

UNIVERSITAT POLITÈCNICA DE VALÈNCIA
DEPARTAMENTO DE INGENIERÍA HIDRÁULICA Y MEDIO
AMBIENTE



**METODOLOGÍAS EPIDEMIOLÓGICAS DE ANÁLISIS DE
DATOS PARA LA OPERACIÓN Y GESTIÓN DE REDES DE
ABASTECIMIENTO URBANO DE AGUA**

TESIS DOCTORAL

PRESENTADA POR

CLAUDIA FERNANDA NAVARRETE LÓPEZ

DIRIGIDA POR

DR. MANUEL HERRERA
PROF. JOAQUÍN IZQUIERDO
DR. BRUNO BRENTAN

VALENCIA, ESPAÑA, JUNIO 2023

DEDICATORIA

*A mis padres por haberme forjado
como la persona que soy en la
actualidad; muchos de mis logros se los
debo a ustedes, su amor e
incondicionalidad siempre me han
motivado para ser mejor cada día y
siempre alcanzar mis metas.*

*A Gabriel José, Miguel Ángel y
Alejandro, mis hombres, quienes son mi
mayor tesoro y mi fuente de motivación
y alegría.*



AGRADECIMIENTOS

Doy gracias a Dios quien ha sido mi guía, mi fortaleza, porque me ha sostenido y me ha permitido llegar hasta el final. Gracias porque dispuso de maravillosos ángeles en mi camino para que me impulsaran y acompañaran en este andar.

A mis directores Dr. Joaquín Izquierdo, Dr. Bruno Melo Brentan y Dr. Antonio Manuel Herrera por todo el apoyo brindado durante el desarrollo de la presente investigación, por su dedicación, acompañamiento e infinita paciencia durante todo el proceso, así como por su constante asesoría académica y profesional.

Al profesor Dr. Rafael Pérez García (†) quien me brindó la oportunidad de iniciar en este proceso doctoral.

A la Universidad Santo Tomás, por la confianza que me ha otorgado durante los años que he estado vinculada como docente en la institución, así como por la oportunidad de hacerme partícipe de la Convocatoria 001 de estudios posgraduales de la Vicerrectoría Académica en convenio con la Universitat Politècnica de València en el Doctorado de Ingeniería del Agua y Medio ambiental.

A la Universitat Politècnica de València, por su atención constante para el buen desarrollo de mis estudios de Doctorado de forma exitosa.

Quiero agradecer a mis compañeros Nidia Isabel Molina, Darwin Mena y Karina Solano, con quienes inicié este proceso doctoral, con quienes nos acompañamos y compartimos experiencias, gracias por su apoyo e impulso siempre.

A mis amigos la Dra. Angela María Jaramillo Londoño y el Dr. José Luis Díaz Arévalo, ya que con su amistad siempre han estado dispuestos a darme la mano y han sido luz y guía desde que me propuse este reto académico.

Agradezco especialmente a Dayam Soret Calderón y José Fabricio Rueda por su acompañamiento en los aprendizajes que se derivaron de este trabajo doctoral, por su discusiones y aportes que enriquecieron muy valiosamente este proceso.

A mi esposo y mis hijos, que con su paciencia y amor siempre han creído en mí y me han apoyado para que alcance mis metas. Gracias por estar siempre, son mi fuerza y mi razón de ser.

Quiero agradecer a mis padres, a mis hermanas, a Víctor Hugo, por su acompañamiento incondicional; en todo mi camino han sido mi gran compañía, gracias por estar siempre cuando más los necesito.

A mi familia y amigos, quienes siempre me han acompañado en el logro de todo lo que me he propuesto.

¡¡¡Gracias totales!!!

RESUMEN

La definición que entrega el diccionario de la Real Academia Española de la Lengua sobre *Epidemiología* indica que ésta es una ciencia y como tal, tiene aquellos elementos propios de un conjunto sistematizado de conocimientos entre los que se destaca la metodología, aquella que analiza los procedimientos usados en el objeto de estudio.

Desde el surgimiento del ser humano se ha podido evidenciar, a lo largo de la historia, cómo se ha hecho uso del agua en el abastecimiento tanto a nivel de sustento y salud, como en el nivel industrial. Igualmente, se puede evidenciar la evolución de la epidemiología para analizar las enfermedades que principalmente se transmiten por aguas contaminadas y cómo la metodología ha hecho avances tan significativos que permiten predecir el comportamiento de los patógenos y mitigar las consecuencias de un posible contagio.

El desarrollo de la epidemiología se ha volcado principalmente en el área de la medicina en la cual ha mostrado una enorme evolución al afrontar grandes retos como la propagación de las enfermedades infecciosas y el replanteamiento continuo de los modelos de análisis. No obstante, esta ciencia se puede adaptar a cualquier área del conocimiento humano como la gestión de los recursos hídricos y más concretamente en la gestión de las redes de abastecimiento de agua urbana.

La puesta en marcha de una red de abastecimiento de agua en una ciudad, cuyas dimensiones y construcción generalmente son monumentales, implica un diseño y una operabilidad que surge de la aplicación de modelos matemáticos y/o estadísticos, los cuales permiten analizar las distintas condiciones de funcionamiento antes de iniciar obras. Ese comportamiento puede caracterizarse a partir de métodos de resolución basados en los procedimientos epidemiológicos y que han sido contrastados ampliamente en forma empírica y funcional.

En toda red de suministro existen dos componentes independientes e interdependientes, como lo son la gestión de la demanda y la gestión de fallos. En ambos hay incertidumbres que, generalmente, provienen de variables externas, aleatorias, que dificultan su cuantificación y por lo mismo, su predicción. Para la gestión de la demanda, resulta importante la aplicación de modelos de estimación de la demanda precisos, pues con ellos se pueden determinar las capacidades y cargas que soporte la red. A la par, para la gestión de fallos en las redes, resultan importantes modelos de estimación precisos que ayuden a mitigar el impacto de contingencias generadas por fallos en cascada y la propagación de éstos hasta un posible colapso.

En los procesos de gestión de demanda se vienen utilizando principalmente los modelos de series temporales, llegando a la aplicación de modelos que impliquen el algoritmo SAX. En los procesos de gestión de fallos se han aplicado métodos como el análisis de supervivencia y más recientemente, las redes neuronales, llegando a los sistemas multiagente con los modelos SIR, SIRS y SEIR.

El desarrollo de los modelos SAX se pueden apreciar en un caso de estudio de la ciudad de Franca en Brasil, en la que se combinan patrones de similitud entre sectores con patrones de las MINDIST que respaldan los métodos predictivos, mejorando su precisión y facilitando la detección de lecturas anormales en los medidores de flujo e incluso la presencia de usos o fugas inesperados.

El Modelo Basado en Agentes (MBA) se puede desarrollar mediante, por ejemplo, la herramienta NetLogo, y su aplicación en una red de suministro resulta muy efectiva para determinar el comportamiento de los posibles fallos en cascada; ejemplo de ello se aprecia en el caso de estudio de la ciudad de Coro en Venezuela en la que se pueden establecer momentos para cada comportamiento: susceptibilidad, infección (fallo) y recuperación; proporcionando así un modelo predictivo mejorado para este tipo de situaciones.

ABSTRACT

The definition of Epidemiology given by the dictionary of the *Real Academia Española de la Lengua* states that it is a science and, as such, it exhibits those elements of a systematized set of knowledge, among which the methodology, which analyzes the procedures used in the object of study, stands out.

Since the emergence of the human being, it has been possible to demonstrate throughout history, how water has been used in supply at the level of both sustenance and health, as well as at the industrial level. Likewise, the evolution of epidemiology can be evidenced to analyze diseases that are mainly transmitted by contaminated water and how the methodology has made such significant advances that allow predicting the behavior of pathogens and mitigating the consequences of possible contagion.

The development of epidemiology has focused mainly on the area of medicine, in which it has shown enormous evolution when facing great challenges such as the spread of infectious diseases and the continuous rethinking of analysis models. However, this science can be adapted to any area of human knowledge, such as the management of water resources and more specifically the management of urban water supply networks.

Implementing a water supply network in a city, whose dimensions and construction are generally monumental, implies design and operability aspects that arise from applying mathematical and/or statistical models, which allow the analysis of the different conditions of operation before starting its operation. This behavior can be characterized by resolution methods based on epidemiological procedures, which have been widely contrasted empirically and functionally.

In any supply network, there are two independent and interdependent components, such as demand management and fault management. In both, there are uncertainties that generally come from external, random variables, which make it difficult to

quantify and therefore to predict. Regarding demand management, the application of accurate demand estimation models is essential, since the capacities and loads supported by the network can be determined by using them. At the same time, for network failure management, accurate estimation models are also key to help mitigate the impact of contingencies generated by cascading failures and their propagation until a possible collapse.

In demand management processes, time series models have been mainly used, including the application of models that involve the SAX algorithm. In failure management processes, methods such as survival analysis and, more recently, neural networks have been applied; reaching multi-agent systems with the SIR, SIRS, and SEIR models.

The development of SAX models can be seen in a case study from the city of Franca in Brazil, in which patterns of similarity between districts are combined with MINDIST patterns that support predictive methods, improving their accuracy and facilitating the detection of errors, abnormal readings on flow meters and even the presence of unexpected uses or leaks.

The Agent-Based Model (MBA) can be developed using, for example, the NetLogo tool, and its application in a supply network is very effective in determining the behavior of possible cascading failures. An example can be seen in the case study of the city de Coro in Venezuela in which moments for each behavior can be established: susceptibility, infection (failure), and recovery. This provides an improved predictive model for this type of situation.

RESUM

La definició que lliura el diccionari de la Reial Acadèmia Espanyola de la Llengua sobre l'Epidemiologia indica que aquesta és una ciència i com a tal té aquells elements propis d'un conjunt sistematitzat de coneixements entre els quals es destaca la metodologia, aquella que analitza els procediments usats en l'objecte d'estudi.

Des del sorgiment de l'ésser humà s'ha pogut evidenciar al llarg de la història, com s'ha fet ús de l'aigua en l'abastament tant pel que fa al suport i la salut, com al nivell industrial. Igualment, es pot evidenciar l'evolució de l'epidemiologia per analitzar les malalties que principalment es transmeten per aigües contaminades i com la metodologia ha fet avanços tan significatius que permeten predir el comportament dels patògens i mitigar les conseqüències d'un possible contagi.

El desenvolupament de l'epidemiologia s'ha abocat principalment a l'àrea de la medicina on ha mostrat una enorme evolució en afrontar grans reptes com la propagació de les malalties infeccioses i el replantejament continu dels models d'anàlisi. Això no obstant, aquesta ciència es pot adaptar a qualsevol àrea del coneixement humà com la gestió dels recursos hídrics i més concretament en la gestió de les xarxes d'abastament d'aigua urbana.

La posada en marxa d'una xarxa d'abastament d'aigua en una ciutat, les dimensions i la construcció de la qual generalment són monumentals, implica un disseny i una operativitat que sorgeix de l'aplicació de models matemàtics i/o estadístics, els quals permeten analitzar les diferents condicions de funcionament abans d'iniciar obres. Aquest comportament es pot caracteritzar a partir de mètodes de resolució basats en els procediments epidemiològics i que han estat contrastats àmpliament en forma empírica i funcional.

En tota xarxa de subministrament hi ha dos components independents i interdependents, com ho són la gestió de la demanda i la gestió de fallades. En tots dos hi ha incerteses que generalment provenen de variables externes, aleatòries, que dificulten la quantificació i, per tant, la seva predicció. Per a la gestió de la demanda, és important l'aplicació de models d'estimació de la demanda precisos, ja que s'hi poden determinar les capacitats i les càrregues que suporti la xarxa. Alhora, per a la gestió d'errors a les xarxes, són importants els models d'estimació precisos que ajudin a mitigar l'impacte de contingències generades per errors en cascada i la propagació d'aquests fins a un possible col·lapse.

En els processos de gestió de demanda, s'utilitzen principalment els models de sèries temporals, arribant a l'aplicació de models que impliquin l'algorisme SAX. En els processos de gestió de fallades s'han aplicat mètodes com l'anàlisi de supervivència i, més recentment, les xarxes neuronals; arribant als sistemes multiagent amb els models SIR, SIRS i SEIR.

El desenvolupament dels models SAX es poden apreciar en un cas d'estudi de la ciutat de Franca al Brasil, en què es combinen patrons de similitud entre sectors amb patrons de MINDIST que donen suport als mètodes predictius millorant-ne la precisió i facilitant la detecció de lectures anormals als mesuradors de flux i fins i tot la presència d'usos o fugues inesperades.

El Model Basat en Agents (MBA) es pot desenvolupar mitjançant l'eina NetLogo i la seva aplicació en una xarxa de subministrament resulta molt efectiva per determinar el comportament de les possibles fallades en cascada. Exemple d'això s'aprecia en el cas d'estudi de la ciutat de Coro a Veneçuela en què es poden establir moments per a cada comportament: susceptibilitat, infecció (fallida) i recuperació, proporcionant així un model predictiu millorat per a aquest tipus de situacions.

