificación del aceite. Algunos ejemplos de estas condiciones son el contenido de agua disuelta, la temperatura de operación, la formación de barnices en el lubricante, entre otras.



**Figura 19:** BoxPlot para el TAN  
**Fuente:** Elaboración Propia

Para la variable P (Phosphorus) el diagrama de cajas y bigotes se presenta en la **Figura 20**. En él no se observa diferencia entre los 3 reportes de estado pues los cajones que representan cada estado se superponen. Esta variable no parece estar relacionada con los estados reportados en el conjunto específico de datos. Sin embargo, en la gráfica logra apreciarse que la mediana de los datos en estado *Normal* tiene un valor más alto que el de los otros dos estados. Lo anterior puede deberse a que los registros *Normal* aparecen en su mayoría luego de intervenciones en el equipo que implican reposiciones parciales o cambios totales del lubricante, razón por la cual la concentración de aditivos metálicos aumenta en ese momento, aumentando por ende el contenido de fósforo en el aceite.

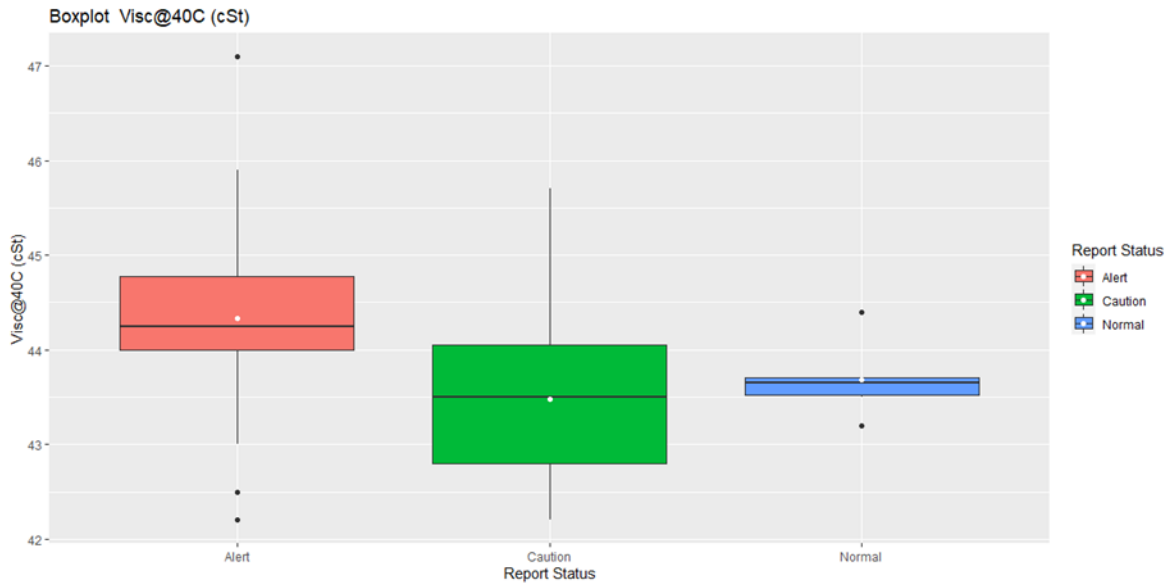


**Figura 20:** BoxPlot para P (Phosphorus (ppm))  
**Fuente:** Elaboración Propia.

Para la viscosidad a 40°C, el boxplot correspondientes se muestra en la **Figura 21**. Para esta variable parece presentarse una diferencia entre el estado *Alert* y *Normal*, pero esto puede obedecer a la poca cantidad de registros pertenecientes a la categoría *Normal*. Además, es importante mencionar que el grado ISO VG del aceite aquí analizado es un ISO VG 46, y que dicha clasificación permite un margen de 10% en ambos límites del intervalo. Lo anterior quiere decir que todo aceite en un rango entre 41.6 cSt y 50.6 Cst cumple con la viscosidad necesaria para la aplicación.

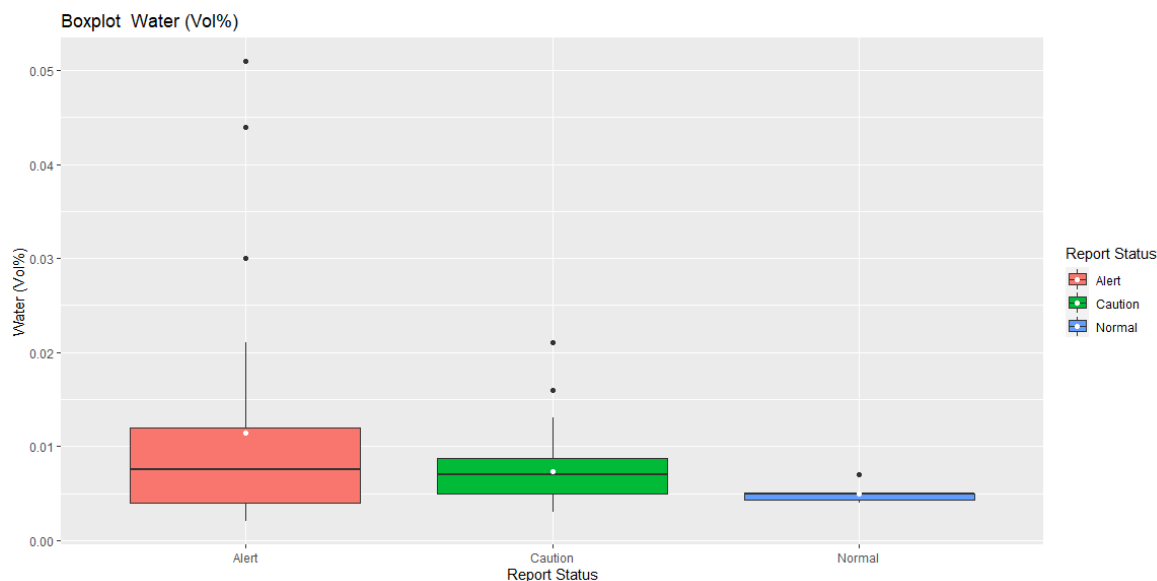
Lo anterior implica que incluso los valores más extremos medidos a las muestras aquí presentadas tienen un valor acorde a los requerimientos de la norma. Por esta razón no hay una relación directa entre los datos de viscosidad y los estados de alerta que se encuentran en el reporte.





**Figura 21:** Boxplot para la Viscosidad a 40°C  
**Fuente:** Elaboración Propia.

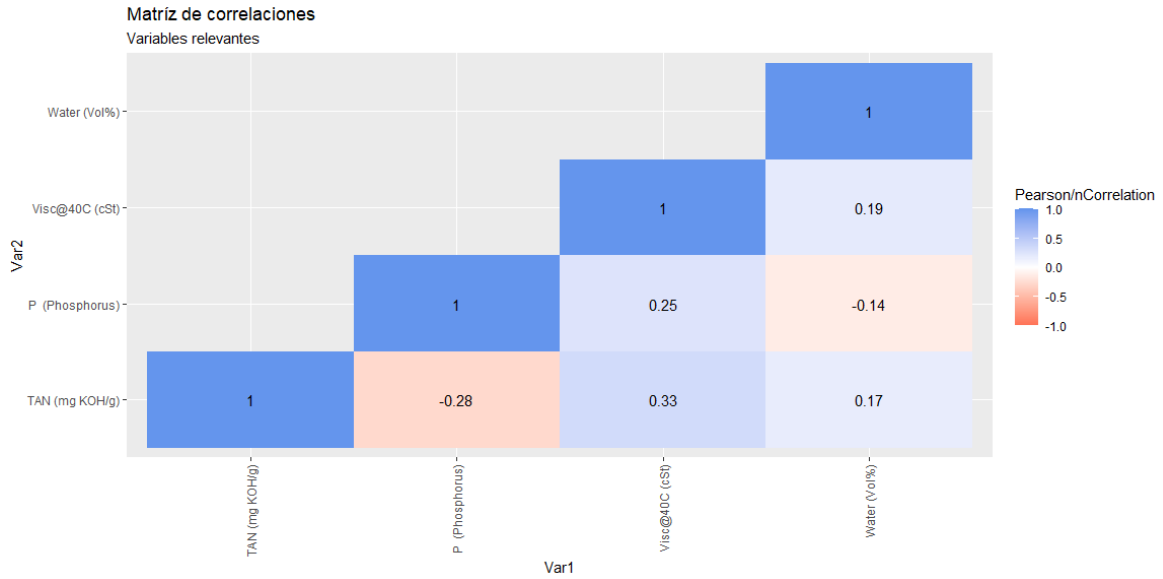
Finalmente, la **Figura 22** muestra el diagrama de cajas y bigotes para el contenido de agua en el aceite (Vol%). No se evidencia diferencia entre los estados así que esta variable no está relacionada con los estados reportados en el conjunto específico de datos. Pese a lo anterior, se logra evidenciar que los registros en estado *Alert* presentan una mayor variabilidad (cajas más amplias), lo anterior puede reforzar la hipótesis de que variaciones en el contenido de humedad en el aceite pueden promover una acidificación en el mismo, generando posibles condiciones de alerta.



**Figura 22:** Boxplot para el contenido de Agua disuelta.  
**Fuente:** Elaboración Propia

De manera general, puede observarse que, para cada una de las 4 variables, la caja correspondiente a los registros en estado *Normal* (caja color azul) es considerablemente más estrecha, lo que implica una menor variabilidad en dicho conjunto de datos. Lo anterior es de esperarse, pues dichos datos se encuentran en general dentro de los rangos recomendados por el fabricante del lubricante, por tanto debe haber una mayor concentración de ellos.

Por último, con el fin de encontrar posibles relaciones de correspondencia entre las variables que fueron seleccionadas para este análisis, se construyó la matriz de correlación que se muestra en la **Figura 23**. Cabe recordar que, si el valor absoluto de los coeficientes de correlación es cercano a 1, hay una fuerte relación lineal entre las variables analizadas. Si el valor absoluto del coeficiente de correlación es cercano a cero se indica independencia entre las variables analizadas. Por otro lado, valores positivos del coeficiente de correlación, indica que la relación entre las variables es directa; mientras que coeficientes de correlación negativos, indican una relación inversa.



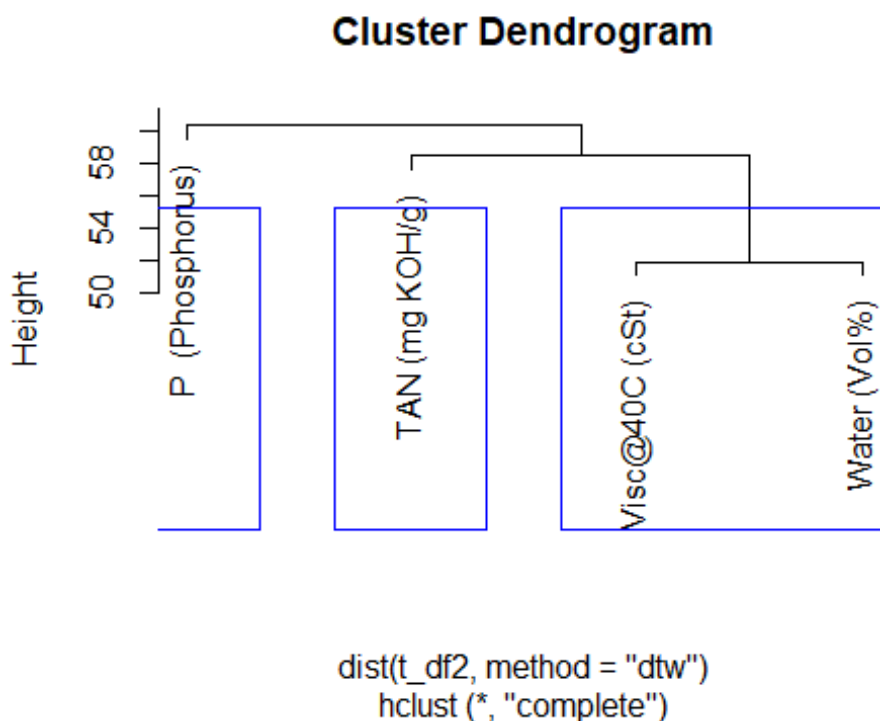
**Figura 23:** Matriz de Correlaciones  
**Fuente:** Elaboración Propia

Se considera que dos variables tienen una correlación lineal cuando el valor absoluto del coeficiente de correlación es mayor que 0.8 (Fernando & Walters, 2021). Como puede apreciarse en la **Figura 23**, ninguno de los coeficientes calculados cumple con dicha condición, por lo que podría considerarse que las variables son independientes. Las variables que tienen una mayor correlación son el TAN y la viscosidad a 40°C. Lo anterior hace sentido pues el TAN es la variable que mide la acidez en el aceite, un aumento en la acidez puede llegar a propiciar una mayor oxidación en el mismo y por tanto un aumento en la viscosidad. Lo anterior, se debe a que un aumento en la acidez, del aceite hace que este sea más propenso a atraer radicales OH y por lo tanto a adquirir una mayor concentración de oxígeno (oxidación). Generalmente, la oxidación en el aceite tiene una relación estrecha con aumentos en la viscosidad del mismo. Sin embargo, lo anterior no debe tomarse como una conclusión puesto que, según la evidencia proporcionada por los datos, dicha correlación es muy débil para poderse considerada real.

## 4.2 Análisis Clasificadorio

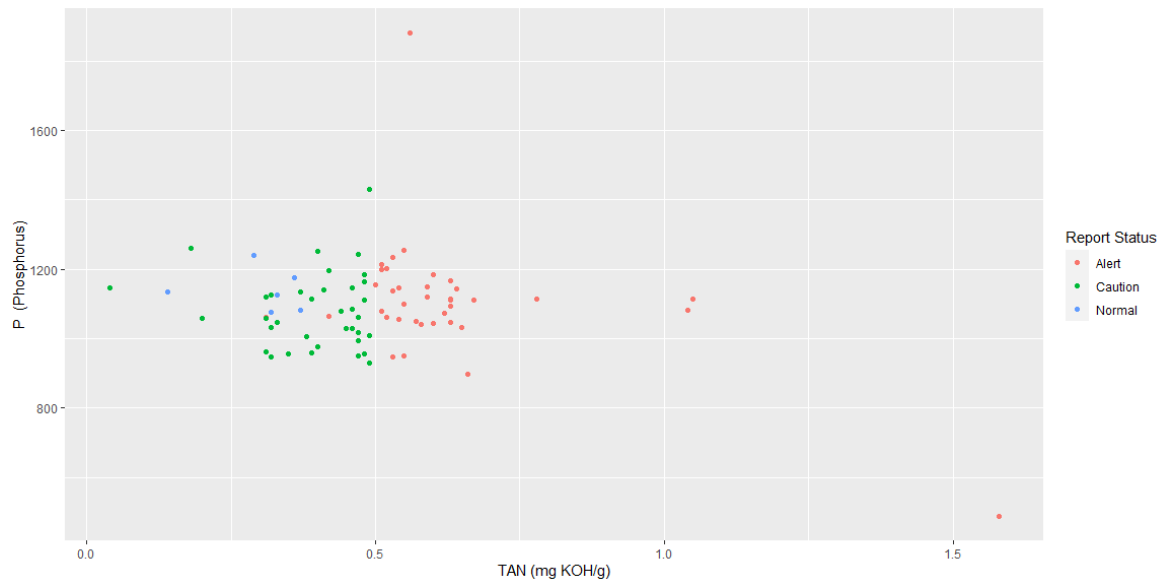
A continuación, se presentará el análisis clasificadorio llevado a cabo sobre los datos de análisis de aceite usado, para tal fin se utilizó el proceso de clusterización jerárquica

reseñado en la sección 2.4.2.1. La **Figura 24** muestra el dendograma realizado sobre las variables seleccionadas. Allí se observa que, dentro del modelo clasificatorio, las variables más íntimamente relacionadas son la viscosidad y el contenido de humedad en el aceite, en orden siguiente se encuentra el número ácido total y por último el contenido de fósforo (aditivos). Lo anterior, es físicamente consistente puesto que las primeras tres variables mencionadas corresponden a propiedades fisicoquímicas. Por otro lado, el fósforo es el principal componente de los aditivos antidesgaste y extrema presión en el aceite. Este tipo de aditivos son los encargados de reducir el desgaste y proteger los mecanismos lubricados en las etapas de arranque.

