

Trabajo Fin de Máster

DESARROLLO DE UN MODELO ESTADÍSTICO PARA LA ESTIMACIÓN DE LA VIDA ÚTIL Y PREDICCIÓN DE FALLOS EN TUBERÍAS DE AGUA POTABLE Y SU APLICACIÓN EN LA GESTIÓN DE ACTIVOS

Intensificación: *HIDRÁULICA URBANA*

Autor:

ROBERTO RAMÍREZ AGUILAR

Tutor:

RICARDO COBACHO JORDAN

Cotutor/es:

AMPARO LÓPEZ JIMÉNEZ

FELIPE SANZ TÁRREGA

JULIO, 2019



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

máster en ingeniería
hidráulica y medio ambiente
mihma



Resumen del Trabajo de Fin de Master

Datos del proyecto

Título del TFM en español: DESARROLLO DE UN MODELO ESTADÍSTICO PARA LA ESTIMACIÓN DE LA VIDA ÚTIL Y PREDICCIÓN DE FALLOS EN TUBERÍAS DE AGUA POTABLE Y SU APLICACIÓN EN LA GESTIÓN DE ACTIVOS

Título del TFM en inglés: DEVELOPMENT OF A STATISTICAL MODEL FOR ESTIMATION OF USEFUL LIFE AND FAILURE PREDICTION OF DRINKING WATER PIPES AND APPLICATION IN ASSET MANAGEMENT

Título del TFM en Valenciano: DESARROLLO DE UN MODEL ESTADÍSTIC PER A L'ESTIMACIÓ DE LA VIDA ÚTIL I PREDICCIÓ DE FALS EN TUBERIES D'AGUA POTABLE, PER A UNA APLICACIÓ EN LA GESTIÓ D'ACTIVOS

Alumno: ROBERTO XAVIER RAMÍREZ AGULAR.

Tutor: RICARDO COBACHO JORDAN

Cotutor/es:

AMPARO LÓPEZ JIMÉNEZ

FELIPE SANZ TÁRREGA

Director experimental:

Fecha de Lectura: JULIO, 2019

Resumen

En español (máximo 5000 caracteres)

En los sistemas de redes de distribución de agua potable es crucial llevar a cabo una gestión eficaz de la operación, el mantenimiento y los recursos con los que cuentan las empresas prestadoras de este servicio, de manera que se garantice el suministro de agua tanto en cantidad y calidad a los usuarios; así como también asegurar la operación y sostenibilidad del sistema a lo largo del tiempo.

Una red de distribución de agua se conforma de diferentes tipos de equipos y elementos, dentro de los cuales las tuberías forman la parte medular de estas infraestructuras. Estas tuberías pueden verse afectadas por diversos factores de tipo externo o propios de la misma operación de la red, lo cual a su vez puede condicionar su tiempo de vida útil de servicio. Por lo tanto, es importante conocer los principales factores que las afectan, así como también poder estimar su frecuencia de fallos en el tiempo, a fin de tomar con anticipación las acciones más adecuadas para prevenir fallos repentinos que ocasionan interrupciones de servicio, y con ello pérdidas económicas tanto a los usuarios como a las empresas operadoras.

En este trabajo se recoge la aplicación de un modelo estocástico de variable simple, el cual se desarrolla en base a la aplicación de una metodología propuesta dentro del campo de la ingeniería de la mantenibilidad y confiabilidad, a partir del cual se estiman tiempos de vida útil y predicción de fallos para muestras de tubería de similares características. Además, se realiza una comparación del método propuesto con uno de vida acelerada de múltiples variables.

De esta manera, se busca proponer un método sencillo y práctico con base técnica y científica del análisis de supervivencia de redes de agua potable, con el fin que pueda ser adoptado por las empresas de agua que busquen dentro de sus objetivos desarrollar su nivel en la gestión de activos.

Como caso de estudio para la aplicación de la metodología se cuenta con los datos de un abastecimiento que cuenta con una estructura de datos básica y un sistema de gestión de información SIG, del cual se tiene además un registro de incidencias de fallos por un periodo menor a tres años. Por lo tanto, se aplicará la metodología a un caso de estudio sobre el cual se tiene un registro de datos a corto plazo, lo cual a nivel general se puede considerar una condición común entre las empresas de agua.

En la metodología propuesta se estima la vida útil de servicio para muestras de tuberías de características homogéneas o similares, tomando en consideración dos escenarios. Uno en el que se considera solo el primer fallo presentado en cada tubería y otro en el que se analizan todos los fallos, esto es, incluyendo fallos repetitivos sobre

un mismo elemento; en ambos casos sobre un periodo de observación determinado. Además, inicialmente se asume y posteriormente se verifica que los tiempos de incidencias de los fallos pueden ser representados mediante ciertas funciones de distribución de probabilidad paramétricas, como por ejemplo la distribución Weibull, a partir de la cual es posible representar la curva de Davies o también conocida como curva de bañera, la cual es comúnmente usada para representar los fallos durante la vida útil de diversos sistemas o equipos.

Finalmente, el análisis e interpretación de las curvas y parámetros que describen la función de distribución contemplada, brindan una orientación en términos de confiabilidad para la toma de decisiones en cuanto al planteamiento de estrategias, tácticas y acciones de mantenimiento recomendadas dentro de un plan de gestión de activos.

En valenciano (máximo 5000 caracteres)

En els sistemes de xarxes de distribució d'aigua potable és crucial dur a terme una gestió eficaç de l'operació, el manteniment i els recursos amb els quals compten les empreses prestadores d'aquest servei, de manera que es garantisca el subministrament d'aigua tant en quantitat i qualitat als usuaris; així com també assegurar l'operació i sostenibilitat del sistema al llarg del temps.

Una xarxa de distribució d'aigua es conforma de diferents tipus d'equips i elements, dins dels quals les canonades formen la part medul·lar d'aquestes infraestructures. Aquestes canonades poden veure's afectades per diversos factors de tipus extern o propis de la mateixa operació de la xarxa, la qual cosa al seu torn pot condicionar el seu temps de vida útil de servei. Per tant, és important conèixer els principals factors que les afecten, així com també poder estimar la seua freqüència de fallades en el temps, a fi de prendre amb anticipació les accions més adequades per a previndre fallades sobtades que ocasionen interrupcions de servei, i amb això perdudes econòmiques tant als usuaris com a les empreses operadores.

En aquest treball es recull l'aplicació d'un model estocàstic de variable simple, el qual es desenvolupa sobre la base de l'aplicació d'una metodologia proposada dins del camp de l'enginyeria de la mantenibilitat i confiabilitat, a partir del qual s'estimen temps de vida útil i predicció de fallades per a mostres de canonada de similars característiques. A més, es realitza una comparació del mètode proposat amb un de vida accelerada de múltiples variables.

D'aquesta manera, es busca proposar un mètode senzill i pràctic amb base tècnica i científica de l'anàlisi de supervivència de xarxes d'aigua potable, amb la fi que pugui ser adoptat per les empreses d'aigua que busquen dins dels seus objectius desenvolupar el seu nivell en la gestió d'actius.

Com a cas d'estudi per a l'aplicació de la metodologia es compta amb les dades d'un proveïment que compta amb una estructura de dades bàsica i un sistema de gestió d'informació SIG, del qual es té a més un registre d'incidències de fallades per un període menor a tres anys. Per tant, s'aplicarà la metodologia a un cas d'estudi sobre el qual es té un registre de dades a curt termini, la qual cosa a nivell general es pot considerar una condició comuna entre les empreses d'aigua.

En la metodologia proposada s'estima la vida útil de servei per a mostres de canonades de característiques homogènies o similars, prenent en consideració dos escenaris. Un en el qual es considera només la primera fallada presentada en cada canonada i un altre en el qual s'analitzen totes les fallades, això és, incloent fallades repetitives sobre un mateix element; en tots dos casos sobre un període d'observació determinat. A més, inicialment s'assumeix i posteriorment es verifica que els temps d'incidències de les fallades poden ser representats mitjançant certes funcions de distribució de probabilitat paramètriques, com per exemple la distribució Weibull, a partir de la qual és possible representar la corba de Davies o també coneguda com a corba de banyera, la qual és comunament usada per a representar les fallades durant la vida útil de diversos sistemes o equips.

Finalment, l'anàlisi i interpretació de les corbes i paràmetres que descriuen la funció de distribució contemplada, brinden una orientació en termes de confiabilitat per a la

presa de decisions quant al plantejament d'estratègies, tàctiques i accions de manteniment recomanades dins d'un pla de gestió d'actius.

En inglés (máximo 5000 caracteres)

In potable water distribution systems it is crucial to carry out effective management of the operation, maintenance and the resources available to the companies providing this service, so that the water supply is guaranteed in both quantity and quality to users, as well ensuring the operation and sustainability of the system through the time.

A water distribution network is composed of different types of equipment and elements, within which the pipes represent the core of these infrastructures. These pipes can be affected by various factors of external type or inherent to the operation of the network itself, which can also condition its service lifetime. Therefore, it is important to know the main factors affecting them, as well as to be able to estimate their frequency of failures over the time, in order to take in advance the most adequate actions to prevent sudden failures that cause service interruptions, and with it, economic losses both to the users and to the operating companies.

In this work, the application of a stochastic model of simple variable is collected, which is developed based on the application of a proposed methodology within the maintainability and reliability engineering field, from which useful life times are estimated and it is also predicted failures of samples of pipes with similar characteristics. In addition, a comparison of the proposed method with one of accelerated life of multiple variables is carried out.

In this way, the aim of this work is to propose a simple and practical method, with technical and scientific basis, of the survival analysis of potable water networks, in order to be adopted by water companies that seek within their objectives to develop their level in asset management.

As a case of study for the application of this methodology, we have the data of a supply which has a basic data structure and an information management system GIS,

which also has a record of failures incidents for a period of less than three years. Therefore, the methodology will be apply to a case of study on which there is a short-term data record, which at a general level can be considered a common condition among water companies.

In the proposed methodology, the service life for pipe samples of homogenous or similar characteristics is estimated taking into consideration two scenarios. In one scenario only the first failure occurred in each pipe is considered and in the other scenario all the failures are analyzed, that is, including repetitive failures on a same element; in both cases over a given observation period. In addition, it is initially assumed and subsequently verified that fault incidence times can be represented by certain parametric probability distribution functions, such as the Weibull distribution, from which it is possible to represent the Davies curve or also known as bathtub curve, which is commonly used to represent failures during useful life of some systems or equipment.

Finally, the analysis and interpretation of the curves and parameters that describe this distribution function, provide an orientation in terms of reliability for decision making regarding the approach of recommended maintenance strategies, tactics and actions within an asset management plan

Palabras clave español (máximo 5): Vida útil; Predicción de fallos; Rotura de tubería; Análisis de supervivencia; gestión de redes de agua

Palabras clave valenciano (máximo 5): Vida útil; Predicció de fallades; Ruptura de canonada; Anàlisi de supervivència; gestió de xarxes d'aigua

Palabras clave inglés (máximo 5): Useful lifetime, Failure prediction, Pipe breakages, survival analysis, water network management

Dedicatoria

A los sueños y esperanzas que nos motivan a superarnos cada día, a la vida que siempre nos presenta nuevas oportunidades, a los errores por las grandes lecciones aprendidas y a todo aquello que hemos dejado atrás en búsqueda de la felicidad.

A mi esposa, a mis padres y seres queridos por su confianza y apoyo incondicional.

Agradecimientos

Este trabajo nace del programa formativo de la Cátedra Aguas de Valencia, correspondiente al periodo 2018-2019, para el desarrollo de temas de investigación que se concreta en la realización de este trabajo de fin de máster, relacionado al estudio de la gestión de activos y vida útil de elementos de redes de agua potable. Doy gracias en general a la Cátedra Aguas de Valencia y a su director Abel Solera, por la oportunidad de haber sido parte de este proyecto de investigación; el cual resultó muy interesante y enriquecedor durante los nueve meses de su desarrollo.

Agradezco a las personas que estuvieron involucradas directamente en el desarrollo de este trabajo: tutor Ricardo Cobacho, profesor asociado al Departamento de Ingeniería Hidráulica y Medio Ambiente - ITA.; cotutora Amparo López, directora del Departamento de Ingeniería Hidráulica y Medio Ambiente de la Universitat Politècnica de València (UPV); cotutor Felipe Sanz, director de Operaciones Agua Potable en Aguas de Valencia (Global Omnium).

Un agradecimiento especial a Ricardo Cobacho por la confianza brindada, y por su valiosa guía, enseñanza, supervisión y sobre todo el apoyo permanente durante todo el proceso que comprendió este trabajo.

Gracias a David Torres y Román Ponz de Global Omnium quienes, con Felipe Sanz a la cabeza, estuvieron atentos compartiendo sus experiencias y ofreciendo el suficiente soporte en el transcurso de este proceso para el desarrollo del caso de estudio.

A los profesores del máster de ingeniería hidráulica y medio ambiente (MIHMA), por los conocimientos que fueron impartidos en cada una de las clases y que motivan a seguir aprendiendo cada día.

A todos quienes conforman el ITA y en general docentes y personal administrativo de la UPV, con quienes he tenido el gusto y dicha de compartir jornadas diarias de trabajo, y aportaron con su guía y ayuda en varias ocasiones.

Por último, pero no menos importante, a Dios y a todos quienes de cerca y también de lejos me han acompañado y alentado durante estos meses de trabajo y en general durante mi estancia en Valencia, España.

Índice de Contenido

Resumen del Trabajo de Fin de Master	i
Dedicatoria	vii
Agradecimientos	viii
Índice de Siglas y Acrónimos	xii
1 Introducción	1
1.1 Planteamiento del problema	1
1.2 Gestión de activos	2
1.3 Gestión infraestructural de redes de agua potable	4
1.3.1 Niveles de planificación	5
1.3.2 Otras referencias	6
1.4 Objetivos del trabajo	7
1.5 Estructura del documento	8
2 Estado del arte	10
2.1 Análisis de fiabilidad	10
2.1.1 Funciones de fiabilidad de fallos básicas	10
2.2 Modelos básicos de fiabilidad	13
2.2.1 Modelos Paramétricos	13
2.2.2 Modelos No Paramétricos	21
2.2.3 Modelos semiparamétricos	22
2.3 El análisis de datos de fallos	24

2.4	Descripción de Síntesis Universal de Medición CMD	28
2.5	Modelos de análisis de redes de agua.....	30
2.5.1	Modelos de estimación de vida útil.....	30
2.5.2	Modelos de predicción de fallos.....	31
3	Metodología.....	38
3.1	Síntesis de evaluación de redes de agua potable. SERAP	38
3.1.1	Variables de análisis y estructura de datos	41
3.2	Modelo de riesgos proporcionales de Weibull. WPHM	47
3.2.1	Descripción general	47
3.2.2	Estimación de parámetros.....	51
3.2.3	Predicción de fallos.....	52
4	Aplicación de los modelos.....	54
4.1	Breve descripción del caso de estudio	54
4.2	Selección de muestras o subgrupos de análisis	55
4.3	Modelo SERAP	57
4.3.1	Planteamiento 1	57
4.3.2	Planteamiento 2	64
4.3.3	Variable tiempo entre fallos TBF	68
4.4	Modelo WPHM	72
4.5	Comparación entre modelos	76
5	Conclusiones	80
5.1	Bases y estructuración de datos.....	81
5.2	Modelos.....	82
5.3	Trabajos y líneas de instigación futuras	84
	Referencias.....	86
	Anexos	

Índice de Siglas y Acrónimos

AL	Age-Left
AWWA	American Water Works Association
CARE	Computer Aided Rehabilitation of Water Networks
CMD	Confiabilidad, Mantenibilidad y Disponibilidad
E	Edad hasta el fallo
FC	Fibro cemento
FD	Fundición Dúctil
FG	Fundición Gris
FDAE	Función de Distribución acumulado Empírica
KM	Kaplan-Meier (estimador)
K-S	Kolmogorov-Smirnov (Test de ajuste)
LSE	Least Squares Estimate
MLE	Maximum Likelihood Estimate
MTTF	Mean Time To Failure
MTBF	Mean Time Between Failure
NOPF	Numbers Of Previous Failure
rc	Right censored
nrc	Non right censored
PE	Polietileno
R	(lenguaje de programación orientado al software de análisis estadístico)
ROCOF	Rate of Occurrence of Failure
SERAP	Síntesis de evaluación de redes de agua potable
SIG	Sistema de Información Geográfica

TBF	Time Between Failures
TTF	Time To Failure
VBA	Visual Basic for Applications
WPHM	Weibull Proportional Hazard Model
WALM	Weibull Accelerate Lifetime Model

1 Introducción

1.1 Planteamiento del problema

En las infraestructuras de servicios públicos, como los sistemas de distribución de agua potable, existen diversos factores de tipo económico, social o ambiental; que exigen realizar una gestión eficiente de estos sistemas, de manera que se pueda asegurar su disponibilidad y garantizar su sostenibilidad en el tiempo. Ello requiere de un direccionamiento apropiado de los recursos con los que cuentan las empresas operadoras del servicio, que muchas veces además de ser limitados pueden resultar insuficientes.

Para llevar a cabo una gestión de activos eficiente en una red de agua potable, es necesario conocer el ciclo de vida de los activos y sus costos asociados. Esto engloba las etapas de diseño, construcción, operación-mantenimiento y posterior eliminación o reemplazo. Durante estas etapas se puede ver afectada la vida útil de los activos, por lo que es importante reconocer los factores de mayor significancia que inciden en los fallos a fin de tomar acciones necesarias para prevenir el desgaste o fallo prematuro. Sin embargo, en la práctica, además de resultar un reto significativo controlar todas las variables que afectan a los elementos de un sistema, los mismos están sujetos a una degradación propia natural del material del cual están conformados, este último es otro factor que afecta la vida útil.

Parte medular de estas infraestructuras está constituida por conductos que pueden estar conformados de diversos materiales y diámetros, que en conjunto forman las

redes de suministro de agua, las cuales merecen una especial atención ya que estas representan, en gran medida, parte principal del patrimonio de las empresas proveedoras de este servicio; y las decisiones que se toman sobre las mismas suponen millonarios montos anuales en inversiones ya sea para su renovación, mantenimiento o instalación de nuevas redes.

En este trabajo se aborda el problema de la gestión de estos activos, específicamente en lo que se refiere a su vida útil, dada la importancia dentro del manejo de recursos y en la planificación de la operación y mantenimiento de redes en abastecimientos urbanos. Para ello, se plantea realizar un estudio estadístico del historial de incidencia de fallos o reparaciones, de cierto abastecimiento urbano.

En la literatura existen diversos estudios en los que se proponen diferentes tipos de modelos para el análisis y evaluación del estado de redes de agua. En general los modelos de predicción de fallos se clasifican en determinísticos, estocásticos o probabilísticos y redes neuronales. El desarrollo de este trabajo planteará la búsqueda de un modelo de tipo estadístico.

Este trabajo supone además otro enfoque distinto al estudio realizado por Mogollón (2017), relacionado también a la gestión de activos, en el cual se usa un modelo de tipo determinístico, desarrollado por Shamir y Howard (1979), para determinar la probabilidad de rotura de tuberías de un abastecimiento.

Se busca, a partir de una revisión bibliográfica, proponer una metodología o modelo que permita estimar la vida útil de servicio de las tuberías de una red de agua, el cual sea de fácil aplicación para empresas de agua potable y que sirva de apoyo en el desarrollo de su gestión de activos.

1.2 Gestión de activos

La gestión de activos dentro del ámbito de la industria y empresa evalúa el ciclo de vida de los bienes activos físicos o de los recursos disponibles para maximizar su valor durante su tiempo de vida útil. La norma ISO 55000 (2014), la define como: “La

coordinación de las actividades de una organización para crear valor a través de sus activos”.

Esto se relaciona con la ingeniería de mantenibilidad y confiabilidad, la cual se encarga de estudiar y analizar los fallos y reparaciones en los sistemas para el mejoramiento de la operación y aumento de la disponibilidad de funcionamiento y de su vida útil. La confiabilidad es definida como la probabilidad de que un elemento o sistema realice una función durante un tiempo determinado bajo ciertas condiciones establecidas. La mantenibilidad es la probabilidad de que un elemento o sistema que ha fallado pueda ser recuperado o reparado bajo un tiempo y condiciones de mantenimiento específicas. La disponibilidad es la probabilidad de que un componente o sistema esté en funcionamiento cuando este sea requerido en un momento dado cuando se opera bajo condiciones establecidas. Estos conceptos son definidos por Ebeling (2010).

Las acciones de mantenimiento se realizan en función de los objetivos seguidos por una empresa y su nivel de desarrollo en la gestión del mantenimiento. En la Tabla 1, Mora (2009) presenta un cuadro donde se aprecia la evolución en el tiempo del mantenimiento y las distintas etapas que conlleva en su desarrollo.

Así se tiene (I) una etapa inicial orientada a acciones correctivas solo atendiendo fallos imprevistos; (II) el desarrollo de acciones planificadas, las cuales sirven para prevenir fallos; (III) aplicación de tácticas de mantenimiento donde ya se comienza a tener un sistema organizado y donde los conjuntos de acciones buscan la maximización de la productividad o explotación y reducción del riesgo; (IV) implementación de estrategias y definición de metas a alcanzar, donde según indicadores, costos de ciclo de vida, análisis de curvas de fallos (curva de bañera), confiabilidad, etc., se miden los resultados de las acciones y tácticas implementadas; (V) desarrollo y fortalecimiento tecnológico, competencias profesionales y aplicación de técnicas de mantenimiento integral como el análisis de modos de fallos y análisis de causa raíz de falla; (VI) finalmente se conforma la gestión integral de activos, la cual aplica las mejores prácticas desarrolladas para

poder anticiparse a los requerimientos que puedan ocasionar fallos, a fin de conseguir la máxima disponibilidad de funcionamiento del sistema.

Tabla 1. Evolución del mantenimiento y sus etapas de desarrollo (Mora, 2009)

Etapa	Sucede aproximadamente	Producción - Manufactura		Mantenimiento e ingeniería de fábricas	
		Orientación hacia...	Necesidad específica	Orientación hacia...	Objetivo que pretende
I	antes de 1950	el producto	generar el producto	hacer acciones correctivas	reparar fallos imprevistos
II	entre 1950 y 1959	la producción	estructurar un sistema productivo	aplicar acciones planeadas	prevenir, predecir y reparar fallos
III	entre 1960 y 1980	la productividad	optimizar la producción	establecer tácticas de mantenimiento	gestar y operar bajo un sistema organizado
IV	entre 1981 y 1995	la competitividad	mejorar índices mundiales	implementar una estrategia	medir costos, CMD, compararse, predecir índices, etc.
V	entre 1996 y 2003	la innovación tecnológica			
VI	desde 2004	Gestión y operación integral de activos en forma coordinada entre ambas dependencias anticiparse a las necesidades de los equipos y de los clientes de mantenimientos - Predicciones - Pronósticos - Gestión de activos			

La gestión de activos se complementa tomando en consideración los temas transversales que afectan a cada etapa o nivel de desarrollo, así se puede nombrar, por ejemplo: la gestión de riesgos, objetivos, planes estratégicos, toma de decisiones, recursos e inversiones, o adaptabilidad.

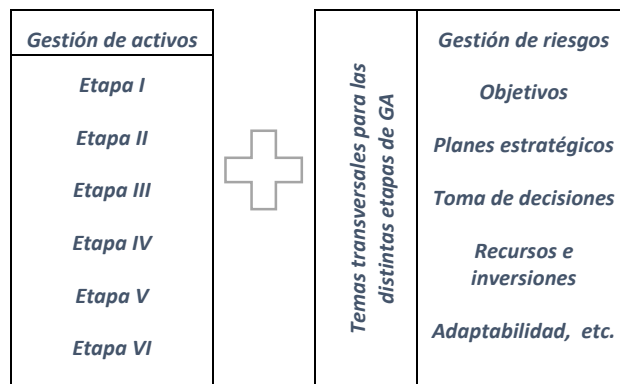


Figura 1. Etapas de gestión de activos y temas transversales a considerar

1.3 Gestión infraestructural de redes de agua potable

Cuando el enfoque de la gestión de activos se realiza sobre los activos físicos se conoce como gestión infraestructural de activos (Alegre y Coelho, 2012) y su aplicación en las redes de agua potable es relativamente reciente. A mediados y finales de los 80, países

como Australia y Reino Unido fueron los primeros en implementar planes para la gestión de activos y a partir de ello se han desarrollado guías y metodologías para su aplicación (van Heck, 2008; Amaral et al., 2017).

Este tema se ha desarrollado por diversas razones como requisitos legislativos gubernamentales, regulatorios o por el desarrollo de mejoras en distintas organizaciones mediante guías técnicas de buenas prácticas.

Esto se refleja en el desarrollo de planes de gestión que tienen como objetivo alcanzar las metas organizacionales involucrando todos los niveles de decisión que las conforman.

1.3.1 Niveles de planificación

Los planes de gestión, si bien integran a toda una organización, se desarrollan en distintos niveles. Así, se tiene tres tipos de planes: estratégico, táctico y operacional.

La planificación estratégica se aplica en un ámbito global y se relaciona con la visión y misión de empresa, donde la toma de decisiones está a cargo de los directivos y el horizonte de las metas es a largo plazo (entre 10 y 20 años)¹. Ejemplo: plan de inversiones.

La planificación táctica actúa sobre todo el sistema de distribución. Los planes están a cargo de los mandos intermedios y tienen como objetivo definir los medios para alcanzar las metas del plan estratégico. La proyección de estos planes se realiza a medio plazo (entre 3 y 5 años). Ejemplo: plan de gestión operacional y mantenimiento.

La planificación operacional comprende componentes, elementos o un área específica de intervención. Está a cargo de los mandos operativos, quienes se encargan de ejecutar

¹ Los plazos de ejecución mostrados para cada plan son referenciales y no deben ser considerados como determinantes

acciones relacionadas al plan al plan táctico, proyectando planes en un tiempo horizonte de corto plazo (1 a 2 años). Ejemplo: plan de mantenimiento de válvulas.

1.3.2 Otras referencias

De acuerdo al manual de entrenamiento para operadores de redes de agua de la American Water Works Association AWWA (Lauer, 2013), la gestión infraestructural en estos sistemas tiene como objetivo entregar de manera confiable y eficiente agua potable de calidad para satisfacer las necesidades de los usuarios, como cumplir con los flujos requeridos contra incendio. Para ello se plantea la necesidad de desarrollar planes de gestión, los cuales especifiquen objetivos a cumplir y que estos a su vez puedan ser valorados. Se destacan la calidad del agua, confiabilidad del sistema y la eficiencia como puntos prioritarios a ser atendidos.

Un factor significativo que actúa sobre estos 3 puntos, y que atañe sobre el problema planteado de este trabajo, es la sustitución o rehabilitación de tuberías. Este factor apunta sobre tuberías que alcanzan una edad avanzada o que presentan constantemente problemas. Un plan de sustitución y rehabilitación de redes debe ser considerado, y en él, priorizadas las redes sobre las que se debe intervenir. Para ello se anota como fundamental contar con un inventario que guarde detalle de las redes, así como también un historial de fallos o reparaciones. Los criterios para la selección de las redes prioritarias a renovar toma en consideración métodos de valoración de activos, análisis económicos, probabilidad de fallos o modelos de predicción de fallos. El manual M28 (AWWA, 2014) recoge varios enfoques para el reemplazo de tuberías que pueden ser consultados.