TABLA DE CONTENIDO

DEDICATORIA	i
AGRADECIMIENTOS	ii
RESUMEN	iii
ABSTRACT	v
RESUM	vii
TABLA DE CONTENIDO	ix
TABLA DE ILUSTRACIONES	xiii
ÍNDICE DE TABLAS	xvii
LISTA DE ABREVIATURAS	xviii
<i>CAPÍTULO 1</i>	20
<i>INTRODUCCIÓN</i>	20
1.1 PRELIMINARES	20
1.1.1 Concepto de Epidemiología	20
1.1.2 Una herramienta de la epidemiología	22
1.1.3 Metodologías de investigación en epidemiología: aplicaciones en ingeniería hidráulica	23
1.1.4 Nuevos desarrollos y regulaciones en cuanto al agua	26
1.1.5 Concepto de agua potable y redes de abastecimiento	27
1.1.6 Gestión inteligente de las redes de abastecimiento	30
1.1.6.1. Gestión de demanda de agua	32
1.1.6.2. Gestión de fallos	35
1.2 JUSTIFICACIÓN	40
1.3 OBJETIVOS	45
1.3.1 Objetivo General	45
1.3.2 Objetivos Específicos	45
1.4 BREVE DESCRIPCIÓN DE ESTE DOCUMENTO	46
<i>CAPÍTULO 2</i>	49
<i>REVISIÓN BIBLIOGRÁFICA</i>	49

2.1 RESEÑA HISTÓRICA	49
2.1.1 Epidemiología	49
2.1.2 Las redes de distribución de agua	56
2.2 ESTADO DEL ARTE	67
2.2.1 El papel de la Epidemiología y sus herramientas de análisis de datos	67
2.2.1.1. Métodos heredados de la epidemiología	73
2.2.1.2. La pandemia de COVID-19	75
2.2.1.3. Epidemiología de la energía	78
2.2.2 El papel del agua urbana y la gestión del recurso hídrico	79
2.2.2.1. El agua urbana dentro de los procesos epidemiológicos	86
2.2.3 La gestión de fallos en redes de distribución de agua	88
<i>CAPÍTULO 3</i>	94
<i>GESTIÓN DE DEMANDA</i>	94
3.1 DEMANDA DE AGUA URBANA	94
3.2 GESTIÓN DE DEMANDA DE AGUA	96
3.3 PATRONES DE CONSUMO PARA LA DEMANDA DE AGUA	98
3.4 MODELOS PARA LA ESTIMACIÓN DE LA DEMANDA DE AGUA	100
3.4.1 Desarrollo de modelos de previsión para la demanda.	101
3.4.1.1. Interoperabilidad y explicabilidad de los modelos	106
3.4.2 Análisis de datos basados en epidemiología para la demanda de agua	108
3.5 ALGORITMO SAX	110
3.5.1 Enfoque simbólico del algoritmo SAX.	113
3.5.1.1. Reducción de dimensionalidad vía PAA.	114
3.5.1.2. Discretización.	115
3.5.1.3. Medida de distancia.	116
3.5.1.4. Reconocimiento de patrones en palabras SAX.	117
3.6 ARBOLES DE SUFIJOS Y ALGORITMO SAX EN DIVERSOS CONTEXTOS	120
3.7 ALGORITMO SAX PARA EL ANÁLISIS DE LA DEMANDA DE AGUA	122
<i>CAPÍTULO 4</i>	124

<i>GESTIÓN DE FALLOS</i>	124
4.1 CONCEPTO DE FALLOS EN REDES DE ABASTECIMIENTO DE AGUA	124
4.2 PRINCIPALES CAUSAS DE FALLOS	126
4.2.1 Síntesis del proceso de gestión de activos.	127
4.3 LA IMPORTANCIA DE LA PREDICCIÓN DE FALLOS EN LOS SISTEMAS DE ABASTECIMIENTO DE AGUA	130
4.3.1 La incertidumbre en la predicción de fallos.	132
4.3.2 Tasa de fallos y frecuencia en la predicción de fallos.	134
4.4 ALGUNOS MÉTODOS FAVORABLES EN LA ESTIMACIÓN DE FALLOS	135
4.4.1 El Análisis de Supervivencia.	135
4.4.2 Redes Neuronales.	138
4.5 FALLOS EN CASCADAS	146
4.5.1 Los desafíos de la simulación de fallos en cascada.	148
4.5.2 El Análisis del árbol de fallos.	149
4.5.3 Modelo de IO de inoperabilidad dinámica: efectos de cascada.	151
4.6 SISTEMA MULTIAGENTE PARA LA SIMULACIÓN DE EPIDEMIAS	155
4.6.1 Concepto de sistemas multiagente y de redes complejas.	156
4.6.2 Redes complejas multiagente.	158
4.6.3 Modelos epidemiológicos.	158
4.6.3.1. El modelo SI.	159
4.6.3.2. El modelo SIS.	159
4.6.3.3. El modelo SIR.	159
4.6.3.4. El modelo SIRS.	159
4.6.3.5. El modelo SEIR.	160
4.6.4 Ejemplo de simulación con sistema multiagente.	160
4.7 APLICACIÓN DE LOS SISTEMAS MULTIAGENTE PARA LA ESTIMACIÓN DE FALLOS EN CASCADA	166
4.7.1 Riesgo y confiabilidad de los sistemas de distribución de agua.	169
4.7.2 Epidemiología para modelar fallos en sistemas de distribución.	172
4.7.2.1. Modelos compartimentales en epidemiología.	172
4.7.2.2. Epidemiología en redes.	173

<i>CAPÍTULO 5</i>	175
<i>CASO DE ESTUDIO PARA GESTIÓN DE DEMANDA</i>	175
5.1 LA CIUDAD DE FRANCA, SAO PAULO, BRASIL	175
5.2 DESCRIPCIÓN DEL CASO DE ESTUDIO	179
5.3 RESULTADOS	180
<i>CAPÍTULO 6</i>	187
<i>CASO DE ESTUDIO PARA GESTIÓN DE FALLOS</i>	187
6.1 LA CIUDAD DE CORO, ESTADO DE FALCON, VENEZUELA	187
6.2 LA HERRAMIENTA DE MODELADO NETLOGO	189
6.3 MODELACIÓN BASADA EN AGENTES CON NETLOGO	190
6.4 APLICACIÓN DEL MODELADO DE NETLOGO A LA RED DE CORO	191
6.5 ANÁLISIS DE RESULTADOS Y DISCUSIÓN	197
<i>CAPÍTULO 7</i>	206
<i>CONCLUSIONES</i>	206
7.1 METODOLOGÍA PREVENTIVA	206
7.1.1 Consideraciones previas.	206
7.1.2 Orientación metodológica.	208
7.2 CONCLUSIONES Y LÍNEAS FUTURAS	211
REFERENCIAS	213
ANEXOS	230
Anexo 1	230
Código fuente del modelo SIR para el caso de estudio de la red de Coro.	230
Anexo 2	237
Tablas de resultados de las simulaciones en el caso de estudio de la red de Coro.	237

TABLA DE ILUSTRACIONES

Ilustración 1. Algunos tipos de grafos.....	22
Ilustración 2. Ciclo integral de Agua.....	27
Ilustración 3. Concepto global de Smart Cities.....	30
Ilustración 4. Partes interesadas y sus intereses relacionados con las aguas subterráneas. Big Data Analytics.....	31
Ilustración 5. Clasificación de los problemas de gestión en las WDNs.....	39
Ilustración 6. Ejemplo de interface del modelo HI-LIFE (NetLogo) y proyección demográfica de agentes (MBA).....	42
Ilustración 7. Mapa Conceptual del documento.....	46
Ilustración 8. Diagrama de flujo de la organización del documento.....	47
Ilustración 9. Portada Epidemiorum. Imagen de Leaf & Stone Books.....	51
Ilustración 10. John Snow en 1856, Cortesía de Renato Sabbatini.....	53
Ilustración 11 Mapa original de John Snow que muestra los grupos de casos de cólera en el brote de Broad Street.....	54
Ilustración 12. Mapa de Nippur.....	57
Ilustración 13 Una sección de probablemente 500 metros de largo del Aqua Appia en los barrios de Centocelle.....	61
Ilustración 14 Tornillo de Arquímedes.....	62
Ilustración 15. Sistemas de agua en Palenque.....	64
Ilustración 16. Templo de Angkor Wat.....	65
Ilustración 17. Métodos epidemiológicos para análisis de datos.....	70
Ilustración 18. Fuerza de asociación de elementos relevantes para métodos epidemiológicos de análisis de datos.....	72
Ilustración 19. Comportamiento de producción para métodos de análisis de datos.....	72
Ilustración 20. Métodos heredados de Epidemiología.....	75
Ilustración 21. Volumen de publicaciones COVID.....	77
Ilustración 22. Países que más publican sobre COVID.....	77
Ilustración 23. Gestión del recurso hídrico.....	82
Ilustración 24. Fuerza de asociación de elementos relevantes para gestión del recurso hídrico.....	84
Ilustración 25. Comportamiento de producción para la gestión del recurso hídrico.....	85
Ilustración 26. Clústeres de Chlorination disinfection by products in municipal drinking water – A review.....	87
Ilustración 27. Gestión de fallos.....	90

Ilustración 28. Fuerza de asociación de elementos relevantes para gestión de fallos.....	91
Ilustración 29. Comportamiento de producción para gestión de fallos.	92
Ilustración 30. Ejemplo de las variaciones (%) para los años 2000 y 2007 en cada distrito de Mitilene de la tasa anual de consumo de agua (a) y el número de conexiones de agua activas (b)	95
Ilustración 31. Curva de modulación para un área residencial.....	98
Ilustración 32. Uso típico diario de agua durante el verano en la zona del centro oeste de Ottawa, (MXDY día de mayor demanda para todo el registro de 1994–2002; promedio MXDY promedio demanda máxima para todo el registro.	102
Ilustración 33. Cálculo de un perfil de distancia a partir de la segunda ventana.	105
Ilustración 34. Clasificación de explicabilidad de modelos.....	107
Ilustración 35. Aplicaciones del algoritmo SAX.....	110
Ilustración 36. (a) Esquema de discretización del componente temporal. (b) Discretización simbólica de series de tiempo	111
Ilustración 37. Ejemplo árbol de sufijos.	112
Ilustración 38. Cadena de Markov en tiempo continuo.....	112
Ilustración 39. La representación PAA se puede visualizar como un intento de modelar una serie de tiempo con una combinación lineal de funciones de base de caja. En este caso, una secuencia de longitud 128 se reduce a ocho dimensiones.....	114
Ilustración 40. Una serie de tiempo se discretiza obteniendo primero una aproximación PAA y luego usando puntos de corte predeterminados para mapear los coeficientes PAA en símbolos SAX. En el ejemplo anterior, con $n = 128$, $w = 8$ y $a = 3$, la serie temporal se asigna a la palabra baabccbc	116
Ilustración 41. Ejemplo de proceso de extracciones en árbol de sufijos.	119
Ilustración 42. Diagrama de flujo del proceso general para mejorar los modelos predictivos de demanda de DMA.	123
Ilustración 43. Ejemplo de Red de Distribución de agua en Madrid en un sector de 1000 nodos.	125
Ilustración 44. Principales causas de fallos.....	130
Ilustración 45. Ejemplo de actualización del modelo bayesiano.....	133
Ilustración 46. Maqueta EPANET de un sistema de agua.	139
Ilustración 47. Una capa convolucional de gráfico para un gráfico con 4 nodos (N=4).....	145
Ilustración 48. Relación de interdependencia entre diferentes infraestructuras.	147
Ilustración 49. Ejemplo de fallos en cascada en una red.....	149

Ilustración 50. Representación gráfica de un árbol de fallos.....	150
Ilustración 51. Metodología típica en fallos en cascada.....	151
Ilustración 52. Diagrama del modelo dinámico de entrada-salida.....	154
Ilustración 53. Ejemplos de redes complejas.....	157
Ilustración 54. Modelos compartimentales.....	160
Ilustración 55. Transmisión espaciotemporal epidémica basada en modelos SWN y MAS.....	166
Ilustración 56. Vista aérea de Franca.....	176
Ilustración 57. IBGE - Pesquisa Nacional de Saneamento Básico - Abastecimento de água e Esgotamento sanitário 2017 a.....	177
Ilustración 58. IBGE - Pesquisa Nacional de Saneamento Básico - Abastecimento de água e Esgotamento sanitário 2017 b.....	177
Ilustración 59. Perímetro Urbano de Franca.....	178
Ilustración 60. Diseño DMA del WDS de Franca.....	180
Ilustración 61. Número de segmentos para la configuración PAA; (a) SS vs. número de segmentos; (b) SS acumulada vs. número de segmentos.....	181
Ilustración 62. Sensibilidad de las reducciones de PAA. Caso de DMA de SL. (a) SL DMA: valores originales de demanda de agua y PAA calculados a través de 50 segmentos; (b) SL DMA: valores originales de demanda de agua y PAA calculados a través de 100 segmentos; (c) SL DMA: valores originales de demanda de agua y PAA calculados a través de 200 segmentos; (d) SL DMA: valores originales de demanda de agua y PAA calculados a través de 400 segmentos.....	182
Ilustración 63.. Box-plots de los errores asociados por configuración de ANN.....	185
Ilustración 64. Red de abastecimiento de Coro.....	188
Ilustración 65. Ejemplo de interfaz de Netlogo.....	190
Ilustración 66. Interfaz de trabajo para el modelo MBA.....	192
Ilustración 67. Pestaña 1 de la interfaz.....	192
Ilustración 68. Pestaña 2 de la interfaz.....	193
Ilustración 69. Pestaña 3 de la interfaz.....	193
Ilustración 70. Area de trabajo de la interfaz.....	194
Ilustración 71. Comandos de ejecución para el modelo.....	195
Ilustración 72. Comportamiento de nodos susceptibles.....	198
Ilustración 73. Comportamiento de los nodos infectados (con fallo).....	199
Ilustración 74. Comportamiento de los nodos recuperados (con fallo).....	201
Ilustración 75. Combinación entre infectados y recuperados (true).....	201
Ilustración 76. Resumen de nodos susceptibles.....	203
Ilustración 77. Resumen de nodos recuperados.....	204

Ilustración 78. Resumen de nodos infectados (fallos)	204
Ilustración 79. Mapas de calor de la Red de Coro.....	205
Ilustración 80. Orientación metodológica.....	208
Ilustración 81. Propuesta metodológica.....	209

ÍNDICE DE TABLAS

Tabla 1. Categorías en las que se aplican métodos epidemiológicos de análisis de datos.....	71
Tabla 2. Categorías en las que se aplica la gestión del recurso hídrico	83
Tabla 3. Categorías en las que se aplica la gestión de fallos	91
Tabla 4. Visión general para la construcción de árbol de sufijos	118
Tabla 5. Principales herramientas involucradas en la gestión de activos.....	129
Tabla 6. MINDIST para los 4 DMA del caso de estudio.....	183
Tabla 7. Error medio de modelos predictivos utilizando DMA SL y DMA combinación SL + SA-3. Comparación por la prueba t de Student para igualdad de medias.....	184
Tabla 8. Error medio de predicción (configuración superior de ANN) para DMA SL y para la combinación de DMA SL + SA-3.	186
Tabla 9. Resultados simulación: estado Susceptibles (true)	197
Tabla 10. Resultados simulación: estado Infectados (true).....	198
Tabla 11. Resultados simulación: estado Recuperados (true)	200