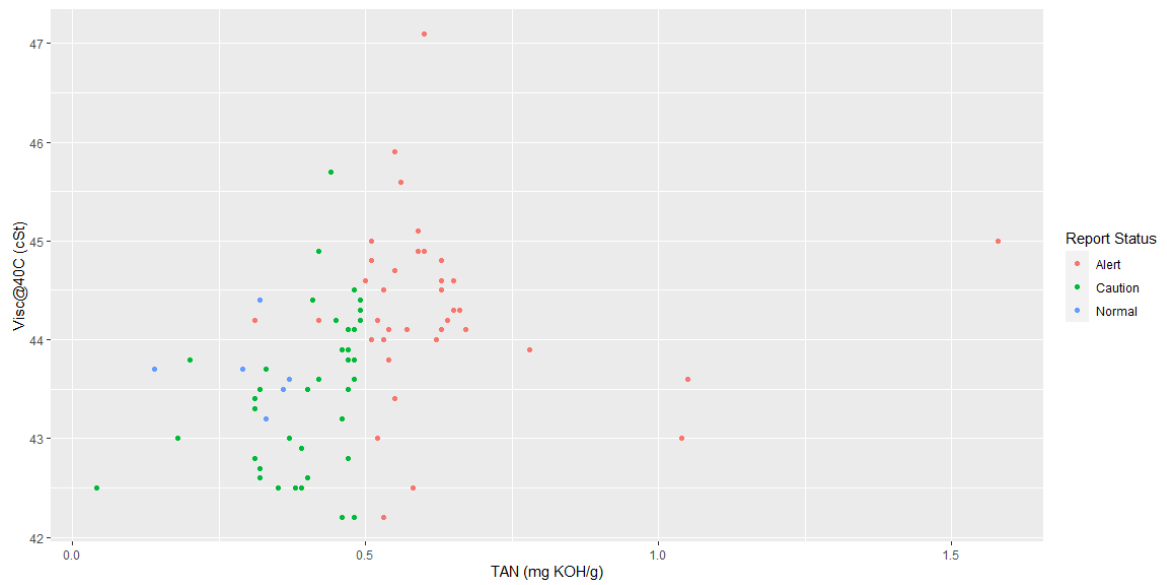


**Figura 24:** Dendograma de clusterización.  
**Fuente:** Elaboración Propia

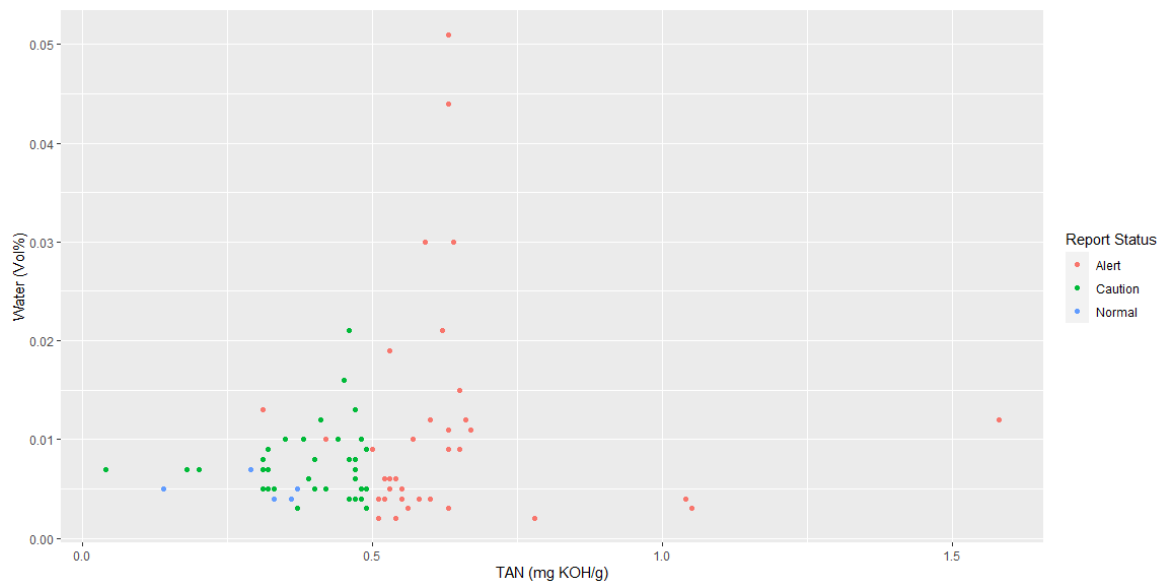
Luego, se generaron gráficos de dispersión relacionales entre las variables seleccionadas, donde se pone nuevamente en evidencia la relación de este conjunto específico de datos con el TAN (mg KOH/g), de modo que se sugiere prestar especial atención a los elementos que pueden afectar esta medida. Las **Figura 25 a laFigura 28** muestran los gráficos de dispersión cruzada entre el TAN y da una de las otras variables analizadas.



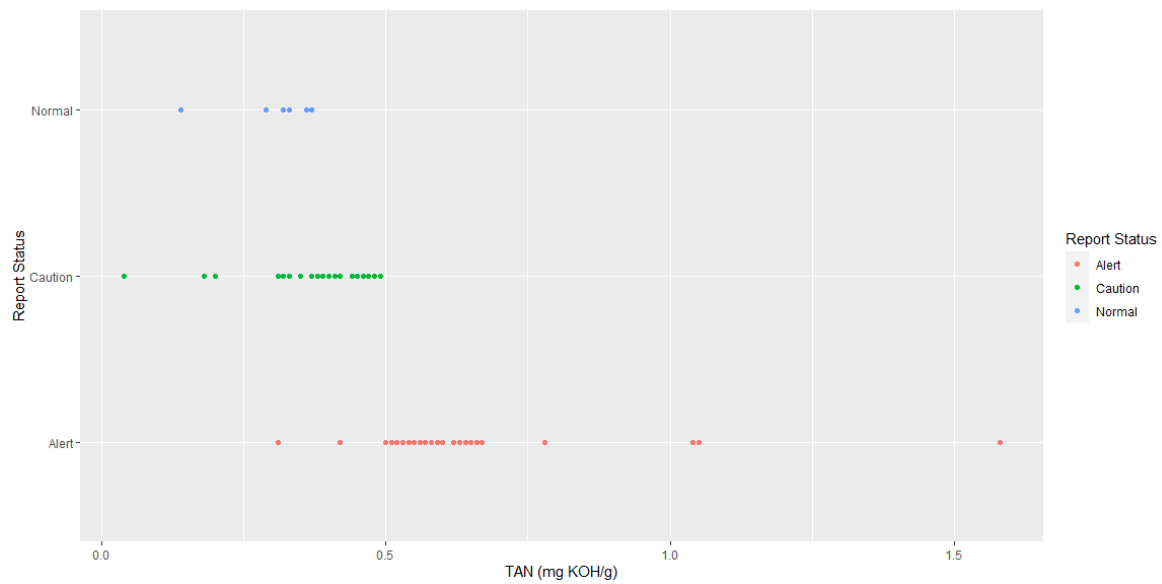
**Figura 25:** Gráfico de dispersión cruzada entre TAN y Fósforo (ppm)  
**Fuente:** Elaboración Propia



**Figura 26:** Gráfico de dispersión cruzada entre TAN y viscosidad a 40°C  
**Fuente:** Elaboración Propia



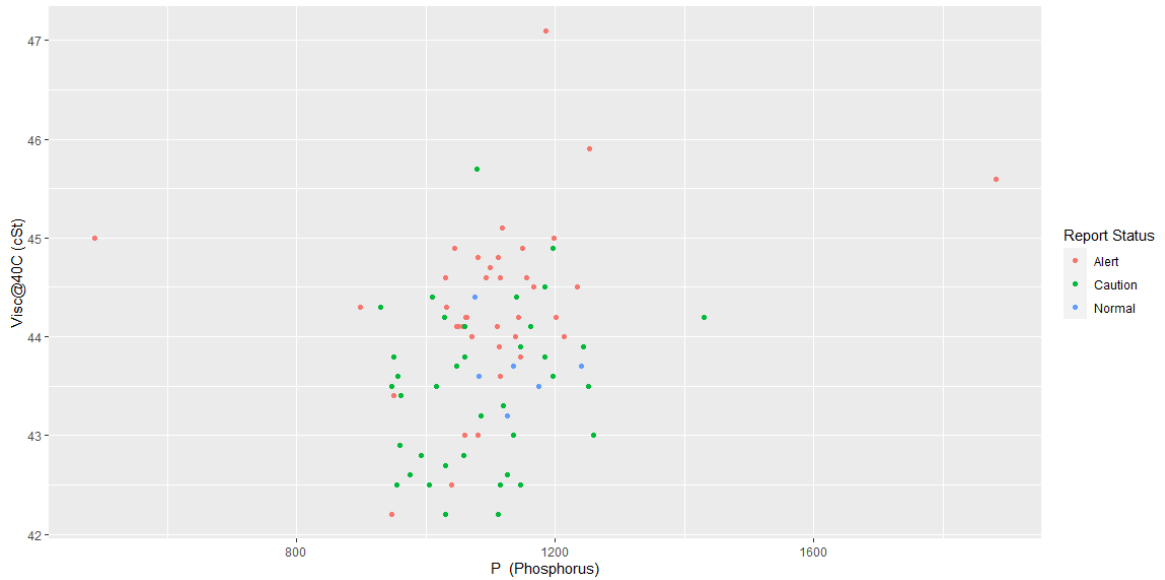
**Figura 27:** Gráfico de dispersión cruzada entre TAN y % de agua disuelta  
**Fuente:** Elaboración Propia



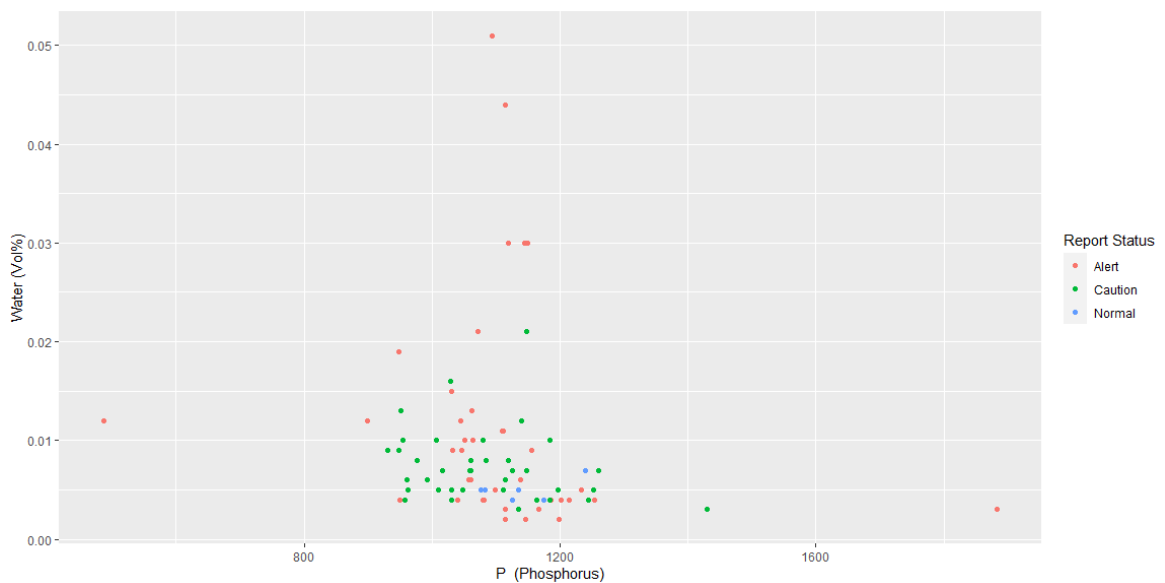
**Figura 28:** Gráfico de dispersión cruzada entre TAN y reporte de status  
**Fuente:** Elaboración Propia

En estas gráficas, puede apreciarse una división marcada en los estados de alerta para la variable TAN con valores superiores a 0.5. Para ninguna de las otras variables se logran apreciar tendencias que permitan discernir entre los estados de normal, precaución y alerta.

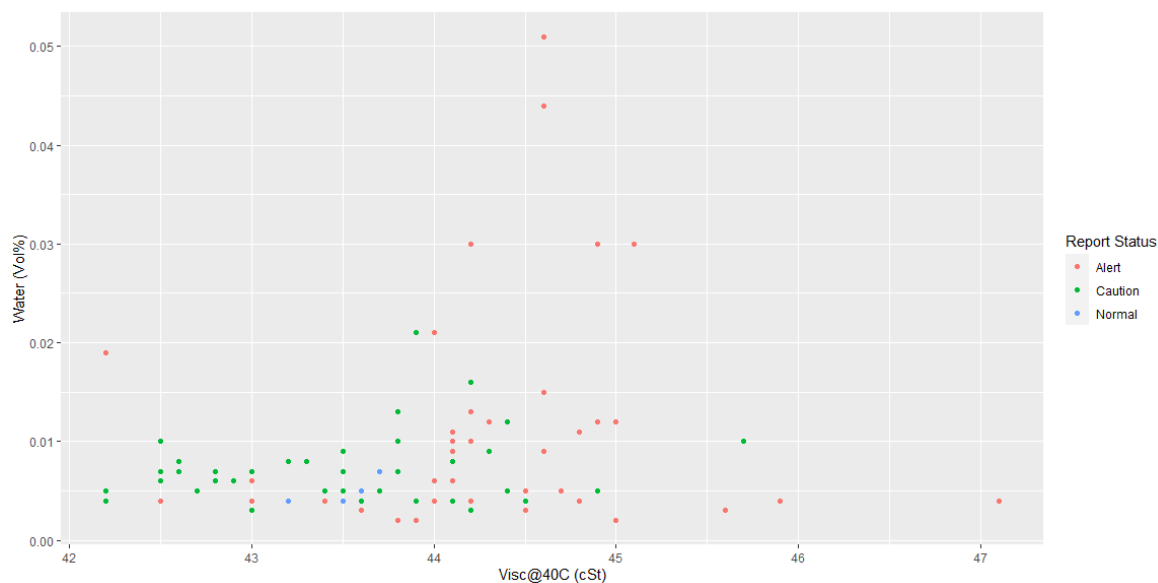
El mismo proceso fue llevado a cabo para cada una de las variables seleccionadas. Las **Figura 29 a la Figura 31** muestran los diagramas de dispersión cruzada, para el resto de las variables analizadas.



**Figura 29:** Diagrama de dispersión cruzada entre P (ppm) y la viscosidad a 40°C.  
**Fuente:** Elaboración Propia.



**Figura 30:** Diagrama de dispersión cruzada entre fósforo (ppm) y % de agua disuelta.  
**Fuente:** Elaboración Propia.



**Figura 31:** Diagrama de dispersión entre viscosidad a 40°C y % de agua disuelta  
**Fuente:** Elaboración Propia

De estos gráficos, puede concluirse que las variables viscosidad a 40°C, % de agua disuelta y contenido de fósforo no presentan una separación clara, ni una relación para los estados *Normal*, *Caution* y *Alert*. Por lo tanto, puede concluirse del análisis clasificatorio que la única variable que aparentemente tiene relación directa con la separación de los diferentes estados es el TAN. Con las otras variables aquí estudiadas no es posible crear un criterio que permita diferenciar los diferentes estados. Es importante resaltar que los estados *Normal*, *Caution* y *Alert* son definidos por el fabricante para referenciar cuando un resultado en cualquiera de las variables medidas sale de determinados límites de control. Es decir, si en un análisis de aceite solo una variable sobre pasa los límites de control y las demás están en rango, el status que el fabricante dará al aceite es de *Alert*. En este caso, y según los resultados discutidos anteriormente, la variable que más comúnmente superó dichos límites de control fue el TAN.