Uno de los primeros proyectos europeos referente a la gestión infraestructural de redes de agua fue el proyecto CARE-W (Computer-Aided Rehabilitation of Water Networks). El mismo concluyó en el 2005 y tuvo como objetivo proveer un enfoque integrado de la rehabilitación de este tipo de infraestructuras. Esto se realizaba según indicadores de desempeño, herramientas estadísticas para el análisis de supervivencia, procedimientos

para planes de rehabilitación y estrategias a largo plazo. Todo ello combinado en un paquete de software (Saegrov et al., 2003).

Este proyecto ha servido de base para el surgimiento de nuevas propuestas como AWARE-P, el cual es una metodología para la gestión de activos con un software de código abierto, producto de un proyecto de I+D para la gestión infraestructural de servicios públicos de agua en Europa.

1.4 Objetivos del trabajo

A partir del estudio de la incidencia de frecuencias de fallos en los tramos de tubería de una red, se pretende estimar la vida útil de las tuberías, analizar su supervivencia y evaluar su estado.

En este trabajo se analizan los tiempos de fallo independientemente de los modos o causas que los generan. Con ello se busca obtener una solución lo más práctica, sencilla y fiable posible que permita a las empresas gestoras del suministro de agua potable que estén en camino a desarrollar su nivel de gestión de activos, implementar el procedimiento propuesto en este trabajo como herramienta para facilitar la toma de decisiones en cuanto a manejo de recursos y planificar estrategias o acciones de mantenimiento, a fin de alcanzar un funcionamiento óptimo de la red.

Dado el enfoque y el objetivo general planteado, se trazan los siguientes objetivos específicos que se espera permitan alcanzar las metas proyectadas:

- Revisar el estado del arte que permita conocer los antecedentes y el desarrollo del campo de estudio del análisis de fallos, particularmente en redes de tuberías
- Definir un procedimiento o metodología a seguir para proponer tiempos de vida de elementos presentes en tuberías
- Construir un modelo que permita determinar la influencia sobre la vida útil de los elementos de una red.
- Aplicar el modelo propuesto sobre un caso de estudio, comparándolo con otros modelos existentes disponibles.

1.5 Estructura del documento

El documento se compone de 6 capítulos, distribuidos de la siguiente manera:

Capítulo 1. Se plantea el problema de estudio, se realiza una breve descripción acerca de la gestión de activos y de la gestión de infraestructuras de redes de agua potable. Posteriormente se anotan los objetivos generales y específicos que se pretenden alcanzar en la realización del presente trabajo.

Capítulo 2. En función del primer objetivo específico se realiza un resumen de conceptos y literatura revisada referente al tema tratado, lo cual servirá de base para la comprensión del tema y de referencia en el desarrollo de la propuesta de este trabajo. En términos generales se revisará: los fundamentos del análisis de fiabilidad, análisis de datos de fallos y finalmente se concluirá realizando un resumen de los modelos estadísticos en redes de agua existentes.

Capítulo 3. En este capítulo se presenta la metodología y el modelo estadístico propuesto para la estimación de la vida útil y predicción de roturas. Se presenta un modelo según el análisis de supervivencia en el que se presentan dos enfoques. Uno en el que solo se realiza el análisis de la primera rotura de una tubería, y otro que contempla fallos sucesivos que se pueden presentar en un tubo, bajo ciertas hipótesis. Adicionalmente, aplica el modelo de riesgos proporcionales de Weibull (WPHM, por sus siglas en inglés), para la predicción de fallos, a fin de contrastar los resultados con el modelo propuesto para posteriormente determinar beneficios y desventajas que presenta cada uno.

Capítulo 4. Se introduce brevemente sobre las características generales del caso de estudio y se aplican los modelos planteados, presentando tablas, gráficos y el análisis de los resultados obtenidos.

Capítulo 5. Se presentan las conclusiones derivadas del análisis de resultados y en general de los modelos presentados en función de los objetivos del trabajo, marcando las limitaciones que se presentaron en el desarrollo del mismo y presentando las

posibles líneas de investigación futuras a desarrollar en base al trabajo realizado y los resultados obtenidos.

2 Estado del arte

2.1 Análisis de fiabilidad

En la ingeniería de la confiabilidad y mantenibilidad se usan modelos matemáticos para el estudio de la fiabilidad² de los elementos de un sistema. Estos modelos pueden ser contruidos a partir de un registro de datos, tomando en consideración distintas funciones de distribución de probabilidad para la representación de fallos. A continuación, brevemente se describen estas funciones de fiabilidad de probabilidad de fallos

2.1.1 Funciones de fiabilidad de fallos básicas

Las funciones básicas en el análisis de fiabilidad son cuatro: función de probabilidad de densidad, función de distribución acumulativa o in fiabilidad, función de confiabilidad o supervivencia y función de tasa de fallos o también conocida como función de riesgo. Estas funciones se relacionan entre si y proporcionan diferente información específica sobre el proceso de fallo.

Su relación se define tomando en consideración una variable aleatoria continua $T [0, \infty)$ que es el tiempo de fallo del sistema o componente, y considerando (t) como el tiempo que funcionará determinado sistema o elemento. La función de densidad representa la probabilidad de fallo de una variable aleatoria (T) y tiene las siguientes propiedades:

² Los términos fiabilidad y confiabilidad son semejantes en el desarrollo de este trabajo.

$$f(t) > 0 \quad \text{y} \quad \int_0^{\infty} f(t)dt = 1$$

La función $F(t)$ es la función de distribución acumulada y se define como:

$$F(t) = \int_0^t f(t) dt = P\{T < t\} \quad (2.1)$$

De manera que, para un tiempo $t=0$, $F(0) = 0$ y cuando t tiende al infinito, $F(\infty) = 1$, siendo $F(t)$ la probabilidad que una falla ocurra antes de un tiempo t .

$R(t)$ es la función de confiabilidad³ y se define como:

$$R(t) = 1 - F(t) = P\{T \geq t\} \quad (2.2)$$

Cumpléndose:

$$f(t) = \frac{dF(t)}{dt} = -\frac{dR(t)}{dt} \quad (2.3)$$

De manera contraria a $F(t)$, para un tiempo $t=0$, $R(0) = 1$ y cuando t tiende al infinito $R(\infty) = 0$. Así, $R(t)$ es la probabilidad de que el tiempo hasta el fallo (T) sea mayor o igual a t .

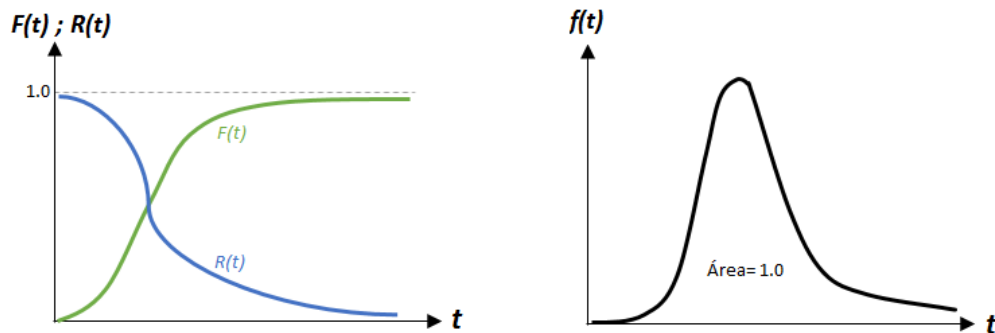


Figura 2. Representación esquemática de funciones de distribución de $F(t)$, $R(t)$ (izquierda) y $f(t)$ (derecha).

³ Dependiendo del autor la confiabilidad o supervivencia puede ser denotada con la letra R o S

Finalmente, la cuarta función conocida como tasa de fallos o función de riesgo, proporciona la tasa de fallo instantánea para un determinado tiempo t , hasta el intervalo de tiempo siguiente Δt lo cual se representa como se muestra a continuación:

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P\{t \leq T \leq t + \Delta t | T \geq t\}}{\Delta t} \quad (2.4)$$

La tasa de fallos, $\lambda(t)$ o $h(t)$, se puede expresar en función de la supervivencia o fiabilidad de la siguiente manera:

$$\lambda(t) = h(t) = -\frac{dR(t)}{dt} \times \frac{1}{R(t)} = \frac{f(t)}{R(t)} \quad (2.5)$$

La forma de esta función puede ser de tipo creciente, decreciente o constante, y a partir de ella es posible representar la curva de Davies, o mejor conocida como “curva de bañera”.

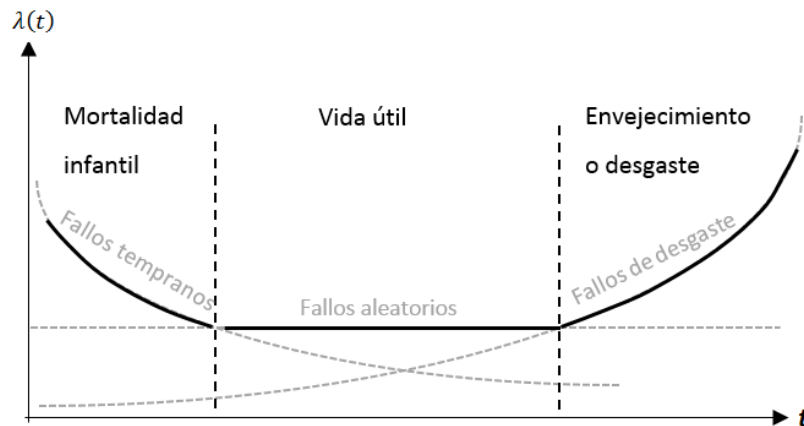


Figura 3. Esquema general de Curva de bañera y sus fases principales.

Esta curva se conforma por tres fases principales. La primera, conocida como mortalidad infantil es característica de equipos o componentes que inician su puesta en marcha y presentan inicialmente una alta cantidad de fallos que se reducen en el tiempo hasta que se estabilizan. La segunda fase, denominada vida útil, corresponde a la ocurrencia de fallos aleatorios con una frecuencia de fallos baja y tasa constante, donde los fallos presentados se deben a accidentes o un estrés excesivo inusual de una variable

explicativa o factor de afectación en el sistema. En la tercera fase de envejecimiento o desgaste, los fallos se producen debido al estrés provocado por alguna variable presente durante toda la vida útil o pérdida de capacidad de resistencia de los elementos debido a características propias del mismo tras cumplir su vida útil de servicio.

En relación a la mantenibilidad, para la primera fase priman las acciones correctivas, para la segunda fase las acciones modificativas y para la tercera fase se usan acciones preventivas o predictivas. Más detalles acerca de la curva de bañera y de las funciones de distribución básicas de fallos pueden ser consultados en Ebeling (2010) y Mora (2009).

2.2 Modelos básicos de fiabilidad

Los modelos para el estudio de la fiabilidad o análisis de supervivencia pueden utilizar distribuciones de probabilidad, según métodos de tipo paramétrico, no paramétrico y semiparamétrico. Según Solano (2008), los métodos paramétricos y semiparamétricos son los más extendidos ya que sus estimaciones dan lugar a análisis estadísticos más elaborados. En particular, estos modelos permiten analizar tiempos de vida, procesos de fallo y su predicción. Los métodos no paramétricos son más sencillos de aplicar, pero son menos eficientes. Si bien este último no permite realizar un análisis predictivo, es útil para obtener una primera aproximación de los procesos de fallos ya que representa de manera más precisa su comportamiento, a diferencia de un modelo paramétrico que podría resultar no muy preciso al seguir una tendencia determinada. A continuación, se explica un poco más acerca de los tres modelos.

2.2.1 Modelos Paramétricos

Existen diversas distribuciones de tipo paramétrico que son útiles para la representación de los procesos de fallo. Entre ellas se mostrarán más en detalle las distribuciones Weibull, normal, lognormal y gamma, que son distribuciones continuas dependientes del tiempo y referenciadas por varios autores para el análisis de fiabilidad.

Antes de ello, cabe mencionar otro tipo de distribución comúnmente usada en el estudio de fiabilidad, la distribución exponencial. Esta es un caso especial de la distribución Weibull y gamma que se describirán posteriormente. Se caracteriza por ser la única en poseer una intensidad constante de fallas y es usualmente utilizada para modelar el tiempo de vida de componentes electrónicos (Mora, 2009). Se define por un parámetro λ que define la tasa de fallos, la cual es constante a lo largo del tiempo y si bien esta distribución permite representar la fase constante de vida útil de la curva de bañera, su aplicación no es práctica ya que λ en la práctica suele variar en el tiempo. En su lugar se utilizan las distribuciones antes mencionadas Weibull o Gamma. La distribución exponencial también es conocida por su característica de ausencia de memoria, lo cual significa que los tiempos de falla no son dependientes del tiempo de operación, envejecimiento o desgaste.

Distribución Weibull

Esta distribución es una de las más importantes en el análisis de supervivencia o tiempos de vida. Tiene la característica de poder adaptarse o seguir la forma de diferentes funciones de distribución, donde además, su función de tasa de fallos $\lambda(t)$ puede emplearse para obtener la curva de la bañera. De esta manera, es posible representar la evolución de la vida útil de los elementos de un sistema (fallos tempranos, fallos durante vida útil y fallos por desgaste o envejecimiento). A continuación, se muestra la representación general de esta distribución, dadas las funciones de fiabilidad vistas anteriormente.

$$f(t) = \frac{\beta}{\eta} \left(\frac{t - \gamma}{\eta} \right)^{\beta-1} e \left[- \left(\frac{t - \gamma}{\eta} \right)^{\beta} \right] \quad (2.6)$$

$$F(t) = 1 - R(t) = 1 - e \left[- \left(\frac{t - \gamma}{\eta} \right)^{\beta} \right] \quad (2.7)$$

$$\lambda(t) = \frac{\beta}{\eta} \left(\frac{t - \gamma}{\eta} \right)^{\beta-1} \quad (2.8)$$

La distribución Weibull se conforma de tres parámetros: forma (β), escala (η) y posición (γ). El parámetro de posición representa el intervalo de tiempo en que la probabilidad de fallo es nula. Debido a la complejidad para estimar este parámetro habitualmente se asume su valor igual a cero, pasando en estos casos a definirse la distribución solo con los parámetros de forma y escala.

El parámetro de β , determina la forma de cada función de distribución de probabilidad. En la función de riesgo representa la pendiente de la curva, la cual es: decreciente para $\beta < 0.95$; llano para valores de β entre 0.95 y 1.05; cóncavo para β entre 1.05 y 2 y creciente para $\beta > 2$. Así es como a través de este parámetro de distribución es posible representar la curva de la bañera (Figura 4).

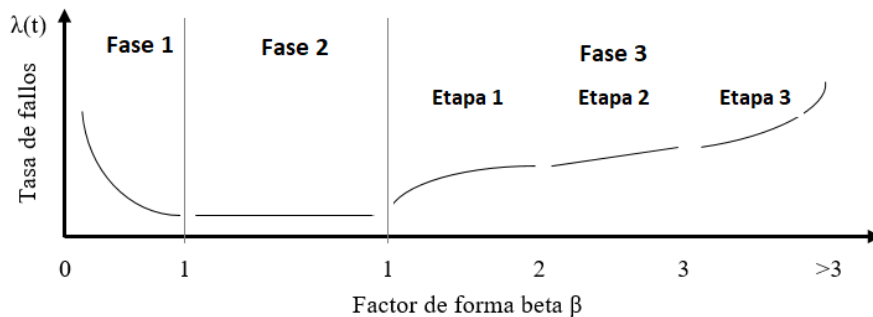


Figura 4. Fases de la Curva de Bañera según el factor de forma.

En la gráfica anterior se observan las diferentes fases de la curva de bañera explicadas previamente. La fase 3 sin embargo, presenta varias etapas descritas a continuación:

- *Etapa 1.* La tasa de fallos se incrementa levemente y los fallos son producto del paso del tiempo y de la intensidad de uso. En esta etapa se asume que existe algún tipo de control sobre las fallas, comenzando a utilizar acciones preventivas.
- *Etapa 2.* La tasa de fallos crece de manera lineal y constante. El crecimiento continuo de los fallos requiere que se comience a aplicar acciones de tipo predictivas.

- *Etapa 3.* Es la fase de envejecimiento en que la vida útil de los elementos se acelera, haciendo que la tasa de fallos incremente de forma exponencial. En este caso es necesario una sustitución como única alternativa.

Beta se considera un indicador de la medida de dispersión del comportamiento de los fallos y es inverso a la duración promedio de ellas (Mora, 2009).

El parámetro η , conocido también como vida característica está relacionado con la vida útil. Representa el tiempo en el que se espera que el 63.2% de los elementos fallen. Esto se deduce al igualar $(t - \gamma) = \eta$ en la función de distribución acumulada, ecuación (2.7), donde independiente mente del valor que tome β , siempre se cumple que la probabilidad es del 0.632. A diferencia de η , la media (μ), si se puede ver afectada para diferentes valores de β .

Aunque la vida característica no es la media de la distribución de Weibull, puede interpretarse como un indicador aproximado de posición (Solano, 2008). Por lo tanto, un alto valor obtenido de este parámetro representará un mayor tiempo de duración o vida útil.

El cálculo o estimación de estos dos parámetros (β, η) se puede realizar mediante métodos como estimación de máxima verosimilitud, regresión lineal con mínimos cuadrados o métodos gráficos, entre otros. De los métodos mencionados, el método gráfico es el más subjetivo, siendo los métodos de estimación de máxima verosimilitud y mínimos cuadrados los más recomendables por su origen de tipo matemático.

El método de mínimos cuadrados (LSE, por su abreviatura en inglés) es una técnica matemática que permite, establecer el mejor ajuste de determinada función para un conjunto de pares ordenados dados, para lo cual se cumple el criterio de mínimo error cuadrático. A partir de una función lineal ($y = bx + a$), se deberá cumplir que la sumatoria del cuadrado de las distancias de los puntos conformados por N pares ordenados (x_i, y_i) a una recta estimada, sea mínima:

$$\sum_{j=1}^N (\hat{a} + \hat{b} \cdot x_i - y_i)^2 \min(a, b) \sum_{j=1}^N (a + b \cdot x_i - y_i)^2 \quad (2.9)$$

donde, el valor de a es la intercepción al eje Y de la recta y b su pendiente.

Por otro lado, se transforma la distribución acumulativa $F(t)$, en este caso de la distribución Weibull ecuación (2.7), de manera que su expresión sea aproximada a la ecuación de una recta. Esto se consigue aplicando un doble logaritmo natural y propiedades de exponencial de los logaritmos, obteniendo:

$$\ln \left[\ln \left(\frac{1}{1 - F(t)} \right) \right] = \beta \ln(t - \gamma) - \beta \ln(\eta) \quad (2.10)$$

De esta manera la ecuación (2.10) puede ser representada como una función lineal de la forma ($y = bx + a$), realizando las siguientes transformaciones:

$$y = \ln \left[\ln \left(\frac{1}{1 - F(t)} \right) \right]; \quad x = \ln(t - \gamma); \quad a = \beta \ln(\eta) \quad (2.11)$$

donde se puede deducir que la pendiente de la recta $b = \beta$ y el parámetro de escala mediante la expresión $\eta = EXP(-a/\beta)$.

Para determinar la calidad de los datos de alineación de la recta obtenida, los cuales sirven además como una prueba previa de la bondad de ajuste, se calculan el coeficiente de correlación r y el coeficiente de determinación muestral r^2 . El coeficiente de correlación r indica que tan débil o fuerte es la relación lineal, y puede tomar valores entre -1 y 1, donde el signo denota la correlación inversa o directa de los datos. Se considera admisible cuando el rango de valores está entre 0.95 y 1. Por su parte, el coeficiente de determinación muestral r^2 indica el porcentaje de puntos que están relacionados linealmente, considerándose admisible dentro de un rango entre el 0.9025 y 1.0 (Mora, 2009).

Por otra parte, el método de estimación por máxima verosimilitud (MLE, por sus siglas en inglés) consiste en obtener el valor más probable o más verosímil de los parámetros de una distribución de un conjunto de datos dado. Esto se realiza maximizando la

función de máxima verosimilitud L aplicada a la función de densidad de la distribución $f(x)$ de la cual se requieren conocer sus parámetros. También es posible manipular la función aplicando logaritmo natural, ecuación (2.13), lo cual facilita los cálculos de los estimadores de máxima verosimilitud.

$$L(\theta_1, \theta_2, \dots, \theta_k | x_1, x_2, \dots, x_k) = L = \prod_{j=1} f(x; \theta_1, \theta_2, \dots, \theta_k) \quad (2.12)$$

$$\Lambda = Ln(L) = \prod_{j=1} \ln[f(x; \theta_1, \theta_2, \dots, \theta_k)] \quad (2.13)$$

donde θ_j son los parámetros a obtener.

Tomando en consideración la censura de los datos (ver Sección 2.3), la ecuación la ecuación (2.13) puede ser reescrita de la siguiente manera,

$$\Lambda = Ln(L) = \prod_{j=1(nrc)} \ln[f(x; \theta_1, \theta_2, \dots, \theta_k)] + \prod_{j=1(rc)} \epsilon[F(x; \theta_1, \theta_2, \dots, \theta_k)] \quad (2.14)$$

Donde los datos no censurados por derecha (nrc, por sus siglas en inglés), contribuyen a la función de densidad y los datos censurados por derecha (rc, por sus siglas en inglés), contribuyen a la función de supervivencia.

Desde el punto de vista estadístico, es un método muy potente, ya que permite trabajar con muestras grandes y posee propiedades como su forma asintótica consistente, naturaleza asintótica y geometría asintótica imparcial, lo cual se resumen en poder obtener con este método resultados muy precisos (Mora, 2009).

Su resolución se realiza aplicando métodos numéricos, como por ejemplo el método de Newton-Raphson y es necesario contar con un tamaño de muestra lo suficientemente grande, como referencia mayor a treinta (Levin et al., 1996), para evitar una mala estimación que pueda llevar a una equivocada interpretación de los resultados, siendo esta una de sus desventajas.

En la práctica existen diferentes herramientas informáticas de análisis estadístico que, por medio de ambos métodos (LSE, MLE), permiten obtener los resultados de estos

cálculos o estimaciones complejas. En este trabajo se ha utilizado el software de análisis estadístico *Rstudio (R)*, el cual permiten obtener los distintos parámetros deseados usando el método de máxima verosimilitud.

Distribución Normal

La distribución normal es una distribución frecuentemente utilizada en estadística. Su función de densidad se caracteriza por tener una forma acampanada simétrica. Los parámetros que describen esta distribución son μ y s , que representan la media y desviación estándar respectivamente. Su función de densidad es la que se muestra a continuación:

$$f(t) = \frac{1}{\sqrt{2\pi}s} \text{EXP} \left[-\frac{1}{2} \frac{(t - \mu)^2}{s^2} \right]; \quad -\infty < t < \infty \quad (2.15)$$

Si bien esta no es una distribución propia para el análisis de fiabilidad, debido a que su rango de valores incluye números negativos, la misma es modificada para que solo pueda tomar valores positivos. A partir de ello, puede ser aproximada para el análisis de procesos de fallo. La forma de su función de riesgo es creciente, por lo tanto, es útil para modelar fallos de la fase de envejecimiento o desgaste.

Mediante la transformación de normalización $Z = (T - \mu)/\sigma$, la distribución normal puede ser representada como sigue.

$$\varphi(Z) = \frac{1}{\sqrt{2\pi}} e^{-Z^2/2} \quad (2.16)$$

A partir de ello, las funciones de distribución acumulada y de riesgo puede ser representadas como se muestra a continuación:

$$F(t) = \Phi((t - \mu)/\sigma) \quad (2.17)$$

$$\lambda(t) = \frac{f(t)}{R(t)} = \frac{f(t)}{1 - \Phi((t - \mu)/\sigma)} \quad (2.18)$$

Los parámetros de (μ, σ) son estimados de la ecuación de la recta haciendo el uso de transformaciones de manera similar a lo descrito en Weibull con el método de LSE, con el uso de métodos gráficos o MLE.

Distribución Lognormal

Dada una variable aleatoria T , su tiempo de fallo, tiene una distribución lognormal, cuando el logaritmo natural de T tiene una distribución normal. Los parámetros que describen esta distribución son (μ, σ) , los cuales son el logaritmo de la media y desviación estándar de la variable aleatoria respectivamente.

La función de densidad está definida por:

$$f(t) = \frac{1}{\sigma t \sqrt{2\pi}} \text{EXP} \left[-\frac{1}{2} \left(\frac{\log t - \mu}{\sigma} \right)^2 \right] \quad (2.19)$$

Y su función de distribución:

$$F(t) = \Phi((\log t - \mu)/\sigma) \quad (2.20)$$

La forma de esta distribución puede variar llegado a un valor máximo y luego va decreciendo a un valor próximo a cero cuando t tiende al infinito, con una cola marcada hacia la derecha.

Al igual que la distribución anterior, sus parámetros pueden ser obtenidos a través de la ecuación de la recta haciendo uso de transformaciones necesarias o aplicando el método MLE o LSE.

Distribución Gamma

Esta distribución, de manera semejante a la distribución Weibull, se describe con dos parámetros, uno de forma α y otro de escala γ . En particular, cuando $\gamma = 1$ la distribución es igual a una exponencial con media α . La función de densidad se representa mediante la siguiente expresión:

$$f(t) = \frac{t^{\gamma-1} e^{-t/\alpha}}{\alpha^\gamma \Gamma(\gamma)}; \gamma, \alpha > 0 \text{ y } t \geq 0 \quad (2.21)$$

donde Γ es la función Gamma.

Esta distribución puede también ajustarse a varias formas y se recomienda para caracterizar fallos durante periodos de rodaje (Mora, 2009). Para estimar sus parámetros se usa el MLE, optimización de funciones o métodos numéricos iterativos.

2.2.2 Modelos No Paramétricos

Son modelos de análisis de inferencia estadística que se definen a partir de datos observados, sin seguir ningún tipo de modelo probabilístico. Esto significa que los modelos no son aptos para realizar predicciones a diferencia de los modelos paramétricos. Sin embargo, permiten realizar un análisis inicial de los datos, que dependiendo del propósito puede resultar suficiente. Además, permite representar la distribución que siguen los datos, donde, a partir de ello, puede ser relacionado a un modelo de tipo paramétrico.

Solano (2008) presenta tres modelos no paramétricos destinados al análisis de supervivencia: tablas de vida, el estimador de Kaplan-Meier (K-M) y el método de Nelson-Aalen extendido (NAE).

A efectos de interés del desarrollo del presente trabajo se hace una descripción del modelo de Kaplan-Meier, que considera datos con censura por derecha y no supone fallos ni censuras previo al periodo de observación, lo cual se ajusta a las consideraciones del modelo propuesto, que se explica y desarrolla en los capítulos 3 y 4.