LISTA DE ABREVIATURAS

ANFIS	<i>Adaptive Neuro-Fuzzy Inference System</i>
ANN	<i>Artificial Neural Networks</i>
AR	<i>Auto Regressive</i>
ARMA	<i>Auto Regressive Moving Average</i>
BSA	<i>Backtracking Search Algorithm</i>
BWPHM	<i>Bayesian Weibull Proportional Hazards Model</i>
CNN	<i>Convolutional Neural Networks</i>
Cox-PHM	<i>Cox-Proportional Hazard Model</i>
CSA	<i>Crow Search Algorithm</i>
DIIM	<i>Dynamic Interoperability Input Model</i>
DMA	<i>District Measurement Area</i>
DWT	<i>Discrete Wavelet Transform</i>
EDAA	<i>Epidemiological Data Analysis Approaches</i>
ELM	<i>Extreme Learning Machines</i>
GCNN	<i>Graph Convolutional Neural Networks</i>
GFT	<i>Graph Fourier Transform</i>
GIS	<i>Geographic Information Systems</i>
GMAO	<i>Gestión de Mantenimiento Asistido por Ordenador</i>
IA	<i>Inteligencia Artificial</i>
IIM	<i>Interoperability Input Model</i>
IIS	<i>Industrial Information Systems</i>
IOM	<i>Input-Output Model</i>
IoT	<i>Internet of Things</i>
IWA	<i>International Water Association</i>
KDD	<i>Knowledge Discovery in Databases</i>
LTP	<i>Long Term Problem</i>
MA	<i>Moving Average</i>
MARS	<i>Multivariate Adaptive Regression Splines</i>
MAS	<i>Multi-Agent System</i>
MBA	<i>Modelo Basado en Agentes</i>
MCDSS	<i>Multi Criteria Decision Support Systems</i>
ML	<i>Machine Learning</i>
MLP	<i>MultiLayer Perceptron</i>
MMR	<i>Mean Motion Resonances</i>
MTP	<i>Medium Term Problem</i>

NHPP	<i>Non-Homogeneous Poisson Process</i>
PAA	<i>Piecewise Aggregate Approximation</i>
PCA	<i>Principal Component Analysis</i>
PPR	<i>Projection Pursuit Regression</i>
RAE	<i>Real Academia Española de la lengua</i>
RF	<i>Random Forests</i>
RNCs	<i>Redes Neuronales Computacionales</i>
SARIMA	<i>Series Auto Regressive Integrated Moving Average</i>
SAX	<i>Symbolic Aggregate approxImation</i>
SCADA	<i>Supervisory Control And Data Acquisition</i>
SGBD	<i>Sistemas de Gestión de Bases de Datos</i>
SIF	<i>Sistema de Información Financiera</i>
SIR	<i>Susceptible-Infected-Recovered</i>
SIRS	<i>Susceptible-Infected-Recovered- Susceptible</i>
SEIR	<i>Susceptible-Exposed-Infected-Recovered</i>
STP	<i>Short Term Problem</i>
SVM	<i>Support Vector Machines</i>
SVR	<i>Support Vector Regression</i>
SWN	<i>Small World Network</i>
UFW	<i>Unaccounted For Water</i>
WDNs	<i>Water Distribution Networks</i>
WDS	<i>Water Distribution System</i>
WNTR	<i>Water Network Tool for Resilience</i>
WPHM	<i>Weibull Proportional Hazards Model</i>

CAPÍTULO 1

INTRODUCCIÓN

“Este capítulo trata brevemente todos los conceptos y los contenidos que se desarrollan en esta investigación: la Epidemiología, los métodos de investigación en epidemiología. La aplicabilidad de dichos conceptos dentro del marco del Agua Potable, su regulación, desarrollos metodológicos y, especialmente, los desarrollos en cuanto a las redes de distribución y/o suministro de agua, aclarando dos aspectos importantes: la gestión de demanda y la gestión de fallos. Adicionalmente, se presenta la justificación y los objetivos de esta investigación, para cerrar el capítulo con una breve descripción de este documento.”

1.1 PRELIMINARES

1.1.1 Concepto de Epidemiología

Hablar de *epidemiología* es tratar un concepto que busca analizar y poder encontrar soluciones a eventos colectivos sanitarios; su punto de partida está en el estudio de los contagios de enfermedades infecciosas, procurando describir cómo se distribuyen en las poblaciones y los factores que influyen en esa distribución dentro de un marco de tiempo determinado. Uno de los primeros referentes al estudio de enfermedades como un padecimiento colectivo, se remonta hacia el año 2000 a. C. con el papiro de Ebers¹, que menciona unas fiebres pestilentes (probablemente malaria) que asolaron a la población de las márgenes del Nilo (López-Moreno et al., 2000). La epidemiología está presente a través de la historia como un denominador común,

¹ Papiro de Ebers: antiguo documento médico que contiene más de 842 remedios para enfermedades y lesiones. Derivó su nombre del famoso egiptólogo Georg Ebers y se estima que fue creado entre 1550-1536 a. C. durante el reinado de Amenophis I.

inicialmente abordaba el estudio de las enfermedades y, actualmente, comprende además el estudio de los determinantes de la salud.

El vocablo epidemiología proviene del término griego *epidemios*. Su raíz *epi* significa sobre, *demikos* relativo a la gente y *logos*, estudio o tratado; etimológicamente significa el estudio de lo que sucede sobre la población (Blanco et al., 2001) y en el campo de la salud se refiere al estudio de la presencia y distribución de eventos relacionados con la sanidad en poblaciones específicas, incluyendo análisis de los factores que los afectan y, también, la posibilidad de aplicar ese raciocinio para controlar los problemas de salud (Porta et al., 2008).

El protagonismo de la epidemiología en el amplio y sinuoso campo de la medicina se ha caracterizado por afrontar grandes retos, romper paradigmas, replantear los modelos y trazar nuevas líneas de pensamiento. Es así que la epidemiología se reinventa y aporta significativamente a la aproximación de los diversos fenómenos en la medicina (Rodolfo et al., 2015). Precisamente, el cambio de esos paradigmas permite expandir los horizontes conceptuales y científicos de la epidemiología considerando simultáneamente los niveles molecular, humano, social y medioambiental, con modelos multinivel que permiten una mayor comprensión del complejo proceso salud-enfermedad (Garzón et al., 2010).

La observación en ese proceso se distribuye de una manera diferencial en cuanto a tipo y frecuencia según características biológicas, psicológicas y/o sociales de los individuos; influyen también variaciones geográficas o regionales y en ocasiones, se pueden apreciar tendencias en el tiempo claramente identificables. Esto lleva a analizar la situación de una manera integral, al planteamiento de posibles explicaciones; es decir, de hipótesis susceptibles de ser abordadas en los distintos tipos de diseños de estudios epidemiológicos (García et al., 1998).

La investigación en epidemiología combina conocimientos de las ciencias biológicas y sociales, además de considerar técnicas tanto cuantitativas como cualitativas para el

análisis de los datos. Las primeras son ampliamente reconocidas y usadas en la investigación tradicional; mientras que las segundas vienen ganando mayor influencia al ser más flexibles y apreciar enfoques socio-ambientales que responden a preguntas dentro de situaciones complejas. Los métodos cualitativos e interpretativos, además de ofrecer otro enfoque (que puede ser complemento de métodos cuantitativos), permiten mejorar la comprensión de los fenómenos sociales en investigación epidemiológica y están presentes en diversos campos de estudio (Silva et al., 2012).

1.1.2 Una herramienta de la epidemiología

Los desarrollos en esta área giran en torno a la utilización de diferentes modelos y métodos para el análisis de datos. Dentro de estos, los modelos de red también conocidos como modelos de grafos (*ilustración 1*), proporcionan una manera natural de describir una población y sus interacciones. Los nodos (vértices) del grafo representan los individuos y los bordes (enlaces) representan las interacciones entre los individuos que potencialmente conducirían a la transmisión de la infección. Representaciones de red similares se utilizan en un número de contextos, como las redes de transporte, las redes de comunicación (por ejemplo, Internet y la *World Wide Web*) y las redes sociales (Lloyd & Valeika, 2007).

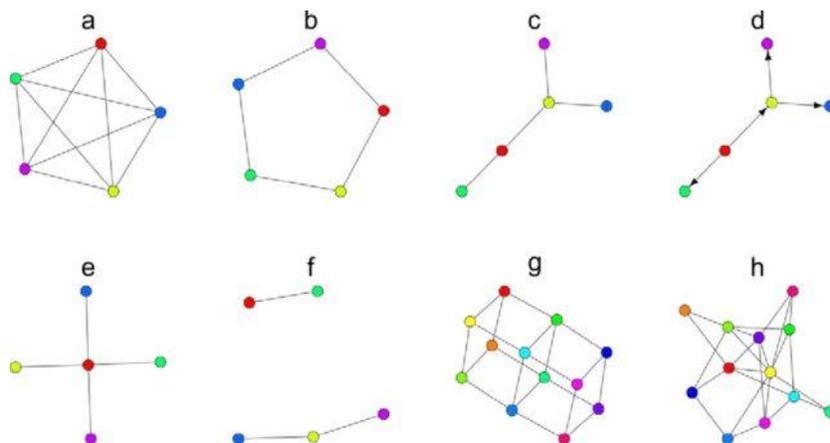


Ilustración 1. Algunos tipos de grafos. a: grafo completo, b: anillo o grafo simple, c: árbol, d: grafo dirigido (acíclico), e: grafo estrella, f: grafo no conexo, g: grafo de celosía, y h: grafo de mundo pequeño. Fuente: (Ruiz Ruano García & López Puga, 2020)

Las redes reales de agua usualmente podrían ser categorizadas como ejemplos de estos grandes grafos (Herrera et al., 2012).

Con la creciente disponibilidad de datos, la potencia de cálculo, y el avance metodológico en las últimas dos décadas, los enfoques de la teoría de redes se han buscado cada vez más para el modelado epidemiológico humano (Bansal et al., 2006), ganadero (Kao et al., 2006) y biológico-ambiental (Perkins et al., 2009), entre otros.

Ya se dijo que la epidemiología posee elementos que provienen de diversas áreas del conocimiento; en términos de teoría de redes, esto la convertiría en lo que podría denominarse un *hub* científico (Reynoso et al., 2008). De esta manera se entrelazan diversos tópicos de investigación que entretejen las ramas del saber con las áreas de aplicación científica. Es claro que la epidemiología resulta un área fundamental de la salud pública, ambiental y de la medicina clínica; pero, por otro lado, es un sólido puente de conexión con saberes como la sociología, la antropología, la economía, la estadística y más recientemente, la ingeniería y las ciencias de la computación. Se presenta así la posibilidad de encontrar semejanzas entre problemas que parecen ajenos y, por lo mismo, resolverlos con técnicas análogas, lo que permite aprender a usar similitudes correctas (Kuhn et al., 2011).

1.1.3 Metodologías de investigación en epidemiología: aplicaciones en ingeniería hidráulica

Estudios del uso de metodologías de investigación en epidemiología, aplicados en el campo de la ingeniería, consideran trabajos en sistemas de distribución de agua urbana, de la misma forma en que se han desarrollado trabajos en epidemiología apoyados en estudios de ingeniería en el mismo campo (Maslia et al., 2000). Algunos ejemplos se citan a continuación:

Bardet et al. (2014) fueron pioneros en esta idea al establecer un paralelismo entre los casos de agrietamiento de tuberías agrupadas y la mortalidad humana observada durante eventos climatológicos extremos como olas de calor o contaminación del aire. Herrera et al. (2016) hacen uso de las series de tiempo y el algoritmo SAX para encontrar patrones relacionados con la mortalidad humana observada durante eventos climáticos extremos similares en las ciudades de Recife y Londres.

Respecto a las redes de distribución de agua se ha logrado tener un modelo híbrido que facilita la toma de decisiones sobre la renovación de tuberías de conducción y que permite analizar los principales factores que inciden en el deterioro de las mismas. En el desarrollo de este trabajo se aplicó la técnica de los Sistemas de Ayuda a la Decisión Multicriterio (MCDSS), así como los Algoritmos Genéticos, Lógica Difusa, y el análisis del riesgo de la probabilidad de fallo y sus consecuencias (Guzmán et al., 2012)

Se ha trabajado con modelos estocásticos del tipo ARIMA y con la metodología de Redes Neuronales para organizar y predecir la demanda de agua urbana en áreas densamente pobladas. Espinoza et al. (2010) propuso una metodología que incorpora implícitamente en un modelo estocástico de predicción, el conjunto de eventos sociológicos perturbadores del proceso de demanda. Esta metodología fue probada en un caso real para la ciudad de Valencia, España, y se evidenció que sí consigue mejorar los resultados de las predicciones que obtienen tanto los modelos ARIMA como las redes neuronales.

Las metodologías de análisis de supervivencia también se han aplicado en algunos estudios para evaluar ese patrón en las redes de suministro de agua potable; los algoritmos de este tipo de análisis constituyen los métodos estadísticos para simular y modelar la rotura de las tuberías. En la investigación de Solano et al. (2008) “Análisis de supervivencia en fiabilidad. Predicción en condiciones de alta censura y truncamiento: el caso de las redes de suministro de agua potable”, se identifica el

patrón de fiabilidad y supervivencia de una red de suministro de agua y sus tuberías; así como los factores de riesgo (físico-ambientales) más influyentes en la red.