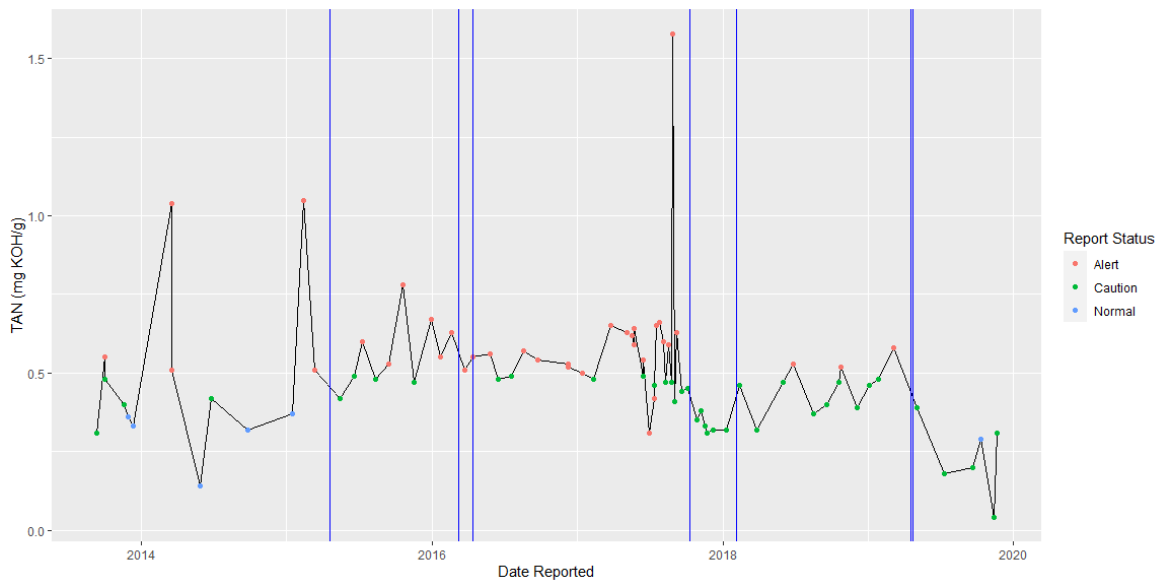
### 4.3 Análisis Predictivo

Para el análisis predictivo se realizaron las gráficas de las series de tiempos para analizar gráficamente la existencia de comportamientos que puedan sugerir modelos predictivos de series de tiempos a los que ajustar las variables. Para facilitar la observación



de los datos, se agregó puntos del color de la categoría reportada. Además, se obtuvo información sobre la adición de aceite o cambios, lo que podría requerir analizar el comportamiento de la serie de manera posterior a cada cambio. Estas adiciones se presentan en las series temporales como líneas verticales de color azul

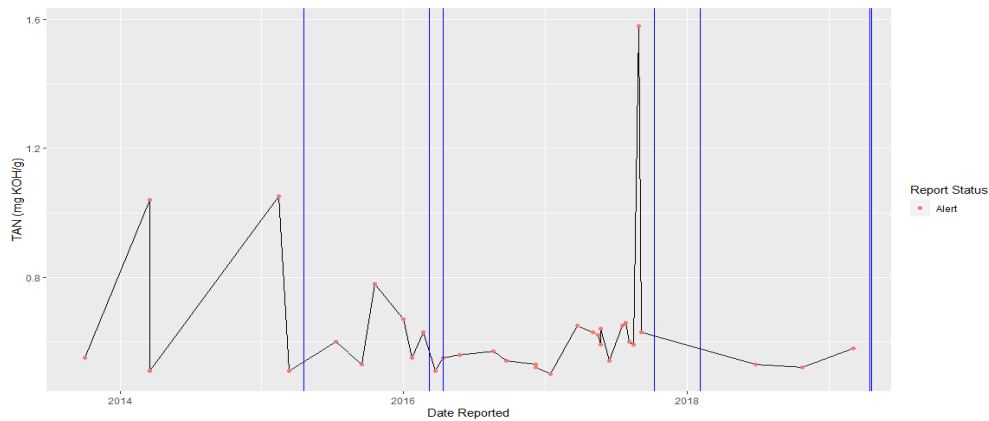
La variable TAN (mg KOH/g), cuya gráfica de serie temporal se muestra en la **Figura 32** muestra cambios relevantes en su comportamiento posteriores a cada adición o cambio de aceite. Sin embargo, parece existir una tendencia creciente que progresa a medida que avanza el tiempo después de cada modificación en el sistema. Una observación importante, es que la gran mayoría de puntos reportados como alerta, están en un valor de TAN superior a 0.5.



**Figura 32:** Gráfica de serie temporal para el TAN

**Fuente:** Elaboración Propia

Si filtramos la base de datos de modo que, los valores incluidos sólo estén por encima o sean iguales a un TAN de 0.5, obtenemos el gráfico mostrado en la **Figura 33**, el cual incluye solo observaciones en alerta, es decir. En total, este gráfico excluye 6 observaciones en las cuales se generó un estado de alerta sin tener un TAN superior a 0.5.



**Figura 33:** Serie temporal para el TAN mayor a 0.5  
**Fuente:** Elaboración Propia

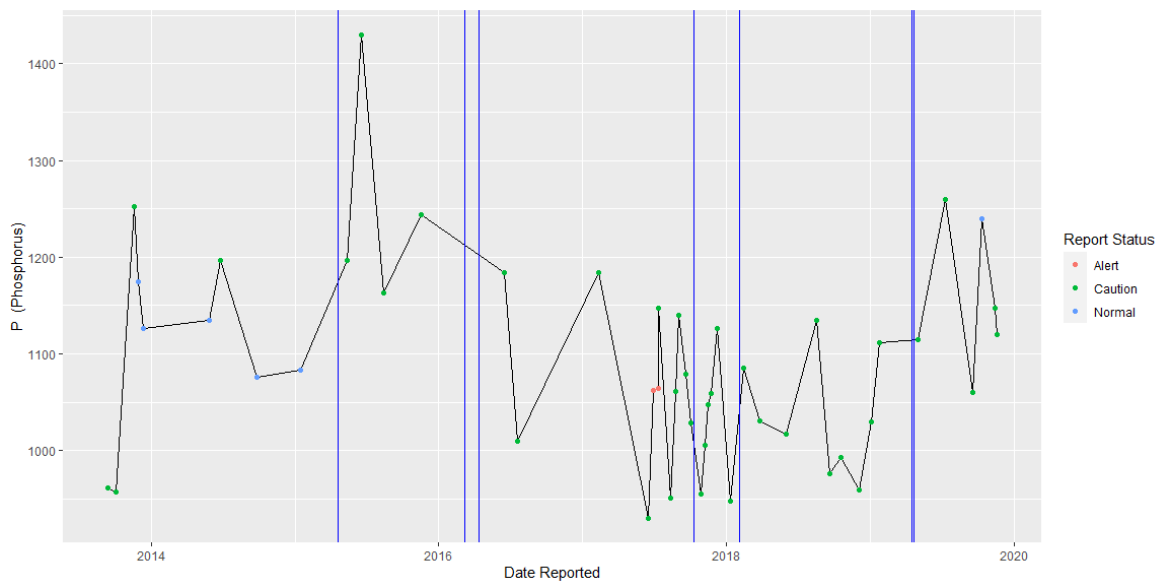
Se procede entonces a analizar las observaciones marcadas como Alert que no corresponden a un TAN igual o superior a 0.5 para verificar el origen del estado de alerta. Esto se analizará en las series temporales de las otras variables principales seleccionadas.

La variable P (Phosphorus), cuya serie temporal se muestra en la **Figura 34**, muestra un comportamiento estacionario en media, es decir, un comportamiento aleatorio que en principio no parece indicar tendencias. Sin embargo, seleccionar un modelo predictivo sin tener en cuenta el comportamiento a mayor plazo del sistema, implicando la variabilidad natural del mismo, puede resultar apresurado. En esta gráfica, inmediatamente después de la tercera adición de lubricante, logra apreciarse un incremento importante (*outlier*) en el contenido de fósforo en el aceite. Lo anterior puede deberse a una mayor concentración de aditivos antidesgaste en el aceite de lubricación que fue adicionado en aquella ocasión.



**Figura 34:** Serie temporal de la variable P (Phosphorus (ppm))  
**Fuente:** Elaboración Propia.

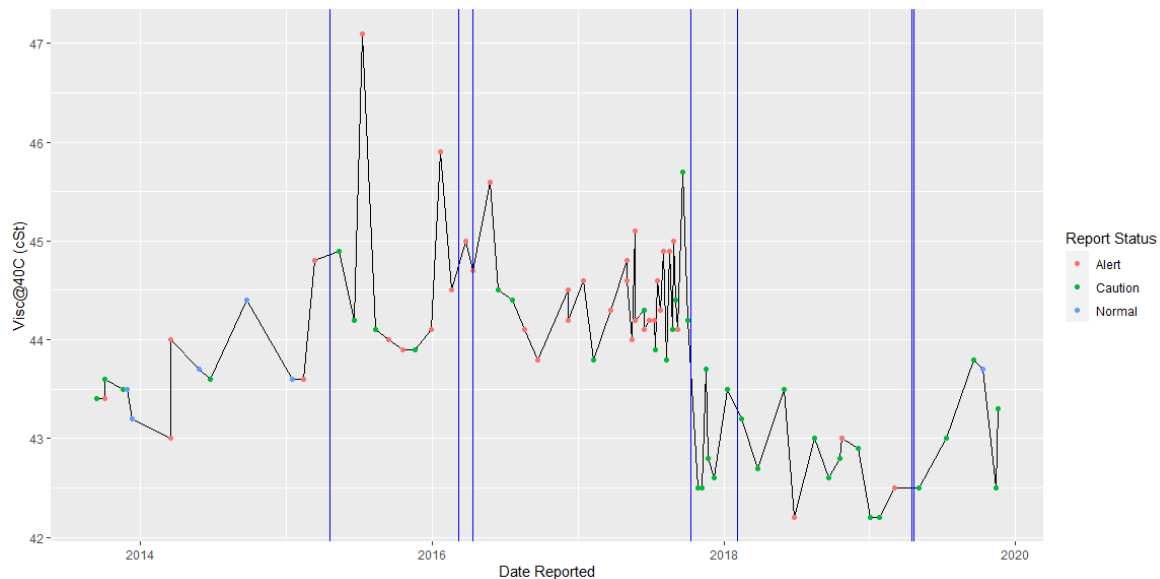
La **Figura 35** muestra esta serie temporal, una vez que se realiza un filtro eliminando las observaciones que tienen valor de TAN mayor a 0.5, observamos que los valores de esta variable no permiten diferenciar las observaciones con señal de alerta.



**Figura 35:** Serie temporal para la variable P (ppm) con filtro de TAN > 0.5  
**Fuente:** Elaboración Propia

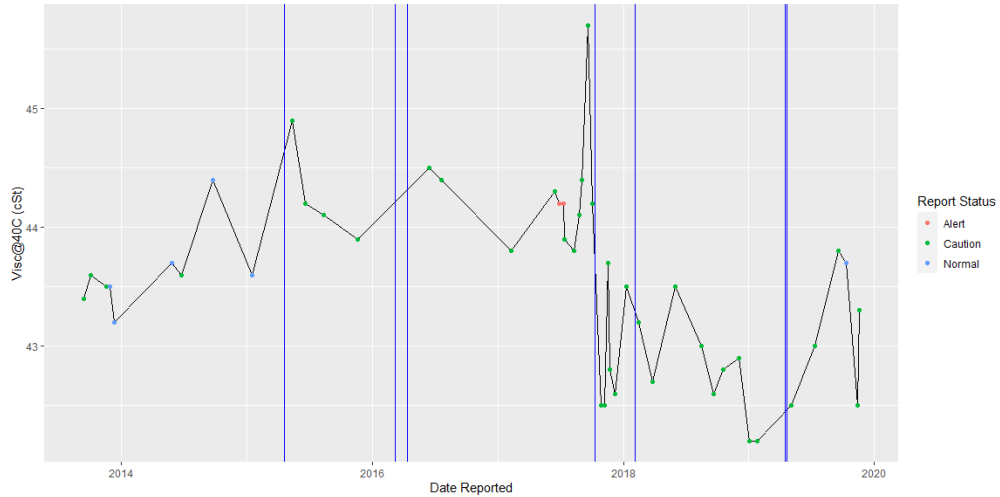
Para la variable relacionada con la viscosidad a 40°C, cuya serie temporal se presenta en la **Figura 36**, se evidencian cambios bruscos en su comportamiento posterior

a las adiciones o cambios en el sistema, presentando tendencias crecientes o decrecientes en diferentes rangos e incluso estacionarias en media. Aquí podemos observar que luego de la adición de lubricante llevada a cabo en el año 2017 se presentó una disminución importante en la tendencia de la viscosidad. Tal y como se ha mencionado anteriormente, las adiciones o reposiciones de lubricante modifican significativamente las propiedades fisicoquímicas del aceite. En este caso particular la viscosidad, aunque sufrió una clara disminución, permaneció dentro de los rangos aceptables de viscosidad recomendados por el fabricante, razón por la cual no se aprecian muchos datos en estado *Alert* en el periodo posterior a la adición.



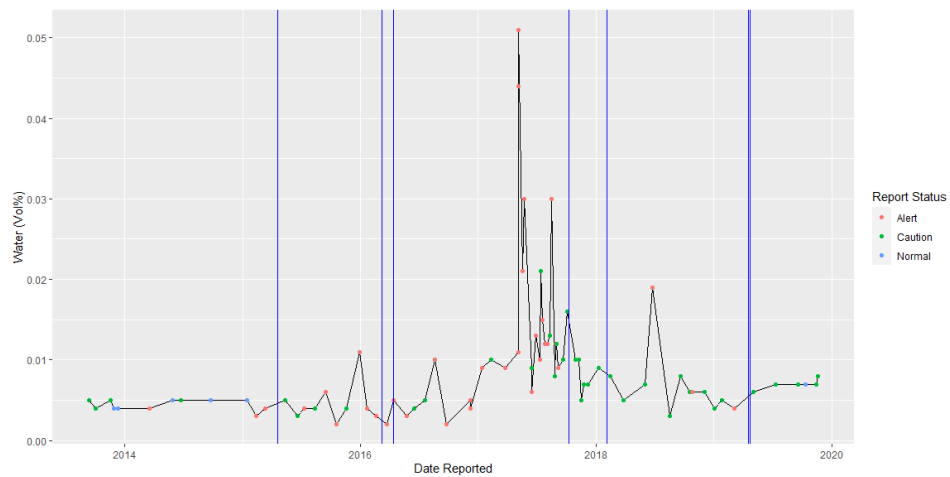
**Figura 36:** Serie temporal para la viscosidad a 40°C  
**Fuente:** Elaboración Propia

Si filtramos la base de datos para visualizar únicamente valores de TAN inferiores a 0.5, observamos que los valores de la variable no permiten diferenciar las observaciones marcadas como alerta. Lo anterior se muestra en la **Figura 37**.