Estimador de Kaplan-Meier

El estimador de Kaplan-Meier conocido también como estimador producto-límite, fue desarrollado en 1958 por Edward L. Kaplan y Paul Meier y se expresa como:

$$\hat{R}(t) = \prod_{j:t_j < t} \frac{n_j - d_j}{n_j} \quad (2.22)$$

donde t_j es el tiempo hasta el fallo ordenado de menor a mayor, n_j es el número de elementos supervivientes en instante t_j y d_j es el número de fallos ocurridos en el instante t_j .

Algunas características de este estimador son las siguientes (Solano, 2008):

- Cuando se trabaja con datos sin censura, este estimador coincide con el estimador no paramétrico de la función empírica de fiabilidad.
- Posee propiedades asintóticas, permitiendo representar problemas con datos censurados (el tema referente a la censura se aborda dentro de la sección 2.3).
- Funciona bien para muestras de tamaño grande, no así para las muestras de tamaño pequeño.
- Su función puede representarse de manera puntual o escalonada.
- Es un estimador de máxima verosimilitud. Sin embargo, su eficiencia asintótica es inferior a la de un estimador paramétrico de máxima verosimilitud.

2.2.3 Modelos semiparamétricos

Se denominan modelos semiparamétricos ya que toman en consideración funciones paramétricas para la modelación de datos, junto a otras funciones básicas o arbitrarias que las afectan. Son utilizados principalmente para describir los efectos de las variables explicativas.

De ellos, el modelo de riesgos proporcionales de Cox es uno de los más utilizados en análisis de supervivencia, este puede expresarse de la siguiente manera.

$$h(t; Z) = h_0(t)e^{b^T Z} = h_0(t)e^{b_1 Z_1 + b_2 Z_2 + \dots + b_p Z_p} \quad (2.23)$$

Donde $h(t; Z)$ es la función de riesgo, t es una variable de tiempo, Z es un vector de covariables, $h_0(t)$ es una función básica o arbitraria, b es un vector de coeficientes que se estima por máxima verosimilitud. En la siguiente figura, extraída de Kleiner y Rajani (2001), se muestra el efecto de las covariables sobre un función de riesgo $h(t)$ y supervivencia $R(t)$.

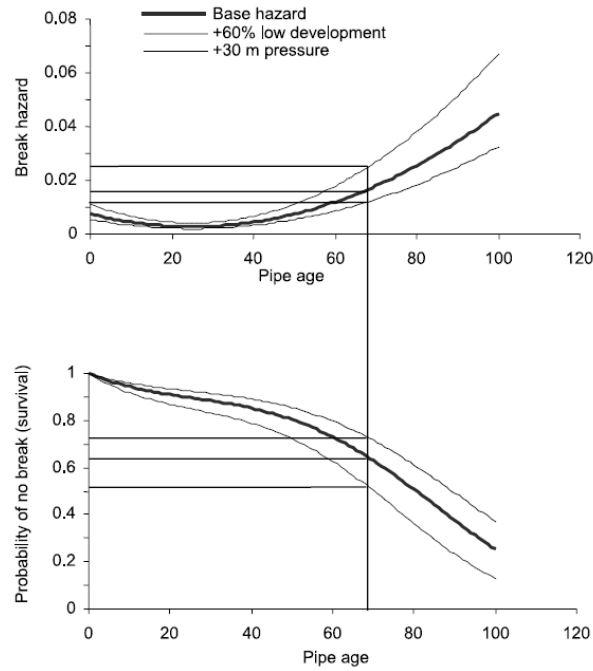


Figura 5. Efecto de las variables de presión en un Modelo de Riesgo proporcional (Kleiner y Rajani, 2001).

Las influencias de las distintas variables son incluidas dentro del modelo de manera multiplicativa. En caso de considerarse que estas no actúan de esta manera, los datos pueden ser estratificados y afectados con otra función de riesgo, que se suponga más adecuada según diversos criterios.

El modelo de riesgos proporcionales de Cox nos permite determinar la influencia de las variables explicativas o covariables, las cuales se consideran que no dependen del tiempo. Pero cuando se trata de estimar el comportamiento de los datos en el tiempo se requiere seleccionar una función paramétrica adecuada que nos permita modelar su comportamiento. Un caso particular es el método de riesgos proporcionales de Weibull, donde la función de riesgo asumida $h_0(t)$ es una distribución Weibull.

En estos modelos, cuando se analiza la proporción de riesgo entre dos individuos (dos tuberías), se comprueba que su relación es constante, de aquí el hecho que este modelo se denomine de *riesgo proporcional*, como se ilustra en la siguiente expresión.

$$\frac{h(t; Z_i)}{h(t; Z_j)} = \frac{h_0(t)e^{b^T Z_i}}{h_0(t)e^{b^T Z_j}} = \frac{e^{b^T Z_i}}{e^{b^T Z_j}} = cte$$

2.3 El análisis de datos de fallos

Hasta el momento se han revisado brevemente las bases del análisis de supervivencia y algunos de los modelos básicos. Sin embargo, antes de aplicar cualquier concepto o método es necesario conocer los datos con los que se cuenta, y a partir de ello seleccionar la ruta de análisis más adecuada o factible.

El proceso de análisis puede estar condicionado a los datos disponibles, los tipos de pruebas de evaluación, y la duración de las pruebas (tiempo de observación), entre otros. El diseño de las pruebas y recolección de datos deben ser determinados con criterios estadísticos y técnicos según objetivos específicos.

El análisis de datos de fallos aborda una gran cantidad de conceptos y definiciones a tomar en consideración. Estos permiten estructurar y desarrollar un análisis confiable de la información. A continuación, se describen los conceptos de elementos reparables y no reparables, censura y pruebas estadísticas, los cuales son de interés en el desarrollo de este trabajo.

Elementos reparables y no reparables

A fines estadísticos es sustancial conocer si los datos de fallo de un estudio provienen de elementos o componentes de tipo reparables o no reparables.

Los de tipo reparables son aquellos que generalmente se componen de partes que pueden ser reparadas, tal es el caso de un vehículo, un ordenador, maquinaria, etcétera. De manera general, los elementos de tipo reparable se componen de elementos no reparables.

Los no reparables son aquellos elementos o componentes que se conocen como de una sola vida, que debido a su naturaleza resulta más práctico reemplazarlo, reciclarlo o desecharlo. Como ejemplo de elementos no reparables se puede mencionar: fusibles, bombillas, botellas de vidrio, etc.

Las tuberías son un caso particular ya que dependiendo del criterio del analista puede considerarse como un elemento de tipo reparable o no reparable.

Para los elementos reparables se toma en consideración el tiempo medio entre fallos (MTBF, por sus siglas en inglés), mientras que para los no reparables se considera el tiempo medio hasta el fallo (MTTF, por sus siglas en inglés), el cual su tiempo de vida desde el inicio del periodo de observación hasta el fallo.

En paralelo, para los elementos reparables se idéntica la tasa de ocurrencia de fallos (ROCOF, por sus siglas en inglés), mientras que para los no reparables, se identifica la tasa de riesgo o tasa de fallo (λ). El ROCOF es la tasa absoluta con la que ocurren los fallos en un sistema, mientras que la función de tasa de fallos es la tasa relativa de fallos de un componente que sobrevive hasta un determinado tiempo.

Datos censurados

Un punto importante a considerar en el análisis de supervivencia es la presencia de datos censurados. Estos pueden corresponder a elementos removidos o cambiados antes de que se produzca su fallo o cuando la prueba (periodo de observación) se completa antes que todos los elementos fallen. A continuación, se presentan algunas definiciones.

Censura por la derecha. Un elemento se define censurado por derecha, cuando se desconoce su tiempo de fallo, pero se sabe que ocurre en un tiempo mayor al de duración de una prueba.

Censura por la izquierda. Un elemento se define censurado por la izquierda, cuando se desconoce el tiempo exacto de fallo, pero se sabe que ha ocurrido antes del comienzo de una prueba.

Censura doble. Cuando se conoce que un elemento ha tenido fallo antes de una prueba y posterior a la misma, pero sin embargo no registra fallos durante el periodo de observación.

Censura Tipo I. Cuando una prueba finaliza después de un periodo de tiempo, donde el tiempo de duración de la prueba es prefijado para todas las muestras.

Censura Tipo II. Cuando una prueba finaliza después de registrarse un determinado número de eventos de fallo prefijado. A diferencia de la censura tipo I, el tiempo en esta prueba es desconocido. Se denominan como ensayos censurados por número de fallos.

Censura múltiple. Cuando los tiempos de censura durante una prueba u operación no es el mismo para todos los elementos debido a su remoción o incorporación en diferentes tiempos.

Truncamiento por la izquierda. Es el escenario en que la edad de los sujetos de análisis precede al periodo de observación de estudio, de manera que los elementos que se incorporan durante un tiempo intermedio del estudio no son tomados en consideración. De igual manera, aquellos sujetos en los que se conoce que el fallo tiene lugar antes del inicio del estudio son ignorados en el análisis.

Pruebas de bondad de ajuste.

Las pruebas estadísticas de bondad de ajuste permiten a través de una hipótesis nula (H_0) y una hipótesis alternativa (H_1), determinar el ajuste de datos entre muestras o grupos de datos. En este trabajo se emplearán para determinar si una distribución teórica asumida, describe correctamente o no un conjunto de datos observados.

Para ello se definen las hipótesis, donde H_0 indica que los datos de tiempos de fallo provienen de una distribución teórica (ejemplo, una distribución Weibull). Mientras que H_1 indica que los datos de tiempos de fallo “NO” provienen de la distribución teórica.

Estas pruebas consisten en la determinación de un valor crítico, según el cual se estipula si la hipótesis nula (H_0) es aceptada o, caso contrario, se rechaza.

El valor crítico puede ser obtenido a través de tablas o calculado dependiendo de la prueba de bondad de ajuste aplicada. Este depende del tamaño de la muestra y el valor del nivel de significancia (α) de la prueba, el cual determina la probabilidad de error de rechazar la hipótesis nula cuando la misma es verdadera. Dependiendo del estudio, el

nivel de significancia puede tomar valores de 0.1; 0.05 y 0.01; siendo el más habitual 0.05.

Las pruebas de hipótesis generalmente están enfocadas a rechazar o desaprobar la hipótesis nula, tomando la hipótesis alternativa como la opción deseable. De esta manera mientras menor sea el nivel de significancia (α), menor será la probabilidad de error de rechazar la hipótesis nula cuando esta es verdadera. En cambio, para el caso de este trabajo lo deseable es verificar que se cumpla la hipótesis nula. Por lo tanto, un valor mayor nivel de significancia será más restrictivo, que uno de menor valor.

Existen una gran cantidad de pruebas que pueden ser aplicadas. A continuación, se describen las pruebas de Kolmogorov-Smirnov y Chi-cuadrado, las cuales han sido empleadas en este trabajo.

Prueba de Kolmogorov-Smirnov.

Esta prueba se basa en la función de distribución acumulada empírica y permite determinar si los datos de una muestra provienen de una distribución específica, siendo esta su hipótesis nula (H_0). Para ello se define el máximo valor de la diferencia absoluta (D_{MAX}) entre dos funciones de distribución acumulada (distribución teórica y distribución empírica), el cual es comparado con un valor ($D_{Crítico}$).

La prueba es aceptada si ($D_{MAX.} < (D_{Crítico})$), para cierto nivel de significancia determinado. En caso contrario, se rechaza la hipótesis nula.

El estadístico de prueba de esta prueba se define por:

$$D = \max_{1 < j < N} |F(t_j) - E_j| \quad (2.24)$$

Donde $F(t_j)$ es la función de distribución acumulada asumida, E_j es la función de distribución acumulada empírica y N el número total de datos.

Los valores para ($D_{crítico}$), de acuerdo al nivel de significancia se obtienen mediante las siguientes expresiones:

Tabla 2. Valores críticos del estadístico D para la prueba de K-S

Tamaño de la muestra	Nivel de significancia α		
	0.1	0.05	0.01
n	$1.22/\sqrt{n}$	$1.36/\sqrt{n}$	$1.63/\sqrt{n}$

Prueba Chi-cuadrado

Al igual que la prueba anterior, se usa para determinar si los datos de una muestra provienen de una distribución específica. Es aplicable a muestras continuas o discretas. Solo es válida para muestras de tamaño grande y los datos deben agruparse en clases. La hipótesis nula y alternativa son las mismas que la prueba de K-S y su estadístico de prueba se define de la siguiente forma:

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} \quad (2.25)$$

Donde k es el número de clase, O_i son los fallos observados para la clase j , E_i son los fallos esperados para la clase j .

Se debe establecer un mínimo de 5 elementos para cada clase j . Esa prueba tiene una distribución chi-cuadrado con $(k - c)$ grados de libertad, donde c es el número estimado de parámetros. La hipótesis nula es rechazada cuando $\chi^2 > \chi_{(\alpha, k-c)}^2$.

2.4 Descripción de Síntesis Universal de Medición CMD

La Síntesis Universal de Medición CMD (Confiability, Maintainability and Disponibility), es una metodología sintética propuesta por Mora (2009), cuya aplicación está dirigida a la industria en general. En su desarrollo, este método emplea conceptos de análisis de fiabilidad, descritos previamente en este trabajo, así como también otros más específicos de esta rama de estudio.

Este modelo de CMD se divide en 6 etapas. A continuación, se resume cada una de ellas (el esquema completo de esta metodología puede ser visto en el Anexo 1).

- *Etapa 1. Obtención y preparación de datos.* Consiste en la recopilación, manipulación y tratamiento de datos, con la finalidad de estructurar una base de datos sólida.
- *Etapa 2. Disponibilidad.* Se relaciona con diferentes niveles de detalle que puede proporcionar el estudio en función de los datos disponibles y los objetivos que se persiguen alcanzar, desde el caso más simple al más complejo y tomando en consideración que a una mayor complejidad, mayor cantidad, calidad y costo de obtención de datos es requerida.
- *Etapa 3. Estimación de parámetros.* Se debe seleccionar el método de estimación de los parámetros de una distribución Weibull, inicialmente asumida.
- *Etapa 4. Validación de ajuste de Weibull o selección de otra función.* Se verifica el ajuste de la función Weibull mediante pruebas de bondad de ajuste, recomendándose el cumplimiento de al menos 2 de 3 pruebas propuestas en el modelo. En caso de cumplirse, se continúa con la etapa 6.
- *Etapa 5. Parametrización de otra función.* Se selecciona un método de estimación para determinar los parámetros de una nueva función de distribución (Gamma, Lognormal o Normal), la cual se selecciona dependiendo del valor del parámetro de forma (β) obtenido en Weibull y luego se verifica la bondad de ajuste de la nueva distribución contemplada. En caso de no cumplirse esta, se debe volver a la etapa 1 de obtención y preparación de datos, y reagrupar las muestras o esperar a obtener un mayor registro de datos.
- *Etapa 6. Cálculos, predicción y estrategias CMD.* Se desarrollan los cálculos de parámetros y coeficientes relacionados con CMD, los cuales posteriormente son analizados junto a curvas de distribución de fallos que pueden ser obtenidas. A partir de la evaluación de todos ellos, se realiza un pronóstico del sistema para luego plantear estrategias, acciones y tácticas de mantenimiento.

Esta metodología propuesta es citada ya que inicialmente ha sido considerada para desarrollo del modelo propuesto en el presente trabajo.

2.5 Modelos de análisis de redes de agua

En este trabajo se busca obtener un modelo que permita estimar el tiempo de vida útil de las tuberías de una red de agua y predecir su número de fallos en el tiempo. Para ello es importante ante todo tener conocimiento del ciclo de vida de las tuberías (St. Clair y Sinha, 2014), lo cual incluye todas sus fases: diseño, construcción, transporte, instalación, servicio, mantenimiento y su reemplazo. Dentro de este proceso existen variables o factores internos, externos o propios del servicio que prestan, lo cual se ve reflejado en ocurrencia de fallos y en consecuencia de la vida útil de una red.

Múltiples modelos y aproximaciones han sido propuestos para determinar la probabilidad de fallos o estimar la vida útil de las tuberías de redes de agua. A continuación, se trata brevemente acerca de los modelos según dos enfoques: estimación de la vida útil y predicción de fallos.

2.5.1 Modelos de estimación de vida útil

Para la estimación de la vida útil generalmente se asume que los fallos que se presentan en las tuberías durante su vida útil pueden ser representados mediante la curva de la bañera (curva de Davies, mostrada previamente en sección 2.1). Esta curva puede ser de dos tipos, una que describe la probabilidad de fallo instantánea (tasa de fallo) para componentes de tipo no reparables, y otro que describe la tasa de ocurrencia de fallos, (ROCOF por sus siglas en inglés), para componentes de tipo reparables (Ascher y Feingold, 1984).

Por su parte, Kleiner y Rajani (2001) anotan varios aspectos a tomar en consideración.

No todas las tuberías cumplen estrictamente las fases que desarrolla esta curva, pudiendo variar drásticamente para varias tuberías bajo distintas condiciones.

Una tubería usualmente se considera como un elemento reparable, así la curva de bañera de ROCOF debería ser asociada con el ciclo de vida. A pesar de ello algunos modelos utilizan la curva de bañera de la tasa de fallo en conjunción con

el fallo de la tubería... Cuando se modela la duración del tiempo de interrupciones, la función de tasa de fallos es usada a menudo por que la “unidad” modelada se convierte, de hecho, en una unidad que solo permite un fallo.

Por lo tanto, es importante, como se comentó anteriormente, definir si los componentes de la red se tratarán como elementos renovables o no renovables y definir el tipo de curva de bañera a utilizar, ya que los resultados dependerán de ello.

La estimación de la vida útil para estimar la fecha óptima de renovación, consiste en identificar el tiempo en que las tuberías en conjunto dejan de cumplir ciertos requisitos para los cuales fueron diseñadas y resulta más conveniente reemplazarlas. Esto puede relacionarse a una tasa alta de roturas, incremento en costes de mantenimiento, pérdida de capacidad hidráulica, afectaciones a la calidad del agua, alto nivel de fugas, etc. Esto implica considerar este tipo de modelos con otros que además incluyan el análisis económico de mantenimiento, renovación y otros tipos de costes relacionados al impacto ocasionados por el estado de las redes de agua.

Sin embargo, los modelos de vida útil no dejan de suponer una herramienta de ayuda para gestores de abastecimientos, relacionado a la planificación del mantenimiento preventivo, planes de renovación de redes y manejo de inversiones; requiriendo contar con una base sólida de datos que permita hacer frente a incertidumbres propias del modelo para así obtener resultados lo suficientemente precisos para basar la toma de decisiones (Martínez, 2017).

2.5.2 Modelos de predicción de fallos

A nivel global estos modelos se pueden dividir en modelos de tipo físicos, estadísticos y heurísticos.

Los físicos son modelos que pueden proporcionar buenos detalles de los fallos que ocurren en una red a través de modelos matemáticos (determinísticos o probabilísticos), pero que a su vez demandan de una base de datos muy detallada, lo cual para casos

prácticos es muy difícil de conseguir, por cuanto su aplicación se limita a objetivos muy específicos.

Los modelos estadísticos son ampliamente usados en la práctica, estos pueden proporcionar información intuitiva de gran utilidad sin requerir una excesiva cantidad de datos y recursos. A partir de datos históricos, permiten identificar patrones y predecir fallos. Estos modelos se pueden clasificar en general en determinísticos (modelo de regresión lineal o exponencial) y estocásticos.

Los modelos determinísticos se conforman de funciones matemáticas relacionadas a variables explicativas, requiriendo de las mismas una larga serie de datos. En ellos se relaciona la dependencia entre una variable dependiente y una o varias independientes. Como referencia se tiene el modelo de Shamir y Howard (1979) que realiza el análisis de una regresión de función exponencial en función de las variables de tiempo y la edad de las tuberías para la predicción del número de fallos por unidad de longitud y año, aplicando la siguiente expresión:

$$N(t) = N(t_0) \cdot e^{A(t+g)} \quad (2.26)$$

donde t es el tiempo observado en años, $N(t)$ el número de fallos en unidades de longitud por año, $N(t_0)$ es el número de roturas por unidad de longitud para t_0 (siendo $N(t_0)$ su valor siempre distinto de cero) lo que implica que la tubería tiene siempre una frecuencia de rotura que puede incrementar en el tiempo, t_0 es el año base de análisis y A un parámetro de regresión o coeficiente de crecimiento. Los autores de este modelo recomiendan la aplicación del mismo para grupo de tuberías de características homogéneas, sin especificar la precisión del mismo.

Por citar algunos modelos posteriores a los de Shamir y Howard se tienen los propuestos por Walski y Pelliccia (1982), Clark et al. (1982) y Karney et al. (1994); los cuales incorporan más variables como diámetro, presión, edad, reparaciones registradas, etc.

En un estudio predecesor a este trabajo (Mogollón, 2017) plantea una metodología haciendo uso del modelo de Shamir y Howard. En ella se evalúan factores como el caudal

máximo transportado y el número de fallos de una tubería, para en base a ello facilitar la toma de decisiones y establecer planes de renovación de activos.

En la Tabla 3, se muestra un resumen de modelos de regresión lineal y exponencial para fallos en tuberías de agua realizado por Kabir (2016).

Tabla 3. Modelos de regresión lineal y exponencial para fallos en tuberías de agua. Adaptado de Kabir, 2016

Referencia	Variables	Modelo (*)
Shamir y Howard 1979	Edad	MRE
Walski y Pelliccia 1982	Edad, diámetro y número previo de fallos	MRE
Clark et al. 1982	Edad, número previo de fallos	MRLM, MREM
Kettler y Goulter 1985	Edad, diámetro	MRL
Goulter y Kazemi 1988	Tiempo, sitio de fallo previo	MRE
Jacobs y Karney 1994	Longitud, edad	MRL
Kleiner y Rajani 1999	Edad	MRE
Rajani y Kleiner 2001	Edad, índice de congelación, déficit de precipitación acumulado, déficit instantáneo de lluvia, longitud acumulada de redes de agua reemplazadas, longitud acumulada de reequipamiento de protección catódica	MREM
Kleiner y Rajani 2000, 2002	Índice de congelamiento, déficit de precipitación, longitud acumulada de redes reemplazadas, longitud acumulada de protección catódica	MREM
Boxall et al. 2007	Diámetro, edad, longitud, material, corrosividad del suelo	MRP
Asnaashari et al. 2009	Edad, longitud, diámetro, espesor de tubería, presión máxima, localización de tubería, profundidad, historial de fallos	MRLM, MRP
Wang et al. 2009	Edad, tamaño, longitud	MRLM
Wood y Lence 2009	Edad, material, diámetro, tipo de suelo	MREM
Yamijala et al 2009	Diámetro, material, longitud, año de instalación, tiempo desde el último fallo, presión, uso de suelo, tipo de suelo, temperatura, precipitación, máxima y mínima humedad del suelo, corrosividad del suelo	MRLM, MREM, MLG, MLG logístico
Bubtiena et al. 2011	Edad, material, longitud, diámetro, profundidad, tipo de suelo, calidad del agua	MLG
Christodoulou 2011	Material, diámetro, tipo de tubería, carga de tráfico en la vecindad, número previo de fallos, tipo de incidente	MRP

(*) MRE: Modelo de Regresión Exponencial
MRLM: Modelo de Regresión Lineal Múltiple
MREM: Modelo de Regresión Exponencial Multivariado
MRL: Modelo de Regresión Lineal
MRP: Modelo de Regresión de Poisson
MLG: Modelo Lineal Generalizado

Los modelos estocásticos son de naturaleza aleatoria. Y pueden ser probabilísticos multivariable o de variable única. Los modelos multivariable, mediante el ajuste de una

distribución de fallos, permiten obtener los parámetros de la función y determinar la influencia de las variables explicativas. Los de variable única requieren de una agrupación de datos de características homogéneas. Dentro de los modelos estadísticos existen los modelos de análisis de supervivencia, los cuales permiten determinar mediante distribuciones el proceso de fallo en base a la fiabilidad. Por citar algunos procesos de análisis relacionados a estos modelos, se menciona: el proceso de regresión de Poisson (Asnaashari et al., 2009), proceso de Poisson no homogéneo (Kleiner y Rajani, 2010), modelo de Weibull – exponencial (Dridi et al., 2009; Mailhot et al., 2003), modelos multivariados exponenciales (Kleiner y Rajani, 2002), regresión de Cox (Andreou, 1986; Park et al., 2011), riesgos proporcionales de Weibull (Le Gat y Eisenbeis, 2000), etc. Dentro de los modelos estocásticos existen también los modelos de análisis o inferencia Bayesiana, (Watson et al., 2004; Martínez, 2017). La Figura 6 muestra la clasificación de modelos de supervivencia, mientras que en la Tabla 4 se presenta un resumen de estos.

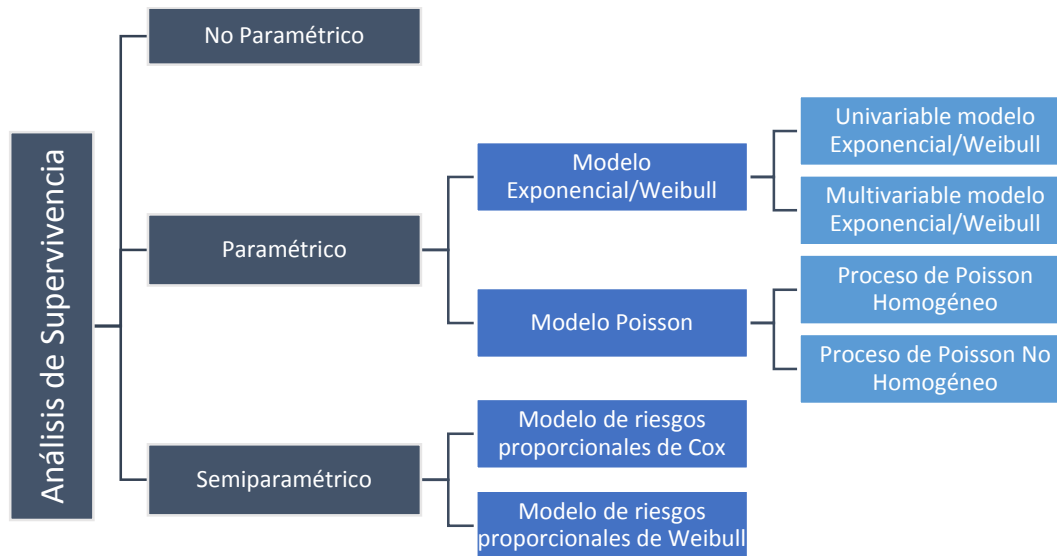


Figura 6. Clasificación de modelos de análisis de supervivencia para el análisis de fallos de tuberías (Kabir, 2016)

En la Figura 6 se puede seguir la procedencia de los modelos a estudiar en este trabajo. El modelo propuesto SERAP (sección 3.1) corresponde así, a un análisis de supervivencia

aplicando un modelo de distribución paramétrica, mientras que el modelo de riesgos proporcionales de Weibull WPHM, por sus siglas en inglés (sección 3.2), corresponde a un modelo de tipo semiparamétrico.