Publicaciones posteriores como la de Snider et al. (2020) en la línea del análisis de supervivencia, presentan una comparación entre dos métodos de estimación de roturas en tuberías: un modelo de aprendizaje automático de árbol de decisiones que aumenta el gradiente y un modelo de análisis de supervivencia de riesgos proporcionales de Weibull para predecir el tiempo de rotura en tuberías de hierro fundido. El análisis de supervivencia ha demostrado tener una buena capacidad de estimación cuando se tienen datos censurados.

El Modelado Basado en Agentes (MBA) es una metodología de análisis que ha sido ampliamente usada para el estudio de sistemas complejos presentes en el ámbito epidemiológico. Actualmente tiene una gran aplicabilidad en áreas como la medicina, la ecología, la economía, las ciencias sociales, las ciencias computacionales entre otras. La hidráulica y el estudio de las dinámicas de redes han recurrido al uso del MBA para la adopción de proyectos de reutilización de agua y evaluación de planes de expansión (Kandiah et al., 2019). Los modelos que se adoptaron simulan las percepciones de los consumidores, los patrones de discusión y la adopción o rechazo frente a la reutilización del agua.

Shafiee et al. (2018) desarrolla en sus trabajos marcos de MBA para probar y evaluar las decisiones de mitigación que un administrador de servicios públicos domiciliarios puede tomar para proteger la salud pública en una amplia gama de eventos de contaminación. También se ha recurrido al MBA para simular la dinámica del suministro de agua urbano, en el que la oferta y la demanda se combinan con modelos de consumidores y de administradores para evaluar la sostenibilidad de los planes de gestión ante eventos de sequías (Ali et al., 2017).

Desde el marco de la investigación interdisciplinaria de ciencias de la salud y en particular de la epidemiología, se han presentado nuevos usos y desarrollos

conceptuales en áreas como la investigación en la demanda final de energía en zonas urbanas. El enfoque sistémico de la “Epidemiología de la energía” (Hamilton et al., 2013), que se centra en la investigación empírica, proporciona un marco metodológico que involucra profesionales en diversas áreas: físicos, ingenieros, sociólogos y economistas, en un trabajo interdisciplinario que permite observar y describir patrones de demanda de energía y demás asociaciones entre los factores que la determinan.

1.1.4 Nuevos desarrollos y regulaciones en cuanto al agua

El agua es un recurso natural vital e indispensable para el desarrollo de la vida. El acceso al servicio de agua potable lo es igualmente y así está considerado en las legislaciones de muchos países y debe cumplir con condiciones de calidad, cantidad y continuidad. En las últimas décadas la cobertura del servicio de agua potable ha mejorado, pero todavía se tienen grandes desafíos a causa del crecimiento poblacional, cambios en los modelos de consumo, el desarrollo socioeconómico y la escasa planificación de algunos servicios propios al agua.

El informe mundial de las Naciones Unidas publicado en el 2020 sobre el Desarrollo de los Recursos Hídricos, plantea que el uso global de agua se ha multiplicado por seis en los últimos cien años y sigue aumentando a un ritmo constante de aproximadamente 1% al año por causa de aquellos desafíos. Si la situación no cambia, el mundo podría enfrentarse a un déficit global del 40% para el año 2030 (Hídricos et al., 2020).

Actualmente es un desafío para muchos países garantizar la cobertura del servicio de agua, situación que se aprecia más en zonas rurales. Según el informe “Avances en agua potable, saneamiento e higiene domiciliaria 2000-2017: especial atención a las desigualdades”, presentado por la Organización Mundial de la Salud (2019), una de cada tres personas en el mundo, carece de agua potable segura y las estimaciones presentaron que, para el año 2017, el 71 % de la población mundial utilizaba servicios

de agua potable gestionados de forma segura; de estos, la cobertura fue menor en las áreas rurales (53 %) que en las áreas urbanas (85 %).

1.1.5 Concepto de agua potable y redes de abastecimiento

Por agua potable se entiende el líquido apto para el consumo humano, aquel que sirve para beber, para preparar alimentos, para higiene y fines domésticos. Llega a los hogares a través de las redes de abastecimiento y distribución de agua. Una red de distribución de agua potable es aquel sistema de obras y/o tecnologías, que permite llevar el agua desde las fuentes naturales (captación del agua) hasta el punto de consumo con la calidad, cantidad y presión requeridas satisfaciendo las necesidades de la población (*ilustración 2*). Según la función que desempeñen, las redes de

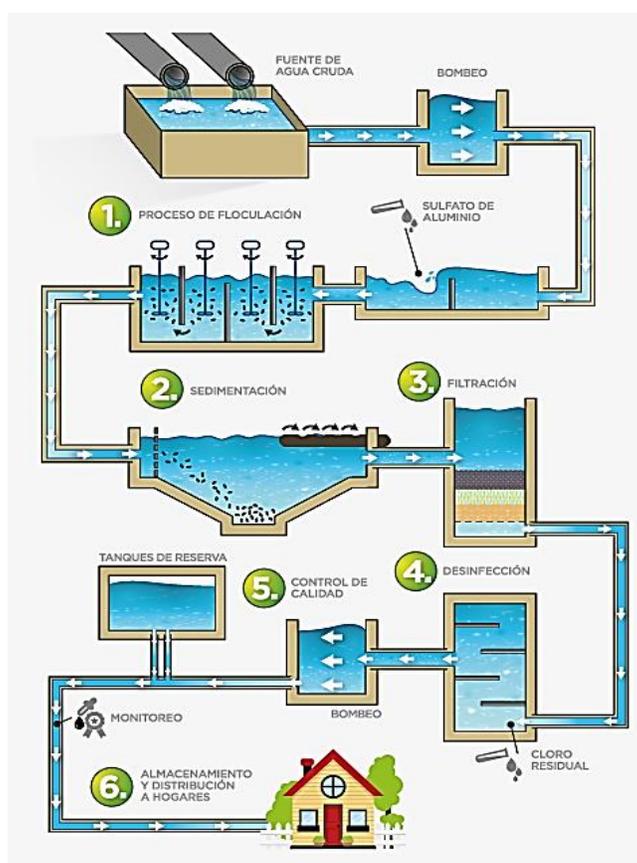


Ilustración 2. Ciclo integral de Agua. Fuente: <https://www.acuacar.com/Acuacar/Gesti%C3%B3n-Ambiental/Ciclo-Integral-del-Agua#/List>

distribución pueden clasificarse jerárquicamente como redes de aducción, redes de distribución urbana y de acometidas.

El estudio de las redes urbanas comprende diferentes procesos que han sido detalladamente analizados con el correr del tiempo y que aún se investigan debido a su importancia dentro del contexto humano, social, cultural, ambiental y económico. La investigación de la demanda de agua, los planes de rehabilitación de tuberías, el reconocimiento y detección de transitorios, la resiliencia de redes y la propagación de contaminantes han sido de gran interés y utilidad para la definición de lineamientos y la formulación de planes, programas y políticas más robustas y eficientes en el manejo del recurso hídrico, brindando mejoras en la calidad en el servicio (usos más eficientes), así como optimizando el recurso económicamente y con un menor impacto ambiental.

La gestión de una red de agua es todo un reto y procura una adecuada operación del sistema en todo su proceso a fin de satisfacer las necesidades de las poblaciones, bajo un enfoque eficiente. Una buena gestión de la red de distribución implica conocer los diferentes puntos de suministro, así como el destino de toda el agua que ingresa al sistema.

Son múltiples los desafíos que se van presentando a diario con el recurso hídrico. Cada vez más empresas y gestores buscan mejorar la eficiencia hidráulica y energética en la operación y gestión diaria de las redes de agua potable; tanto así que recientes desarrollos de conocimiento, de la mano de la innovación tecnológica, han llevado al establecimiento de plataformas web interactivas como WatEner® (Zarza et al., 2018). Dicha plataforma integra y potencia herramientas como los modelos GIS², modelos

² GIS (*Geographic Information Systems*): un sistema informático que analiza y muestra información geográficamente referenciada.

hidráulicos, SCADA³ entre otros, con soporte a la toma de decisiones a través de técnicas de Inteligencia Artificial y aprendizaje automático.

Con el gradual crecimiento poblacional y la expansión de territorios, se ha hecho necesaria la ampliación o re adecuación, de muchas de las redes de distribución urbana para lograr satisfacer las demandas requeridas y la sustentabilidad en el servicio, a tal punto que algunos sistemas pueden quedarse cortos para cumplir con los criterios de cantidad debido a que la demanda puede exceder la oferta existente. Realizar un correcto *pronóstico de la demanda de agua* con exactitud y confiabilidad es un proceso esencial en la planificación, operación y gestión de los sistemas de agua, ofreciendo no solo eficiencia técnica, sino también económica y social.

El creciente consumo de agua, así como la competencia por adquirir y distribuir el recurso, obligan a un análisis permanente en el desempeño de las redes de abastecimiento que esté acorde con los requerimientos de las comunidades. Lo anterior significa promover la gestión del recurso hídrico, así como buscar nuevas alternativas en metodologías eficientes para el manejo de los procesos de operación y gestión de dichas redes. La operación, planificación y gestión de los sistemas hídricos tienen fuerte injerencia frente a los escenarios y los cambios actuales, y la necesidad de afrontar dichos escenarios en el contexto mundial. Esto se logra mediante un enfoque holístico para la gestión del agua, permitiendo así herramientas más robustas para la toma de decisiones que favorecen a empresas, gestores y administradores (Sillero, 2021).

³ SCADA (*Supervisory Control and Data Acquisition*): herramienta de automatización y control industrial utilizada en los procesos productivos que puede controlar, supervisar, recopilar datos, analizar datos y generar informes a distancia mediante una aplicación informática.

1.1.6 Gestión inteligente de las redes de abastecimiento

En el año 2004 surge el concepto de *Ciudades Digitales* haciendo referencia a la necesidad de estructurar y organizar las ciudades bajo un enfoque de desarrollo urbano basado en la *sostenibilidad* y con el objetivo de proyectar ciudades más eficientes y con mejor calidad de vida. Ese concepto evoluciona al de *Smart Cities* (*ilustración 3*) aplicable en ciudades que apuestan por un desarrollo económico sostenible, por la gestión de los recursos naturales a través de la acción participativa y comprometida entre la administración pública y los ciudadanos y, por último, en la responsabilidad y compromiso con su entorno medioambiental (Zabaleta et al., 2015).

Se ha indicado que el agua es un recurso muy valioso y que, además, juega un papel muy importante dentro del desarrollo de los asentamientos; es por ello que su gestión inteligente “*Smart Water*” es un gran desafío para las Smart Cities pues es uno de los seis componentes que definen a una ciudad inteligente; los otros son energía, movilidad, edificios, servicios públicos e integración.

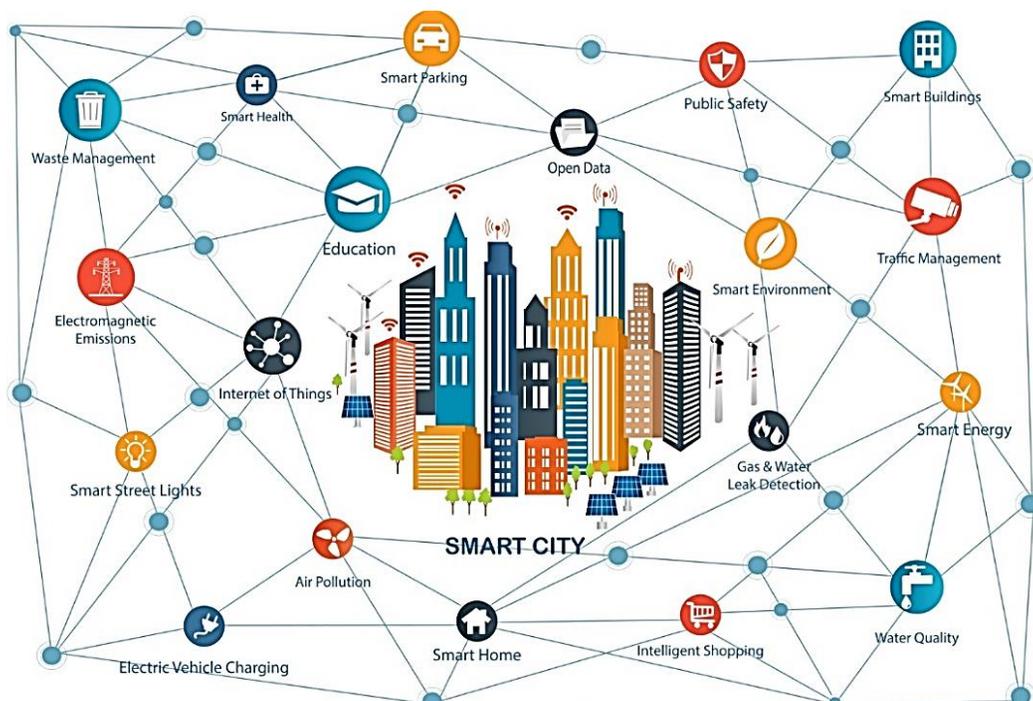


Ilustración 3. Concepto global de Smart Cities. Fuente: http://infraroi.com.br/wp-content/uploads/2018/02/smart_cities.jpg

Durante el periodo 2016-2020, la gestión inteligente del agua creció un 19% cada año, en promedio, como resultado de sus importantes beneficios para el sector. Diversos estudios han demostrado que las tecnologías inteligentes del agua podrían ahorrar a las empresas de servicios públicos de agua entre 12.5 y 15 mil millones de dólares al año en todo el mundo (Donoso et al., 2021).

Los sistemas inteligentes de distribución de agua potable deben ser cada vez más eficientes y responder a múltiples restricciones de abastecimiento (Pandolfi et al., 2018), como lo retoma el Instituto de Estudios Urbanos (2021) en atención a sus dificultades; la planificación y la gestión futura deben considerar una evaluación amplia de los recursos hídricos y detallar la importancia de los ecosistemas con miras a la resiliencia hídrica y climática.

El uso actual de las tecnologías (sensores, teledetección), el internet de las cosas (IoT), la inteligencia artificial, los sistemas informáticos apoyando el análisis de datos de los grandes volúmenes de información (Big Data) (ilustración 4), conducen a nuevos desarrollos y avances en los procesos tradicionales al interior de las gestión de redes, proporcionando nuevas alternativas y herramientas integradas más eficaces y robustas para el análisis de información así como, también, para soportar la toma de decisiones.