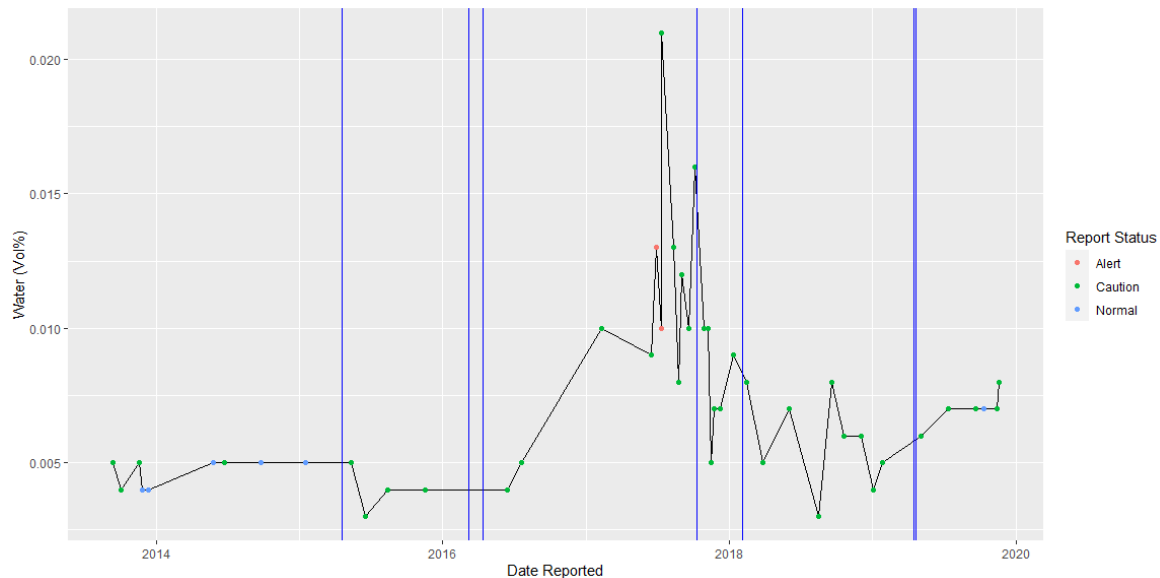
**Figura 37:** Serie de Tiempo para la viscosidad a 40°C. Filtro TAN > 0.5  
**Fuente:** Elaboración Propia

Para la variable **Water (Vol%)** no se observan se evidencia un comportamiento atípico que comienza a mediados de 2017. La gráfica de serie temporal para esta variable se presenta en la **Figura 38**. Aquí, puede apreciarse que hacia finales del año 2017 se presentó un evento de alto contenido de agua en el aceite, el cual al parecer fue tratado con algún tipo de proceso de deshidratación externo que no implicó reemplazo del lubricante (la gráfica disminuye antes de alcanzar la próxima línea azul).



**Figura 38:** Serie temporal para el porcentaje de agua disuelta.  
**Fuente:** Elaboración Propia

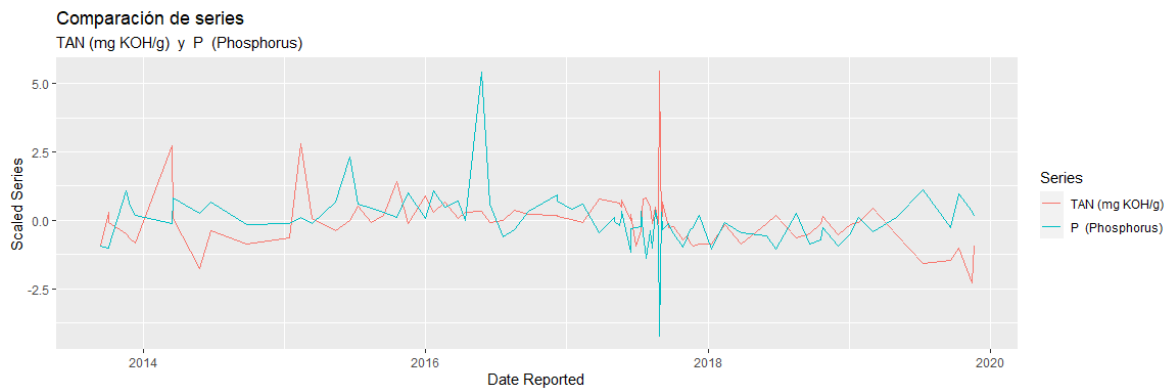
Si filtramos la base de datos para valores de TAN inferiores a 0.5, observamos que los valores de la variable no permiten diferenciar las observaciones marcadas como alerta. Esto puede observarse en la **Figura 39**.



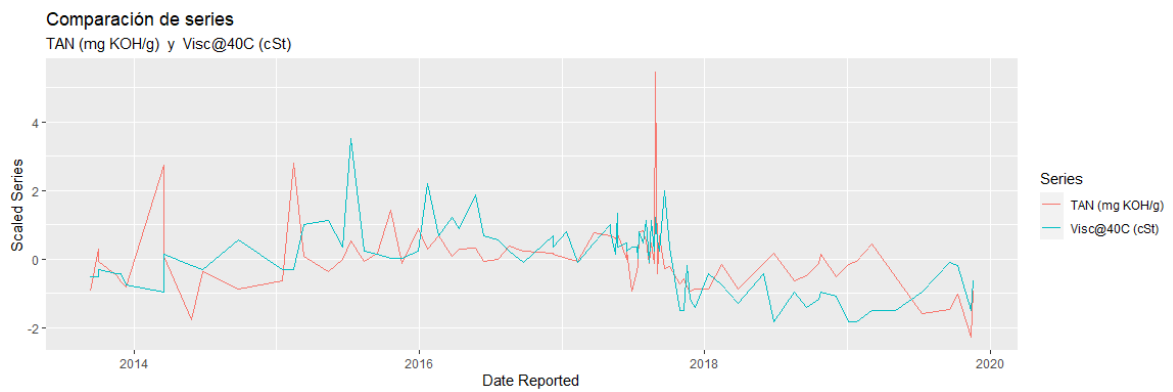
**Figura 39:** Serie Temporal para porcentaje de agua. Filtro TAN > 0.5  
**Fuente:** Elaboración Propia

De manera general, puede observarse que la implementación de modelos predictivos, basados en el conjunto de datos y las variables actuales, puede resultar compleja en cuanto a su implementación en la práctica pues hay variabilidad asociada, no a la naturaleza del sistema, sino a los cambios provocados por elementos externos como adiciones parciales o reemplazos totales de lubricante, implementación de nuevos sistemas de filtración, eventos de contaminación, entre otros. Es posible implementar sistemas de control del comportamiento basados en alertas tempranas ante cambios de tendencia. Por ello se decidió implementar la estrategia de cartas de control para el análisis de dichos datos, permitiendo generar una herramienta de toma de decisiones, ante alertas generadas con base en el conocimiento de datos anteriores.

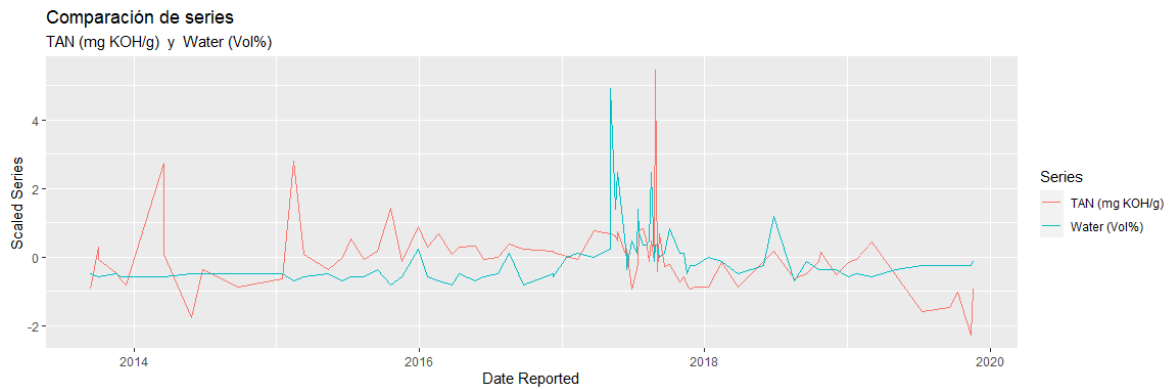
Se realizan gráficos de manera conjunta para pares de variables, usando la función scale para escalar y centrar las series, evitando problemas de escala. Lo anterior se realiza con el propósito de identificar si existe alguna relación aparente entre las tendencias de las variables estudiadas. Las **Figura 40 a la Figura 45** presentan dicho análisis comparativo.



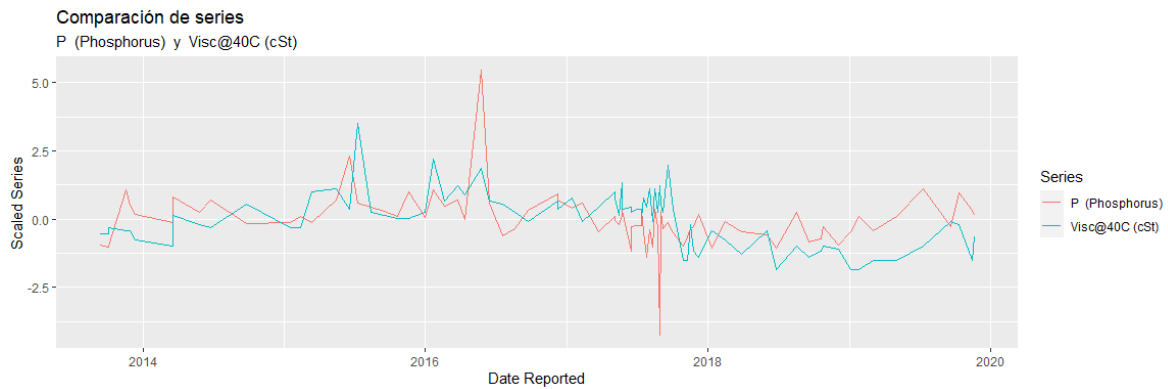
**Figura 40:** Comparativo series de tiempo TAN y P (ppm)  
**Fuente:** Elaboración Propia



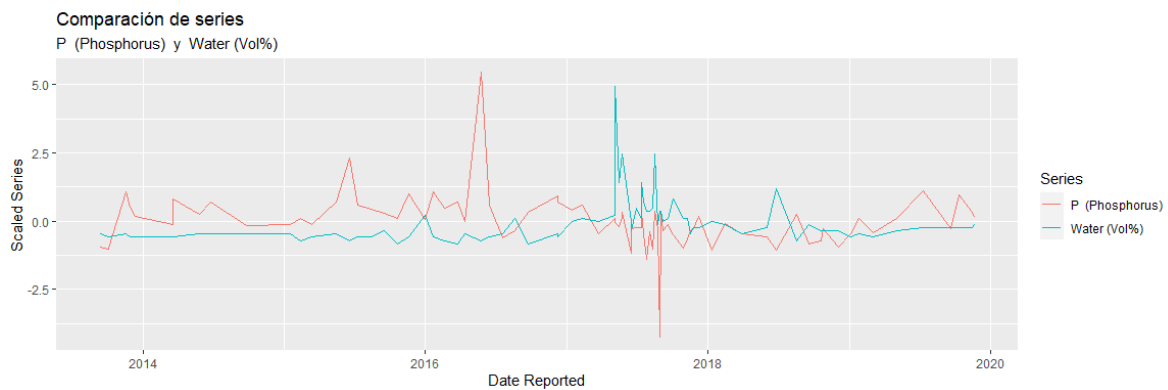
**Figura 41:** Comparativo Series de Tiempo TAN y Viscosidad a 40°C  
**Fuente:** Elaboración Propia.



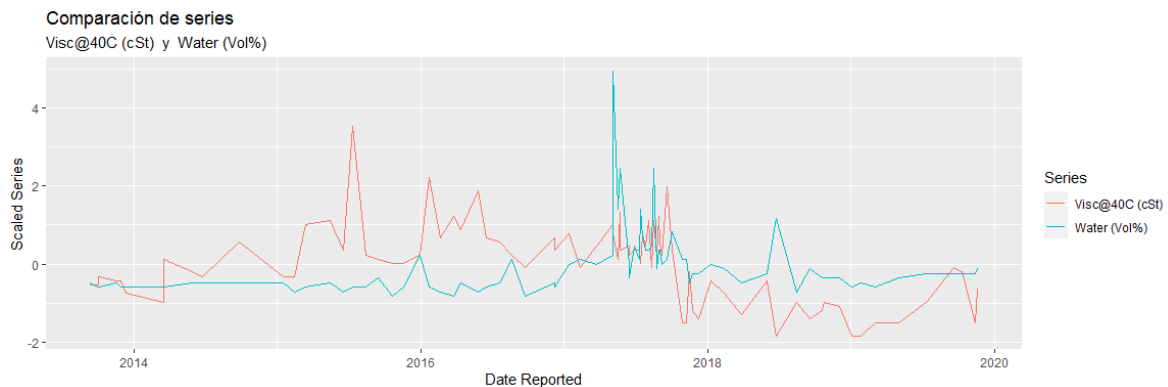
**Figura 42:** Comparativo Series de Tiempo TAN y % de Agua Disuelta  
**Fuente:** Elaboración Propia



**Figura 43:** Comparativo Series de Tiempo P (ppm) y Viscosidad a 40°C  
**Fuente:** Elaboración Propia



**Figura 44:** Comparativo Series de Tiempo P (ppm) y % de agua disuelta.  
**Fuente:** Elaboración Propia



**Figura 45:** Comparativo Viscosidad a 40°C y % de agua disuelta  
**Fuente:** Elaboración Propia.

Del anterior análisis se observa cierta similitud en la tendencia generada en las variables que representan propiedades fisicoquímicas en el aceite, es decir, TAN, viscosidad a 40°C y porcentaje de agua disuelta. Se observa además que hacia mediados



del año 2017 existió una desestabilización del sistema que generó alteración en los valores de casi todas las variables de interés en el análisis.

En las variables TAN y viscosidad a 40°C puede observarse una tendencia a la baja luego de la adición de aceite. Esta tendencia corresponde a la instalación de dos sistemas de filtración fuera de línea adicionales en el sistema. Estos elementos permiten la remoción de material particulado, mejorando las condiciones del aceite de lubricación y extendiendo la vida útil de los equipos. Cabe resaltar que la presencia de partículas contaminantes de características ácidas, como el material orgánico, pueden generar un aumento en el TAN y por lo tanto en la viscosidad (ver sección 4.1.1). De esta manera se explica el hecho de que ante la instalación de esta filtración se haya logrado una reducción del TAN y la viscosidad simultáneamente.

Es importante resaltar que debido a que la frecuencia de muestreo de aceite es mensual, y solo se cuenta con histórico de datos desde el año 2014, la cantidad de observaciones con las cuales se cuenta en la base de datos es insuficiente para la implementación de algún modelo de aprendizaje de máquina. Existen algunos instrumentos que permiten realizar la medición de dichas variables en tiempo real, generando el flujo de datos necesario para la implementación de modelos más robustos.

A pesar de lo anterior, en el presente análisis se implementó un modelo de medias móviles como el propuesto por (Raposo et al., 2019) y reseñado en la sección 2.5.3 del presente trabajo. Allí, es importante encontrar el valor de  $k$  (número de datos anteriores que se toman para realizar la predicción) que minimice el error porcentual absoluto medio (MAPE). Se llevaron a cabo 20 pruebas variando  $k$  y analizando diferentes resultados para el MAPE, estableciendo así los modelos más adecuados para las variables principales. La **Tabla 1** muestra los resultados de  $k$  óptimos hallados para cada una de las variables analizadas. El MAPE se calcula a partir del error absoluto de cada período, dividiendo éste entre el valor real observado el mismo período y promediando estos errores porcentuales absolutos (Hanke & Wichern, 2010).

Variable analizada	$k$ óptimo	MAPE
TAN (mg KOH/g)	3	0.335028
P (Phosphorus)	6	0.079587
Viscosidad @40°C	4	0.0118214
Water (%Vol)	3	0.434431

Tabla 1:  $k$  óptimo para modelo de medias móviles  
Fuente: Elaboración Propia.