Tabla 4. Modelos estocásticos. Adaptado de Kabir,2016

Referencia	Covariables	Grupo	Método (*)
Kimutai et al. 2015	Longitud, diámetro, material, resistividad del suelo, índice de congelamiento, déficit de precipitación	material, número de fallos previos	Cox-PHM, W-PHM, Modelo Poisson
Economou et al 2012	Edad, numero de fallos previos	material	NHPP inflado de ceros
Fuchs-Hanusch et al. 2012	Diámetro, longitud, antigüedad	material	Cox-PHM
Rajani et al. 2012	Temperatura del agua y del aire	material	NHPP inflado de ceros
Toumbou et al. 2012	Número de fallos en tubería, diámetro, longitud y fecha de instalación	Material, número de fallos previos	MEWM
Christodoulou 2011	Material, diámetro, tipo de tubería, carga de tráfico en la vecindad, numero previo de fallos, tipo de incidente	Material, número de fallos previos	RNA, estimador de Kaplan-Meier, MRP
Clark y Thurnau 2011	Material, diámetro, factor de fragilidad	Material, diámetro	Cox-PHM, modelo de fragilidad compartida
Osman y Bainbridge 2011	Edad, longitud, índice de congelamiento	material	MEM
Park et al. 2011	Longitud, presión interna en tubería, uso de suelo, número de clientes de la red	Material, número de fallos previos	Cox-PHM
Renaud et al. 2011	Material, longitud, año de instalación, historial de fallos	Material, número de fallos previos	Extensión proceso lineal de Yule
Rogers 2011	Edad, material, diámetro, tipo de fuga, número previo de fallos previos	Material, número de fallos previos	NHPP
Alvisia y Franchinia 2010	Diámetro, material, año de instalación, número de fallos	Año de instalación, número de fallos previos	MEWM, W-PHM
Clark et al. 2010	Material, diámetro, factor de fragilidad	Material, diámetro	Cox-PHM
Debón et al. 20110	Edad, longitud, diámetro, presión de tráfico bajo una acera, material	Sin especificar	Cox-PHM, W-PHM, MRM
Fadace y Tabatabaci 2010	Tipo de fallo, número acumulado de fallos	Tipo de fallo	Modelo de ley de potencia
Kleiner y Rajani 2010	Edad, diámetro, material, longitud, conexión de servicio, antigüedad, tipo de suelo, índice de congelamiento, déficit de precipitación acumulada, déficit de precipitación instantánea, protección catódica, reequipamiento de protección catódica	Diámetro	NHPP
Asnaashari et al. 2009	Edad, longitud, diámetro, espesor de tubería, presión máxima, localización de tubería, profundidad de cobertura	material	MRM, MRP

Referencia	Covariables	Grupo	Método (*)
Asnaashari et al. 2009	Edad, longitud, diámetro, espesor de tubería, presión máxima, localización de tubería, profundidad de cobertura	material	MRM, MRP
Dridi et al. 2009	edad	Número de fallos previos	MEWM Bayesiano
Rogers y Grigg 2009	Edad, material, diámetro, número previo de fallos	Material, número de fallos previos	NHPP, ADMC
Gorji-Bandpy y Shateri 2008	Material, diámetro, longitud, año de instalación, tipo de suelo, uso de suelo	número de fallos previos	MEWM
Park et al. 2008b	Material, longitud, diámetro, tiempo de instalación, presión interna en tubería, uso de suelo	Material, número de fallos previos	Cox-PHM
Park et al. 2008a	Edad, tiempo de fallos previos	material, diámetro	Modelo de tasa de ocurrencia de fallos
Boxall et al. 2007	Diámetro, edad, longitud, material, corrosividad del suelo	material	MRP
Park et al. 2007	Material, diámetro	Tiempo de supervivencia	Cox-PHM
Economou et al. 2007	Edad, longitud, diámetro, presión, presión absoluta	Sin especificar	NHPP Bayesiano
Dridi et al. 2005	edad	Número de fallos previos	NEWM Bayesiano
Varenterghem-Raven 2004	Edad, longitud, diámetro, material, tráfico, suelo, paso subterráneo, localización en la calle, presencia de zonas hidráulicas antiguas	Material, número de fallos previos	Cox-PHM, W-PHM
Watson et al. 2004	edad	Sin mencionar	NHPP Bayesiano
Pelletier et al. 2004	Diámetro, longitud, material, año de instalación, tipo de suelo, uso de suelo	número de fallos previos	MEWM
Kleiner y Rajani 2002	Longitud acumulada de redes reemplazadas, longitud acumulada de protección catódica, índice de congelamiento, déficit de precipitación	material	MEM
Le gat y Eisenbeis 2000	Edad, longitud, diámetro, tipo de instalación, presión, tipo de suelo, nivel de tráfico, tipo de suministro	Material, número de fallos previos	W-PHM
Mailhot et al. 2000	Material, longitud, diámetro, año de instalación, tipo de suelo, uso de suelo sobre la tubería	Año de instalación	MEWM

(*) Cox-PHM: Modelo de riesgos proporcionales de Cox
W-PHM: Modelo de riesgos proporcionales de Weibull
NHPP: Proceso de Poisson No Homogéneo
MEWM: Modelos Exponencial/Weibull Multivariable
RNA: Redes Neuronales Artificiales
MRP: Modelo de Regresión de Poisson
MEM: Modelo Exponencial Multivariable
MRM: Modelo de Regresión Múltiple
ADMC: Análisis de Decisión Multicriterio

Los modelos heurísticos por su parte consisten en la aplicación de métodos de técnicas de aprendizaje aplicando algoritmos genéticos mediante técnica de redes neuronales o métodos de lógica difusa que permiten establecer respuestas a partir de datos parciales o imprecisos. Para ello, en función de la experiencia y evidencia científica del efecto de las variables explicativas sobre las redes, se determinan las causas y efectos de la probabilidad de fallo y sus consecuencias, para lo cual la experiencia y juicio de expertos deben ser considerados para determinar el riesgo de fallos. Por lo tanto, son aplicables a redes con bases de datos limitadas o de baja calidad y fiabilidad. Como referencia de estos procesos de análisis se cita el trabajo de Yan y Vairavamoorthy (2003) en el cual se aplica la combinación de técnicas de decisión multicriterio y teoría de conjuntos difusos; y el trabajo de Jafar et al. (2010) donde se presenta la aplicación de redes neuronales para determinar tasa de fallo y tiempo óptimo de reemplazo de las tuberías de una red de distribución de agua.

3 Metodología

En este capítulo se describe la metodología a seguir para el desarrollo del método propuesto, el cual se identificará en este trabajo como “Síntesis de evaluación de redes de agua potable” o “SERAP”, el cual consta de dos planteamientos de acuerdo a la estructuración de las bases de datos. También se describe el “Modelo de Riesgos Proporcionales de Weibull” WPHM (por sus siglas en inglés), presentado en el trabajo de Le Gat y Eisenbeis (2000), explicando los pasos y consideraciones para su aplicación y exponiendo también el motivo del por qué se ha seleccionado este modelo para la comparación con el modelo propuesto.

3.1 Síntesis de evaluación de redes de agua potable. SERAP

Como se indicó previamente, el método propuesto sigue de cerca la metodología de Síntesis Universal de Medición CMD (Confiability, mantenibilidad y disponibilidad) propuesta por Mora (2009). Dentro de la clasificación de modelos de análisis de redes de agua visto dentro de la sección 2.4, este modelo se puede clasificar como un modelo estadístico de análisis de supervivencia paramétrico univariable.

Para su aplicación es necesario considerar los siguientes puntos e hipótesis,

- a) Para diferenciar los distintos tipos de tuberías estas requieren ser agrupadas en grupos de características homogéneas utilizando algún criterio (diámetro, material, etc.), ya que se asume que los procesos de fallo dependen de las propiedades o características de las tuberías. Hay que tomar en consideración que la agrupación de los datos puede resultar en grupos con un número limitado

o insuficiente de registro de fallos lo cual puede conducir a obtener predicciones no significativas en estos grupos.

- b) Se requiere que el número de observaciones de fallos para un grupo de análisis sea de un tamaño significativo ya que, a mayor cantidad de datos disponibles, la estimación de los parámetros es más cercana al valor verdadero. Para ello se requiere un registro mayor de 30 observaciones de fallos registrados (por teorema de limite central donde se demuestra la bondad estadística de tener tamaños de muestras mayor a 31 elementos (Levin et al., 1996)).
- c) Se asume que los datos de las variables de tiempo analizadas siguen un patrón de fallos que puede ser descritos mediante distribuciones de tipo paramétricas.
- d) Se asume que los datos de las variables son truncados por la izquierda y al final de la observación pueden o no ser censurados por derecha.

A continuación, se presenta el método que consta de diferentes etapas:

- *Etapas 1.* Consiste en la obtención y preparación de datos para el análisis. A partir de esto se estructura una nueva base de datos donde cada una de las tuberías que contiene la red se encuentran identificadas y acompañadas de diferentes variables explicativas. Adicionalmente se seleccionan los grupos o subgrupos de análisis que consisten en tuberías de características homogéneas agrupadas bajo cierto criterio de análisis. Para ello, resulta útil realizar un análisis estadístico descriptivo previo de la red, el cual permite identificar características de interés general.
- *Etapas 2.* Comprende el análisis estadístico de las variables de tiempo a analizar para los grupos o subgrupos seleccionados. Se calculan los parámetros de las distribuciones paramétricas por métodos de estimación de mínimos cuadrados o estimación por máxima verosimilitud y se verifica el ajuste, mediante pruebas de bondad de ajuste con un intervalo de confianza del 95%, de la distribución paramétrica contra una empírica, no paramétrica, obtenida a partir del estimador de K-M. En caso de no ajustarse, se deberán reconfigurar los grupos o subgrupos de análisis y reiniciar el proceso, o esperar a recolectar una mayor

cantidad de datos que se ajusten a una función paramétrica. La función predilecta a ser empleada es la distribución Weibull, debido a su versatilidad para adaptarse a distintas funciones de distribución, donde además su función de riesgo permite representar las distintas fases de la curva de bañera, como se ha comentado previamente. Otras distribuciones como la lognormal, gamma y normal, pueden ser consideradas en caso de que la distribución Weibull no se ajuste apropiadamente. Para la estimación de los parámetros de las funciones paramétricas se ha utilizado el software de análisis estadístico *R*.

- *Etapa 3.* Se evalúan los resultados obtenidos de los grupos de tuberías analizados. A partir de los resultados de las variables de tiempo analizadas, se estima la vida útil, se valora el estado de las tuberías mediante los parámetros o medidas de tendencia central de las distribuciones y se predice el número de fallos esperados. Estos resultados deben ser analizados finalmente con el suficiente criterio técnico para, a partir de ellos, proponer estrategias o acciones de mantenimiento más idóneas a aplicar.

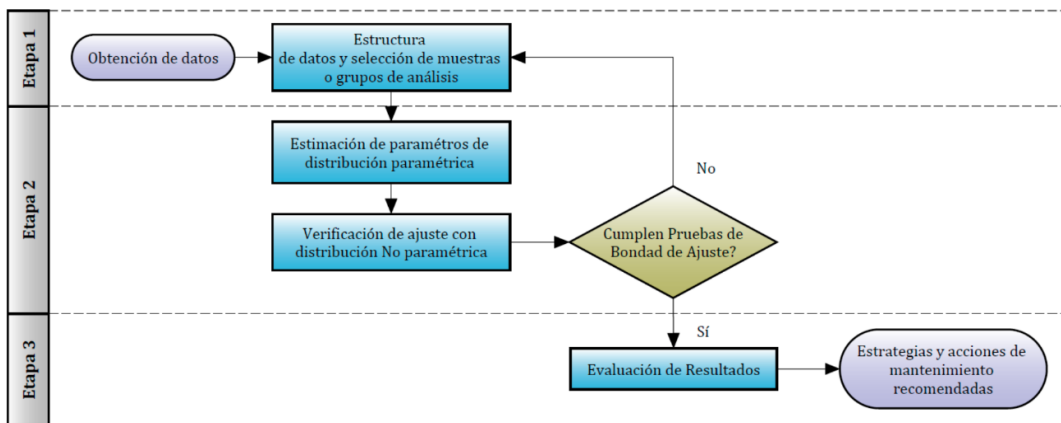


Figura 7. Esquema de la metodología del modelo propuesto y sus etapas

3.1.1 Variables de análisis y estructura de datos

Como se mencionó en la sección 2.3, es necesario conocer los datos con que se cuenta, así como los que pueden ser deducidos a partir de los datos existentes, en función de los objetivos a alcanzar, lo cual además es importante para la estructuración de datos.

Variables de análisis

En el esquema de la Figura 8, se representan las observaciones realizadas sobre distintas tuberías y sobre distintos casos posibles, donde a partir de ello se definen las variables de análisis de edad hasta el fallo (E), tiempo hasta el fallo (TTF , por sus siglas en inglés) y tiempo entre fallos (TBF , por sus siglas en inglés). Donde $Tub1$, $Tub2$ y $Tub3$ es la identificación de diferentes tuberías.

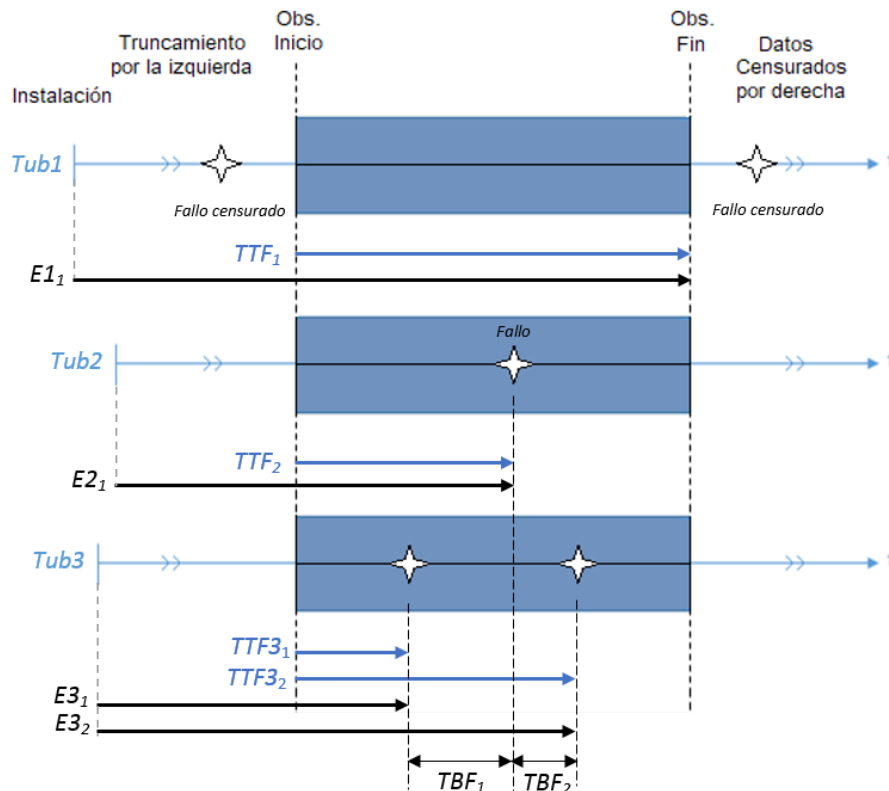


Figura 8. Definición de variables de análisis E , TTF y TBF .

A partir de la figura anterior se puede describir lo siguiente:

- Tub1, representa el caso de una tubería que no registra fallos durante el tiempo de observación, donde los fallos previos al tiempo de observación son desconocidos e ignorados. Para este caso las variables de análisis E y TTF se definen respecto a la fecha del fin de la observación e identificando al elemento como censurado por derecha. Para ello se definirá la censura por derecha de cada elemento con un valor escalar (0,1), siendo 0 para tuberías que no presentan fallos durante el periodo de observación y 1 cuando estas los presentan.
- Tub2, es el caso cuando la tubería presenta un primer fallo durante el periodo de observación. En este caso las variables E y TTF se definen respecto al registro de esta observación. La censura de este elemento ya se definirá con un valor igual a 1.
- Tub3, muestra el caso de un elemento que presenta múltiples fallos durante el periodo de observación, donde las variables E y TTF se definen de manera individual para cada evento de fallo, como es el caso de $E3_1$, $E3_2$, $TTF3_1$ y $TTF3_2$ en la Figura 8.
- En referencia a la variable TBF , esta corresponde al tiempo entre fallos registrado para todo el grupo de tuberías analizado, tomando a todo el sistema como un solo elemento, como es el caso representado en la Figura 8 (TBF_1, TBF_2).

Un punto de inicio importante es definir si los datos de las variables de análisis parten de elementos o componentes de tipo reparable o no reparable.

De esta manera, las variables de edad E y TTF , debido a cómo se encuentran definidas, pueden considerarse como datos provenientes de elementos no reparables, dado que registran cada evento de manera individual, donde el tiempo de vida de un elemento se relaciona con el tiempo medio hasta el fallo MTTF (por sus siglas en inglés), y donde la tasa de fallos o función de riesgo (λ) es la probabilidad condicional de que en un instante de tiempo el elemento (tubería) falle. Por ello, también se la conoce como probabilidad de fallo instantánea.

Por otra parte, la variable *TBF* puede considerarse como proveniente de elementos reparables, a partir del cual puede estimarse el tiempo medio entre fallos MTBF (por sus siglas en inglés). La función de intensidad, también conocida como tasa de ocurrencia de fallos ROCOF (por sus siglas en inglés), mide la probabilidad incondicional que un sistema reparable falle en un instante de tiempo.

Estructura de datos

Se conforma a partir de dos bases de datos. Una que comprende las características del total de tuberías existentes y otra con el registro de fallos (reparaciones). Ambas bases de datos deben contar con una característica común y única que permita relacionarlas, para lo cual se utiliza un código para cada tubería.

En la Tabla 5 se ilustra cómo podrían ser representadas las bases de datos del historial de reparaciones y de los tramos de tuberías (siendo “Id.Tub” el nexo entre ambas). A partir de ello, se conforma una nueva base de datos en la cual se deducirán nuevas variables, como por ejemplo las variables de tiempo *E*, *TTF*, *TBF* o la agrupación de diámetros por clases, que será útil para el análisis.

Tabla 5. Ilustración de tabla base de datos de historial de reparaciones (Izq), y de elementos de red y sus características (Der)

Id.Tub.	Fecha Fallo	Id.Tub.	Fecha de Inst.	Material	Diámetro	longitud	X
...
i-1	t(i-1, n _{i-1})	i-1	t _{inst} (i-1)	M(i-1)	D(i-1)	L(i-1)	X(i-1)
i	t(i, 1)	i	t _{inst} (i)	M(i)	D(i)	L(i)	X(i)
...	...	i+1	t _{inst} (i+1)	M(i+1)	D(i+1)	L(i+1)	X(i+1)
i	t(i, n _i)
i+1	t(i+1, 1)						
...	...						

Donde *i* es un código o número identificador de cada tubería, la letra *t* representa a una variable de tiempo, *n* el fallo correspondiente al *i*-ésimo elemento, y *X* puede ser otra covariable de interés agregada o deducida a partir de datos existentes.

3.1.1.1 Estructura de datos. Planteamiento 1

Para el modelo propuesto SERAP, se han definido dos formas de estructurar los datos para las variables de análisis edad (E) y tiempo hasta el fallo (TTF). El análisis para la variable tiempo entre fallos (TBF) se aborda más adelante. En el primer planteamiento o escenario se considera que cada elemento o tramo de tubería puede presentar solo un fallo. Esto se representa en el siguiente esquema,

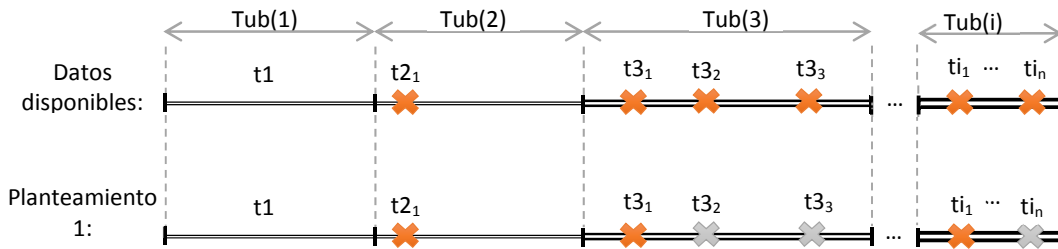


Figura 9. Esquema de datos para planteamiento 1

Donde i representa a cada una de las tuberías y n es el número de la secuencia de fallos que pueden presentar en cada una. En el esquema se dibujan las tuberías de diferentes grosores haciendo alusión a que pueden tratarse de tuberías de diferentes características. Así mismo, los primeros fallos registrados en cada tramo y que se toman en consideración para el análisis se pintan en color naranja, mientras los que no entran en el análisis se pintan en color gris.

De esta manera, la nueva estructura de datos toma como nexo la variable de identificación de cada elemento acompañado del resto de variables, a partir de las cuales es posible estratificar muestras o definir subgrupos. En la Tabla 6 se muestra cómo podrían quedar representados los datos en la nueva base de datos para este planteamiento.

Tabla 6. Ilustración de referencia de base de datos estructurada para planteamiento 1

Id.	Fecha 1º Fallo	Fecha Inst.	Obs. inicio	Obs. fin	C	E	TTF	Mat.	Clase Diám.	Long	X
...
$i-1$	$t(i-1)$	$t_{inst}(i-1)$	$t_a(i-1)$	$t_b(i-1)$	0	$t_b(i-1)-t_{inst}(i-1)$	$t_b(i-1)-t_a(i-1)$	$M_j(i-1)$	$D_j(i-1)$	$L(i-1)$	$X(i-1)$
i	$t(i)$	$t_{inst}(i)$	$t_a(i)$	$t_b(i)$	1	$t(i)-t_{inst}(i)$	$t(i)-t_a(i)$	$M_j(i)$	$D_j(i)$	$L(i)$	$X(i)$
$i+1$	$t(i+1)$	$t_{inst}(i+1)$	$t_a(i+1)$	$t_b(i+1)$	1	$t(i+1)-t_{inst}(i-1)$	$t(i-1)-t_a(i-1)$	$M_j(i+1)$	$D_j(i+1)$	$L(i+1)$	$X(i+1)$
...

El inicio y fin del periodo de observación (Obs. Inicio y Obs. fin) es el mismo para todas las tuberías (t_a, t_b). Los datos que no registran fallo se estiman según t_a y t_b , y se especifica el valor de censura (C) igual a cero. Las variables E y TTF son deducidas a partir de los datos existentes. X representa a otras variables que pueden ser agregadas. El subíndice j presente en algunas covariables denota alguna agrupación o estratificación por clases.

3.1.1.2 Estructura de datos. Planteamiento 2

El segundo planteamiento consiste en segmentar los tramos de tuberías en tantos segmentos como unidades de longitud posea cada tubería, donde cada uno de estos tramos segmentados, como en el caso anterior, permite o puede presentar solo un fallo. De esta manera los fallos de cada tramo de tubería se asignan de uno en uno a tantos segmentos como fallos se presenten en el tramo, siendo los segmentos restantes de tipo censurado por derecha. Cada segmento generado hereda el identificador del tramo de tubería original.

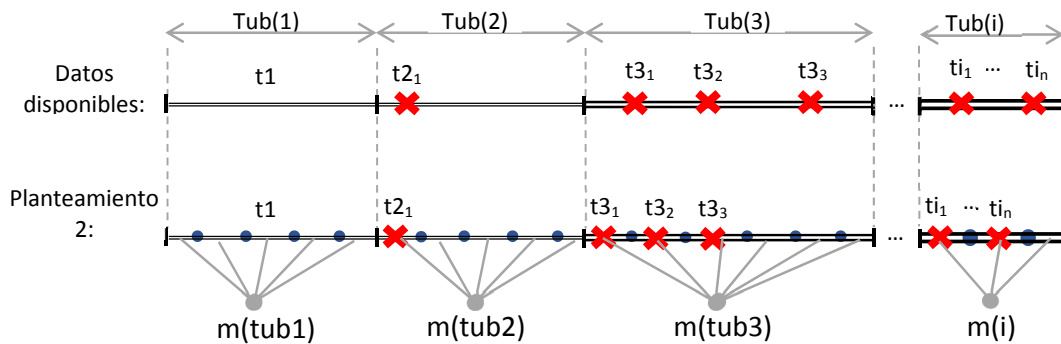


Figura 10. Esquema de datos para planteamiento 2

donde m corresponde al número de segmentos en que se divide cada tramo de tubería.

Para la estructura de datos de este segundo planteamiento, la variable nexa principal para el ordenamiento de los datos sigue siendo la identificación de cada elemento, los cuales ya no serán únicos y se verán incrementados tantas veces como grande sea la longitud de cada tubería del grupo de análisis.

Los tramos de tubería con decimales en su longitud son redondeados al valor entero más cercano. Aquellos que registran valores menores a uno son redondeados a la unidad de longitud mínima (1 metro).

La estructuración de estos datos es un poco más compleja en referencia al planteamiento 1, pero permite aprovechar toda la información del historial de fallos disponible. Para la formación de la nueva base de datos con este planteamiento es conveniente realizar a través de una secuencia de programación, que permita optimizar el proceso y evitar errores. En el presente trabajo esto se lo ha realizado mediante VBA para MS Excel.

A continuación, se muestra la estructura de la nueva base de datos para este segundo planteamiento.

Tabla 7. Ilustración de referencia de base de datos estructurada para planteamiento 2

Id.	Fecha Fallo	Ce	Edad	TF	Mat.	Clase Diám.	Long.	X
...
$i-1(m-1)$	$t(i-1,m-1)$	0	$t_b(i-1,m-1)-t_{inst}(i-1)$	$t_b(i-1,m-1)-t_a(i-1)$	$M_j(i-1)$	$D_j(i-1)$	$L_{i-1(m-1)}(1)$	$X(i-1)$
$i-1(m)$	$t(i-1,m)$	1	$t(i-1,m)-t_{inst}(i-1)$	$t(i-1,m)-t_a(i-1)$	$M_j(i-1)$	$D_j(i-1)$	$L_{i-1(m)}(1)$	$X(i-1)$
$i(1)$	$t(i, 1)$	1	$t(i,1)-t_{inst}(i)$	$t(i,1)-t_a(i)$	$M_j(i)$	$D_j(i)$	$L_{i(1)}(1)$	$X(i)$
$i(2)$	$t(i, 2)$	1	$t(i,2)-t_{inst}(i)$	$t(i,2)-t_a(i)$	$M_j(i)$	$D_j(i)$	$L_{i(2)}(1)$	$X(i)$
...
$i(m-1)$	$t(i,m-1)$	0	$t_b(i,m-1)-t_{inst}(i)$	$t_b(i,m-1)-t_a(i)$	$M_j(i)$	$D_j(i)$	$L_{i(m-1)}(1)$	$X(i)$
$i(m)$	$t(i,m)$	0	$t_b(i,m)-t_{inst}(i)$	$t_b(i,m)-t_a(i)$	$M_j(i)$	$D_j(i)$	$L_{i(m)}(1)$	$X(i)$
$i+1$	$t(i+1,1)$	1	$t(i+1,1)-t_{inst}(i+1)$	$t(i+1,1)-t_a(i+1)$	$M_j(i+1)$	$D_j(i+1)$	$L_{i+1(1)}(1)$	$X(i+1)$
$i+1$	$t(i+1,2)$	0	$t_b(i+1,2)-t_{inst}(i+1)$	$t_b(i+1,2)-t_a(i+1)$	$M_j(i+1)$	$D_j(i+1)$	$L_{i+1(2)}(1)$	$X(i+1)$
...