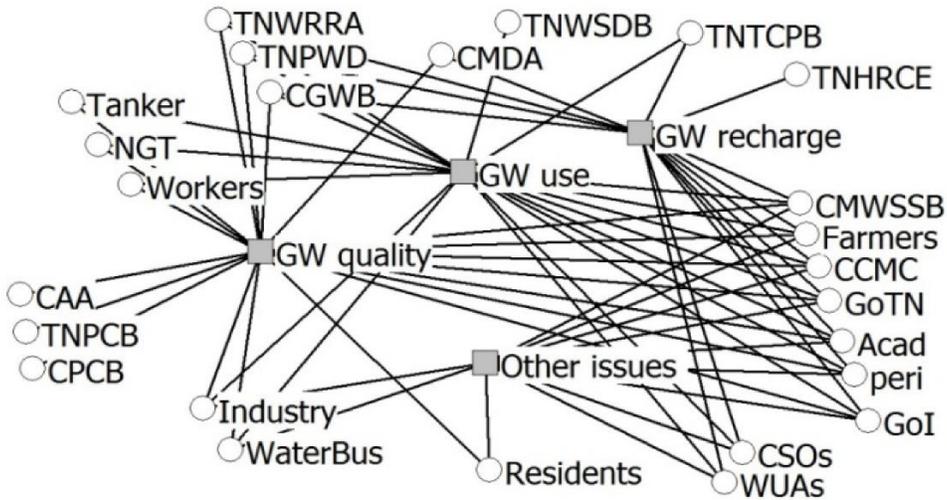


Ilustración 4. Partes interesadas y sus intereses relacionados con las aguas subterráneas. Big Data Analytics. (Brunner et al., 2014).

Apoyados en el uso de esta tecnología y combinando el análisis inteligente de datos con la epidemiología, es posible identificar nuevas formas de conocimiento (KDD: *Knowledge Discovery in Databases*) que permiten abordar los diferentes enfoques de gestión de demanda y gestión de fallos. A partir de esto, resulta conveniente realizar un análisis de las ventajas y riesgos asociados a la implementación de estas nuevas tecnologías, siempre con la meta de continuar progresando en la garantía del acceso sostenible al agua.

En este trabajo se tocan dos de las principales problemáticas que presentan los procesos de gestión de las redes de abastecimiento de agua potable: *la gestión de la demanda y la gestión de fallos*. La gestión de la demanda de agua es el procedimiento usado para mantener un equilibrio entre el consumo de agua y la disponibilidad del recurso; dicha disponibilidad está enmarcada en muchos países dentro de las políticas que buscan garantizar la disponibilidad con suficiencia y calidad. La gestión de fallos ayuda a identificar, evitar y resolver los problemas que puedan dificultar el rendimiento de la red buscando mantener y garantizar la disponibilidad, la continuidad y los rendimientos óptimos.

1.1.6.1. Gestión de demanda de agua

La demanda de agua se define como el “volumen de agua, en cantidad y calidad, que los ciudadanos o los usuarios están dispuestos a adquirir o desean recibir de la correspondiente entidad suministradora para satisfacer un determinado objetivo de producción o consumo” (RAE).

A nivel mundial son diversos los sectores que requieren grandes cantidades de agua para desarrollarse o mantenerse; la agricultura (incluida la irrigación, la ganadería y la acuicultura) es el mayor consumidor de agua, dado que representa el 69% de las extracciones anuales de agua a nivel global. La industria (incluyendo la generación de energía) representa el 19%, y los hogares el 12% (AQUASTAT, 2021).

En el año 2010 la Asamblea General de las Naciones Unidas aprobó la resolución 64-292 que reconoce el derecho de todos los seres humanos a acceder a servicios de agua y saneamiento, a tener acceso a una cantidad de agua suficiente para el uso doméstico y personal (entre 50 y 100 litros de agua diarios por persona), segura, aceptable y asequible (el coste del agua no debería superar el 3% de los ingresos del hogar), y accesible físicamente (la fuente debe estar a menos de 1.000 metros del hogar y su recogida no debería superar los 30 minutos) (Riera et al., 2019).

Para tratar este tema es necesario hacer una aproximación a la cuestión de *por qué es importante estimar la demanda de agua*. Básicamente, a partir de la determinación de la población servida se procede a estimar el número de conexiones por cada categoría de usuario, esto en virtud del volumen requerido por cada grupo de usuarios, el cual determinará la demanda por el servicio de agua potable que enfrentará el gestor en los próximos años.

Se han realizado múltiples estudios enfocados en la estimación de la demanda de agua; los primeros estudios de estimación a corto plazo involucran modelos matemáticos basados en la correlación entre los datos de la demanda y factores demográficos y ambientales (Maidment et al., 1985; Ruefenacht et al., 1997); por otro lado también se han realizado estimaciones a partir de la modelación de la relación de datos presentes y pasados de la demanda (análisis estocástico de series temporales) (Molino et al., 1996; Nel et al., 1996; Saporta et al., 1994). El análisis de series temporales (modelos ARIMA), la regresión múltiple y las Redes Neuronales Computacionales (RNCs) fueron los métodos usados para predecir la demanda de agua bajo un enfoque agrícola; las mejores estimaciones fueron las obtenidas con las RNCs. Las RNCs han sido ampliamente utilizadas en el ajuste de funciones y en el modelado de sistemas altamente no lineales, ya que realizan una transformación de este tipo entre los datos de entrada y salida, lo que supone importantes ventajas frente a las metodologías estadísticas convencionales tales como la regresión lineal múltiple o la regresión en

componentes principales en las que la relación entre los datos ha de ser lineal (Luque et al., 2002).

Otros trabajos han realizado la comparación de modelos predictivos para la estimación de demanda de agua. Se destaca el trabajo de Herrera et al., (2010) en el que se presenta el uso de Redes Neuronales Artificiales (*Artificial Neural Networks: ANN*), Proyección para la Búsqueda de Regresiones (*Projection Pursuit Regression: PPR*), Splines de Regresión Adaptativa Multivariante (*Multivariate Adaptive Regression Splines: MARS*), Regresión Vectorial de Soporte (*Support Vector Regression: SVR*), Bosques Aleatorios (*Random Forests: RF*) y un modelo ponderado para la predicción de la demanda de agua en series altamente no lineales. Los resultados de esta comparación identificaron a los SVR como los modelos más precisos, seguidos de cerca por MARS, PPR y RF.

Los campos de aplicación de la inteligencia artificial (IA) también han aportado en el análisis de la predicción de demanda de agua; son múltiples las metodologías estudiadas tales como sistemas de lógica difusa, máquinas de vectores soporte, máquinas de aprendizaje extremo, Redes Neuronales Artificiales y modelos ARIMA. En el trabajo de revisión titulado “Enfoques de inteligencia artificial para el pronóstico de la demanda de agua urbana” Muhammad et al. (2019) realiza un análisis de diferentes enfoques de IA e híbridos, demostrando la utilidad y las aplicaciones de cada modelo como solución para la previsión de la demanda de agua. La revisión muestra que no existe un único modelo híbrido o de inteligencia artificial que parezca ser el mejor; asimismo, la IA se puede aplicar con éxito para la estimación de dichas predicciones.

Nuevos algoritmos a los tradicionalmente usados han sido probados para predecir la demanda de agua. En el trabajo de Navarrete-López et al. (2019) se presentan varias herramientas de reducción para análisis de series de tiempo basadas en una técnica de aproximación del agregado simbólico (*SAX: Symbolic Aggregate Approximation*) capaz de manejar versiones simples de conjuntos de datos. Se muestra que un modelo

basado en redes neuronales que utiliza la generación de conocimiento apoyado en SAX a partir de varias series temporales puede mejorar las capacidades de pronóstico. Este nuevo marco de análisis muestra varios beneficios en términos de precisión y rendimiento de los modelos de redes neuronales para la demanda de agua, mostrando también las ventajas de SAX a la hora de identificar patrones, distancias y correlaciones entre series de tiempo.

Continuando con las redes neuronales, estas también se han usado como parte de métodos combinados para estimar demandas mensuales en la provincia de Gauteng de la República de Sudáfrica (Zubaidi et al., 2020). La investigación realizada aplica una metodología novedosa que incluye preprocesamiento de datos y una Red Neuronal Artificial (ANN) optimizada con el algoritmo de búsqueda de retroceso (BSA: *Backtracking Search Algorithm*) para estimar la demanda de agua mensual en relación con el consumo de agua anterior. El modelo presentado BSA-ANN arrojó mejores resultados demostrando ser más eficiente y confiable frente al algoritmo de búsqueda de cuervo (CSA: *Crow Search Algorithm* - ANN).

En un trabajo paralelo Zubaidi, et al. (2020) se propone otra metodología que combina la transformada discreta de ondículas (DWT: *discrete wavelet transform*) con un sistema de inferencia neuro-difuso adaptativo (ANFIS: *Adaptive Neuro-Fuzzy Inference System*) para predecir la demanda mensual de agua urbana en función de varios intervalos de consumo histórico de agua. Las conclusiones de esta investigación muestran que las dos metodologías analizadas son estadísticamente equivalentes y capaces de predecir con fidelidad la demanda de agua urbana mensual.

1.1.6.2. Gestión de fallos

La infraestructura de las redes de abastecimiento de agua es una de las temáticas que más se ha estudiado a través del tiempo y que seguirá teniendo un papel protagónico en la investigación, no solo por ser un tema asociado al progreso y desarrollo sino

también al estar relacionado de manera estrecha y directa en la gestión del agua. La red de agua es un sistema complejo, dinámico e interconectado en el que intervienen e interactúan múltiples factores de naturaleza heterogénea. Cualquier variación mínima entre sus elementos componentes puede modificar, de forma imprevisible, las interrelaciones y, por lo tanto, el comportamiento de todo el sistema.

Así, la evolución de esta clase de sistemas se caracteriza por la fluctuación, situación en la que el orden y el desorden se alternan constantemente (Caparrini et al., 2015).

Desafortunadamente, las redes de distribución de agua modernas son complejas y difíciles de administrar debido al mayor nivel de urbanización, las diferentes demandas de los consumidores y los recursos limitados (Bello et al., 2019). Diferentes patologías pueden afectar a las redes de agua; con el paso del tiempo la infraestructura envejece y las tuberías pueden sufrir procesos de deterioro lo que afecta la capacidad de transporte hidráulico y hace que se presenten fisuras, fugas o roturas generando una pérdida en volumen del agua del sistema, una pérdida de energía y por consiguiente una pérdida económica. La investigación realizada por la Asociación Internacional de Abastecimiento de Agua (IWA), mostró que la cantidad de agua perdida o “no contabilizada” (UFW, por sus siglas en inglés, *unaccounted for water*) suele oscilar entre el 20 y el 30 % de la producción (Daoudi et al., 2005).

Las fugas dependen en existencia y cantidad de diversos factores como lo son las malas conexiones entre tuberías, corrosión interna o externa de la tubería, daño mecánico causado por una carga excesiva en la tubería, movimiento de tierra, altas presiones en la red, daños en las tuberías por excavaciones, la edad de la tubería, temperaturas extremas, defectos en las tuberías y baja calidad de la mano de obra (Puust et al., 2010).

Diferentes mecanismos pueden generar fallos en las redes de tuberías; Rajani & Kleiner (2001) consideran tres aspectos principales: (i) propiedades estructurales de tuberías, material, interacción entre el suelo y la tubería, y la calidad de la instalación; (ii) cargas internas debido a la presión de operación y cargas externas, sobrecargas del

suelo, tráfico pesado, cargas de heladas y daños accidentales o intencionales; y (iii) deterioro del material debido a cambios internos y externos y por ambientes químicos.

Un fallo o rotura en el sistema de la red de agua se puede presentar de manera espontánea en un punto de la red y esto implica un mantenimiento de reparación. Según (Ortega et al., 2014) el mantenimiento correctivo incluye cinco períodos de tiempo:

- Tiempo transcurrido desde que se produce la avería hasta que se detecta, o tiempo de respuesta
- Tiempo necesario para contar con quien efectúa la reparación (disponibilidad del personal técnico encargado de la reparación)
- Tiempo que transcurre hasta disponer de los recambios necesarios para reparar el fallo.
- Tiempo necesario para subsanar la avería
- Tiempo necesario para la nueva puesta en servicio de la red (limpieza, desinfección...).

Un suceso inicial presente en una red de agua, puede desencadenar una serie de fallos sucesivos en el interior del sistema denominado “*fallos en cascada*”. Es común que, si se presenta un fallo en una parte del sistema, las otras partes, deben subsanar el componente defectuoso de manera ágil y eficiente. No siempre se consigue realizar rápidamente dicho arreglo lo que puede suponer una sobrecarga en otros nodos de la red conduciendo al colapso de estos nodos uno tras otro o finalmente comprometiendo la integridad física y operacional de todo el sistema. Este tipo de fallos en cascada se presenta comúnmente en las redes eléctricas; en las redes de agua pueden darse como consecuencia de depresiones y sobrepresiones en el sistema.

Existen algunas situaciones en la naturaleza en los que se puede apreciar la dinámica del fallo en cascada, por ejemplo: a nivel biológico la cascada isquémica (Felberg et al., 2000) en la que un pequeño ataque isquémico libera toxinas que matan muchas

más células que el daño inicial, lo que resulta en la liberación de más toxinas. En electrónica, el generador Cockcroft-Walton ya que un diodo fallado puede provocar la falla de todos los diodos en una fracción de segundo (Cockcroft et al., 1932) en finanzas la quiebra de una institución financiera puede provocar la quiebra de otras instituciones financieras lo que se conoce como riesgo sistémico (Huang et al., 2013).