Las **Figura 46 a la Figura 49** muestran la estimación realizada de cada serie de tiempo mediante el modelo de medias móviles. Allí se observa que para las variables que presentan un mejor ajuste al modelo son el porcentaje de agua disuelta y la viscosidad a 40°C.

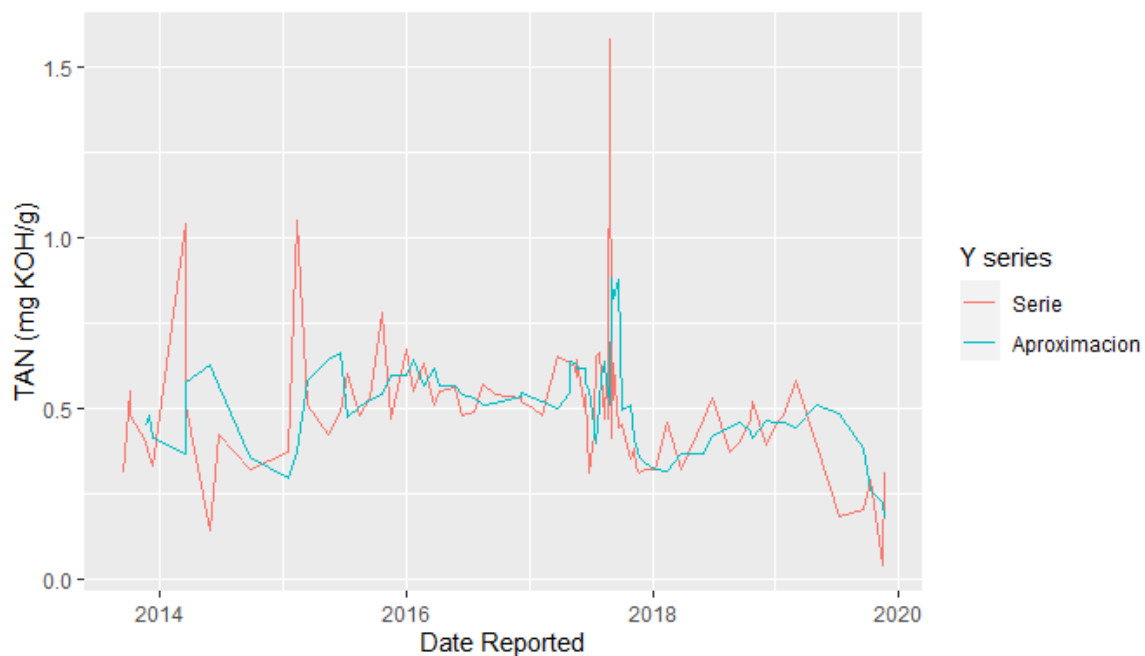
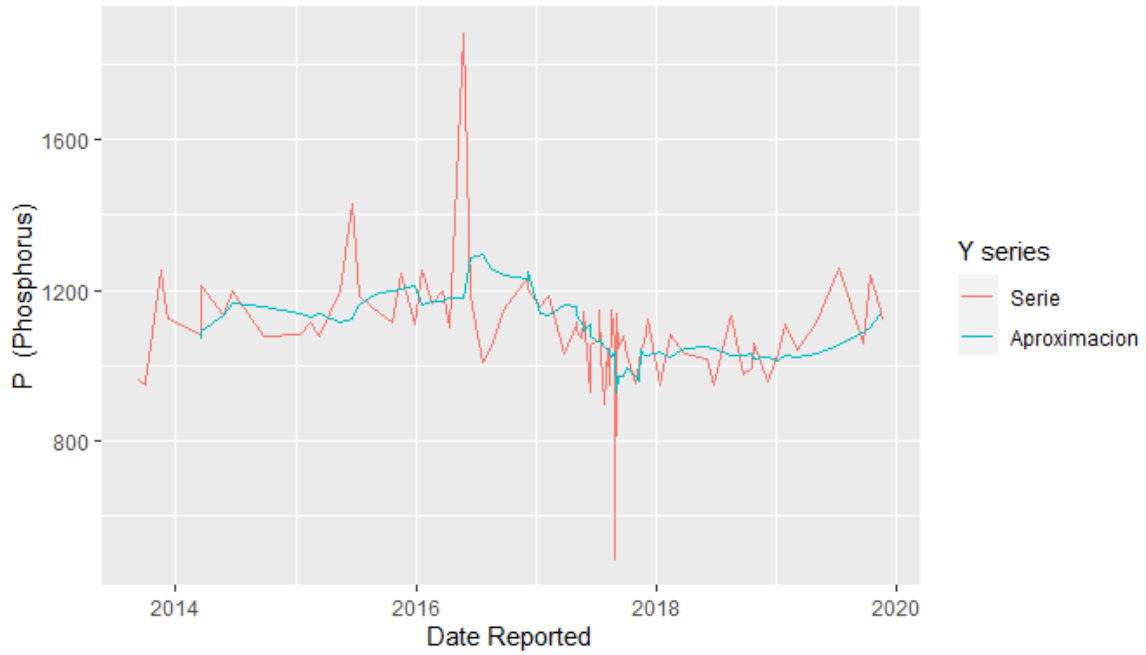
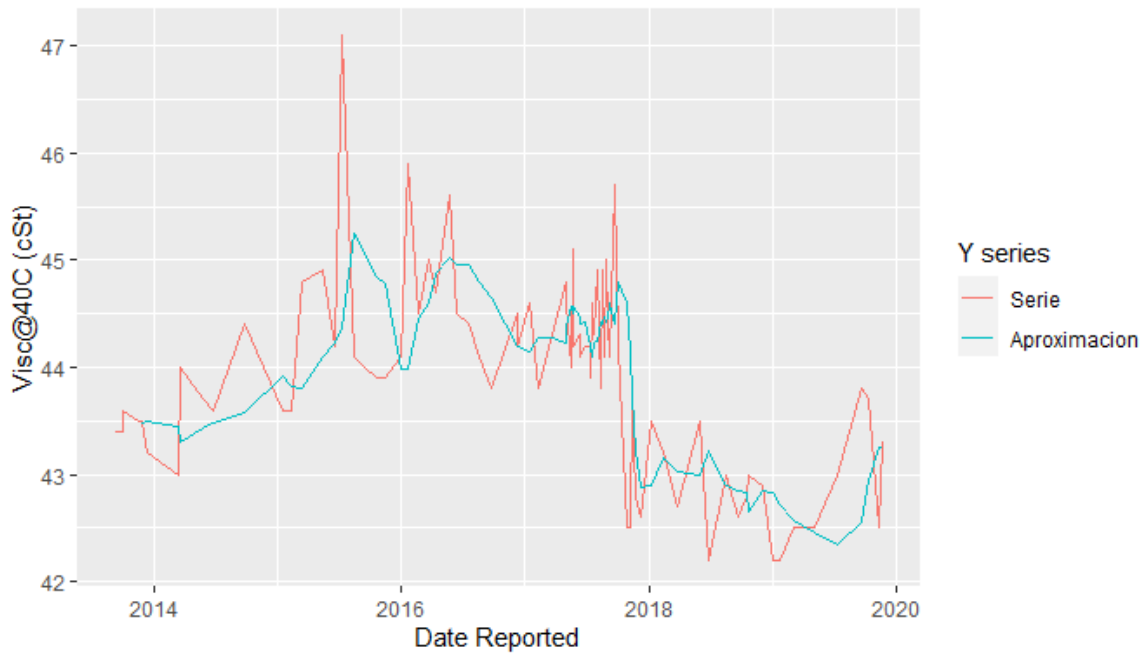


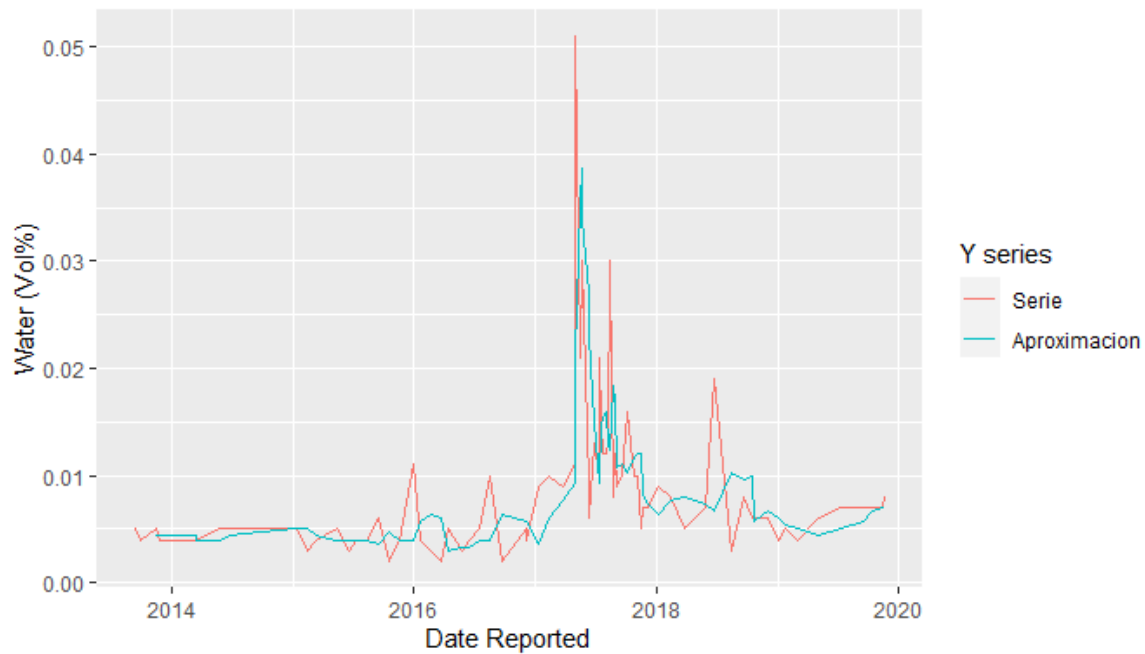
Figura 46: Aproximación mediante medias móviles para el TAN  
Fuente: Elaboración Propia



**Figura 47:** Aproximación mediante medias móviles para el Fósforo (ppm)  
**Fuente:** Elaboración Propia



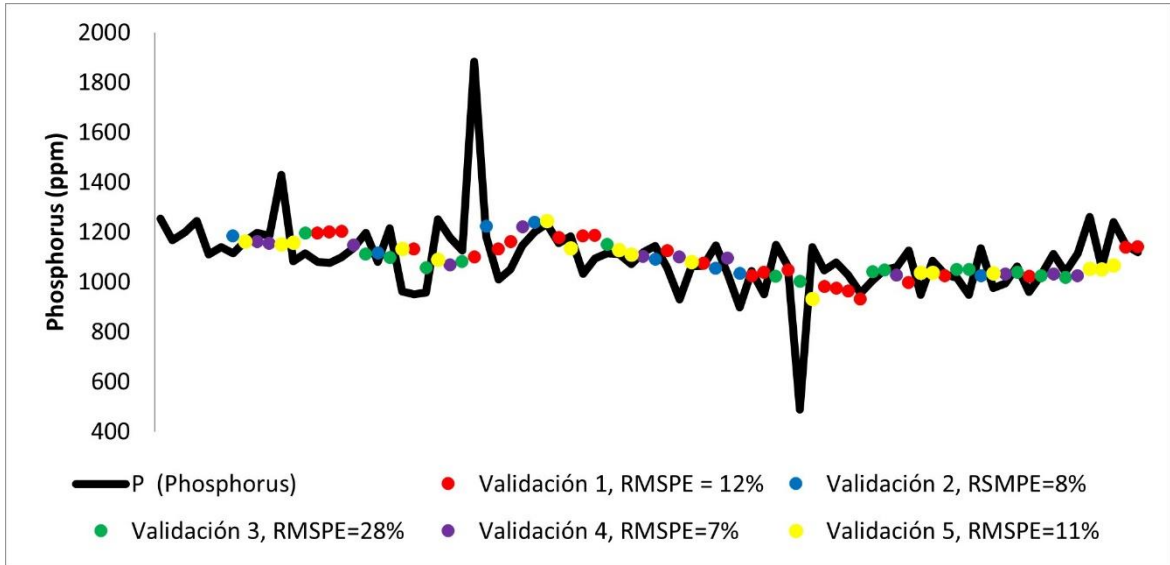
**Figura 48:** Aproximación mediante medias móviles para Viscosidad a 40°C  
**Fuente:** Elaboración Propia



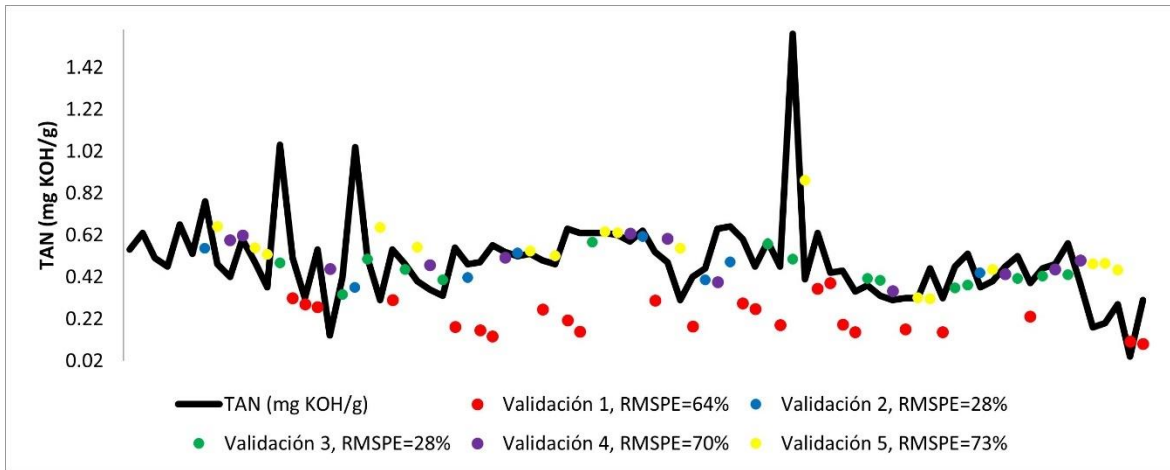
**Figura 49:** Aproximación mediante medias móviles para % de agua disuelta  
**Fuente:** Elaboración Propia.

### 4.3.1 Validación del análisis predictivo.

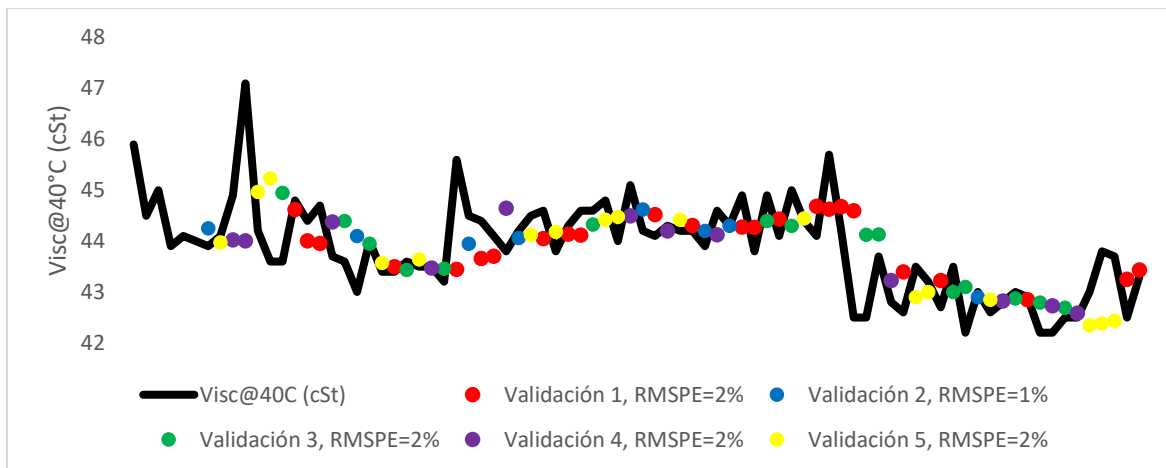
Tal y como se presentó en la sección 3.2.1 se procedió a realizar una validación del modelo de medias móviles mediante la estrategia de validación cruzada de orden 5. De la **Figura 50** a la **Figura 53** se muestran los resultados obtenidos para los datos de validación en cada una de las etapas del proceso y para cada variable principal revisada en el presente trabajo.