Donde las abreviaturas tienen la misma definición antes vistas. En este caso la nueva abreviatura introducida es la correspondiente a (m), la cual representa el número de segmentos en que se divide cada tramo de tubería, el cual depende de la longitud original del tramo de tubería, como ya se mencionó previamente.

3.1.1.3 Análisis de variable de tiempo entre fallos TBF

El análisis de la variable de tiempo entre fallos *TBF* es aplicable tanto al planteamiento 1 y como al planteamiento 2, y a diferencia de las variables *E* y *TTF*, no se toma en consideración los datos censurados.

Se comienza definiendo la variable TBF , el cual corresponde a la diferencia entre cada uno de los tiempos entre fallos, donde el número total de registro de datos para el análisis de esta variable será igual a $(n-1)$, siendo n el número de fallos registrados durante el tiempo de observación (número total de primeros fallos para el caso del planteamiento 1).

A diferencia del planteamiento 1 y 2, donde cada tubería tiene un valor definido para las variables E y TTF , el TBF no se relaciona directamente con un determinado elemento, sino con datos temporales de tiempos entre fallos.

En la Figura 11, se muestra un esquema de lo explicado anteriormente, en donde sobre una la línea temporal (t), se muestra el registro de tiempos de fallo (t_1, t_2, t_3, t_4), correspondiente a todas las tuberías que conforman un grupo de análisis, entre los cuales se define la variable TBF .

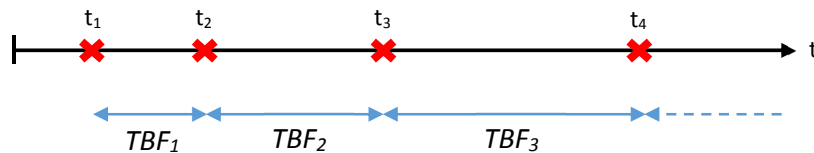


Figura 11. Esquema para representación de la variable TBF .

3.2 Modelo de riesgos proporcionales de Weibull. WPHM

3.2.1 Descripción general

Este es un modelo de tipo estadístico semiparamétrico perteneciente a la categoría de modelos de procesos de renovación (Cox y Isham, 1980), el cual estudia los eventos recurrentes en los componentes de un sistema. El primero en aplicarlo fue Andreou, (1986) y posteriormente Eisenbeis, (1994) realizó varias mejoras al mismo.

Este modelo se conoce también como modelo de vida acelerada Weibull (WALM, por sus siglas en inglés) y se basa en el análisis de supervivencia a partir de datos de históricos de fallos de cada tubería, las cuales pueden ser estratificadas por

características que guarden relación con el proceso de fallo como por ejemplo el material, diámetro, etc. La probabilidad de fallos se relaciona con diferentes factores como las variables relacionadas a la operación de la red o factores ambientales. Para este modelo se recomienda al menos 5 años de registros de mantenimiento (Eisenbeis et al., 2002).

Este modelo también es un modelo estadístico de análisis de supervivencia que supone el ajuste de los datos de análisis a una distribución de tipo Weibull. Podría considerarse una versión más compleja y detallada respecto al modelo propuesto (SERAP), ya que permite relacionar la influencia de diferentes factores de riesgo mediante una matriz de covariables y además obtener una predicción de fallos detallada por cada elemento de red. A partir de estos antecedentes, así como también la repercusión que ha tenido, reflejado en los trabajos posteriores de autores (Alvisi y Franchini, 2010; da Costa, 2011; Kimutai et al., 2015); es que se ha considerado realizar una comparativa respecto a este modelo.

La Figura 12 (Alvisi y Franchini, 2010) muestra un esquema genérico para una tubería en la cual se presentan varios fallos durante su periodo de observación. También se observa el criterio adoptado para definir la variable de tiempo entre fallos. Otras variables que aparecen en la figura se definirán más adelante.

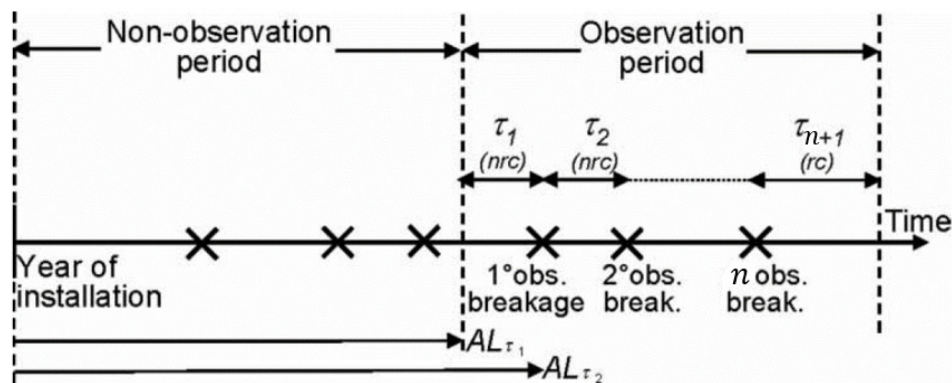


Figura 12. Esquema utilizado por el modelo de riesgo proporcional de Weibull. Alvisi y Franchini, 2010

La principal variable de análisis es el tiempo entre fallos, la cual es descrita mediante una distribución Weibull. Siendo $T(\tau_1, \tau_2, \dots, \tau_n)$ una variable de tiempo de vida entre n fallos, se asume que el logaritmo natural de T se relaciona con un vector X a través de un modelo lineal, válido para todas las tuberías (Solano, 2008; Le Gat y Eisenbeis, 2000), esto se puede definir con las siguientes expresiones.

$$Y = \ln(T) = X\beta + \sigma W \quad (3.1)$$

$$T = e^{X\beta} e^{\sigma W} \quad (3.2)$$

Donde Y es un vector del logaritmo de enésimos tiempos de fallo con o sin censura, X es una matriz de covariables, β es un vector⁴ de coeficientes transpuesto ($\beta = [\beta_0, \beta_1, \beta_2, \dots, \beta_p]$ para $(p + 1)$ covariables, σ es un parámetro de regresión desconocido a estimar. En particular β_0 y σ son conocidos como *Intercept* y *Scale*, los cuales no son iguales, pero se relacionan a los parámetros de forma y escala del modelo propuesto presentado en la sección 3.1, por cuanto hay que tener cuidado de no confundirlos. Finalmente W es un vector de errores aleatorios, independiente e idénticamente distribuido acorde a una distribución de valores extremos, con una función de supervivencia denotada por S :

$$S(W) = \exp[-\exp(W)] \quad (3.3)$$

Despejando W de la ecuación (3.1) y sustituyendo la expresión resultante dentro de la ecuación (3.3), se obtiene la función de supervivencia de WPHM:

$$S(t, \beta, x) = \exp \left[-\exp \left(\frac{\ln t - x\beta}{\sigma} \right) \right] = \exp \left[-t^{1/\sigma} \exp \left(-\frac{x\beta}{\sigma} \right) \right] \quad (3.4)$$

El nombre de modelo de vida acelerada se debe al efecto de aceleración o desaceleración que generan las variables explicativas sobre la variable T . Así, las variables explicativas actúan multiplicativamente sobre la variable T o aditivamente sobre la variable $\ln(T)$, ecuaciones (3.1) y (3.2) (Solano, 2008).

⁴ A diferencia del modelo propuesto en este caso beta β (sin estilo de cursiva) corresponde a un vector de coeficientes y no al valor específico de un parámetro.

Esquema del modelo y estructura de datos

Al igual que el modelo anterior, la estructura de datos se compone principalmente de dos bases de datos, una que describe las tuberías de la red y otro con el registro de sus fallos. La variable principal de análisis es el tiempo entre fallos, junto con las variables explicativas a considerarse en el estudio. En la Figura 13 (Eisenbeis et al., 2002), se muestra los diferentes procesos que sigue este modelo.

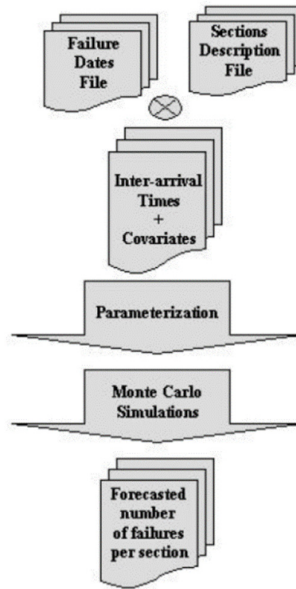


Figura 13. Esquema funcional general. Fuente: Eisenbeis et al., 2002

La estructura de la nueva base de datos tiene como principal nexo el identificador de las tuberías, donde cada una se lista $n+1$ veces su número de fallos como se muestra en la siguiente tabla.

Tabla 8. Fusión de base de datos de descripción de tuberías e historial de reparaciones. Fuente: Eisenbeis et al., 2002

Id.	U	C	R	Age	Mat.	Diám.	Long.	X
...
$i-1$	$t(i-1, n_{i-1}) - t(i-1, n_{i-1}-1)$	1	n_{i-1}	$t(i-1, n_{i-1}-1) - t(i-1, 0)$	$M(i-1)$	$D(i-1)$	$L(i-1)$	$X(i-1)$
$i-1$	$t_b(i-1) - t(i-1, n_{i-1})$	0	$n_{i-1}+1$	$t(i-1, n_{i-1}) - t(i-1, 0)$	$M(i-1)$	$D(i-1)$	$L(i-1)$	$X(i-1)$
i	$t(i, 1) - t_a(i)$	1	1	$t_a(i) - t(i, 0)$	$M(i)$	$D(i)$	$L(i)$	$X(i)$
...
i	$t(i, n_i) - t(i, n_i-1)$	1	n_i	$t(i, n_i-1) - t(i, 0)$	$M(i)$	$D(i)$	$L(i)$	$X(i)$
i	$t_b(i) - t(i, n_i)$	0	n_i+1	$t(i, n_i) - t(i, 0)$	$M(i)$	$D(i)$	$L(i)$	$X(i)$
$i+1$	$t(i+1, 1) - t_a(i+1)$	1	1	$t_a(i+1) - t(i+1, 0)$	$M(i+1)$	$D(i+1)$	$L(i+1)$	$X(i+1)$
...

Donde U corresponde al tiempo entre fallos, C es la censura, R es el número o rango de fallo, Age es la edad de la sección en la fecha de fallo anterior.

La variable de tiempo entre fallos considerada en este modelo (τ , Figura 12) se estructura de manera distinta a la mostrada en el modelo propuesto (TBF , Figura 8). Ya que el tiempo entre fallos considerado en este modelo, corresponde al tiempo del primer fallo registrado y los subsecuentes registrados para una misma tubería, de esta manera se van listando todos los elementos en la base de datos (Tabla 8). Mientras que para el modelo propuesto (SERAP) el tiempo entre fallos no distingue la tubería en que se origina el fallo y se descarta el primer intervalo de tiempo, entre el inicio de la observación y el primer fallo. De esta manera, los tiempos entre fallos no quedan asociados por tubería, sino por el conjunto de tuberías total (TBF_1 y TBF_2 de Figura 8).

3.2.2 Estimación de parámetros

Los parámetros de regresión se estiman por medio del método de máxima verosimilitud. Se aplica el logaritmo natural a la función de verosimilitud, que se encuentra en función de la variable de tiempo entre fallos t_n ($i=1,2,\dots,n$), los coeficientes de regresión β y las variables explicativas asociadas X :

$$\ln[L(t_1, t_2, \dots, t_n, \beta, X)] = \sum_{i, nrc} \ln(f(t_i, \beta, X)) + \sum_{i, rc} \epsilon(F(t_i, \beta, X)) \quad (3.5)$$

Para la estimación de los parámetros de este modelo se ha utilizado el software R. En el caso de la función utilizada en el software R, ésta parametriza la distribución de una forma que puede llevar a confusiones sobre los parámetros Intercept y Scale, antes mencionados (Therneau y Lumley, 2015). A continuación, se muestra la equivalencia de los parámetros proporcionados por R, respecto a los parámetros del modelo propuesto (Sección 3.1), donde β (en estilo cursivo) es el parámetro de forma, y η (en estilo cursivo) es el parámetro de escala.

$$Scale = 1/\beta \quad (3.6)$$

$$intercept = \log(\eta) \quad (3.7)$$

Por otro lado, durante el periodo de observación todas las tuberías son truncadas por la izquierda y pueden o no ser censuradas por derecha. Conforme a la Figura 12, los datos no censurados por derecha (nrc, por sus siglas en inglés) contribuyen a la probabilidad de la función de densidad, mientras que los datos censurados por derecha (rc, por sus siglas en inglés) contribuyen a la función de supervivencia, tal como se muestra en ecuación (2.14).

El resultado obtenido de los coeficientes de las covariables permite determinar la influencia de cada una sobre el modelo, lo cual puede resultar muy útil en un estudio de análisis de fallos en función de factores de riesgo.

3.2.3 Predicción de fallos

El modelo de riesgos proporcionales de Weibull que se presenta se basa en modelo presentado en el trabajo de Le Gat y Eisenbeis (2000). El proceso analítico para calcular el número acumulado de fallos puede ser calculado teóricamente por una distribución de convolución de sucesivos tiempos de fallos, lo cual resulta un proceso considerablemente complicado. Le Gat y Eisenbeis (2000), proponen un procedimiento de cálculo aproximado mediante el uso del proceso de simulación de Monte Carlo. Para ello, a partir de la función de supervivencia, ecuación (3.4) para cada tubería se generan tiempos de fallos aleatorios, para posteriormente proceder a predecir su número de fallos dentro de un tiempo horizonte. A continuación, se muestra la expresión para determinar el tiempo de fallos.

$$t = \left(\ln \left(\frac{1}{S} \right) \exp \left(\frac{x\beta}{\sigma} \right) \right)^\sigma \quad (3.8)$$

El proceso de simulación de Monte Carlo se lleva a cabo de la siguiente manera:

- Se genera un número aleatorio mayor a cero y menor que 1 , el cual representa la supervivencia S , con lo cual es posible obtener el tiempo de fallo (t) de la expresión (3.8).

- Si el tiempo de fallo obtenido es menor al tiempo horizonte de predicción, vuelve a generar un nuevo tiempo de fallo utilizando una versión actualizada de la función de supervivencia, lo cual implica actualizar los datos de las covariables dinámicas. Esto se repite hasta que el tiempo horizonte de predicción es alcanzado, guardando registro del número de tiempos de fallos generados, lo cual representa el número de fallos acumulado.
- Lo anterior se repite 1000 veces. Siendo el valor promedio del número de fallos de estas mil repeticiones, el número de fallos estimado para la tubería. Este mismo procedimiento se repite para cada tubería que conforma la muestra o grupo de análisis.

Como ejemplo de una variable dinámica se tiene la “*edad izquierda*” (abreviada *AL* por sus siglas en inglés), la cual es una variable introducida en la propuesta del modelo de Le Gat y Eisenbeis (2000), en un intento por considerar la edad como una covariable del modelo.

Se toma en consideración *AL* y no la edad real de la tubería al momento del fallo porque, durante el proceso de simulación, este dato debe ser ingresado como una covariable, y esto no es posible dado que el tiempo en que se produce el fallo es la incógnita a estimar. Así, se utiliza la covariable *AL* que representa la edad (AL_{i-1}) conocida de una tubería antes de la ocurrencia de un fallo (*i*), según el esquema de la Figura 12.

4 Aplicación de los modelos

4.1 Breve descripción del caso de estudio

El caso de estudio corresponde a un abastecimiento de dimensiones considerables, con una población servida mayor a los 500,000 habitantes y una longitud de red mayor a 1,000 kilómetros. Se cuenta con dos bases de datos, una con el histórico de reparaciones correspondiente a los años 2017 y 2018, y otra que almacena los datos de cada tubería y sus características obtenidas a través de un sistema de información geográfica (SIG).

De estas bases de datos es importante reconocer aquellos datos que son útiles para el análisis. A propósito de esto, como referencia para la construcción o estructuración de un sistema de información, existen varias propuestas como la expuesta por St. Clair y Sinha (2014), en la cual plantean varias estructuras de datos enfocados a la predicción de roturas y análisis de modos y causas de fallos en redes de agua, dependientes de factores como el tamaño del sistema y personal técnico calificado, etc. En la base de datos del histórico de reparaciones, a efectos prácticos, la fecha de una reparación se asume como la fecha de ocurrencia del fallo.

Para el caso de estudio, la base de datos del SIG cuenta con datos generales de las tuberías, teniendo entre sus principales características: el material, diámetro, longitud, fecha de instalación, identificador de tramos de red (tuberías), entre otros. El dato del identificador de los tramos de red se encuentra presente en ambas bases de datos (histórico de reparaciones y SIG), lo que permite relacionar ambas bases de datos.

4.2 Selección de muestras o subgrupos de análisis

Un vistazo a los diferentes materiales que conforman la red de distribución de estudio es mostrado en la Figura 14, donde si bien, existe una diversa cantidad de materiales, se puede apreciar que los más significativos en cuanto a su frecuencia por longitud de red son: fibrocemento (FC), fundición dúctil (FD), fundición gris (FG) y Polietileno (PE).

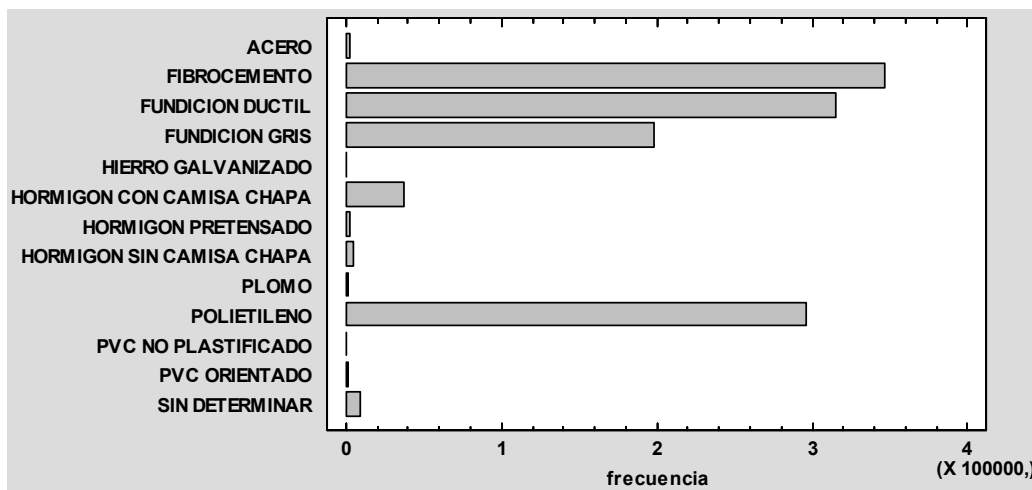


Figura 14. Frecuencia del tipo de material de la red de estudio en función de la longitud.

De esta manera, se definen estos cuatro materiales como diferentes grupos de análisis, que además deben cumplir con una cantidad mínima de ocurrencia de fallos, lo cual se muestra en la Tabla 9.

Tabla 9. Registro de número de fallos totales por cada tipo de material.

Grupo Mat.	Nº Fallos Totales
Fibrocemento (FC)	116
Fundición dúctil (FD)	79
Fundición gris (FG)	125
Polietileno (PE)	432
Total	752

La estratificación por materiales responde también a la necesidad de diferenciar los procesos de rotura que ocurren en cada tipo de material, debido a la variedad de su resistencia mecánica y química (Eisenbeis et al., 2002).

Dado que para el modelo propuesto (SERAP) se requiere agrupar tuberías con características homogéneas, en este trabajo se ha decidido además del tipo material, tomar en consideración el diámetro, siendo inicialmente estos dos los principales criterios para selección de los grupos de análisis. Los diámetros de las tuberías se han agrupado en seis clase o rangos:

Tabla 10. Grupos de tuberías según su diámetro (exterior)

Clase	Rango Diámetro(mm)	Nº Fallos Totales
1	[20, 63)	192
2	[63, 140)	452
3	[140, 200)	71
4	[200, 280)	35
5	[280, 400)	2
6	>400	0

De las clases de diámetros definidas, la *Clase 2* es la que cuenta con un mayor registro de fallos y la que se tomará en consideración para el análisis. Por lo tanto, agrupando las tuberías por estos dos factores (Material y Clase de Diámetro), los subgrupos quedan definidos como se muestran en la Tabla 11.

Tabla 11. Definición y características de subgrupos de análisis seleccionados.

Subgrupo	Material	Clase Diámetro	Nº de elementos	Longitud (Km)	Fallos Registrados	
					1er Fallo	Nº Fallos
1	FC	2	6149	165.7	62	75
2	FD	2	8278	121.4	56	58
3	FG	2	5242	120.1	70	87
4	PE	2	5382	138.1	166	232

Dado que el planteamiento 1 del modelo SERAP toma en consideración solo el registro del primer fallo, el número de registros con los que trabaja puede ser significativamente menor registro total de fallos disponible (dos últimas columnas de Tabla 11). A pesar de ello, el número de registros de primer fallo cumple con el número mínimo de registros necesarios (>30) para el análisis. Por otro lado, el planteamiento 2 del modelo SERAP y el modelo de riesgos proporcionales de Weibull (WPHM), permiten tomar en consideración el total del registro de fallos (última columna de Tabla 11).

4.3 Modelo SERAP

4.3.1 Planteamiento 1

A partir de una base de datos estructurada y tomando como referencia la representación realizada en la Tabla 6, se analizan las variables de tiempo: edad (E) y tiempo hasta el fallo (TTF). La variable de tiempo entre fallos (TBF) será analizada de manera diferenciada posteriormente.

La Tabla 12, presenta la estimación de los parámetros de la distribución paramétrica (Weibull), para la variable E . A su vez, la Tabla 13 muestra los resultados equivalentes a la tabla que le antecede, pero para el análisis de la variable TTF .

Tabla 12. Parametrización Weibull para la edad (E).

Subgr.	Grupos Mat. / Diám.	Distribución Weibull			
		β	η	μ	s
1	FC / 2	6.93	88.27	82.53	13.99
2	FD / 2	9.83	84.48	80.31	9.82
3	FG / 2	5.08	161.45	148.37	33.49
4	PE / 2	6.96	77.90	72.85	12.30

Tabla 13. Parametrización Weibull para el tiempo hasta el fallo (TTF).

Subgr.	Grupos Mat. / Diám.	Distribución Weibull			
		β	η	μ	s
1	FC / 2	0.91	311.84	326.16	358.91
2	FD / 2	1.33	83.48	76.76	58.28
3	FG / 2	1.02	135.58	134.48	131.84
4	PE / 2	1.12	44.42	42.61	38.11

En la Tabla 12 y Tabla 13 (y sobre el resto de este documento), los símbolos μ y s representan la media y desviación estándar respectivamente. La media μ se relaciona con la vida útil media (MTTF), tiempo en que se espera que el 50% de las tuberías fallen.

Las siguientes tablas muestran la evaluación de las pruebas de bondad de ajuste para E y TTF . Con ello se verifica si los datos de los diferentes subgrupos pueden o no ser representados con la distribución Weibull.

Tabla 14. Pruebas de bondad de ajuste para la Edad E

Subgr.	Grupos Mat. / Diám.	Prueba de bondad de ajuste K-S		
		D_{MAX} .	D_{CR}	Evaluación [$D_{MAX} < D_{CR}$]
1	FC / 2	0.059	0.017	No cumple
2	FD / 2	0.003	0.014	Cumple
3	FG / 2	0.304	0.018	No Cumple
4	PE / 2	0.010	0.018	Cumple

Tabla 15. Pruebas de bondad de ajuste para el tiempo hasta el fallo TTF

Subgr.	Grupos Mat. / Diám.	Prueba de bondad de ajuste		
		D_{MAX} .	D_{CR}	Evaluación [$D_{MAX} < D_{CR}$]
1	FC / 2	0.001	0.017	Cumple
2	FD / 2	0.0008	0.014	Cumple
3	FG / 2	0.002	0.018	Cumple
4	PE / 2	0.003	0.018	Cumple

La prueba de bondad de ajuste de Kolmogorov-Smirnov (K-S) se cumple cuando el valor D_{MAX} resulta ser menor que el valor D_{CR} . En tal caso, se acepta la hipótesis nula, que supone que la distribución de datos puede ser representada mediante una distribución específica (Mora, 2009; Luís y Maya, 1999).

Conforme a evaluación de la Tabla 14 los resultados obtenidos para E en los subgrupos 1 y 3, correspondientes a los materiales FC y FG respectivamente, no cumplen las pruebas de bondad de ajuste. De este modo, el análisis de la variable E de estas muestras debe descartarse. Por otro lado, los resultados obtenidos para la variable TTF de estos mismos subgrupos (Tabla 15), sí cumplen las pruebas de bondad de ajuste. Así, para esta variable, sí podrían analizarse estos últimos resultados obtenidos.

Si bien en las tablas presentadas sólo se muestran los resultados obtenidos con la distribución Weibull, se verificó también que para la variable E el resto de distribuciones tampoco cumple las pruebas de bondad de ajuste para los subgrupos 1 y 3.

Las siguientes gráficas presentan la función de distribución acumulada empírica (FDAE) obtenida con el estimador de Kaplan-Meier, contra la distribución paramétrica de Weibull, para las variables E y TTF de cada subgrupo. En ellas se puede verificar gráficamente el ajuste de las dos distribuciones.