Grandes infraestructuras como las centrales eléctricas, el transporte, el combustible, el suministro de agua están conectadas y requieren de interdependencia para su funcionamiento. Debido a este acoplamiento, las redes interdependientes son extremadamente sensibles a fallas aleatorias y, en particular, a ataques dirigidos, de modo que una falla de una pequeña fracción de nodos en una red puede desencadenar una cascada iterativa de fallas en varias redes interdependientes (Rinaldi et al., 2001).

El estudio de la identificación, detección y diagnóstico de posibles fallos en una red de agua cobra cada vez más importancia e involucra modelos matemáticos más robustos y tecnologías informáticas que facilitan su análisis optimizando el rendimiento de la infraestructura y sus costos operativos.

Recientemente Bello et al. (2019) presentan una revisión detallada de los problemas de gestión de las redes modernas de distribución de agua (WDNs: *Modern water distribution networks*) y los modelos matemáticos utilizados en la resolución de estos problemas (*ilustración 5*) según la clasificación temporal a corto plazo (STP: *short term problem*), a mediano plazo (MTP: *medium term problem*) y a largo plazo (LTP: *long term problem*), en el estudio titulado: Resolviendo Problemas de Gestión en Redes de Distribución de Agua: Un Estudio de Enfoques y Modelos Matemáticos.

Diversos enfoques referentes a la gestión de fallos han sido abordados en nuevos estudios como los citados a continuación:

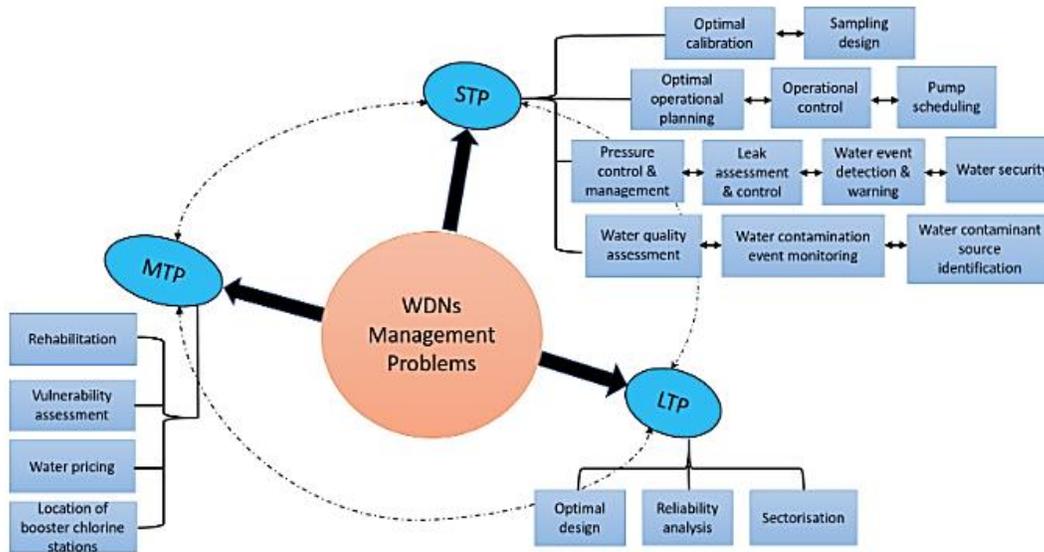


Ilustración 5. Clasificación de los problemas de gestión en las WDNs. Bello, O., et al. (2019)

El modelo de riesgos proporcionales ha servido como punto comparativo en el estudio realizado por (Cobacho et al., 2019) en el que se hace uso de la distribución Weibull y registros históricos de reparaciones para presentar un modelo de predicción de fallos relativamente más simple frente a los usados actualmente con mayor nivel de complejidad. El modelo propuesto se fundamenta en un requerimiento limitado y básico de datos, una complejidad baja de cálculo y la obtención de unos resultados de interpretación clara y directa.

Las redes neuronales han sido usadas y con buena predicción en estudios de detección de fugas de agua. En su estudio Aramane et al. (2019) hacen uso de las redes neuronales y de herramienta de código abierto como Neuroph Studio para plantear un modelo que consiste en una red neuronal de perceptrón multicapa y que identifica fugas simultáneas en múltiples regiones. Su modelo fue útil para detectar regiones únicas y múltiples de fugas con buena precisión. También Khorasgani et al. (2019) hace uso de las redes neuronales convolucionales gráficas (GCNN, *Graph Convolutional Neural Networks*) para la detección y aislamiento de fallos en redes

industriales a gran escala con el mismo tipo de componentes. Los resultados obtenidos presentan que las GCNN superan a algunos algoritmos de referencia.

Hashim et al. (2020) propone un método para diagnóstico y detección de fallos para redes de agua de edificios no residenciales a partir de la combinación de Análisis de Componentes Principales (PCA: *Principal Component Analysis*) y Máquinas de Vector Soporte (SVM: *Support Vector Machines*). El enfoque propuesto detectó y diagnosticó más alarmas reales del sistema frente a un enfoque convencional.

Investigaciones recientes como lo publicado por (Umoh et al., 2022), sugiere la posibilidad de usar la teoría del caos para modelar y controlar la dinámica de desastres en cascada desde la perspectiva de la ingeniería de sistemas de control. El trabajo llega a la conclusión de que el enfoque adoptado tiene el potencial de dilucidar estrategias de control viables para reducir el riesgo de desastres en infraestructuras críticas durante eventos climáticos extremos como huracanes e inundaciones.

1.2 JUSTIFICACIÓN

El agua es un elemento vital para el ser humano y para todo aquello que signifique “vida”. Su importancia se extiende no sólo en comprender este criterio, sino para aquello que implique la supervivencia humana. El agua, al ser un recurso finito, requiere de una gestión sustentable que garantice la disponibilidad del mismo, asegurando un uso y abastecimiento continuo, cubriendo las necesidades presentes sin comprometer las necesidades de las generaciones futuras. La gestión de los recursos hídricos ha evolucionado de manera significativa en los últimos 100 años, pasando de la misión hidráulica a la Gestión Integrada de los Recursos Hídricos (Santos et al., 2019).

Tradicionalmente, la gestión de los recursos hídricos se ha realizado de manera “unidimensional”, respondiendo a necesidades muy concretas, tales como el abastecimiento urbano o el regadío (Nhapi et al., 2005). La historia moderna de la gestión de los recursos hídricos se inicia a partir de la revolución industrial, que genera un cambio de paradigma en el uso del agua en el mundo occidental (Molle et al., 2009). Las investigaciones más recientes abordan la temática desde una perspectiva más holística, en la que se otorga una participación más activa de otros contextos como el social, el económico y el ambiental. La gestión del recurso hídrico en el contexto actual debe estar orientada a la anticipación de las posibles problemáticas que se pueden presentar en el interior del sistema; esto conlleva actuaciones oportunas y eficientes con un impacto directo sobre el componente económico.

La gestión del agua se ha ido complicando con el tiempo debido al aumento de la población y al aumento del consumo, tanto familiar como industrial, que vienen creciendo en una progresión geométrica; mientras que las alternativas de manejo y adecuación de los sistemas de suministro crecen como una progresión aritmética. Como consecuencia de esto, se hace imprescindible contar con un mayor control y tener más rigurosidad en los procedimientos de gestión de las redes de suministro urbanas.

La mayoría de estos problemas de gestión son desafiantes y difíciles de resolver utilizando técnicas tradicionales de resolución de problemas. La aplicación reciente de modelos matemáticos y tecnologías informáticas en las redes de distribución moderna ha ayudado en gran medida a los directivos a abordar la mayoría de estos problemas que no eran triviales en el pasado (Bello et al., 2019).

La práctica de la investigación lleva constantemente a replantear los límites de todo aquello que la conforma, desde los planteamientos a las técnicas. El concepto de *blue sky project* o la expresión en castellano “*fronteras del conocimiento*” remite a esa idea (casi geográfica) de expansión continua propia de la ciencia (García et al., 2020). Hay áreas de conocimiento dentro de la gestión de agua, que han sido estudiadas y

sobre las cuales se posee un acervo de investigación científica bastante robusto: el agua propiamente dicha, los procesos de gestión y la epidemiología. No obstante, hay otras aún poco exploradas como la aplicación de herramientas de análisis de datos en el contexto epidemiológico, que puede aportar significativamente en la generación de nuevo conocimiento en virtud de la fusión de metodologías análogas.

La epidemiología es una ciencia dinámica que se ha reinventado a la luz de las nuevas facetas sociales y las nuevas tecnologías. Es evidente que existe un gran potencial de estudio a partir de la evolución de sus metodologías de análisis y la aplicabilidad en otras áreas de conocimiento como la hidráulica (*ilustración 6*). Esta ciencia ofrece una serie de herramientas de análisis que pueden ser aplicadas a momentos, procedimientos o procesos generales de la gestión del agua en su integridad.

En este sentido, la presente investigación tantea la problemática de la gestión de las redes de abastecimiento buscando aplicar procedimientos de análisis epidemiológico en la estimación de demanda de agua para el reconocimiento de patrones y tendencias cuando se tienen series de datos de gran tamaño. También se aborda la gestión de

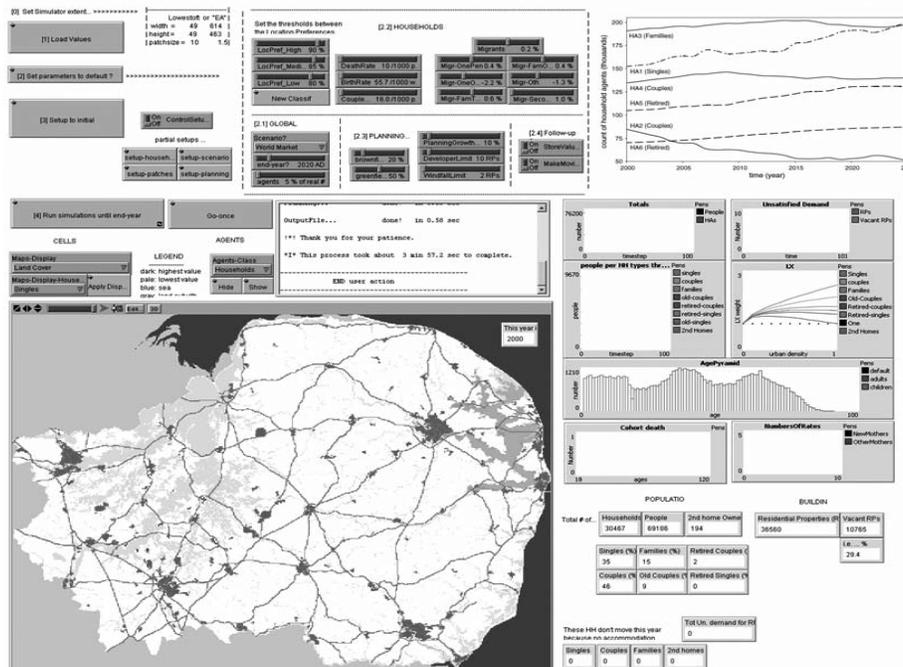


Ilustración 6. Ejemplo de interface del modelo HI-LIFE (NetLogo) y proyección demográfica de agentes (MBA). (de Carvalho Cantergiani & Colegios, 2011)

fallos desde el enfoque de los modelos compartimentales usados en epidemiología para analizar y evaluar la ocurrencia de un fallo, su diseminación de un punto a otro y la identificación de zonas de riesgo en la red. Actualmente se cuenta con software y metodologías fiables para el reconocimiento e identificación de un fallo en una red de agua (*ilustración 6*), pero son pocas las investigaciones que han estudiado las implicaciones de fallos sucesivos o en cascada en las redes de agua.

Para entender el enfoque dado a esta investigación, resulta procedente identificar el problema que se ha de investigar. Para ello, el uso de algunos métodos epidemiológicos en la hidráulica y más puntualmente en la gestión de abastecimiento de agua urbana, es todavía un área poco explorada; el agua y la epidemiología siempre han estado relacionadas y dicha relación se ha intentado principalmente desde una perspectiva unidireccional (Agua-Epidemiología), se ha volcado más hacia el estudio del agua y su calidad como un factor determinante o influyente en procesos epidemiológicos; sin embargo, el enfoque desde la otra perspectiva (Epidemiología - Agua) encierra un gran potencial con múltiples alternativas para aportar en la adaptación y establecimiento de nuevos métodos, que puedan llegar a proporcionar beneficios en la gestión integral del agua urbana.

Es necesario identificar y reconocer nuevas herramientas que el área de la epidemiología ofrezca y que sirvan como apoyo al estudio de las complejas interacciones entre factores relacionados con las prácticas de la operación y la gestión de redes de abastecimiento urbano y las estructuras sociales, procesos físicos, sistemas de ingeniería y factores ambientales.

Aún son pocos los trabajos que conjugan o que combinan las técnicas de epidemiología con la ingeniería. En función de lo anterior, es importante profundizar en la identificación y el desarrollo de estas técnicas brindando alternativas novedosas a las técnicas actualmente existentes.

A partir de los estudios de investigación y empíricos en epidemiología humana, se introduce un contexto para analizar las variaciones en el tiempo de roturas de tuberías que pueden ayudar a las agencias de agua a entender mejor los fallos de tubos agrupados e instaurar medidas para minimizar las interrupciones causadas por ellos. Se postula que, en cualquier momento, una cohorte de la población de tuberías que comprende el sistema de distribución de agua estará en un estado debilitado debido a la fatiga del metal y la corrosión. Esta cohorte frágil se vuelve vulnerable durante el curso de las operaciones normales y, en última instancia, se rompe debido al rápido aumento de longitudes de grieta inducidos por factores de estrés anormales (Bardet et al., 2014).

El modelo de “cosecha” epidemiológica desarrollado en este trabajo simula una serie de tiempo observada de roturas de tuberías mensuales y tiene tanto poder explicativo como predictivo, siendo análogo, bajo el contexto de la epidemiología humana, a individuos vulnerables o frágiles, expuestos a la contaminación del aire o eventos de calor extremo (Bardet et al., 2014).

Comprender el modelo de cosecha para los sistemas de distribución de agua beneficiaría a las agencias y podría ayudar a prevenir y mitigar los brotes de roturas de tuberías de una manera comparable a las medidas de salud pública que han mitigado con éxito los efectos de las olas de calor en la población más vulnerable (Bardet et al., 2014).