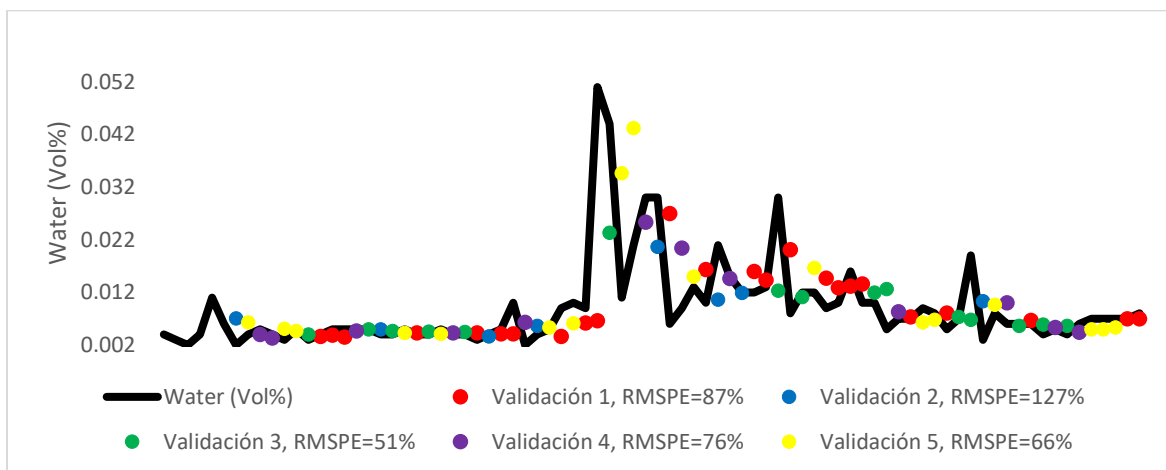
**Figura 50:** Validación cruzada de orden 5 para el fósforo (ppm).  
**Fuente:** Elaboración Propia.



**Figura 51:** Validación cruzada de orden 5 para el TAN (mg KOH/g)  
**Fuente:** Elaboración Propia.



**Figura 52:** Validación cruzada de orden 5 para la Viscosidad @40°C (cSt)  
**Fuente:** Elaboración Propia.



**Figura 53:** Validación cruzada de orden 5 para el % de Agua (Vol%)  
**Fuente:** Elaboración Propia.

Los colores de los datos de validación mostrados en la **Figura 50** a la **Figura 53** corresponden con los colores de la división por grupos mostrados en la **Figura 18**. La línea color negro representa la serie de tiempo real para cada una de las variables.

Posteriormente, se procedió a realizar el cálculo del porcentaje de error cuadrático medio. La **Tabla 2** a la **Tabla 5** muestran los resultados obtenidos para cada etapa de validación en cada una de las variables. Luego, se halló la media geométrica de cada una de las etapas en cada variable, con el fin de hallar el porcentaje de error cuadrático medio general del modelo. Dicho indicador permite comparar la efectividad del modelo de medias

móviles en cada una de las variables estudiadas y validar si la hipótesis presentada en la sección 1.3 es válida para las variables en las cuales se centró el presente trabajo.

<b>Phosphorus (ppm)</b>	
<b>Etapa de Validación</b>	<b>RMSPE</b>
<b>1</b>	12%
<b>2</b>	8%
<b>3</b>	28%
<b>4</b>	7%
<b>5</b>	11%
<b>Media Geométrica</b>	12%

**Tabla 2:** RMSPE obtenido en validación cruzada para el fósforo (ppm)  
Fuente: Elaboración Propia.

<b>TAN ( mg KOH/mg)</b>	
<b>Etapa de Validación</b>	<b>RMSPE</b>
<b>1</b>	64%
<b>2</b>	28%
<b>3</b>	28%
<b>4</b>	70%
<b>5</b>	73%
<b>Media Geométrica</b>	48%

**Tabla 3:** RMSPE obtenido en validación cruzada para el TAN (mg KOH/mg)  
Fuente: Elaboración Propia.

<b>Visc @ 40°C (cSt)</b>	
<b>Etapa de Validación</b>	<b>RMSPE</b>
<b>1</b>	2%
<b>2</b>	1%
<b>3</b>	2%
<b>4</b>	2%
<b>5</b>	2%
<b>Media Geométrica</b>	2%

**Tabla 4:** RMSPE obtenido en validación cruzada para la Visc @40°C (cSt)  
Fuente: Elaboración Propia.

Water (%Vol)	
Etapa de Validación	RMSPE
1	87%
2	127%
3	51%
4	76%
5	66%
<b>Media Geométrica</b>	78%

**Tabla 5:** RMSPE obtenido en validación cruzada para el agua (%Vol)  
**Fuente:** Elaboración Propia.

Tal y como puede observarse, el modelo predictivo de medias móviles tuvo un mejor desempeño en la concentración de fósforo en el aceite y en la viscosidad @40°C. Además el desempeño de dicho modelo en variables como el TAN (mg KOH/mg) y la humedad en el aceite (%Vol) deja bastante que desear, en tanto que sus porcentajes de error cuadrático medio son superiores al 40%.

Una posible explicación para este comportamiento es que los modelos de medias móviles tienen un mejor desempeño en conjuntos de datos donde el rango de los datos es estrecho y no existen *outliers* pronunciados que hagan variar las medias temporales (Hanke & Wichern, 2010). Existen varias razones para que los rangos del TAN y el porcentaje de humedad sean amplios y por tanto los modelos de medias móviles no tengan la efectividad esperada en la predicción de variables.

En primer lugar, es importante mencionar que tanto el TAN como el % de humedad, son variables que se ven fuertemente influenciadas por condiciones externas del aceite, como lo son la presencia de fugas de vapor desde los cojinetes de la turbina, contaminación en el depósito por falla en los filtros de aireación, adiciones de aceite nuevo y procesos de filtración y diálisis llevados a cabo en el aceite.

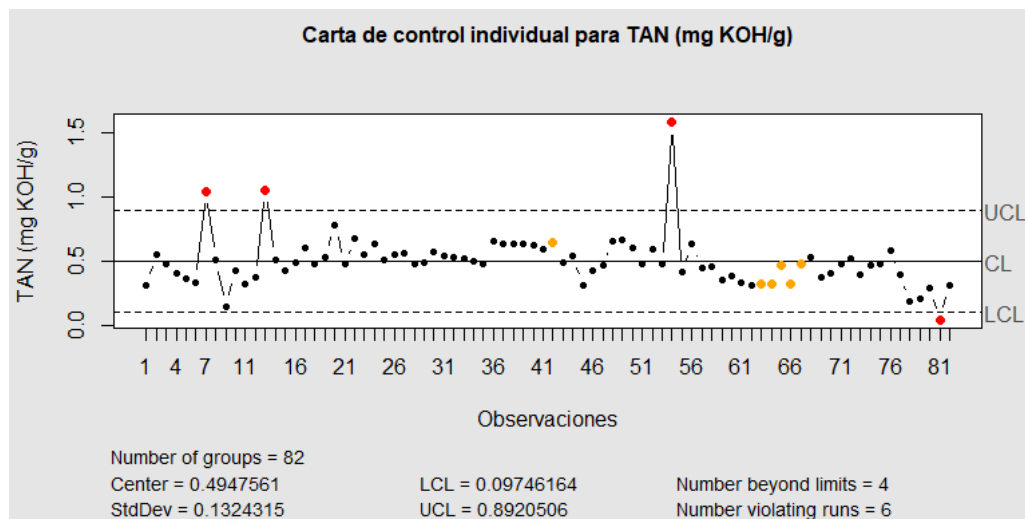
Por otro lado, la concentración de fósforo y la viscosidad del aceite son propiedades mucho más estables y variaciones en las mismas, solo podrían ser ocasionadas por desgaste natural de la vida útil del lubricante.



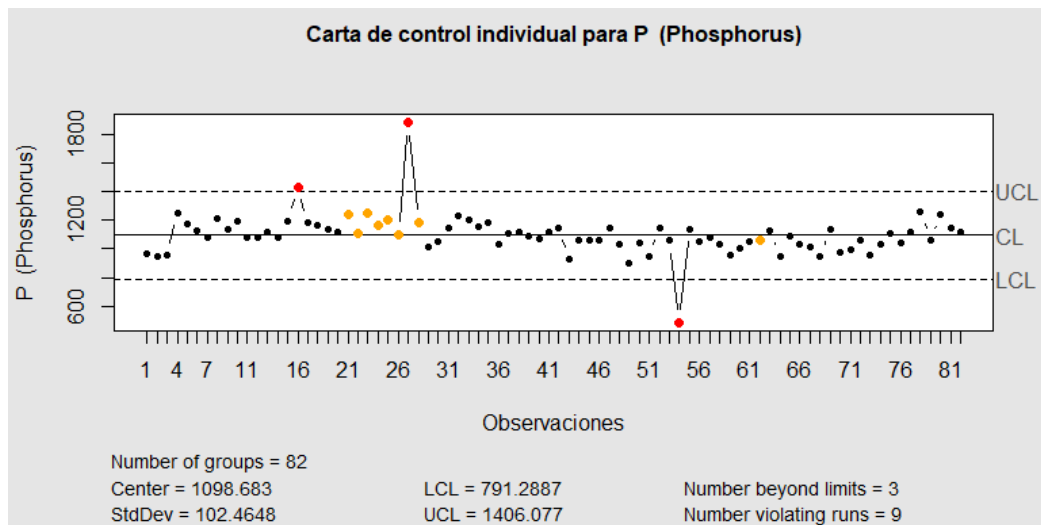
En conclusión, puede decirse que el modelo de medias móviles, confirma la hipótesis propuesta en la sección 1.3, solo para variables medidas en el análisis de aceite que no puedan ser fácilmente afectadas por el entorno del lubricante o por actividades de mantenimiento que se realicen sobre el equipo.

## 4.4 Análisis mediante Cartas de Control

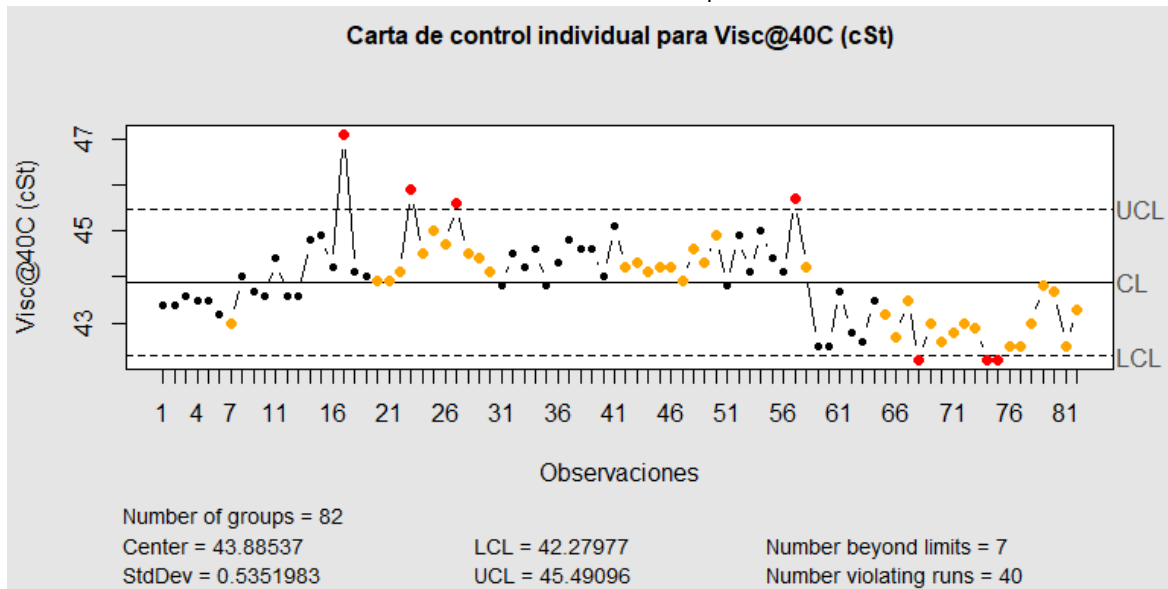
El análisis mediante cartas de control fue realizado para las variables seleccionadas como variables de interés. Se siguió el procedimiento reseñado en la sección 2.4.2.2 hallando los límites de control correspondientes a la variabilidad de las diferentes propiedades estudiadas. Para obtener las cartas de control se utilizó el paquete qcc. Las **Figura 54 a la Figura 57** muestran las cartas de control obtenidas.



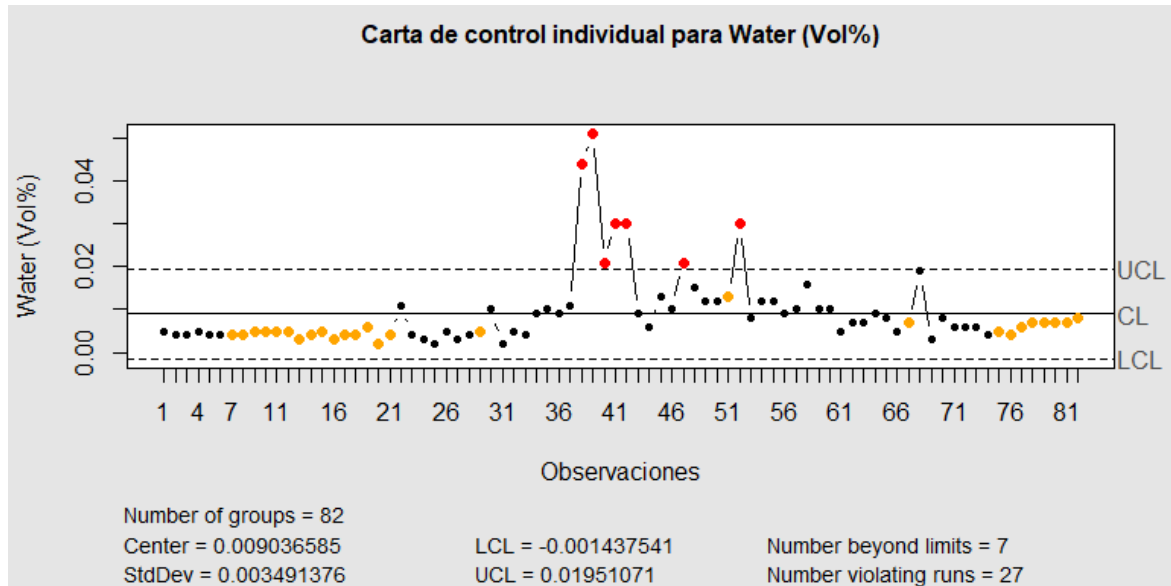
**Figura 54:** Cartas de control para el TAN  
Fuente: Elaboración Propia



**Figura 55:** Carta de control para la variable P (Phosphorus)  
**Fuente:** Elaboración Propia



**Figura 56:** Carta de Control para la viscosidad a 40°C  
**Fuente:** Elaboración Propia



**Figura 57:** Carta de control para % de agua disuelta  
**Fuente:** Elaboración Propia

Los estados de alarma y precaución determinados en los reportes de resultados de análisis de aceite usado, como el mostrado en la **Figura 16** están basados en límites recomendados por el fabricante del lubricante para aplicaciones estándar. Estos no contemplan los datos históricos de cada equipo y las posibles variaciones que pueden presentarse dependiendo de las condiciones de operación. Por otro lado, las alertas generadas a partir de las cartas de control permiten tener una indicación real basado en el histórico de operación del equipo sobre cuando un valor puede considerarse fuera de rango y requerir de alguna intervención sobre la condición.