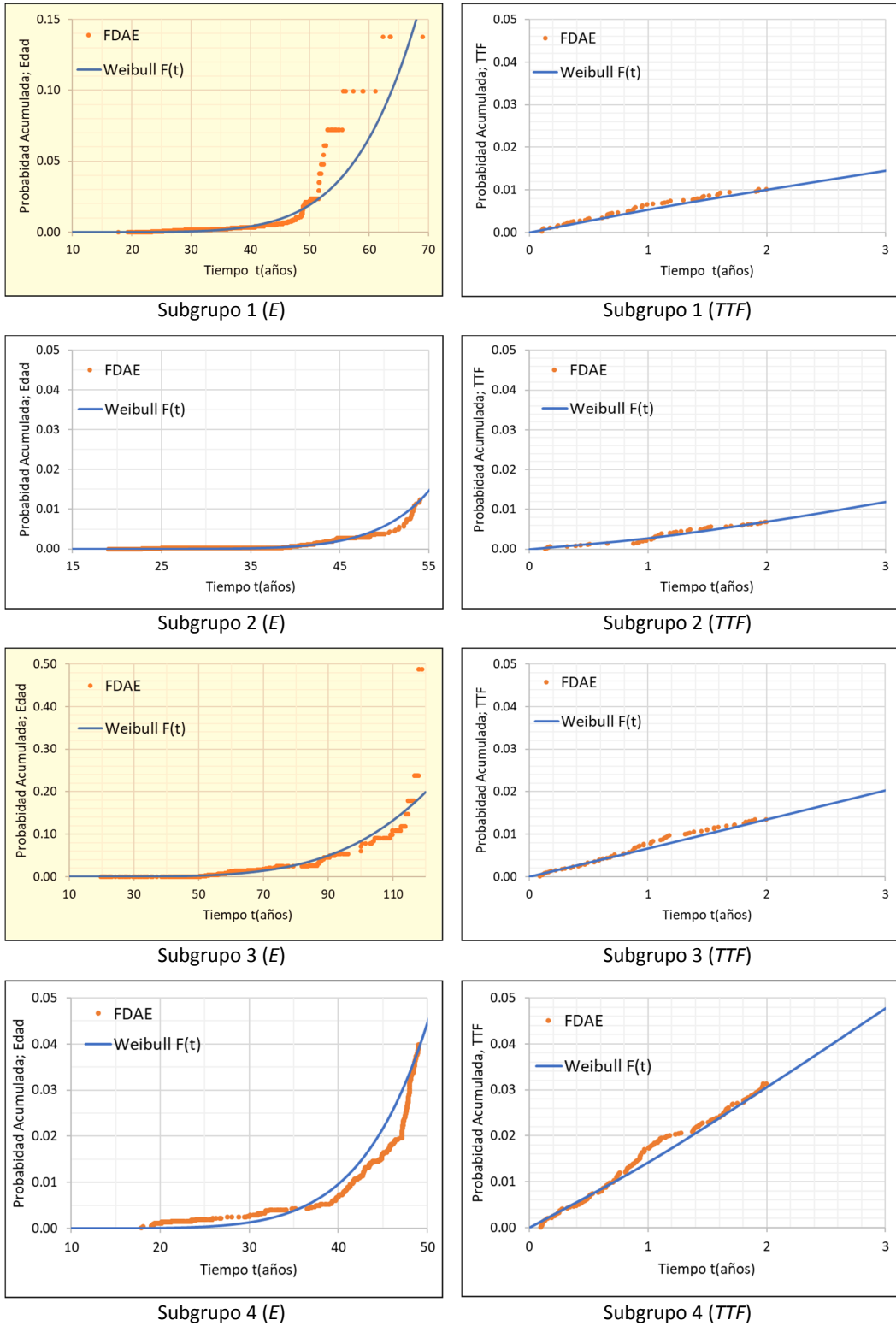


Figura 15. FDAE vs. Función de distribución Weibull para cada subgrupo. Modelo SERAP (Planteamiento 1)

Como se puede apreciar en los gráficos de *Subgrupo 1 (E)* y *Subgrupo 3 (E)* (resaltados en la Figura 15), existen puntos dispersos al final de la cola derecha, que no se ajustan a la función paramétrica. Si se asumen los datos de mayor edad, como datos marginales, es posible tener un mejor ajuste para estos subgrupos, como se aprecia en la Figura 16.

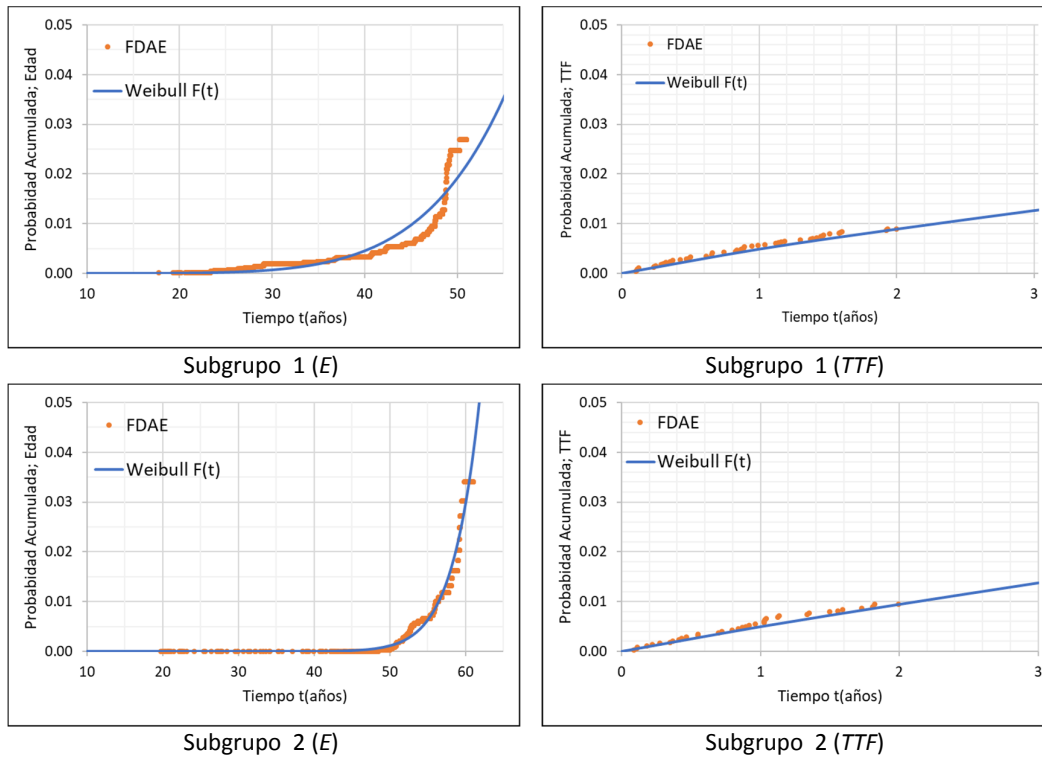


Figura 16. FDAE vs. Función de distribución Weibull para subgrupos 1 y 3. Modelo SERAP (Planteamiento 1)

De esta manera, incluyendo un nuevo criterio de selección, se ha redefinido los subgrupos de análisis 1 y 3, a fin que la variable *E* cumpla con las pruebas de bondad de ajuste, quedando finalmente los grupos de análisis como se muestra en la Tabla 16.

Tabla 16. Subgrupos de análisis definitivos. (*) subgrupos redefinidos⁵

Subgrupo	Material	Clase Diámetro	Edad (años)	Nº de elementos	Longitud (Km)	Fallos Registrados	
						1er Fallo	Nº Fallos
1*	FC	2	<50	5851	156.3	52	64
2	FD	2	-	8278	121.4	56	58
3*	FG	2	<60	3812	84.6	36	51
4	PE	2	-	5382	138.1	166	232

⁵ En adelante el (*) junto a los subgrupos 1 y 3, denota que se trata de los subgrupos redefinidos.

A causa de la redefinición de los subgrupos, se puede apreciar que el registro de fallos para los subgrupos 1* y 3* quedan afectados debido a la restricción realizada. Sin embargo, el registro de fallos para ambos subgrupos (1er Fallo y Nº Fallos) sigue siendo mayor que el mínimo estipulado para la aplicación del modelo (>30 registros). Con ello, se repite el proceso de obtención de parámetros de la distribución Weibull (Tabla 17 y Tabla 18) y se verifican las pruebas de bondad de ajuste (Tabla 19 y Tabla 20), donde en esta ocasión sí cumplen para todos los subgrupos de análisis.

Tabla 17. Parametrización Weibull para E. Modelo SERAP (Planteamiento 1)

Subgr.	Grupos Mat. / Diám. / Edad	Distribución Weibull			
		β	η	μ	s
1*	FC / 2 / <50	6.49	91.84	85.57	15.42
2	FD / 2	9.83	84.48	80.31	9.82
3*	FG / 2 / <60	17.94	72.91	70.78	4.87
4	PE / 2	6.96	77.90	72.85	12.30

Tabla 18. Parametrización Weibull para TTF. Modelo SERAP (Planteamiento 1)

Subgr.	Grupos Mat. / Diám. / Edad	Distribución Weibull			
		β	η	μ	s
1*	FC / 2 / <50	0.87	455.08	488.49	563.89
2	FD / 2	1.33	83.48	76.76	58.28
3*	FG / 2 / <60	0.93	297.38	307.38	330.42
4	PE / 2	1.12	44.42	42.61	38.11

Tabla 19. Pruebas de bondad de ajuste para E. Modelo SERAP (Planteamiento 1)

Subgr.	Grupos Mat. / Diám. / Edad	Prueba de bondad de ajuste K-S		
		D_{MAX}	D_{CR}	Evaluación [$D_{MAX} < D_{CR}$]
1*	FC / 2 / <50	0.007	0.017	Cumple
2	FD / 2	0.003	0.014	Cumple
3*	FG / 2 / <60	0.006	0.020	Cumple
4	PE / 2	0.010	0.018	Cumple

Tabla 20. Pruebas de bondad de ajuste para TTF. Modelo SERAP (Planteamiento 1)

Subgr.	Grupos Mat. / Diám. / Edad	Prueba de bondad de ajuste		
		D_{MAX}	D_{CR}	Evaluación [$D_{MAX} < D_{CR}$]
1*	FC / 2 / <50	0.001	0.017	Cumple
2	FD / 2	0.0008	0.014	Cumple
3*	FG / 2 / <60	0.0003	0.020	Cumple
4	PE / 2	0.003	0.018	Cumple

Estimación de vida útil

En este trabajo se ha establecido definir la vida útil de las tuberías según el valor obtenido del parámetro de escala η de la distribución Weibull, a partir del análisis de la variable E para cada subgrupo. Como se ha indicado anteriormente, el parámetro η , también conocido como “vida característica”, representa el tiempo en que se espera que el 63.2% de los elementos fallen y no depende del parámetro de forma β que describe la distribución Weibull. A diferencia del parámetro de escala η , la media μ sí depende y puede variar en función del parámetro β , por este motivo se ha contemplado considerar al parámetro de escala η para definir la vida útil.

En este trabajo se ha considerado la variable E y no TTF para definir la vida útil de las tuberías, debido a:

- Dado que se cuenta con el dato de fecha de instalación de las tuberías, resulta más verosímil definir el tiempo de vida útil en función de la edad en que se registran los fallos.
- Las estimaciones de η para la variable TTF , pueden resultar en valores irreales (subestimados o sobrestimados) de tiempo de vida útil. Por ejemplo, se puede ver los valores obtenidos en la Tabla 18 para los subgrupos 1* y 3*.
- Las medidas de dispersión de la desviación estándar son mucho mayores en los resultados para TTF , que para los obtenidos para E .

La Tabla 21 presenta la estimación de la vida útil de cada subgrupo.

Tabla 21. Estimación de la vida útil. Modelo SERAP (Planteamiento 1)

Subgr.	Grupos Mat. / Diám. / Edad	Estimación Vida útil (años)
1*	FC / 2 / <50 años	91.8
2	FD / 2	84.5
3*	FG / 2 / < 60 años	72.9
4	PE / 2	77.9

Estado de las tuberías

Para evaluar el estado de los subgrupos, se han empleado los parámetros de forma β obtenidos de la variable de TTF . No se utiliza el parámetro de forma de la variable E , ya

que éste suele estar sobrestimado o ser muy tendencioso. Además, partiendo del hecho de que se cuenta con datos acotados (truncados por la izquierda y censurados por derecha), resulta sensato evaluar el estado de las tuberías en función de las incidencias que se presentan durante el periodo de observación, dado que es incierto lo ocurrido antes de dicho periodo.

Atendiendo al análisis del parámetro β , obtenido a través de la variable *TTF*, los subgrupos 1* y 3* con $\beta = 0.87$ y 0.93 correspondientes a los materiales FC y FG respectivamente, se ubicarían en la fase 1 de la curva de bañera, presentando un comportamiento de mortalidad infantil o fallos tempranos, propio de tuberías más nuevas. Por otro lado, los subgrupos 2 y 4 con $\beta = 1.33$ y 1.12 correspondientes a los materiales FD y PE respectivamente, presentan un comportamiento de tuberías que están iniciando su fase de envejecimiento (fase 3, etapa 1). A partir de ello resultan las acciones genéricas de mantenimiento recomendadas para cada subgrupo. En la Tabla 22 se resume lo comentado en este párrafo.

Tabla 22. Evaluación del estado de tuberías y acciones recomendadas. Modelo SERAP (Planteamiento 1)

Subgrupo	Grupos Mat. / Diám. / Edad	β	Ubicación en Curva de bañera		Acción Táctica recomendada
			Fase	Etapas	
1*	FC / 2 / <50 años	0.87	1	-	Correctiva
2	FD / 2	1.33	3	1	Preventiva
3*	FG / 2 / < 60 años	0.93	1	-	Correctiva
4	PE / 2	1.12	3	1	Preventiva

Los resultados obtenidos para los subgrupos 1 y 3 resultan no verosímiles, ya que los materiales que los conforman corresponden a los grupos de tubería más longevos del sistema y teóricamente con mayor desgaste. Más adelante junto a los resultados de los otros modelos, se analiza y se da una explicación posible de este resultado.

Predicción de fallos

La predicción de fallos para los subgrupos de tuberías seleccionadas (Tabla 23) se realiza tomando en consideración las observaciones de la variable *TTF* y asumiendo que cada subgrupo, para el tiempo horizonte de predicción, mantiene el comportamiento que

describe la distribución paramétrica. O, en otras palabras, que conserva la fase que describe su estado en la curva de la bañera.

Para su cálculo se evalúa la función de riesgo $\lambda(t)$ para distintos años por el número de tuberías del subgrupo.

Tabla 23. Predicción de fallos. Modelo SERAP (Planteamiento 1)

Subgrupo	Grupos Mat. / Diám. / Edad	Número de fallos				
		Año 1	Año 2	Año 3	Año 4	Año 5
1*	FC / 2 / <50 años	25	23	21	21	20
2	FD / 2	31	38	44	48	52
3*	FG / 2 / < 60 años	18	17	16	16	16
4	PE / 2	86	94	98	102	104

Debido a que el *TTF* puede comportarse en el tiempo de una forma diferente a la contemplada durante el periodo de observación, no es recomendable realizar predicciones a largo plazo con esta metodología. Por lo tanto, sería recomendable realizarlo para el corto plazo y procediendo con continuas actualizaciones para verificar el ajuste de las predicciones y un posible cambio de fase de los subgrupos.

4.3.2 Planteamiento 2

Se aplica la estructura de datos presentada en la Tabla 7. Para ello, cada tubería se divide en tantos segmentos como unidades de longitud posea la misma, con lo cual el número de elementos para cada subgrupo se ve incrementado como se muestra en la Tabla 24. Si una tubería presenta uno o más fallos, estos se reparten en los segmentos de la tubería generados, donde cada segmento admite solo un fallo, de esta manera es posible tomar en consideración el total de fallos de cada tubería original.

Tabla 24. Características de subgrupos de análisis. Modelo SERAP (Planteamiento 2)

Subgrupo	Material	Clase Diámetro	Edad (años)	Nº de elementos	Longitud (km)	Fallos Registrados
						Nº Fallos totales
1	FC	2	<50	156,397	156.3	64
2	FD	2	-	121,585	121.5	58
3	FG	2	<60	84,694	84.6	51
4	PE	2	-	138,171	138.1	232

El proceso de seccionamiento de los tramos de tubería puede provocar ligeras variaciones en la longitud total de la red analizada. Sin embargo, se considera que tales variaciones son despreciables y no afectan al análisis.

En las siguientes tablas se muestra la estimación de los parámetros de la distribución Weibull para las variables E y TTF de la estructura de datos correspondiente a este planteamiento.

Tabla 25. Parametrización Weibull para E . Modelo SERAP (Planteamiento 2)

Subgr.	Grupos Mat. / Diám. / Edad	Distribución Weibull			
		β	η	μ	s
1*	FC / 2 / <50	7.48	126.34	118.58	18.74
2	FD / 2	9.96	108.80	103.49	12.49
3*	FG / 2 / <60	16.36	88.86	86.04	6.47
4	PE / 2	7.084	115.03	107.67	17.89

Tabla 26. Parametrización Weibull para TTF . Modelo SERAP (Planteamiento 2)

Subgr.	Grupos Mat. / Diám. / Edad	Distribución Weibull			
		β	η	μ	s
1*	FC / 2 / <50	0.9	11,600	12,205.33	13,584.90
2	FD / 2	1.33	635.78	584.84	444.99
3*	FG / 2 / <60	1.07	2,070	2,016.35	1,885.64
4	PE / 2	1.18	436.09	412.04	350.43

Las siguientes tablas presentan las pruebas de bondad de ajuste, las cuales cumplen la hipótesis nula para todos los subgrupos, y a continuación la Figura 17 muestra los gráficos de las funciones de distribución empírica (FDAE) y teórica asumida (Weibull).

Tabla 27. Pruebas de bondad de ajuste para E . Modelo SERAP (Planteamiento 2)

Subgr.	Grupos Mat. / Diám. / Edad	Prueba de bondad de ajuste K-S		
		D_{MAX}	D_{CR}	Evaluación [$D_{MAX} < D_{CR}$]
1*	FC / 2 / <50	0.0004	0.0034	Cumple
2	FD / 2	0.0002	0.0038	Cumple
3*	FG / 2 / <60	0.0005	0.0046	Cumple
4	PE / 2	0.0005	0.0036	Cumple

Tabla 28. Pruebas de bondad de ajuste para TTF . Modelo SERAP (Planteamiento 2)

Subgr.	Grupos Mat. / Diám. / Edad	Prueba de bondad de ajuste		
		D_{MAX}	D_{CR}	Evaluación [$D_{MAX} < D_{CR}$]
1*	FC / 2 / <50	0.00005	0.0034	Cumple
2	FD / 2	0.0001	0.0039	Cumple
3*	FG / 2 / <60	0.0001	0.0046	Cumple
4	PE / 2	0.0001	0.0036	Cumple

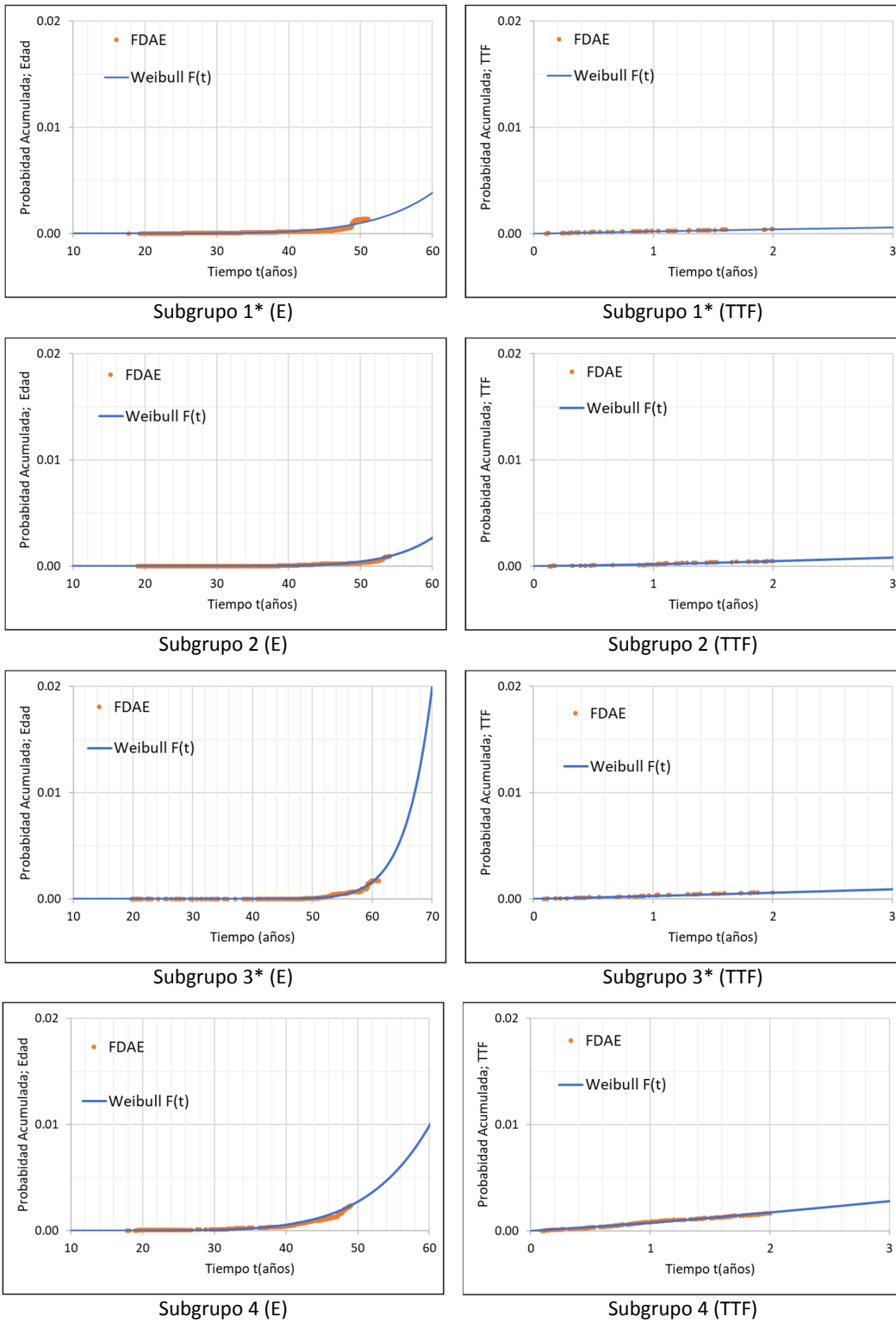


Figura 17. FDAE vs. Función de distribución Weibull para cada subgrupo. Modelo SERAP (Planteamiento 2)

En contraste con los gráficos del planteamiento 1, se puede apreciar que las pendientes de las curvas son mucho menos pronunciadas, esto incluso cuando se ha aumentado la escala del eje de las ordenadas para los gráficos del planteamiento 2.

Estimación de vida útil

La estimación de la vida útil para el planteamiento 2 se realiza de igual manera al planteamiento 1, a partir del parámetro de escala η del análisis de la variable E . En la siguiente tabla se muestra la vida útil estimada para cada subgrupo.

Tabla 29. Estimación de la vida útil. Modelo SERAP (Planteamiento 2)

Subgrupo	Grupos	Estimación Vida útil (años)
	Mat. / Diám. / Edad	
1*	FC / 2 / <50 años	126.3
2	FD / 2	108.8
3*	FG / 2 / < 60 años	88.9
4	PE / 2	115.0

De esta manera, la estimación de vida útil en el planteamiento 2, respecto al planteamiento 1, aumenta de la siguiente manera para los distintos subgrupos.

Tabla 30. Incremento de vida útil de planteamiento 2 vs. planteamiento 1.

Subgrupo	Incremento (años)
1*	34.5
2	24.3
3*	16
4	37.1

Estado de las tuberías

El estado de las tuberías se evalúa, de igual manera que el planteamiento 1, en función del parámetro de forma β estimado para la variable TTF .

Analizando el valor de los parámetros obtenidos se deduce que el subgrupo 1* con $\beta = 0.90$, se ubica en la fase 1 de la curva de bañera, siendo muy similar a la respuesta obtenida en el planteamiento 1. Los subgrupos 2, 3 y 4 con $\beta = 1.33; 1.07; 1.18$ respectivamente, se ubican en la fase 3 (etapa 1) de la curva de bañera; siendo los resultados de los subgrupos 2 y 4 prácticamente los mismos a los obtenidos en el

planteamiento 1. Sin embargo, el subgrupo 3 si presenta cambios, pasando de la fase 1 (en el planteamiento 1) a ubicarse en la fase 3, etapa 1, la cual indica el inicio de la fase de envejecimiento. En la Tabla 31, se presenta un resumen de la evaluación.

Tabla 31. Evaluación del estado de tuberías y acciones recomendadas. Modelo SERAP (Planteamiento 2)

Subgrupo	Grupos Mat. / Diám. / Edad	β	Ubicación en		Acción Táctica recomendada
			Curva de bañera		
			Fase	Etapa	
1*	FC / 2 / <50 años	0.90	1	-	Correctiva
2	FD / 2	1.33	3	1	Preventiva
3*	FG / 2 / < 60 años	1.07	3	1	Preventiva
4	PE / 2	1.18	3	1	Preventiva

Predicción de fallos

La predicción de fallos se realiza de forma análoga al planteamiento 1 y bajo las mismas premisas. En la siguiente tabla se muestra el resultado de la predicción de fallos aplicando el planteamiento 2.

Tabla 32. Predicción de fallos. Modelo SERAP (Planteamiento 2)

Subgrupo	Grupos Mat. / Diám. / Edad	Número de fallos				
		Año 1	Año 2	Año 3	Año 4	Año 5
1*	FC / 2 / <50 años	31	29	28	27	26
2	FD / 2	30	38	43	48	51
3*	FG / 2 / < 60 años	26	27	28	28	29
4	PE / 2	125	142	153	161	167

En general, a excepción de subgrupo 2, se aprecia un incremento en la predicción del número de fallos del planteamiento 2 respecto al planteamiento 1. Para el subgrupo 2 se obtienen valores muy similares.

4.3.3 Variable tiempo entre fallos TBF

Al igual que los casos anteriores, se asume en este apartado que los datos de la variable *TBF* pueden ser representados mediante una distribución paramétrica. En las siguientes tablas se muestran los resultados obtenidos de la estimación de los parámetros y de las pruebas de bondad de ajuste, aplicado tanto para el planteamiento 1 y para el planteamiento 2.

Tabla 33. Parametrización Weibull para TBF. Modelo SERAP (Planteamiento 1)

Subgr.	Grupos Mat. / Diám. / Edad	Distribución Weibull			
		β	η	μ	s
1*	FC / 2 / <50	0.756	0.030	0.035	0.048
2	FD / 2	0.782	0.029	0.034	0.043
3*	FG / 2 / <60	1.020	0.050	0.050	0.048
4	PE / 2	0.857	0.0105	0.011	0.013

Tabla 34. Parametrización Weibull para TBF. Modelo SERAP (Planteamiento 2)

Subgr.	Grupos Mat. / Diám. / Edad	Distribución Weibull			
		β	η	μ	s
1*	FC / 2 / <50	0.771	0.025	0.030	0.038
2	FD / 2	0.783	0.028	0.033	0.042
3*	FG / 2 / <60	0.867	0.033	0.036	0.042
4	PE / 2	0.832	0.007	0.008	0.009

Tabla 35. Pruebas de bondad de ajuste para TBF. Modelo SERAP (Planteamiento 1)

Subgr.	Grupos Mat. / Diám. / Edad	Prueba de bondad de ajuste		
		D_{MAX}	D_{CR}	Evaluación [$D_{MAX} < D_{CR}$]
1*	FC / 2 / <50	0.05	0.19	Cumple
2	FD / 2	0.07	0.18	Cumple
3*	FG / 2 / <60	0.07	0.23	Cumple
4	PE / 2	0.11	0.10	No Cumple

Tabla 36. Pruebas de bondad de ajuste para TBF. Modelo SERAP (Planteamiento 2)

Subgr.	Grupos Mat. / Diám. / Edad	Prueba de bondad de ajuste		
		D_{MAX}	D_{CR}	Evaluación [$D_{MAX} < D_{CR}$]
1*	FC / 2 / <50	0.08	0.17	Cumple
2	FD / 2	0.07	0.18	Cumple
3*	FG / 2 / <60	0.07	0.19	Cumple
4	PE / 2	0.14	0.08	No Cumple

En la Tabla 33 y Tabla 34 el valor de la media μ representa el tiempo medio entre fallos (MTBF), este es un valor útil dentro del campo de estudio del análisis de fiabilidad y mantenibilidad, que brinda información importante acerca del estado de los elementos o componentes de un sistema.