En virtud de lo planteado hasta ahora, se formula la pregunta que orienta el presente estudio:

¿En qué medida el uso de métodos epidemiológicos de análisis de datos respalda la operación, gestión y toma de decisiones en la red de distribución de un sistema de abastecimiento de agua urbana?

Este trabajo bosqueja dos hipótesis de investigación, una para cada uno de los procesos de gestión incoados:

H₁_Gestión de demanda: Es posible una mejora en el proceso de estimación de demanda de agua a partir de la aplicación del algoritmo SAX y redes neuronales.

H₂_Gestión de fallos: Se puede lograr una gestión eficiente de fallos en la distribución de agua usando métodos de análisis compartimental combinado con la simulación hidráulica de una red de abastecimiento urbano.

1.3 OBJETIVOS

1.3.1 Objetivo General

El objetivo principal de la tesis doctoral es el de identificar y adaptar herramientas y metodologías de análisis de datos, usadas en el área de la Epidemiología, en la operación y gestión de redes de abastecimiento urbano, con el fin de favorecer procedimientos enmarcados en la gestión de los sistemas de distribución de agua urbana.

1.3.2 Objetivos Específicos

- Realizar un estudio detallado de la literatura relacionada con los procesos de gestión de las redes de abastecimiento de agua y con los procedimientos de análisis de datos usados en el área de la epidemiología.
- Establecer posibles nexos y evaluar el estado del arte de las metodologías de análisis epidemiológicos con los procesos de gestión de demanda y gestión de fallos en redes de abastecimiento de agua urbana.
- Demostrar la aplicación y eficacia del algoritmo SAX, usado tradicionalmente para los modelos de reconocimiento de patrones asociados con series temporales, en la estimación de la gestión de demanda de agua urbana.

- Aplicar el modelo de epidemiología compartimental (SIR: Susceptible-Infectado-Recuperado) para analizar la propagación de fallos en sistemas de distribución de agua urbana.
- Modelar los fallos en cascada que se pueden presentar en una red de distribución de agua urbana a través del Modelamiento Basado en Agentes (MBA).

1.4 BREVE DESCRIPCIÓN DE ESTE DOCUMENTO

El desarrollo de la investigación que se presenta en esta tesis doctoral se encuentra organizada en siete capítulos cuyo contenido se obtiene a partir del siguiente mapa conceptual en la *ilustración 7*

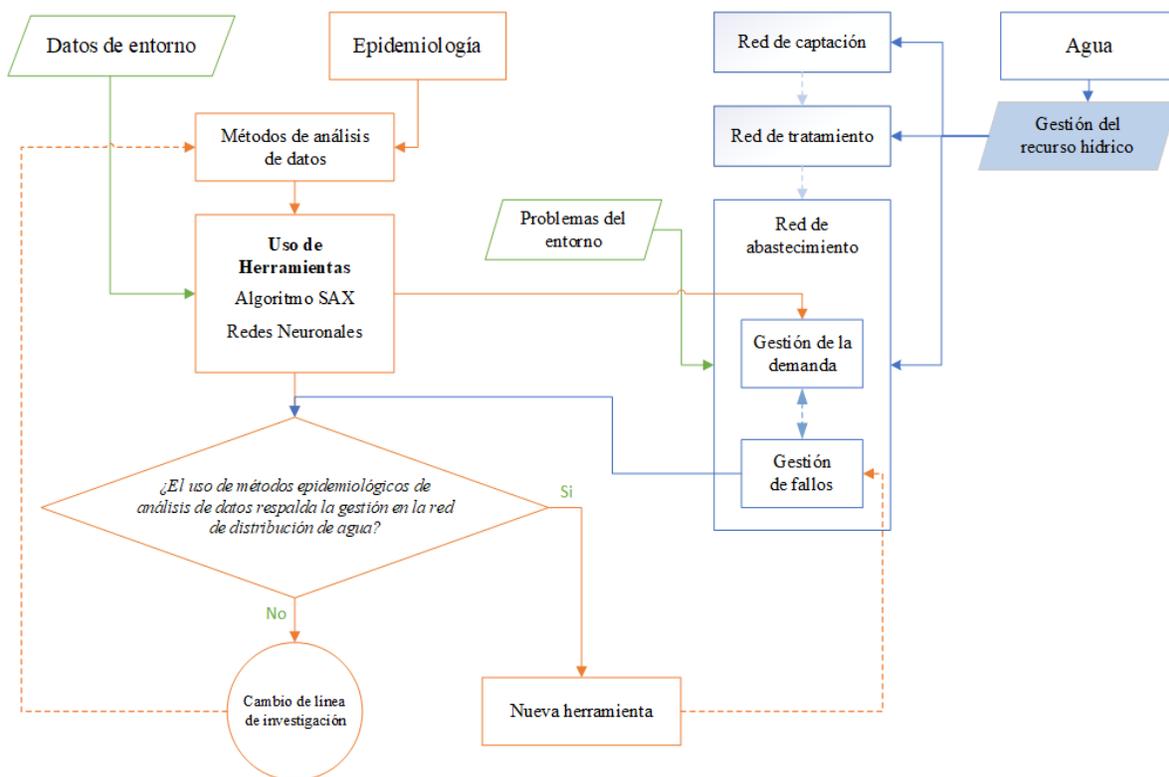


Ilustración 7. Mapa Conceptual del documento.

Se pueden apreciar dos bloques importantes: uno en azul referido al agua y su gestión y el otro en naranja referido a la epidemiología y sus metodologías de análisis de

datos. Con este mapa, las temáticas a desarrollar son: una introducción, una revisión bibliográfica, una gestión de demanda, una gestión de fallos, casos de estudio y unas conclusiones. A continuación, se presenta en la *ilustración 8* un flujograma que resume la organización del actual documento.

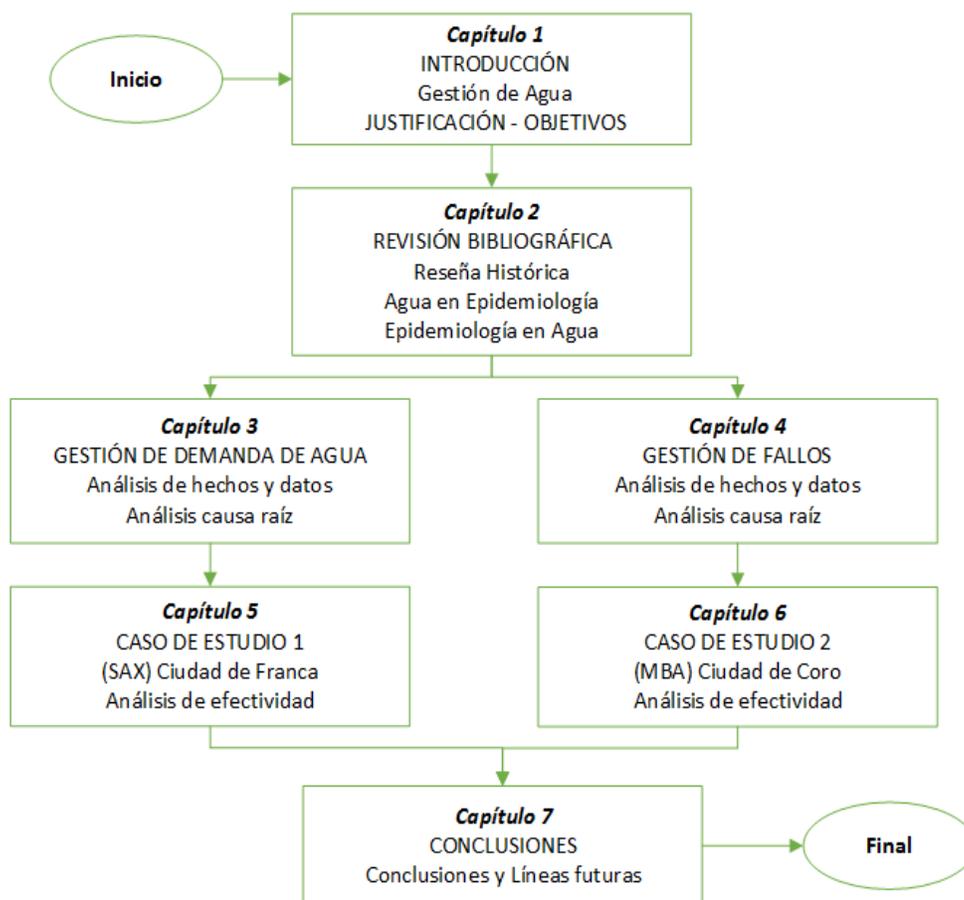


Ilustración 8. Diagrama de flujo de la organización del documento

En el Capítulo 1 se muestra una introducción general del trabajo desarrollado en la tesis. Se plantea la problemática principal y las ideas que dan origen al desarrollo de la investigación, se hace una contextualización de las temáticas que serán abordadas en los capítulos posteriores y la justificación de la investigación. También se presenta

el objetivo general y los objetivos específicos que han servido como punto de partida para la consecución de resultados.

En el Capítulo 2 se enseña una reseña histórica enfocada en la evolución de la epidemiología relacionada con la gestión del agua y, por otra parte, la evolución de la gestión del agua hasta nuestros días. Además, se muestra el estado del arte (revisión bibliográfica): detallado, selectivo y crítico, centrado en las líneas de investigación seguidas en la tesis.

En el Capítulo 3 se trata la temática de gestión de demanda en redes de agua urbana junto con un análisis de métodos en especial del SAX.

En el Capítulo 4 se expone la temática de gestión de fallos en redes de agua urbana, también con un análisis de métodos en especial de MBA.

En el Capítulo 5 se desarrolla el caso de estudio de la ciudad brasilera de Franca, en el cual se aplican las metodologías de análisis de demanda usadas en la red de abastecimiento de agua urbana. Aplicación de SAX. Se realiza su análisis de efectividad.

En el Capítulo 6 se desarrolla el caso de estudio de la ciudad venezolana de Coro, en el cual se aplican las metodologías de análisis de fallos usadas en la red de abastecimiento de agua urbana. Aplicación de MBA. Se realiza su análisis de efectividad.

En el Capítulo 7 se argumentan las conclusiones finales tras la realización de este trabajo y se plantean algunas líneas futuras de investigación en aspectos y temáticas potencialmente interesantes para desplegar en trabajos complementarios.

Como capítulo final, se relaciona una la lista de referencias propias elaboradas en el marco de esta investigación doctoral y la lista de referencias bibliográficas utilizadas para esta tesis doctoral.

CAPÍTULO 2

REVISIÓN BIBLIOGRÁFICA

“Este capítulo aborda una breve reseña histórica de la epidemiología en relación con el impacto del agua en ella, su evolución desde las primeras civilizaciones que acometieron el estudio de las enfermedades y su relación con el agua que consumían. A la par, se presenta de igual forma la evolución de la gestión del agua desde la prehistoria hasta nuestros días y cómo la metodología usada en epidemiología va tomando un lugar en la gestión del recurso. Se concluye con una revisión detallada al Estado del Arte en la epidemiología y en la gestión del recurso hídrico.”

2.1 RESEÑA HISTÓRICA

2.1.1 Epidemiología

El concepto de epidemiología es bastante amplio y se remonta a varios miles de años atrás. Los primeros documentos que trataban asuntos sobre enfermedades y sus posibles curaciones provenían del antiguo pueblo egipcio; algunos de ellos son del año 2900 a. C. y han llegado a nuestro tiempo bastante fragmentados; no obstante, uno se ha logrado conservar aceptablemente y contiene un muy buen tratado de medicina antigua que corresponde al periodo del año 1500 a. C. aproximadamente, se trata del denominado *Papiro Ebers*, nombre dado por su divulgador, el egiptólogo alemán George Ebers en 1872.

La traducción de este documento muestra un tratado principalmente de medicamentos; su fuente era el estudio de las enfermedades y dolencias acaecidas en su tiempo a los habitantes de la ribera del río Nilo, generalmente causadas por picaduras de mosquitos y algunas patologías más complejas que se recopilaban bajo

el criterio de *pestilencias*. Las curas que se detallaban en este documento se obtenían por la observación combinada con el pensamiento “mágico” de la época.

Precisamente, uno de los hechos descritos en este documento era la narración del primer padecimiento colectivo de fiebres pestilentes, que hoy se asocia con la malaria, luego de los cambios en el caudal del río Nilo, el cual dejaba reservorios para la gestación de los vectores transmisores (Ebbell et al., 1937).

La aparición periódica de plagas en la antigüedad es indiscutible y así se registró posteriormente cerca del 1224 a. C. en libros sagrados como el Talmud, la Biblia o el Corán y es precisamente en estos textos en los que se comienzan a indicar las primeras normas para prevenir enfermedades contagiosas. Ya en el 450 a. C., en la antigua Grecia, Hipócrates plantea una definición de estos fenómenos que afectaban a la población de un lugar al referirse a los padecimientos propios del sitio o fuera de aquel. Hipócrates no convalidó la creencia popular sobre el contagio, y más bien atribuía la aparición de las enfermedades al ambiente malsano (la denominada “contaminación” que en griego era acuñada bajo el término **μῑασμα** que traduce como *miasmas*) y a la falta de moderación en la dieta y las reducidas actividades físicas (López-Moreno et al., 2000).

Es en esta época que se acuñó la expresión *epidemia* para describir los acontecimientos que padecía una población y cuyo origen provenían de fuera de su lugar de hábitat; en contraposición, se usó la expresión *endemia*, para describir los acontecimientos que padecía una población y que se originaban en su propio hábitat. Surge así el nuevo vocablo *epidemiología* que provenía del término **επίδημος** (*epidémos*), el cual a su vez emerge de **επι** (*epi*: sobre) y **δημος** (*démos*: pueblo); al generalizarse su uso, se le agregó el vocablo **λόγος** (*lógos*: estudio, tratado) y así se construye su etimología, la que se puede traducir como estudio o tratado de lo que sucede sobre la población.