El ejemplo más claro de lo anteriormente descrito es la carta de control realizada para la variable TAN, la cual se presenta en la **Figura 54**. En ella, podemos ver que únicamente tres datos superan en todo el universo de muestreo el límite de control superior calculado a partir de las cartas de control. Por otro lado, cuando se tiene en cuenta el límite de control propuesto por el fabricante (0.5) se obtienen 21 registros en alerta, lo cual puede convertirse en una falsa indicación de que existe una tendencia indeseada en el comportamiento del lubricante.

El análisis mediante cartas de control permite entonces tomar decisiones de mantenimiento basadas en el comportamiento real de los datos y generan una alerta real sobre el status de determinada variable. Sin embargo, es importante ser cuidadoso a la

hora de interpretar la información derivada de esta herramienta pues es fácil perder de vista una posible tendencia al alza en determinada variable, una vez se establecen y estandarizan dichos límites.

Es importante resaltar que el estudio de cartas de control debe acompañarse con los datos de intervenciones o eventos ocurridos al equipo. Lo anterior pues los datos históricos del comportamiento de la variable capturan como un evento de curso natural, las variaciones producidas por una intervención de mantenimiento o una falla funcional del equipo. Lo anterior, genera *outliers* en la carta de control que deben ser debidamente justificados al contrastarse con la hoja de vida del equipo.

# 5. Conclusiones y recomendaciones

## 5.1 Conclusiones

El presente trabajo de investigación tuvo como objetivo principal la implementación de diferentes modelos de analítica de datos sobre los resultados obtenidos del histórico de muestreo de aceite usado en un turbogenerador de vapor marca Siemens, generando una herramienta de toma de decisiones de mantenimiento que permita optimizar los recursos invertidos en intervenciones preventivas, garantizando la confiabilidad del equipo.

Para tal fin, se recopiló la información de aproximadamente 7 años de datos históricos del programa de análisis de aceite usado del equipo, estructurando una base de datos en la cual pudieron identificarse 4 variables consideradas claves para la caracterización del estado del lubricante. Estas variables fueron seleccionadas de acuerdo con la consistencia de los datos y a la relevancia de las propiedades que representan. Se concluye que las variables que representan en una mejor manera el estado actual del lubricante son la viscosidad a 40°C, el número ácido total (TAN), el porcentaje de agua disuelta en el aceite y el contenido de fósforo.

Luego, se realizaron diferentes análisis sobre los datos recopilados para dichas variables. Se realizaron análisis descriptivos, análisis clasificadorios mediante clustering jerárquico, análisis predictivos mediante series de tiempo y el modelo de medias móviles. Por último, se empleó el modelo de cartas de control para establecer límites de alerta basados en los históricos de datos obtenidos.

Una de las principales conclusiones que pueden extraerse de los modelos anteriormente mencionados es que en los modelos de carácter predictivos cobra gran

importancia, contar con una cantidad de datos lo suficientemente grande para detectar algún tipo de estacionalidad o patrón en las series de tiempo y generar predicciones con un nivel de confianza aceptable. A partir de los 81 datos presentes en la serie temporal, no fue posible la generación de un modelo predictivo lo suficientemente robusto como para generar una herramienta de toma de decisiones confiable.

De los modelos predictivos implementados el modelo de medias móviles fue el modelo que mejor desempeño tuvo con los datos disponible. Se realizó validación del modelo mediante la técnica de validación cruzada de orden 5. Allí se encontró un buen desempeño del modelo en variables cuyas entradas presentan rangos de variación estrechos, esto son variables que no son fácilmente influenciadas por condiciones externas del equipo o actividades de mantenimiento. Este modelo puede ser utilizado para la predicción del comportamiento de variables en periodos cercanos a los últimos datos obtenidos. No es recomendable el uso de modelos de medias móviles para realizar predicciones a largo plazo.

Para hacer la comparación de modelos se utilizó el RMSPE (Porcentaje de error cuadrático medio. Se obtuvo un RMSPE de 12% para la concentración de fósforo en el aceite, 48% para el TAN, 2% para la viscosidad @40 °C y 78% para el % de humedad en el aceite.

Finalmente, se implementó un modelo de cartas de control para cada una de las variables seleccionadas como variables críticas. Mediante este modelo pudieron establecerse límites de control diferentes a los establecidos por el proveedor del lubricante en estudio. Estos límites de control están basados en el histórico de datos y por lo tanto en las condiciones de operación del equipo. Lo anterior, permite tomar decisiones más objetivas y por lo tanto tener un ahorro significativo en intervenciones innecesarias.

## 5.2 Recomendaciones

A continuación, se presentarán algunas recomendaciones para investigaciones futuras que pueden ser tenidas en cuenta con el fin de ampliar los resultados obtenidos en el

---

presente trabajo y generar mejores estrategias para la comprensión y análisis de los resultados derivados de pruebas al aceite usado de activos productivos.

En primer lugar, es importante reconocer que una de las principales limitaciones con las cuales se contó en el presente trabajo fue la baja cantidad de datos disponibles para el análisis. Por lo tanto, se recomienda la instalación de un monitor continuo de condiciones en el aceite, de tal manera que logren recolectarse datos en tiempo real, permitiendo la implementación de modelos más robustos de analítica de datos y machine learning.

Por otro lado, se recomienda la implementación de un software computacional el cual permita el uso a nivel industrial de las herramientas aquí discutidas. Para la implementación de dicho software puede emplearse el mismo lenguaje computacional que fue utilizado para el desarrollo de los conceptos aquí comprendidos.

Además, se recomienda aplicar los modelos aquí desarrollados a otras variables críticas medidas en los análisis de aceite del turbogenerador aquí estudiado. Es importante hacer una implementación crítica de dichos modelos en las otras variables con el fin de obtener información confiable sobre la salud del activo.

Por último, se recomienda hacer un análisis estadístico de los costos asociados a adiciones y reposiciones de aceite realizadas durante el periodo analizado. Lo anterior, con el objetivo de tener una valoración económica del impacto que puede generar un análisis estadístico de los resultados de pruebas al aceite usado. También es importante valorar los costos de las posibles fallas promovidas por deficiencias en las propiedades del lubricante. Este costo debe incluirse según una valoración de riesgos con el fin de justificar inversiones en la implementación de modelos de ingeniería de datos para la predicción de comportamientos y toma de decisiones de mantenimiento relacionadas con el equipo analizado.

**A. Anexo: Ficha Técnica Aceite  
Mobil SHC 825**





## Mobil SHC™ 800 Series

Mobil Industrial , Colombia

Aceite para turbinas

### Descripción de producto

Los aceites para turbinas Mobil SHC™ 800 Series están diseñados específicamente para satisfacer las necesidades de las más severas aplicaciones de turbinas industriales a gas con un vida útil TOST nominal de 10.000 horas. Se recomiendan para la lubricación de las turbinas a gas asentadas en tierra, especialmente las unidades de menos de 3000 hp que se utilizan como unidades de energía en "standby", y en algunos tipos de sistemas de ciclo combinado (gas/ vapor) y de energía total. Mobil SHC™ es una familia innovadora de productos que se formula a partir de hidrocarburos sintetizados sin cera y un sistema único de aditivos. Ésta fórmula proporciona una fluidez excepcional a baja temperatura y una resistencia sin igual a la degradación a altas temperaturas. Los lubricantes de esta serie brindan además excelentes propiedades contra el desgaste, al igual que protección contra el óxido y la corrosión, como así también, un buen desempeño de liberación de aire y resistencia a la formación de espuma.

Estas características brindan una resistencia superior a la degradación térmica/ oxidativa durante el periodo de mayor calor después de parar el equipo y permiten la rápida circulación del aceite a bajas temperaturas durante la puesta en marcha. La resistencia a la degradación es un atributo clave para evitar la formación de depósitos que puedan interferir con el suministro de aceite a los cojinetes o ensuciar las servo válvulas críticas. Esta es una cuestión particular cuando las turbinas de gas están operando en el modo cíclico y experimentan ciclos múltiples de esfuerzo térmico. Debido a que la fluidez a baja temperatura y el alto índice de viscosidad son características inherentes de estos fluidos, resisten cambios durante el servicio como resultado del corte mecánico o los ciclos repetidos de temperaturas altas a bajas. Los aceites Mobil SHC™ 800 Series son totalmente compatibles con los aceites minerales, pero mezclarlos perjudicará las propiedades de desempeño superior de estos productos.

### Características y beneficios

La marca de lubricantes Mobil SHC™ es reconocida y apreciada en todo el mundo por su innovación y desempeño sobresaliente. Estos productos sintéticos Mobil SHC™, elaborados por primera vez por nuestros científicos de investigación, simbolizan el compromiso continuo de utilizar tecnología de avanzada para suministrar productos excepcionales. Los productos de la marca Mobil también han sido la elección de los operadores de turbinas en todo el mundo, ya que se comercializaron por primera vez hace más de cien años. Durante este periodo, nuestros expertos técnicos han permanecido en estrecho contacto con los fabricantes de equipos (OEM) para asegurar que nuestros productos ofrecieran un desempeño excepcional en los diseños de turbinas, los cuales evolucionaban constantemente. Conocer los diseños en constante evolución y las condiciones operativas son clave en la aplicación de la mejor tecnología de lubricantes para desarrollar productos que brinden el desempeño que los usuarios esperan.

Una tendencia a lo largo de estos años ha sido tener diseños con mayor producción de energía, lo que produce mayores esfuerzos térmicos en el lubricante. Esta exposición térmica se ve aumentada por la operación cíclica que emplean los operadores de turbinas a gas para manejar el equilibrio de suministro/ demanda de generación de energía eléctrica, lo que produce temperaturas elevadas en cada secuencia de parada del equipo. En consecuencia, resistir la degradación térmica es una propiedad clave que se requieren en un lubricante para las turbinas a gas modernas.

Para combatir la alta exposición térmica del aceite, nuestros científicos de formulación de productos eligieron los aceites de base sintética (propiedad privada de la compañía) para los aceites Mobil SHC 800 Series, debido a sus capacidades excepcionales de resistencia térmica y oxidativa. Nuestros especialistas eligieron aditivos específicos que maximizan los beneficios de los aceites bases para brindar una vida útil excepcional del aceite y control de los depósitos, como así también, resistencia a la degradación térmica y química, y el equilibrio de todas las características de desempeño. El aceite base sintético también proporciona características excepcionales de fluidez a baja temperaturas que los aceites minerales para turbinas no pueden igualar y es un beneficio clave para las aplicaciones alejadas en ambientes de temperaturas bajas. Entre los numerosos beneficios y características de la serie Mobil SHC 800 se encuentran:

Características	Ventajas y beneficios potenciales
Sobresaliente estabilidad térmica y oxidativa, y control de depósitos	Alto nivel de resistencia al calor después de la parada de la turbina Menos formación de depósitos y mayor confiabilidad. Menores costos de mantenimiento Larga vida útil del aceite y menores costos de producto
Excelente fluidez a baja temperatura	Flujo y lubricación confiables durante las puestas en marcha en frío, aún con muy bajas temperaturas
Índice de viscosidad naturalmente alto	Mayor protección del equipo a temperaturas elevadas
Muy buena resistencia a la formación de espuma y buena liberación de aire	Operación eficiente de los sistemas y menos paradas no programadas

Características	Ventajas y beneficios potenciales
Excelente desempeño anti desgaste	Excelente protección del equipo y menores costos de cambio de equipo

## Aplicaciones

Los aceites para turbinas Mobil SHC 800 Series están diseñados específicamente para satisfacer las necesidades de las aplicaciones más severas de turbinas industriales a gas y equipos auxiliares

- Aplicaciones severas de turbinas fijas a gas, especialmente las unidades de menos de 3000 hp, para la generación de energía "standby"
- Sistemas de energía total

## Especificaciones y aplicaciones

Mobil SHC 800 Series tiene las siguientes aprobaciones de fabricantes:	824	825
Siemens TLV 9013 04	X	X
MHI MS04-MA-CL003	X	

Mobil SHC 800 Series atiende o excede los requerimientos de:	824	825
Solar ES 9-224 Clase I	X	X
GE GEK 32568J	X	

Mobil SHC 800 Series es recomendado por ExxonMobil para el uso en aplicaciones que requieren:	824	825
GE GEK 101941A	X	
GE GEK 28143B	X	

## Propiedades típicas

Mobil SHC 800 Series	824	825
Viscosidad grado ISO	32	46
Viscosidad, ASTM D 445		
cSt @ 40°C	31.5	43.9
cSt @ 100°C	5.9	7.9
Índice de viscosidad, ASTM D 2270	135	145
Punto de escurrimiento, °C, ASTM D 97	<-54	-45
Punto de inflamación, °C, ASTM D 92	248	248
Gravedad específica @ 15°C kg/k, ASTM D 4052	0.83	0.83
Vida útil TOST, ASTM D 943, hrs a 2NN	9500	9500
Ensayo de espuma, ASTM D 892, Seq I, Tendencia/ Estabilidad, ml/ml	10/0	20/0
Liberación de aire	1	1

## Salud y seguridad

Con base en información disponible, no es de esperar que este producto cause efectos adversos en la salud mientras se utilice en las aplicaciones a las que está destinado y se sigan las recomendaciones de la Ficha de Datos de Seguridad (FDS). Las Fichas de Datos de Seguridad están disponibles a través del Centro de Atención al Cliente. Este producto no debe utilizarse para otros propósitos distintos a los recomendados. Al deshacerse del producto usado, tenga cuidado para así proteger el medio ambiente.

Todas las marcas comerciales utilizadas en este documento son marcas comerciales o marcas registradas de Exxon Mobil Corporation o de una de sus subsidiarias a menos que se indique lo contrario.

12-2019

Organización Terpel S.A.

Address: Carrera 7 N° 75-51, Bogotá – Colombia

Phone: (57) 1 3267878

Typical Properties are typical of those obtained with normal production tolerance and do not constitute a specification. Variations that do not affect product performance are to be expected during normal manufacture and at different blending locations. The information contained herein is subject to change without notice. All products may not be available locally. For more information, contact your local ExxonMobil contact or visit [www.exxonmobil.com](http://www.exxonmobil.com)

ExxonMobil is comprised of numerous affiliates and subsidiaries, many with names that include Esso, Mobil, or ExxonMobil. Nothing in this document is intended to override or supersede the corporate separateness of local entities. Responsibility for local action and accountability remains with the local ExxonMobil-affiliate entities.