En la Tabla 35 y Tabla 36 se puede apreciar que para el subgrupo 4, tanto para el planteamiento 1 y 2, no se cumple la prueba de bondad de ajuste. Esto es así, a pesar de que la mayoría de los datos de este subgrupo sí se ajustan a la curva paramétrica de la función de distribución acumulada, como se puede ver en las gráficas correspondientes de la Figura 18.

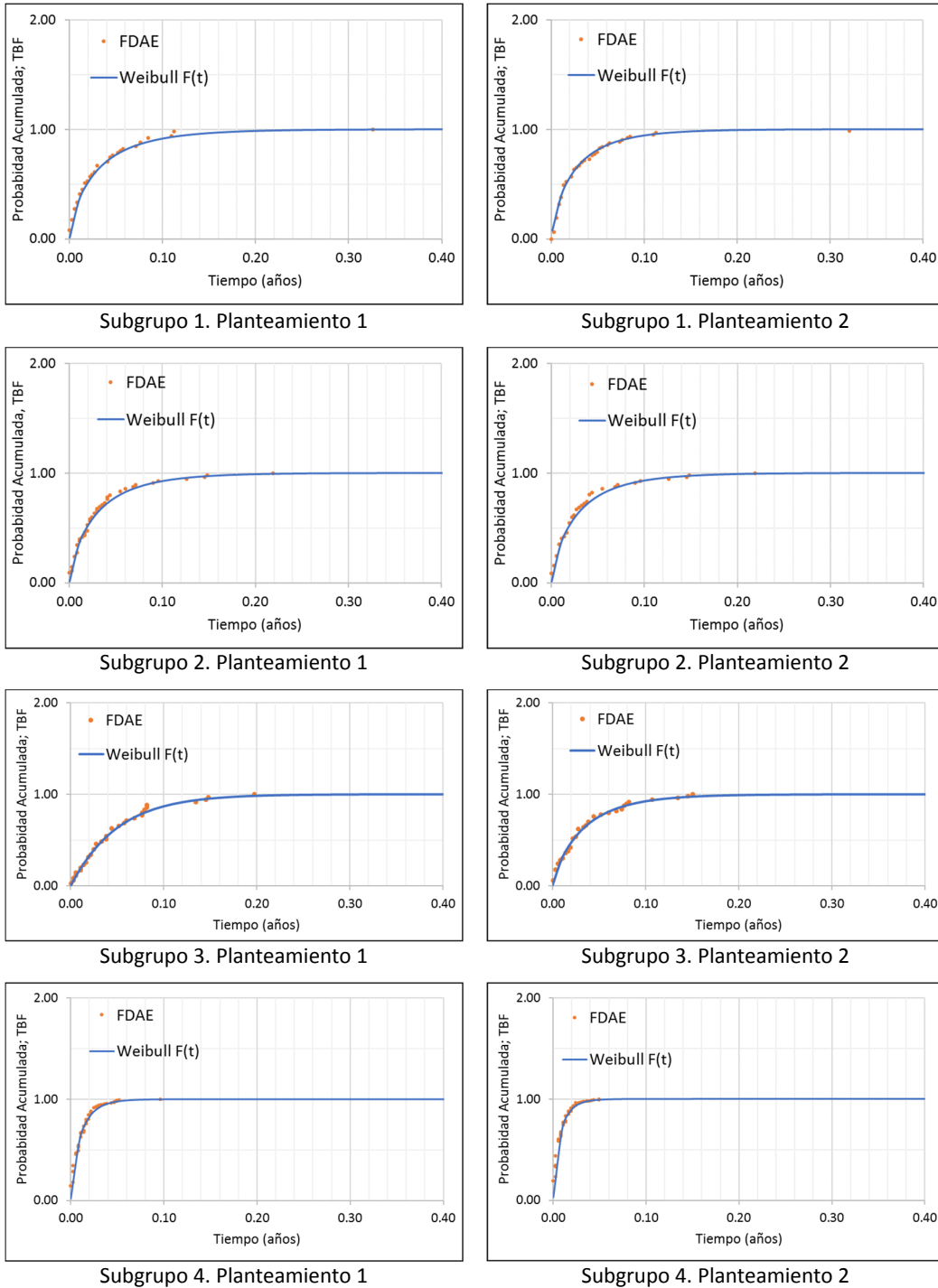


Figura 18. FDAE vs. Función de distribución acumulada Weibull para el análisis de la variable TBF.

En el caso del subgrupo 4, donde no se cumple la prueba de bondad de ajuste, se ha optado por considerar como datos marginales a aquellos datos que no se ajustan a la distribución acumulada paramétrica, a fin de poder analizar los resultados obtenidos.

La interpretación de los parámetros estimados a partir de esta variable (*TBF*) es similar a lo visto previamente. De esta manera, el parámetro β permite reconocer la forma de la función de supervivencia y de la función de riesgo, mientras que el parámetro η nos proporciona el tiempo cuando la probabilidad de fallo es del 63.2%.

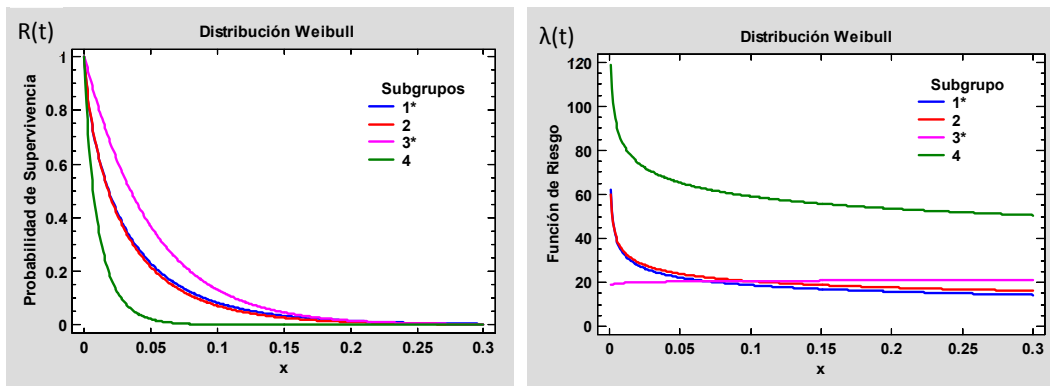


Figura 19. Distribución Weibull función $R(t)$ y $\lambda(t)$ de la variable *TBF*. Planteamiento 1

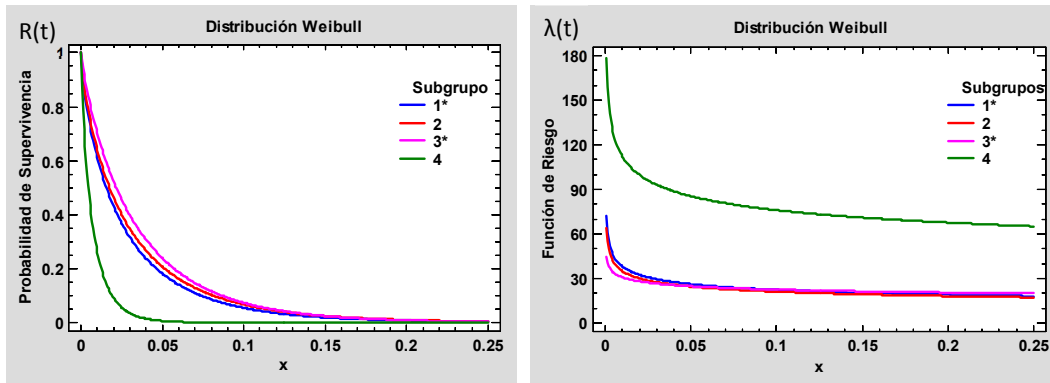


Figura 20. Distribución Weibull función $R(t)$ y $\lambda(t)$ de la variable *TBF*. Planteamiento 2.

Para la curva $R(t)$, el área bajo la curva representa la supervivencia de cada subgrupo de análisis. Así, se puede apreciar que el subgrupo 4, correspondiente al material PE, tanto para el planteamiento 1 y 2, es el que posee una menor supervivencia en relación a la variable *TBF* y, por consiguiente, el riesgo de fallo es mayor para este subgrupo tal como se observa en las curvas de $\lambda(t)$. Por su parte, los subgrupos 1*, 2 y 3* tienen curvas más semejantes, sobre todo en el planteamiento 2.

Lo anterior se puede ver también reflejado en la siguiente tabla donde se muestra el MTBF expresado en días y, a partir de ello, se estima el número de fugas anuales para los planteamientos 1 y 2. Como se puede ver en misma tabla, los valores para los subgrupos 1*,2 y 3*, para el planteamiento 2 son similares, lo cual se relaciona con los gráficos de las dos figuras anteriores.

Tabla 37. Tiempo entre fallos en días y estimación de promedio de fugas anuales.

Subgrupos	Planteamiento 1		Planteamiento 2	
	MTBF (días)	Fugas/año	MTBF (días)	Fugas/año
1*	13	28	10.7	34
2	12.3	30	11.9	31
3*	18.1	20	13.2	28
4	4.1	88	2.9	122

Puede verificarse también que los resultados de la tabla anterior son similares a los obtenidos para el primer año de la predicción realizada para el planteamiento 1 (Tabla 23) y el planteamiento 2 (Tabla 32), lo cual implica cierta relación entre estos resultados.

4.4 Modelo WPHM

Este modelo se ha desarrollado tomando en consideración los mismos subgrupos de tuberías analizadas en el modelo SERAP. El modelo requiere la construcción de una estructura de datos como la mostrada en la Tabla 8, en la cual se analizan los tiempos de fallos sucesivos observados en una tubería, así como también de las tuberías que no presentan fallos, las cuales son incluidas especificando que se tratan de datos censurados por la derecha, para ello se asigna el valor de 0 en la variable de censura. Para estas mismas tuberías, la variable de tiempo entre fallos (τ) se define como la diferencia entre el tiempo de inicial y final del periodo de observación.

Estimación de parámetros y evaluación de covariables

Los principales argumentos ingresados en el software *R* para la estimación de los parámetros son las variables: tiempo entre fallos (τ), censura y las covariables de influencia consideradas. En este modelo se han considerado las covariables logaritmo natural de la longitud, Diámetro, Número de fallos y la edad izquierda (*AL*). En el caso

de la covariable del diámetro, el análisis se realiza en función de los distintos diámetros que conforman la Clase 2. Los resultados del ajuste son los siguientes:

Tabla 38. Parámetros y coeficientes estimados. Modelo WPHM

	Subgrupo 1		Subgrupo 2		Subgrupo 3		Subgrupo 4	
	β	Exp(β)	β	Exp(β)	β	Exp(β)	β	Exp(β)
Intercept β_0	5.38	-	8.15	-	20.07	-	3.81	-
Scale	1.18	-	0.58	-	1.3	-	0.88	-
Ln Long. β_1	-0.206	0.813	-0.287	0.750	-0.174	0.839	-0.236	0.789
Diámetro β_2	0.006	1.006	-0.022	0.977	-0.065	0.937	0.016	1.016
Nº fallos β_3	-3.319	0.036	-2.353	0.095	-2.497	0.082	-0.728	0.482
AL β_4	0.041	1.042	-0.025	0.974	-0.109	0.896	-0.024	0.975

Acerca de los coeficientes, Kimutai et al. (2015) anota lo siguiente:

- Los coeficientes que más se alejan de cero, tanto positiva como negativamente, indican un mayor riesgo de fallo.
- El exponencial de los coeficientes (β) indican la contribución de la covariable sobre el riesgo de fallo de una tubería, donde valores cercanos a 1 implican una contribución marginal al riesgo.

A continuación, se describen algunos puntos importantes considerados en este trabajo para el desarrollo y aplicación del modelo WPHM.

- La variable longitud ha sido introducida en el modelo aplicando el logaritmo natural sobre su valor, asumiendo que la tasa de fallos es proporcional a la longitud de la tubería, tal como menciona da Costa (2011). Así, los coeficientes de la covariable “Ln.long β_1 ”, con valores próximos a -1 indican que la tasa de riesgo es proporcional a la longitud de la tubería. Para los casos de tuberías con dimensiones menores a 1 metro, donde el logaritmo natural toma valores negativos, se ha restringido el valor mínimo que puede tomar esta covariable a 1×10^{-6} .
- Le Gat y Eisenbeis (2000), recomiendan la incorporación de variables relacionadas a la edad izquierda (AL) y el número de fallos previos (NOPF, por sus siglas en inglés), ya que en su estudio estas se consideran de alta significancia. Sin embargo, en el caso de la covariable (NOPF), este puede derivar en

problemas de cómputo al momento de realizar la simulación para la predicción del número de fallos. Así, los tiempos entre fallos que se generan son cada vez más cortos, tendiendo a no alcanzar el tiempo horizonte de predicción, produciendo un número de fallos excesivamente alto o provocando que la simulación entre en un ciclo infinito. Para evitar este problema Alvisi y Franchini (2010) omiten el uso de esta variable, mientras que da Costa (2011) propone dos artificios para conseguir tomar en consideración el número de fallos previos que registran las tuberías.

- Respecto al punto anterior, en este trabajo se ha planteado el uso de la covariable, N° fallos, la cual representa el número de fallos durante el periodo de observación. Esto en un intento por tomar en consideración la influencia del número previo de fallos, pese que a diferencia de NOPF esta es una variable de tipo constante.

A continuación, se realiza la evaluación de las covariables, de acuerdo a los resultados obtenidos.

- Conforme a lo explicado en la sección 3.2.2, a través del parámetro de regresión *Scale*, es posible obtener el parámetro de forma β . Para los subgrupos 1 y 3, $\beta = 0.84$ y 0.76 , respectivamente, lo cual indica que sus tuberías se ubican en la fase 1 de la curva de bañera, siendo estos resultados similares a los obtenidos en el planteamiento 1 del modelo SERAP. Mientras que para los subgrupos 2 y 4, $\beta = 1.72$ y 1.13 , respectivamente, ello ubica a estos subgrupos de análisis en la fase 3, etapa 1, de la curva de bañera. Este resultado es similar a los obtenidos en los planteamientos 1 y 2 del modelo SERAP, para estos dos subgrupos.
- Respecto al parámetro *intercept*, relacionado con el logaritmo natural del factor de escala (η), se obtienen lo siguiente $\eta_{1*} = 217.0$; $\eta_2 = 3,463.3$; $\eta_{3*} = 5.2 \times 10^8$; $\eta_4 = 45.1$, respectivamente para cada subgrupo. De manera similar a lo ocurrido en el modelo SERAP, se podría considerar que la estimación de este parámetro a partir de la variable de tiempo entre fallos resulta en

subestimaciones, sobrestimación o valores irreales no verosímiles en la estimación del tiempo de vida útil.

- Para la covariable Ln Long. (β_1), el signo negativo en el coeficiente se presenta en todos los subgrupos. Esto se puede interpretar como que cuanto mayor sea la longitud de una tubería, menor resultará su tiempo entre fallos.
- Para la covariable Diámetro (β_2), los coeficientes de los subgrupos con signo positivo, indican que el tiempo entre fallos es mayor para tuberías de mayor diámetro, mientras que si el signo es negativo indica lo contrario. Esto es en referencia al rango del diámetro de las tuberías de clase 2 analizadas (de 63 a 140 mm).
- La covariable N° fallos (β_3), es la covariable que implica un mayor riesgo de fallo en las tuberías. El exponencial de sus coeficientes señalan la contribución al riesgo, donde los valores resultantes de esta covariable se pueden observar que son los más lejanos de la unidad (1= contribución marginal).

Predicción de número de fallos.

Para la predicción de fallos se toman los valores de los parámetros y coeficientes obtenidos para cada subgrupo (Tabla 38), y haciendo uso de la ecuación (3.8) se aplica la simulación de Monte Carlo siguiendo las directrices presentadas por Le Gat y Eisenbeis (2000). Para ello es necesario preparar una hoja de cálculo donde se realizarán las predicciones, en la cual conste cada una de las tuberías del subgrupo de análisis y sus covariables.

La consideración de covariables de tipo dinámico merece especial atención ya que deben estar actualizadas para el tiempo de inicio de la predicción y durante el proceso de cómputo de la simulación para que su valor sea actualizado continuamente, como es el caso de la variable relacionada a la edad izquierda (AL).

La edad izquierda (AL), como se ha mencionado previamente, comprende el tiempo en años entre la fecha de instalación y el inicio de la predicción. De ser el caso, esta covariable se debe actualizar, acumulando su valor, en cada iteración de la simulación

de Monte Carlo hasta que se alcanza el tiempo horizonte de predicción en las simulaciones.

A continuación, se muestra la predicción del número total de fugas para cada subgrupo de análisis considerado aplicando este modelo.

Tabla 39. Predicción de fallos. Modelo WPHM.

Subgr.	Grupos Mat. / Diám. / Edad	Número de fallos				
		Año 1	Año 2	Año 3	Año 4	Año 5
1*	FC / 2 / <50 años	35	30	29	28	26
2	FD / 2	18	34	44	52	59
3*	FG / 2 / < 60 años	35	30	30	31	34
4	PE / 2	113	130	138	144	147

La expresión utilizada para estimar los fallos en 5 años consecutivos es:

$$m(T_{i-1}, T_i) = \sum_{p=1}^{n_p} m_p(T_0, T_i) - \sum_{p=1}^{n_p} m_p(T_0, T_{i-1}) \quad (4.1)$$

donde m es el número de predicción de fallos, $T_0 = 0$ y representa el inicio del tiempo horizonte de predicción de fallos (momento inmediato al fin del periodo de observación), T_i es el tiempo horizonte de predicción de fallos, y p es número de tuberías. En el caso de realizar la predicción solo para el primer año la expresión anterior se reduce a:

$$m(T_0, T_1) = \sum_{p=1}^{n_p} m_p(T_0, T_1) \quad (4.2)$$

4.5 Comparación entre modelos

Los modelos aplicados presentan dos características comunes. Ambos tienen la capacidad de:

- Valorar el estado de las tuberías en función del parámetro β .
- Predecir el número total de fallos a futuro.

Adicionalmente, el modelo SERAP tiene la capacidad de estimar una vida útil de las tuberías mediante el análisis de la variable E y estimar el promedio de fallos anuales a

partir del análisis de la variable *TBF*. Mientras que el modelo WPHM, permite además evaluar factores de riesgo o variables de influencia y obtener una predicción de fallos por cada tubería.

La predicción del número de fallos mediante el modelo SERAP es para todo el subgrupo y se considera para un tiempo horizonte de corto plazo, de 1 hasta 5 años. Mientras que para el modelo WPHM no se especifica el tiempo horizonte de predicción, sin embargo, se asume también que es capaz de proporcionar un tiempo horizonte de predicción de corto plazo.

En la Tabla 40 y la Tabla 41 se comparan los resultados de predicción de fallos obtenidos para tiempos horizonte de 1 a 5 años, para el modelo SERAP (planteamiento 1 y 2), respecto al modelo de WPHM.

Tabla 40. Error absoluto entre modelo SERAP Planteamiento 1, respecto al modelo WPHM

Subgr.	Grupos Mat. / Diám. / Edad	Error absoluto				
		Año 1	Año 2	Año 3	Año 4	Año 5
1*	FC / 2 / <50 años	-0.28	-0.23	-0.27	-0.24	-0.23
2	FD / 2	0.68	0.11	0.00	-0.08	-0.11
3*	FG / 2 / < 60 años	-0.49	-0.44	-0.48	-0.49	-0.54
4	PE / 2	-0.24	-0.28	-0.29	-0.29	-0.29

En la tabla anterior se observa que, en general, el planteamiento 1 del modelo SERAP subestima el número de fallos en relación al modelo de WPHM. Esto a excepción del subgrupo 2, donde para los dos primeros años la predicción es sobrestimada, mientras que para los tres últimos años la predicción es más próxima entre los dos modelos.

Tabla 41. Error absoluto entre modelo SERAP Planteamiento 2, respecto al modelo WPHM

Subgr.	Grupos Mat. / Diám. / Edad	Error absoluto				
		Año 1	Año 2	Año 3	Año 4	Año 5
1*	FC / 2 / <50 años	-0.11	-0.04	-0.03	-0.02	0.00
2	FD / 2	0.62	0.11	-0.03	-0.08	-0.13
3*	FG / 2 / < 60 años	-0.26	-0.11	-0.08	-0.11	-0.16
4	PE / 2	0.11	0.09	0.11	0.12	0.13

En el planteamiento 2, en relación al modelo WPHM, los subgrupos 1 y 3 subestiman el número de fallos, el subgrupo 4 los sobrestima, mientras que para el subgrupo 2 se sobrestiman los fallos de los dos primeros años y para los siguientes 3 años existe una

subestimación. Sin embargo, la diferencia numérica entre los resultados obtenidos por estos dos modelos es mucho más cerrada o próxima.

Para el planteamiento 2, las predicciones del subgrupo 2 son muy similares a los resultados obtenidos en el planteamiento 1. Mientras que para el resto de subgrupos de ambos planteamientos la diferencia es más significativa.

Si bien las predicciones del número de fallos en los modelos planteados en su mayoría difieren de lo que se podría considerar un error admisible (por ejemplo, un error relativo de $\pm 10\%$), sus valores tampoco son desproporcionados por cuanto se requiere verificar la eficacia de predicción de los modelos ante datos reales, lo cual no ha sido posible en este caso de estudio debido al corto registro de fallos históricos (2 años).

Se debe también tomar en consideración el número de fugas anuales obtenido a través del análisis de la variable *TBF*, el cual se aproxima a algunos valores de predicción obtenidos en los modelos. Una comparación entre estas aproximaciones puede ser consultada en el Anexo 2.

Por otra parte, el factor de forma β resulta muy similar en los modelos, correspondiéndose en su mayoría a las mismas fases dentro de la curva de bañera. Esto denota que, para este caso de estudio, ambos modelos describen de manera próxima la función de riesgo de fallos.

Tabla 42. Comparación de Parámetro de forma beta para los modelos.

Subgrupo	Grupos Mat. / Diám./ edad	Parámetro de forma β		
		SERAP	SERAP	Modelo WPHM
		Planteamiento 1	Planteamiento 2	
1*	FC / 2 / ≤ 50	0.86	0.90	0.84
2	FD / 2	1.33	1.33	1.72
3*	FG / 2 / ≤ 60	0.93	1.07	0.76
4	PE / 2	1.12	1.18	1.14

El subgrupo 3* es el único que sufre un cambio significativo en relación al valor del parámetro de forma β , lo cual determina su ubicación en la curva de bañera. Para el caso del modelo SERAP-planteamiento 1 y el modelo WPHM, el valor del parámetro es menor de 1. Mientras que para el modelo SERAP-planteamiento 2, el valor es

ligeramente mayor a 1. Dado que este subgrupo comprende a las tuberías de material FG (material más antiguo), podría suponerse que el planteamiento 2 del modelo SERAP podría representar en mejor medida el estado de este subgrupo.

Respecto a los valores del parámetro $\beta < 1$, para el caso del FC y FG (subgrupos 1 y 3), pueden existir varias explicaciones para describir este comportamiento en los modelos. Una de ellas, tal como lo explica Le Gat y Eisenbeis (2000), (haciendo alusión a los coeficientes obtenidos para la covariable AL , en el modelo de su trabajo), puede deberse a que las tuberías más sensibles a los fallos de estos materiales ya han sido sustituidas y han ido desapareciendo de la base de datos, siendo las tuberías restantes las más antiguas y las menos sensitivas a los fallos. En el caso del FC podría pensarse que durante su tiempo de servicio estas tuberías pudieron haber sido parcialmente reemplazadas, eliminando las tuberías más sensibles sin que se haya actualizado la base de datos, si bien esto no se podría asegurar, explicaría en algo los resultados obtenidos. En todo caso, resultaría razonable pensar, debido a la antigüedad de estos materiales, que las tuberías de FC y FD podrían realmente ubicarse en una etapa más avanzada de desgaste de su vida útil, con un $\beta > 1$.

5 Conclusiones

En este trabajo se ha logrado alcanzar el objetivo planteado de proponer una metodología para estimar la vida útil de las tuberías a través de un modelo estadístico, que además permite realizar la predicción del número de fallos futuros, así como evaluar el estado de las tuberías. Los modelos presentados han partido del análisis de la edad hasta el fallo (E), el tiempo hasta el fallo (TTF) y el tiempo entre fallos (TBF). También, se ha presentado y aplicado un modelo de predicción de fallos existente ($WPHM$), el cual ha sido contrastado con el modelo propuesto.

Una de las principales limitaciones de la aplicación de la metodología propuesta en este estudio ha sido el corto historial de reparaciones, de apenas dos años, por cuanto todo esto debe ser tomado con una cierta precaución. Sin embargo, la conformación de los modelos presentados en este trabajo se respalda en el uso de la estadística y conceptos relacionados a la ingeniería de la mantenibilidad y confiabilidad. Además, se han tomado como referencia diversos trabajos relacionados al tema de investigación que han permitido dar consistencia al trabajo realizado.

A continuación, se describen varias conclusiones y recomendaciones sobre el desarrollo de este trabajo que comprenden aspectos relacionados a la estructuración de los datos, el modelo propuesto ($SERAP$) y el modelo existente presentado ($WPHM$). Finalmente se asientan las líneas de investigación y trabajos futuros propuestos a partir del estudio realizado.

5.1 Bases y estructuración de datos

Si bien las bases de datos del caso de estudio han permitido la construcción de los modelos desarrollados en este trabajo, ha sido necesario una depuración previa de las mismas al encontrar datos inconsistentes, lo que provocó en ciertos casos la pérdida de datos. Los aspectos de mejora recomendados entorno a la conformación de la base de datos en este trabajo son:

- Asegurar la confiabilidad de los datos, ya que una pobre calidad de los mismos puede generar distorsión de los resultados de los modelos, de los cuales finalmente depende la toma de decisiones.
- Eficiencia de recolección de datos, registrando aquellas variables que son más significativas a la hora de construir modelos, para así no agotar recursos en el registro de datos que puedan terminar no siendo útiles.

Adicionalmente a las variables significativas que se recomienda en la bibliografía (presión, tipo de suelo, profundidad, etc.), se mencionan, de manera particular, algunas variables que pudieron resultar de gran interés:

- Clase de fuga. Definir si la fuga es comunicada (visible), o no comunicada (detectada por tareas de control activo de fugas) a fin de conocer el origen o proveniencia de las fugas y conocer cómo influye este factor en los modelos.
- Modo de fallo. Diferenciar entre los diferentes tipos de fallo que pueden producirse en una tubería (transversal, longitudinal, agujero, aplastamiento, explosiva). Esto con el objeto de poder conocer las posibles causas de los fallos y en base a ello plantear acciones de mejora.
- Factor de impacto de roturas. En función de criterios operacionales, sociales, ambientales o económicos. De manera que este factor pueda ser incluido en los modelos y sirva para determinar las tuberías más prioritarias a renovar o rehabilitar en función de su impacto.