Energy lives here™

**ExxonMobil**



© Copyright 2003-2019 Exxon Mobil Corporation. All Rights Reserved

## Bibliografía

- Ahmad, R., & Kamaruddin, S. (2012). An overview of time-based and condition-based maintenance in industrial application. *Computers & Industrial Engineering*, 63(1), 135–149. <https://doi.org/10.1016/j.cie.2012.02.002>
- Albarracín, P. R. (2015). *Tribología y Lubricación* (T. Ingeniería, Ed.). Medellín.
- AXA Risk Consulting. (2020). *Steam Turbine Lubricating Oil Systems*. 1–4.
- Ayvaz, S., & Alpay, K. (2021). Predictive maintenance system for production lines in manufacturing: A machine learning approach using IoT data in real-time. *Expert Systems with Applications*, 173, 114598. <https://doi.org/10.1016/j.eswa.2021.114598>
- Bahari, A. (2017). *Investigation into Tribological Performance of Vegetable Oils as Biolubricants at Severe Contact Conditions*. (October), 292.
- Banaszkiewicz, M. (2014). Steam turbines start-ups. *Transactions of the Institute of Fluid-Flow Machinery*, 126(126), 169–198.
- Barrios, R. (2015). 3 Medidas De Tendencia Central Y De Dispersión. *Slideshare*, 59. Retrieved from <https://es.slideshare.net/rbarriosm/3-medidas-de-tendencia-central-y-de-dispersion-49942466>
- Bergmeir, C., & Benítez, J. M. (2012). On the use of cross-validation for time series predictor evaluation. *Information Sciences*, 191, 192–213. <https://doi.org/10.1016/j.ins.2011.12.028>
- Bergmeir, C., Hyndman, R. J., & Koo, B. (2018). A note on the validity of cross-validation for evaluating autoregressive time series prediction. *Computational Statistics and Data Analysis*, 120, 70–83. <https://doi.org/10.1016/j.csda.2017.11.003>
- Bousdekis, A., Papageorgiou, N., Magoutas, B., Apostolou, D., & Mentzas, G. (2017). A Proactive Event-driven Decision Model for Joint Equipment Predictive Maintenance and Spare Parts Inventory Optimization. *Procedia CIRP*, 59(TESSConf 2016), 184–189. <https://doi.org/10.1016/j.procir.2016.09.015>
- Brnabic, A., & Hess, L. M. (2021). Systematic literature review of machine learning

- methods used in the analysis of real-world data for patient-provider decision making. *BMC Medical Informatics and Decision Making*, 21(1), 54.  
<https://doi.org/10.1186/s12911-021-01403-2>
- Brown, P., & Sondalini, M. (n.d.). *Asset Maintenance Management - The Path toward Defect Elimination*. 1–10. Retrieved from [www.lifetime-reliability.com](http://www.lifetime-reliability.com)
- Capuano, G., & Rimoli, J. J. (2019). Smart finite elements: A novel machine learning application. *Computer Methods in Applied Mechanics and Engineering*, 345, 363–381. <https://doi.org/https://doi.org/10.1016/j.cma.2018.10.046>
- Cave, A. (2017). What Will We Do When The World's Data Hits 163 Zettabytes In 2025? Retrieved March 8, 2019, from Forbes website:  
<https://www.forbes.com/sites/andrewcave/2017/04/13/what-will-we-do-when-the-worlds-data-hits-163-zettabytes-in-2025/#694cc76c349a>
- Chiu, S., & Tavella, D. (2008). Introduction to Data Mining. *Data Mining and Market Intelligence for Optimal Marketing Returns*, 137–192. <https://doi.org/10.1016/b978-0-7506-8234-3.00007-1>
- Coleman, W. (1981). Water Contamination of Steam Turbine Lube Oils - How to Avoid It. *Journal of Chemical Information and Modeling*, 53(9), 1689–1699.
- Deloitte Brazil. (2021). Strategic asset management. Retrieved April 30, 2021, from <https://www2.deloitte.com/br/en/pages/finance/solutions/gestao-estrategica-ativos.html>
- Elshawi, R., Sakr, S., Talia, D., & Trunfio, P. (2018). Big Data Systems Meet Machine Learning Challenges: Towards Big Data Science as a Service. *Big Data Research*, 14, 1–11. <https://doi.org/10.1016/j.bdr.2018.04.004>
- Espino Timón, C., & Martínez Fontes, X. (2017). “Análisis predictivo: técnicas y modelos utilizados y aplicaciones del mismo - herramientas Open Source que permiten su uso. 26/27, I(Principio activo y prestación ortoprotésica), 67. Retrieved from <http://openaccess.uoc.edu/webapps/o2/bitstream/10609/59565/6/caresptimTFG0117memoria.pdf>
- Exposito, C. (2020). Clustering jerárquico. *Universidad de La Laguna*.
- Exxon Mobil. (2009). *Turbine Oil System Care & Maintenance*.
- Exxon Mobil. (2020). Mobil SHC 825. Retrieved April 2, 2021, from <https://www.mobil.com.mx/es-mx/lubricantes/industrial/lubricants/products/products/mobil-shc-825>
- Faraldo, P. (2013). Estadística y metodología de la investigación. *Universidad Santiago*

- De Compostela*, 15. Retrieved from [http://eio.usc.es/eipc1/BASE/BASEMASTER/FORMULARIOS-PHP-DPTO/MATERIALES/Mat\\_G2021103104\\_EstadisticaTema1.pdf](http://eio.usc.es/eipc1/BASE/BASEMASTER/FORMULARIOS-PHP-DPTO/MATERIALES/Mat_G2021103104_EstadisticaTema1.pdf)
- Fernando, J., & Walters, T. (2021, February). Correlation Coefficient Definition. Retrieved May 1, 2021, from <https://www.investopedia.com/terms/c/correlationcoefficient.asp>
- Fortune Business Insight. (2020). Lubricants Market Size, Share, Report. Retrieved April 30, 2021, from Market Research Report website: <https://www.fortunebusinessinsights.com/industry-reports/lubricants-market-101771>
- Guevara Carazas, F. J., & Martha daSouza, G. F. (2010). Risk-based decision making method for maintenance policy selection of thermal power plant equipment. *Energy*, 35(2), 964–975. <https://doi.org/10.1016/j.energy.2009.06.054>
- Gupta, G., & Mishra, R. P. (2016). A SWOT analysis of reliability centered maintenance framework. *Journal of Quality in Maintenance Engineering*, 22(2), 130–145. <https://doi.org/10.1108/JQME-01-2015-0002>
- Gutierrez Pulido, H., & De la Vara Salazar, R. (2009). *Control Estadístico de la Calidad y Seis Sigma* (M. Hill, Ed.). Guanajuato.
- Han, X., Wang, Z., Xie, M., He, Y., Li, Y., & Wang, W. (2021). Remaining useful life prediction and predictive maintenance strategies for multi-state manufacturing systems considering functional dependence. *Reliability Engineering & System Safety*, 210, 107560. <https://doi.org/10.1016/j.ress.2021.107560>
- Hanke, J. E., & Wichern, D. W. (2010). Pronósticos en los negocios. In 9 Edición (Ed.), *□□□□ □□□□* (Vol. 4). Retrieved from <http://marefateadyan.nashriyat.ir/node/150>
- Hausermann, A. (1961). *Particular Problems of Steam Turbine Lubrication*. (Cv), 125–132.
- Hejnowicz, Z., Burian, A., Dobrowolska, I., & Kolano, E. (2006). Orientational variability of parallel arrays of cortical microtubules under the outer cell wall of the *Helianthus* hypocotyl epidermis. *Acta Societatis Botanicorum Poloniae*, 75(3), 201–206. <https://doi.org/10.5586/asbp.2006.023>
- Jiang, G., & Wang, W. (2017). Markov cross-validation for time series model evaluations. *Information Sciences*, 375, 219–233. <https://doi.org/10.1016/j.ins.2016.09.061>
- Jiménez Rodríguez, C., & Arias Aguilar, D. (2004). Distribución de la biomasa y densidad de raíces finas en una gradiente sucesional de bosques en la Zona Norte de Costa Rica. *Revista Forestal Mesoamerica Kurú*, 1(2), pág. 44-63.

- Joseph Omosanya, A., Titilayo Akinlabi, E., & Olusegun Okeniyi, J. (2019). Overview for Improving Steam Turbine Power Generation Efficiency. *Journal of Physics: Conference Series*, 1378(3). <https://doi.org/10.1088/1742-6596/1378/3/032040>
- Lahura, E. (2003). El Coeficiente De Correlación Y Correlaciones Espúreas. *Universidad Católica Del Perú*, 1–64.
- Lazovic, T., & Marinkovic, A. (2015). *A case study of turbogenerator journal bearing failure*. (February).
- Luo, J. (2013). Thin Film Lubrication. In *Encyclopedia of Tribology* (pp. 3663–3667). [https://doi.org/10.1007/978-0-387-92897-5\\_682](https://doi.org/10.1007/978-0-387-92897-5_682)
- Macián, V., Tormos, B., Ruíz, S., & Ramírez, L. (2015). Potential of low viscosity oils to reduce CO2 emissions and fuel consumption of urban buses fleets. *Transportation Research Part D: Transport and Environment*, 39, 76–88. <https://doi.org/10.1016/j.trd.2015.06.006>
- Martha deSouza, G. F. (2012). Fundamentals of Maintenance. In Springer (Ed.), *Thermal Power Plant Performance Analysis* (pp. 123–146). Sao Paulo.
- Maxell, D. (1996). The History of the Steam Turbine. *Pacific Turbines*, (1629).
- McCoy, J. T., & Auret, L. (2019). Machine learning applications in minerals processing: A review. *Minerals Engineering*, 132, 95–109. <https://doi.org/https://doi.org/10.1016/j.mineng.2018.12.004>
- Michalke, B., & Nischwitz, V. (2017). Speciation and element-specific detection. In *Liquid Chromatography* (pp. 753–767). <https://doi.org/10.1016/B978-0-12-805392-8.00023-2>
- Montoya-Restrepo, N. E., & Correa-Morales, J. C. (2009). Estadístico de Procesos en el Monitoreo de la Mortalidad Perinatal. *Revista de Salud Pública*, 11(1), 92–99.
- Neale, M. J. (1973). The Tribology Handbook. In *Notes and Queries* (Vol. s8-VI). <https://doi.org/10.1093/nq/s8-VI.151.385-b>
- Nicholson, K. F., Richardson, R. T., van Roden, E. A. R., Quinton, R. G., Anzilotti, K. F., & Richards, J. G. (2019). Machine learning algorithms for predicting scapular kinematics. *Medical Engineering & Physics*, 65, 39–45. <https://doi.org/https://doi.org/10.1016/j.medengphy.2019.01.005>
- Palladino, A. C. (2011). Gráfico de caja. *Atención Primaria de Salud, Epidemiología e Informática II*, 7–10.
- Patiño-Rodríguez, C. E., & Guevara Carazas, F. J. (2020). Maintenance and Asset Life Cycle for Reliability Systems. *Reliability and Maintenance - An Overview of Cases*.

- <https://doi.org/10.5772/intechopen.85845>
- Phillips 66. (2019). *Next-Generation Turbine Oils Combat Oxidation , Thermal Degradation and Varnish*.
- Pintelon, L., & Parodi-Herz, A. (2008). Maintenance: An Evolutionary Perspective. *Springer Series in Reliability Engineering*, 8, 21–48. [https://doi.org/10.1007/978-1-84800-011-7\\_2](https://doi.org/10.1007/978-1-84800-011-7_2)
- Pourahmadi, M. (2002). A Course in Time Series Analysis. *The American Statistician*, 56(1), 77–77. <https://doi.org/10.1198/tas.2002.s131>
- R: The R Project for Statistical Computing. (n.d.). Retrieved April 15, 2021, from <https://www.r-project.org/>
- Raadnui, S., & Kleesuwan, S. (2005). Low-cost condition monitoring sensor for used oil analysis. *Wear*, 259(7–12), 1502–1506. <https://doi.org/10.1016/j.wear.2004.11.009>
- Raposo, H., Farinha, J. T., Fonseca, I., & Galar, D. (2019). Predicting condition based on oil analysis – A case study. *Tribology International*, 135(January), 65–74. <https://doi.org/10.1016/j.triboint.2019.01.041>
- Reddy, a S., Ahmed, I., Kumar, T. S., Reddy, a V. K., & Bharathi, V. V. P. (2014). Analysis Of Steam Turbines. *International Refereed Journal of Engineering and Science*, 3(2), 32–48.
- Sander, J. (2012). Steam Turbine Oil Challenges. *LE White Paper*, 1–10.
- Scientific Spectro. (2000). Guide to Measuring TAN and TBN in Oil. *Spectro Scientific*, Spectro Sci. Retrieved from <https://www.spectrosci.com/resource-center/lubrication-analysis/literature/e-guides/guide-to-measuring-tantbn/>
- Scopus. (2021a). Analyze search results for “lubricant” and “analytics.” Retrieved April 30, 2021, from <https://www.scopus.com/term/analyzer.uri?sid=a091b7a8e94702c690c7bc7598eeaf06&origin=resultslist&src=s&s=TITLE-ABS-KEY%28%22lubricant%22+and+%22analytics%22%29&sort=plf-f&sdt=b&sot=b&sl=42&count=26&analyzeResults=Analyze+results&txGid=417060f79e9e194be4ae35a560a96415>
- Scopus. (2021b). Analyze search results for “Machine Learning” and “maintenance.” Retrieved April 30, 2021, from <https://www.scopus.com/term/analyzer.uri?sid=de8fd338098c34bee445c5b4839b29cd&origin=resultslist&src=s&s=TITLE-ABS->



KEY%28%22Machine+Learning%22+and+%22maintenance%22%29&sort=plf-f&sdt=b&sot=b&sl=51&count=3416&analyzeResults=Analyze+results&txGid=ed806b  
eff

Scopus - Analyze search results. (n.d.). Retrieved March 19, 2019, from <https://www-scopus-com.ezproxy.unal.edu.co/term/analyzer.uri?sid=6bce59022a605e33370d56242a5ba5fe&origin=resultslist&src=s&s=TITLE-ABS-KEY%28machine+learning%29&sort=plf-f&sdt=b&sot=b&sl=31&count=171267&analyzeResults=Analyze+results&txGid=da2bb524285aa00>

Shahbazi, N., Bortoluzzi, B., Raghubar, C., An, A., Fok, R., & McArthur, J. J. (2018). Machine learning and BIM visualization for maintenance issue classification and enhanced data collection. *Advanced Engineering Informatics*, 38(October 2017), 101–112. <https://doi.org/10.1016/j.aei.2018.06.007>

Shimadzu. (2003). Elemental Analysis of Additives in Lubricant Oils Using ICPE-9820. *Application News*, J111.

Sibata. (2020). Viscosity Measurement Series. Sibata. Retrieved from [https://www.sibata.co.jp/wpcms/wp-content/themes/sibata/en/pdf/viscosity\\_measurement\\_series.pdf](https://www.sibata.co.jp/wpcms/wp-content/themes/sibata/en/pdf/viscosity_measurement_series.pdf)

Spakovszky, Z. (2007). Enhancements of Rankine Cycles. Retrieved March 28, 2021, from Unified: Thermodynamics and propulsion website: <https://web.mit.edu/16.unified/www/FALL/thermodynamics/notes/node66.html>

Spectro Scientific. (2015). Guide to Measuring Water in Oil. *Spectro Scientific*, 5. Retrieved from <https://www.spectrosci.com/resource-center/lubrication-analysis/literature/whitepapers/guide-to-measuring-water-in-oil/%5Cnhttp://www.spectrosci.com/resource-center/lubrication-analysis/literature/e-guides/guide-to-measuring-water-in-oil/>

Syan, C. S., Ramsoobag, G., Mahabir, K., & Rajnauth, V. (2020). *A Case Study for Improving Maintenance Planning of Centrifugal Pumps Using Condition-Based Maintenance*. 42(2), 17–24.

Troyer, D., & Fitch, J. (2004). *Oil Analysis Basics*. León: Noria.

Ucci, D., Aniello, L., & Baldoni, R. (2019). Survey of machine learning techniques for malware analysis. *Computers & Security*, 81, 123–147. <https://doi.org/https://doi.org/10.1016/j.cose.2018.11.001>

- Vališ, D., Žák, L., & Pokora, O. (2015). Failure prediction of diesel engine based on occurrence of selected wear particles in oil. *Engineering Failure Analysis*, 56, 501–511. <https://doi.org/10.1016/j.engfailanal.2014.11.020>
- Vellido, A., Martín-Guerrero, J. D., & Lisboa, P. J. G. (2012). Making machine learning models interpretable. *ESANN 2012 Proceedings, 20th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, (April), 163–172.
- Zhang, Z., & Sejdić, E. (2019). Radiological images and machine learning: Trends, perspectives, and prospects. *Computers in Biology and Medicine*. <https://doi.org/https://doi.org/10.1016/j.compbimed.2019.02.017>
- Zhao, Y. (2017). The Importance of Lubricant and Fluid Analysis in Predictive Maintenance. *Spectro Scientific*, (Figure 1), 1–6. Retrieved from <https://www.spectrosci.com/blog/the-importance-of-lubricant-and-fluid-analysis-in-predictive-maintenance/>