5.2 Modelos

Modelo SERAP

El modelo propuesto denominado SERAP se caracteriza por requerir una reducida cantidad de variables (básicamente las características básicas de las tuberías), y por contar con un proceso de desarrollo del modelo sencillo de aplicar. Las predicciones para este modelo se han estimado para tiempos horizontes de corto plazo, de 1 a 5 años.

El planteamiento 1, que toma información solo del registro del primer fallo de las tuberías, supone el modelo más simple de los analizados en este trabajo. Además, resulta de interés particular en este estudio, dada la posibilidad de poder conseguir en un futuro más información referente a años previos, que se relacionan a la estructura de datos que sigue este planteamiento.

El planteamiento 2, en principio soluciona la falencia del planteamiento 1 de solo poder analizar el primer fallo. Este modelo supone una mejor caracterización del fallo de las tuberías, ya que las mismas en la práctica fallan de manera parcial o no a lo largo de toda su longitud. En referencia a la estructura de los datos, este planteamiento requiere un poco más de trabajo respecto al planteamiento 1. Sin embargo, esto ha sido solucionado aplicando una secuencia de programación, mediante VBA, lo cual optimiza el tiempo del desarrollo del modelo.

Modelo WPHM

En la aplicación del modelo WPHM se han seguido las directrices presentadas por Le Gat y Eisenbeis (2000) para la estimación de coeficientes y predicción de número de fallos, pero se ha omitido el uso de la covariable *NOPF* y en su lugar se ha propuesto el uso de la covariable *Nº de fallos*. Otras covariables complementarias como la presión, tipo de suelo, etc., no han sido consideradas debido a la ausencia de estos datos. Aunque la aplicación de este modelo ha sido recomendada en diferentes casos de estudio debido a la aproximación de sus predicciones, especialmente para abastecimientos con

registros cortos de mantenimiento, debe verificarse su capacidad de predicción, considerando las covariables que han sido empleadas en este trabajo.

Comparación SERAP vs. WPHM

Los modelos presentados (SERAP y WPHM) tienen como puntos en común la capacidad de obtener el número total de fallos para un determinado tiempo y valorar el estado de las tuberías en base al parámetro de forma de la distribución Weibull que permite evaluar el estado de las tuberías, tomando en consideración que los mismos siguen la conocida curva de bañera.

Como particularidades especiales de cada modelo, el modelo SERAP permite estimar la vida útil de un grupo de tuberías y determinar el tiempo medio entre fallos (MTBF), mientras que el modelo WPHM permite evaluar factores de riesgo y obtener una predicción que detallada el número de fugas estimado por cada tubería.

En relación a la predicción del número de fallos, el planteamiento 1 del SERAP en general subestimó la cantidad de fallos, en relación a los resultados obtenidos por el planteamiento 2 del SERAP y el modelo WPHM, donde entre estos dos últimos existe una mayor similitud en resultados. Una predicción muy similar fue observada para el modelo SERAP, planteamientos 1 y 2, del subgrupo 2 correspondiente al material FD para los 5 años. En relación al modelo WPHM, para este caso el número de fugas fue sobrestimado para los dos primeros años y similares para los tres últimos años.

La evaluación del estado de las tuberías fue en general muy próxima en los tres modelos, describiendo las mismas fases dentro de la curva de la bañera, esto a excepción del subgrupo 3 correspondiente al material FG, donde para el caso del planteamiento 2 del modelo SERAP se ubica en otra fase dentro de la curva de bañera, que a propósito definiría de mejor manera su estado.

Dadas las características del modelo WPHM, podría considerarse como una versión más robusta o más potente que el modelo SERAP, sin embargo, su desarrollo y análisis son

mucho más complejos a diferencia del modelo propuesto (SERAP), el cual en términos generales es mucho más sencillo de aplicar.

Debido a que ambos modelos permiten obtener resultados particulares y comunes entre sí, puede resultar conveniente el uso de ambos de manera complementaria.

5.3 Trabajos y líneas de instigación futuras

Desarrollos futuros, producto de la realización de este trabajo pueden derivarse en la mejora específica de los modelos presentados y en su aplicación dentro de un plan de gestión de activos.

En ambos casos, es necesario primero determinar la bondad de los modelos en cuanto a la precisión de los resultados que proporcionan. Ello requiere una validación de los modelos, lo cual se puede realizar aplicándolos sobre un abastecimiento que tenga un periodo de prueba, o tiempo de observación de registro de fallos, mínimo necesario para comparar datos reales contra datos estimados (5 años mínimo conforme a lo recomendado en otros estudios). Esto, acompañado de una base de datos confiable que contenga las características básicas de las tuberías como ya se ha detallado en el desarrollo de los modelos.

Respecto a puntos específicos para conseguir mejorar, calibrar u optimizar los modelos presentados, se pueden listar lo siguiente:

1. Para el modelo SERAP en general, si bien la distribución Weibull se ha ajustado a los datos de los subgrupos estudiados, sería razonable variar el análisis explorando con otro tipo de distribuciones paramétricas (p. ej. lognormal, gamma y normal) y comparar resultados.
2. Para el planteamiento 2 del modelo SERAP, se puede analizar la longitud óptima de los seccionamientos que deberían considerarse para optimizar el ajuste y validación del modelo. En este trabajo se ha prefijado el valor en 1 metro (tantas unidades como longitud de red existente), sin embargo, para distintos sistemas podría resultar conveniente redefinir esta longitud a fin de obtener una mejor

estimación en los resultados que proporcione el modelo, siendo en este caso, una de las opciones para su calibración

3. Para el modelo WPHM, podría adaptarse el seccionamiento de las tuberías utilizado en el planteamiento 2 del SERAP, con el objetivo suprimir la dependencia de los fallos relacionados a la covariable longitud, de manera que el modelo pueda simplificarse o en su lugar puedan ser consideradas otras covariables dentro del análisis.

Por otra parte, para el desarrollo de la gestión de activos, se pueden definir los siguientes puntos:

1. Estudio económico de las tuberías en función de los resultados obtenidos por los modelos presentados, a fin de obtener un modelo integrador que permita incluir una valoración económica y facilitar la toma de decisiones.
2. Desarrollar una aplicación de software que permitan automatizar los procesos de los modelos desarrollados.
3. Relacionar modelos según otros criterios que permitan determinar el cumplimiento de la vida útil de las tuberías, como por ejemplo la pérdida de capacidad hidráulica, afectación de la calidad, nivel de fugas o impacto de riesgo de fallo, entre otros.
4. Desarrollar otro modelo de análisis de supervivencia de tipo paramétrico, pero en esta ocasión utilizando un modelo de proceso no homogéneo de Poisson (NHPP), el cual resulta de interés, ya que este modelo muestra otro enfoque en el análisis de los fallos, considerando que el sistema se conforma de componentes o elementos de tipo reparable; y además, permite considerar el estado de los elementos reparados como “igual de buenos o igual de malos” que eran antes de la ocurrencia del fallo (Røstum, 2000).

Referencias

- Alegre, H.; Coelho, S. T. (2012) Infrastructure asset management of urban water systems. En: *Water Supply System Analysis-Selected Topics*; IntechOpen.
- Alvisi, S.; Franchini, M. (2010) Comparative analysis of two probabilistic pipe breakage models applied to a real water distribution system. *Civ. Eng. Environ. Syst.*, **27** (1), 1-22; Taylor & Francis.
- Amaral, R.; Alegre, H.; Matos, J. S. (2017) Highlights of key international water infrastructure asset management initiatives, and trends, challenges and developments in Portugal. *Water Policy*, **19** (1), 128-146; IWA Publishing.
- Andreou, S. (1986) *Predictive models for pipe failures and their implications on maintenance planning strategies for deteriorating water distribution systems*.
- Ascher, H.; Feingold, H. (1984) *Repairable systems reliability: modeling, inference, misconceptions and their causes*; M. Dekker New York.
- Asnaashari, A.; McBean, E. A.; Shahrour, I.; Gharabaghi, B. (2009) Prediction of watermain failure frequencies using multiple and Poisson regression. *Water Sci. Technol. Water Supply*, **9** (1), 9-19; IWA Publishing.
- Clair, A. M. St.; Sinha, S. (2014) Development of a Standard Data Structure for Predicting the Remaining Physical Life and Consequence of Failure of Water Pipes. *J. Perform. Constr. Facil.*, **28** (1), 191-203. Recuperado a partir de <http://ascelibrary.org/doi/10.1061/%28ASCE%29CF.1943-5509.0000384>

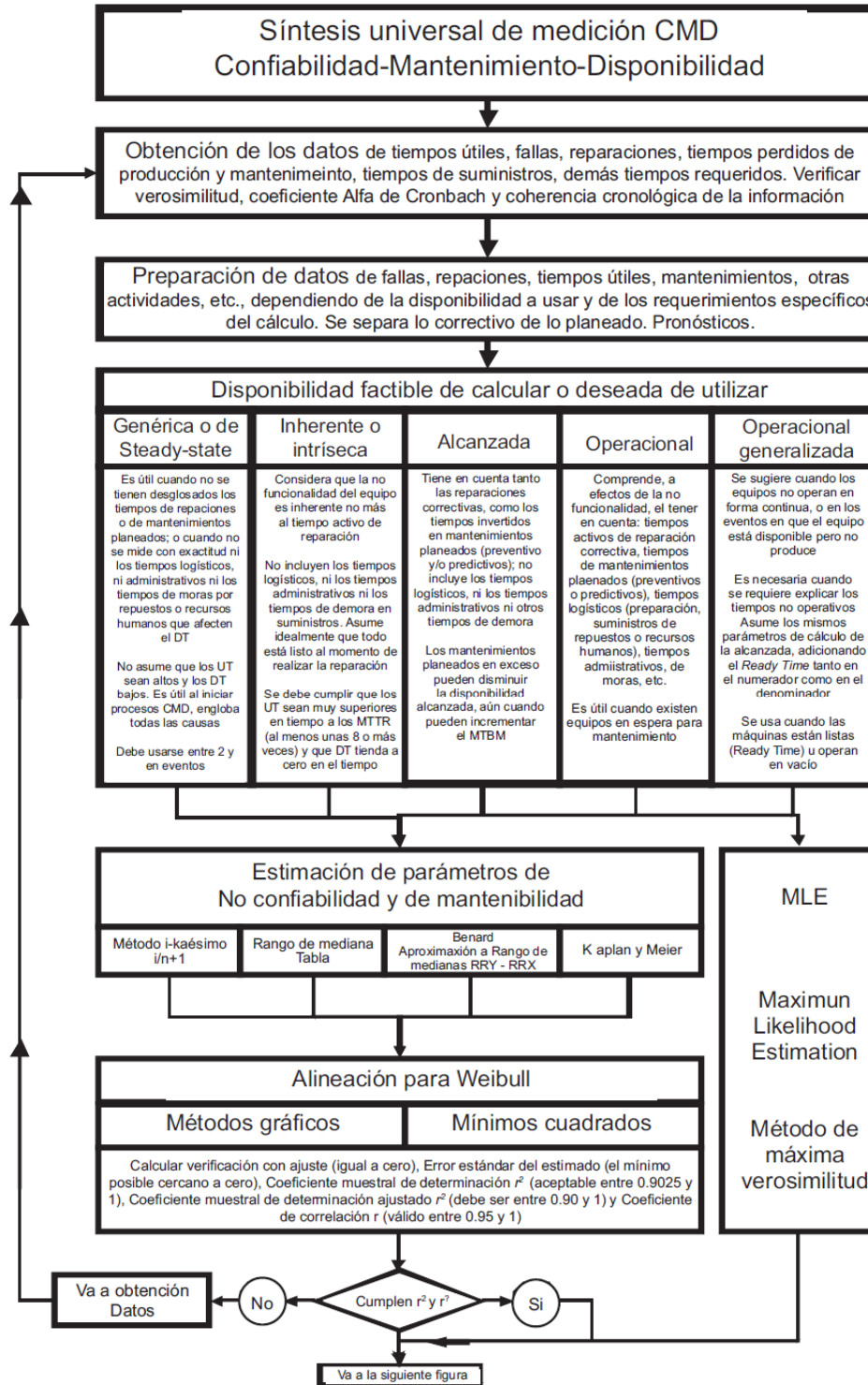
- Clark, R. M.; Stafford, C. L.; Goodrich, J. A. (1982) Water distribution systems: A spatial and cost evaluation. *J. Water Resour. Plan. Manag. Div.*, **108** (3), 243-256; ASCE.
- Costa Martins, A. D. da. (2011) Stochastic models for prediction of pipe failures in water supply systems. *Inst. Super. Tec. Lisboa, Port.*
- Cox, D. R.; Isham, V. (1980) *Point processes*, Vol. 12; CRC Press.
- Dridi, L.; Mailhot, A.; Parizeau, M.; Villeneuve, J.-P. (2009) Multiobjective approach for pipe replacement based on bayesian inference of break model parameters. *J. Water Resour. Plan. Manag.*, **135** (5), 344-354; American Society of Civil Engineers.
- Ebeling, C. E. (2010) *An Introduction to Reliability and Maintainability Engineering*; Waveland Press. Recuperado a partir de <https://books.google.es/books?id=bX0iQgAACAAJ>
- Eisenbeis, P. (1994) Modelisation statistique de la prevision des defaillances des conduites d'eau potable; Strasbourg 1.
- Eisenbeis; Gat, Y. Le; Laffréchine, K.; Gauffre, P. Le; König, A.; Rostum, J.; Tuhovcak, L.; et al. (2002) CARE-W WP2 Description and Validation of Technical Tools: D3-Models Description. *Rep. N°2*, **1**.
- Gat, Y. Le; Eisenbeis, P. (2000) Using maintenance records to forecast failures in water networks. *Urban Water*, **2** (3), 173-181; Elsevier.
- Heck, G. J. van. (2008) Asset management frameworks for (drinking water) infrastructures around the world: A (not exhaustive) overview of different initiatives and developments. *2008 First Int. Conf. Infrastruct. Syst. Serv. Build. Networks a Bright. Futur.*, 1-6; IEEE.
- ISO 55000. (2014) Asset Management; International Organisation for Standardization.
- Jafar, R.; Shahrour, I.; Juran, I. (2010) Application of Artificial Neural Networks (ANN) to model the failure of urban water mains. *Math. Comput. Model.*, **51** (9-10), 1170-1180; Elsevier.

- Kabir, G. (2016) Planning repair and replacement program for water mains: a Bayesian framework; University of British Columbia.
- Karney, B. W.; McInnis, D. A.; Jacob, P. (1994) 'Comprehensive hydraulic analysis of water distribution systems. *Proc., 2nd Int. Conf. Water Pipeline Syst.*, 425-434; BHR Group, Edinburgh, Scotland.
- Kimutai, E.; Betrie, G.; Brander, R.; Sadiq, R.; Tesfamariam, S. (2015) Comparison of Statistical Models for Predicting Pipe Failures: Illustrative Example with the City of Calgary Water Main Failure. *J. Pipeline Syst. Eng. Pract.*, **6** (4), 4015005. Recuperado a partir de <http://ascelibrary.org/doi/10.1061/%28ASCE%29PS.1949-1204.0000196>
- Kleiner, Y.; Rajani, B. (2010) I-WARP: Individual water main renewal planner. *Drink. Water Eng. Sci.*, **3** (1), 71-77; Copernicus GmbH.
- Kleiner; Rajani. (2001) Comprehensive review of structural deterioration of water mains: statistical models. *Urban water*, **3** (3), 131-150; Elsevier.
- Kleiner; Rajani. (2002) Forecasting variations and trends in water-main breaks. *J. Infrastruct. Syst.*, **8** (4), 122-131; American Society of Civil Engineers.
- Lauer, W. C. (2013) *Water Distribution Operator Training Handbook*; American Water Works Association. Recuperado a partir de <https://books.google.es/books?id=yaZLnwEACAAJ>
- Levin, R. I.; Rubin, D. S.; Samaniego, A. H. F. (1996) *Estadística para administradores*; Prentice-Hall Hispanoamericana.
- Luís, R.; Maya, P. (1999) *Fundamentos de Inferencias Estadísticas*. (E. AC, Ed.); Madrid.
- Mailhot, A.; Poulin, A.; Villeneuve, J. (2003) Optimal replacement of water pipes. *Water Resour. Res.*, **39** (5); Wiley Online Library.
- Martínez, P. G. (2017) Análisis de variables explicativas en modelos de predicción de roturas en redes de tuberías. Recuperado a partir de <http://oa.upm.es/id/eprint/47857>

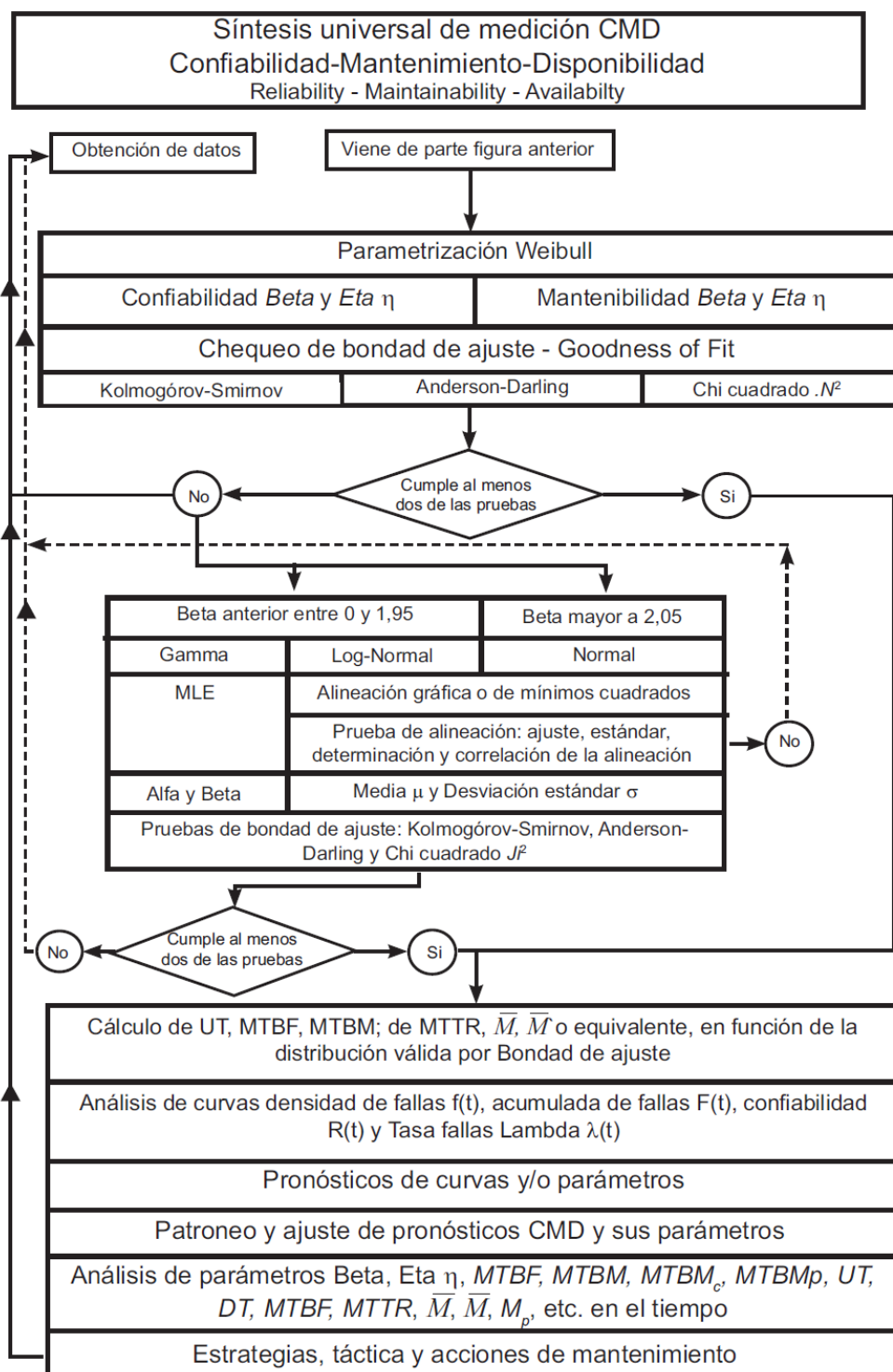
- Mogollón, D. J. (2017) *Estudio de la probabilidad de rotura de tuberías en abastecimientos urbanos con la ayuda de modelos estadísticos e hidráulicos. Estrategias de renovación de activos.* (M. Á. Jiménez Bello & U. P. de V. D. de I. H. y M. A.-D. d'Enginyeria H. i M. Ambient, Eds.).
- Mora, L. A. (2009) *Mantenimiento planeación, ejecución y control*, Primera Ed.
- Park, S.; Jun, H.; Agbenowosi, N.; Kim, B. J.; Lim, K. (2011) The proportional hazards modeling of water main failure data incorporating the time-dependent effects of covariates. *Water Resour. Manag.*, **25** (1), 1-19; Springer.
- Røstum, J. (2000) Statistical modelling of pipe failures in water networks; Fakultet for ingeniørvitenskap og teknologi.
- Saegrov, S.; Schilling, W.; Røstum, J.; Tuhovcak, L.; Eisenbeis, P.; Herz, R.; LeGauffre, P.; et al. (2003) Computer-aided rehabilitation of water networks (CARE-W). *Water Sci. Technol. Water Supply*, **3** (1-2), 19-27; IWA Publishing.
- Shamir, U.; Howard, C. D. D. (1979) An analytic approach to scheduling pipe replacement. *Journal-American Water Work. Assoc.*, **71** (5), 248-258; Wiley Online Library.
- Solano, H. (2008, noviembre 19) *Análisis de supervivencia en fiabilidad. Predicción en condiciones de alta censura y truncamiento: el caso de las redes de suministro de agua potable.* Riunet; Universitat Politècnica de València, Valencia (Spain). Recuperado a partir de <https://riunet.upv.es/handle/10251/3792>
- Therneau, T. M.; Lumley, T. (2015) Package 'survival'. *R Top Doc*, **128**.
- Walski, T. M.; Pelliccia, A. (1982) Economic analysis of water main breaks. *Journal-American Water Work. Assoc.*, **74** (3), 140-147; Wiley Online Library.
- Watson, T. G.; Christian, C. D.; Mason, A. J.; Smith, M. H.; Meyer, R. (2004) Bayesian-based pipe failure model. *J. Hydroinformatics*, **6** (4), 259-264; IWA Publishing.
- Yan, J. M.; Vairavamoorthy, K. (2003) Fuzzy approach for pipe condition assessment. *Pipeline Eng. Constr. Int. Conf. 2003 American Soc. Civ. Eng.*

ANEXOS

Anexo 1



Continúa Anexo 1,



Anexo 2

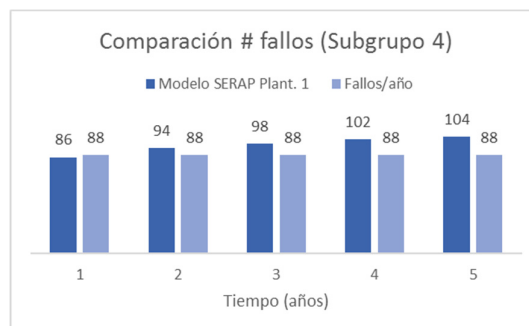
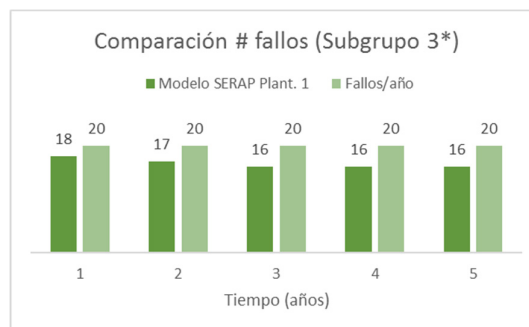
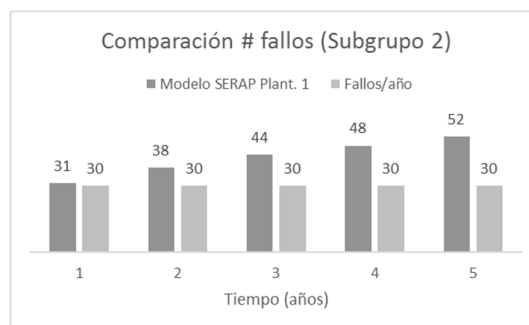
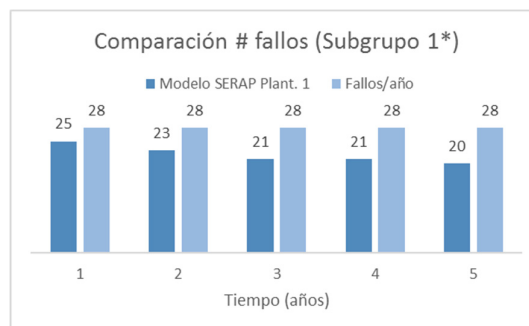
Resultados de Variable TBF

Planteamiento 1

Subgr.	Fallos/año
1*	28
2	30
3*	20
4	88

Modelo SERAP - Planteamiento 1

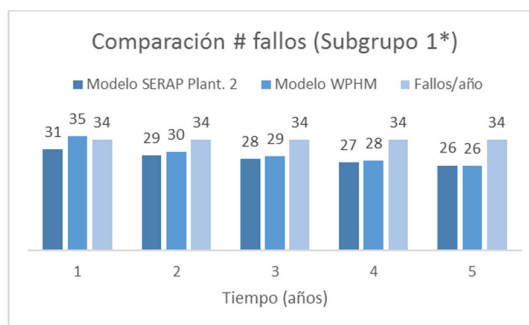
Subg.	Número de fallos				
	Año 1	Año 2	Año 3	Año 4	Año 5
1*	25	23	21	21	20
2	31	38	44	48	52
3*	18	17	16	16	16
4	86	94	98	102	104



Continúa Anexo 2,

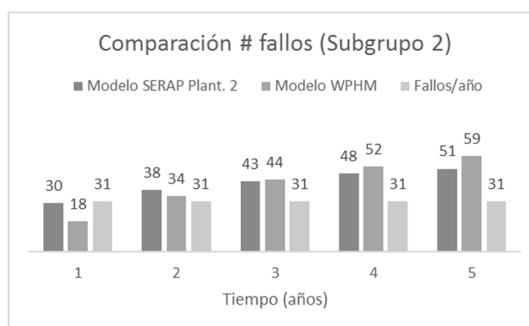
Resultados de Variable TBF
 Planteamiento 2

Subgr.	Fallos/año
1*	34
2	31
3*	28
4	122



Modelo SERAP - Planteamiento 2

Subg.	Número de fallos				
	Año 1	Año 2	Año 3	Año 4	Año 5
1*	31	29	28	27	26
2	30	38	43	48	51
3*	26	27	28	28	29
4	125	142	153	161	167



Modelo WPHM

Subgrupos	Número de fallos				
	Año 1	Año 2	Año 3	Año 4	Año 5
1*	35	30	29	28	26
2	18	34	44	52	59
3*	35	30	30	31	34
4	113	130	138	144	147