Desde ese momento, muchos escritores griegos y posteriormente los latinos, se referían a menudo al surgimiento de lo que se había popularizado como “pestilencias”,

siendo la más famosa la plaga de Atenas que asoló la ciudad durante las guerras del Peloponeso por allá en el 430 a. C. y que fue descrita magistralmente por Tucídides (López-Moreno et al, 2000).

Muchos de esos procesos colectivos de enfermedad se podían considerar como fenómenos epidémicos, a los que la población, curiosamente, creía firmemente que eran contagiosos, contrario a lo que los médicos de la época pensaban y que no se habían ocupado de conceptualizar. Con el tiempo, esos procesos colectivos se fueron amplificando en el planeta y ya en el 401 d. C. se presentó una plaga que azotó al mundo conocido de la época; es entonces cuando dicho proceso fue formalmente denominado epidemia.

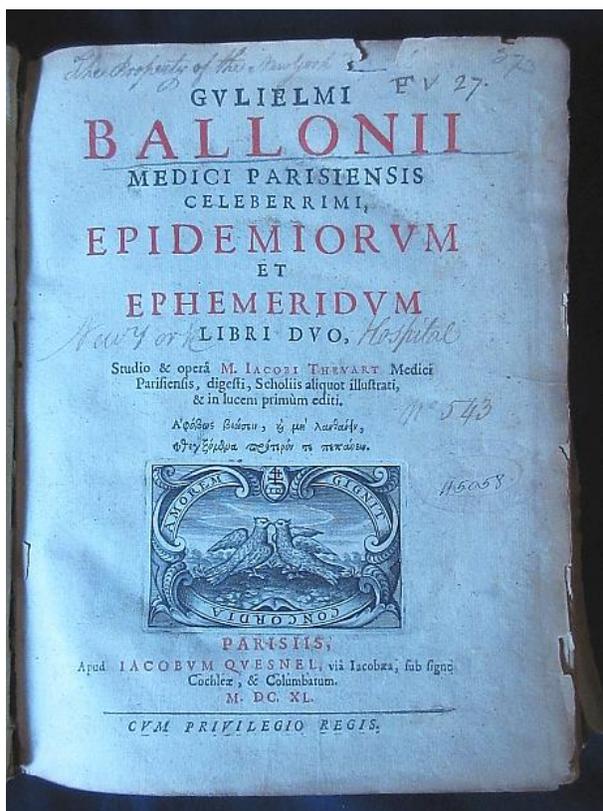


Ilustración 9. Portada *Epidemiorum*. Imagen de Leaf & Stone Books. https://www.leafandstonebooks.com/assets/images/product/9004_2.jpg

En ese momento de la baja Edad Media, el término se refería a la aparición de un número inesperado de casos de enfermedad y al comportamiento que presentaban algunas infecciones que cada cierto tiempo devastaban a las poblaciones.

Avanzando un poco más en el tiempo, por el 1580, el médico francés Guillaume de Baillou publicó el libro *Epidemiorum*, (ilustración 9) el cual contenía una relación sistemática y muy completa de las epidemias de sarampión, difteria y peste bubónica ocurridas en Europa entre 1570 y 1579, e incluía, además, las características de cada una y,

especialmente, su modo de propagación.

Unos veinte años después, el término *epidemiología* como tal se encontraría plenamente en el libro “Epidemiología” publicado por Quinto Tiberio Angelario en Madrid en el 1598. En 1606 se incorporarían al idioma español las expresiones epidémico y endémico, entendiendo en esa época a este último como aquellos eventos ocasionados a los residentes permanentes de un sitio causados por aires, aguas y/o el mismo lugar. Por el contrario, el término epidémico refería a los eventos ocasionados a quienes temporalmente residían en un lugar (López-Moreno et al., 2000).

Precisamente la aparición de la peste bubónica que azotó Europa llevó a reconocer globalmente la teoría del contagio (aún en el ámbito popular). La necesidad de comprender la naturaleza de las enfermedades y su desarrollo en la población llevaron a la elaboración de diversos estudios médicos en la época posterior al Renacimiento.

Conforme se presentaban estos estudios, el nacimiento de la estadística sanitaria coincidió con la búsqueda de un sistema lógico de clasificación botánica, lo cual sirvió para que Thomas Sydenham realizara una muy cuidadosa descripción clínica de enfermedades como la disentería, la malaria, la viruela, la gota, la sífilis y la tuberculosis entre 1650 y 1676. El trabajo de Sydenham se vio complementado y, en cierta forma, fortalecido por otro investigador inglés John Graunt quien por el 1662 comenzó a analizar los reportes de nacimientos y fallecimientos en la ciudad de Londres en un periodo de sesenta años desde 1600, logrando identificar un patrón constante en las causas de los decesos. Graunt daría así inicio a los modelos de análisis de datos para construir las actuales tablas de vida (López-Moreno et al, 2000).

Ya en el siglo XIX se inició el hábito de realizar trabajos para medir la distribución de una enfermedad en grupos determinados de la población; esos trabajos no solo se constituyeron en el comienzo formal de la epidemiología, sino que demostraron uno de sus logros más resonantes como lo fue la asociación que encontró el Dr. John Snow (*ilustración 10*) entre la ingesta de agua pública suministrada y el brote de la enfermedad del cólera. Snow había nacido en 1813 en York, Inglaterra, y en 1827 comenzó su aprendizaje médico; ya por el año 1832 Snow fue testigo de un contagio

de cólera en Killingworth un pueblo cercano, donde trató a muchos pacientes y se familiarizó con la presentación clínica de esta enfermedad y cómo parecía propagarse. El primer caso de cólera se había informado un año antes, en 1831, y se pensaba que su propagación se daba por “miasmas” (una forma venenosa de "mal aire" que surgía a partir de la materia orgánica en descomposición). Posteriormente, Snow se graduó en la Universidad de Londres en 1844 y fue admitido en el Royal College of Physicians en 1850. Snow estableció su práctica en Soho, un barrio de Londres, como cirujano y médico general (Ramsay et al., 2006).

Estando en Soho ocurrió el episodio de cólera en 1854. Para ese entonces ya se mostraba escéptico sobre la teoría del “miasma” y teorizaba firmemente que las aguas residuales vertidas en los ríos y pozos negros cerca de las ciudades podrían contaminar los suministros de agua y causar brotes de cólera. El trabajo de Snow comenzó hablando con los residentes locales y rápidamente comenzó a sospechar que la fuente del brote era la bomba de agua pública en Broad Street. Usó información del hospital local y registros públicos y preguntó específicamente a los residentes si habían bebido agua de la bomba. Con esta información, pasó a crear un mapa de puntos (*ilustración 11*) para ilustrar el grupo de casos alrededor de la mencionada bomba.

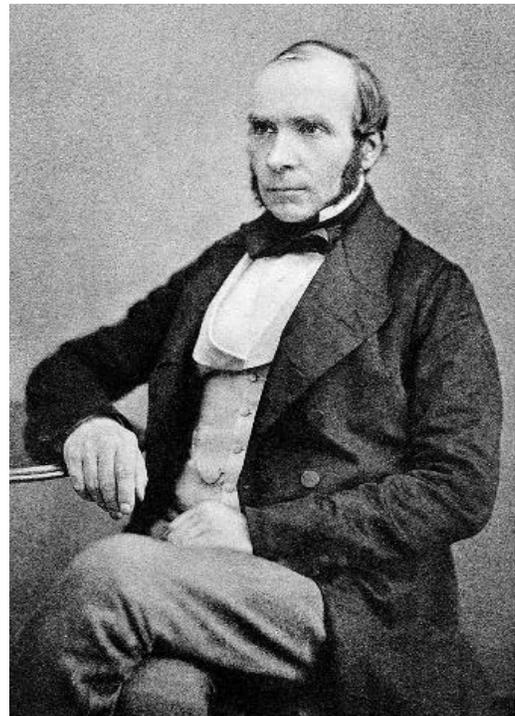


Ilustración 10. John Snow en 1856, Cortesía de Renato Sabbatini CC BY-SA 4.0

Escribiría entonces:

“Dentro de 250 yardas del lugar donde Cambridge Street se une a Broad Street hubo más de 500 ataques fatales de cólera en 10 días... Tan pronto como me familiaricé con la situación y el alcance de esta irrupción (sic) de cólera, sospeché alguna contaminación del agua de la bomba de la calle muy frecuentada en Broad Street.”

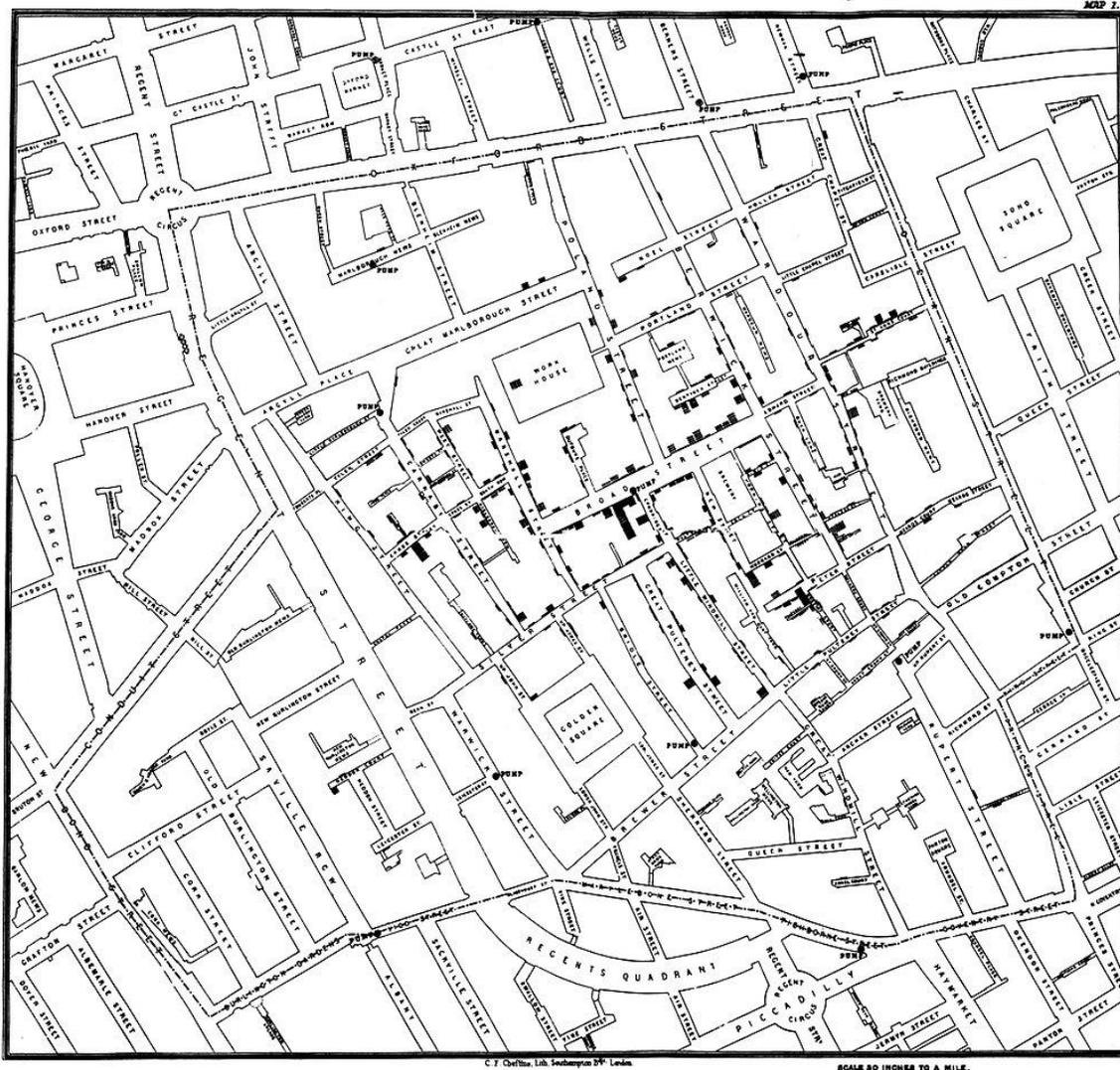


Ilustración 11 Mapa original de John Snow que muestra los grupos de casos de cólera en el brote de Broad Street, dibujado y litografiado por Charles Cheffins

En 1854, Snow presentó sus hallazgos ante las autoridades locales y los convenció de imposibilitar extraer agua de la bomba. Luego de esta acción, el brote de cólera disminuyó y llegó a su fin; se descubrió que el pozo público del que se extraía el agua estaba solo a unos metros de un pozo de aguas residuales al que había llegado un pañal de tela de un bebé que había contraído la enfermedad en otro sitio. He ahí el origen del brote (Barton et al., 2018)

Es importante resaltar este trabajo de John Snow, pues utilizó el análisis de datos estadístico mediante la comparación de las muertes por cólera en los diferentes puntos de suministro de agua, logrando de esta manera demostrar que el número de fallecimientos y, por ende, la tasa de mortalidad, era proporcional en las personas que recibían el abastecimiento de agua de esa bomba ubicada por la compañía Southwark de suministro de agua pública. Con este empleo del análisis de datos muy meticuloso, por cierto, Snow elaboró su teoría sobre el contagio de las enfermedades infecciosas en general, sugiriendo que, como el cólera, se propagaba a través del agua contaminada. Posteriormente en 1883, el médico alemán Robert Koch aisló la bacteria causante del cólera y determinó que se propagaba a través de fuentes de suministro de agua pública o de alimentos insalubres, respaldando así las teorías de Snow. Estos hallazgos llevaron a realizar cambios en el saneamiento y suministro de agua que finalmente terminaron con las epidemias de cólera tanto en Europa como en los Estados Unidos durante el final del siglo XIX (Barton et al., 2018).

Como se puede apreciar, el trabajo de Snow fue el primero en asociar una red de distribución de agua con una epidemia focalizada, y su “método” es parte de la denominada *observación numérica*, la cual representó un extraordinario avance para la época (siglo XIX); otros trabajos destacados son los de Robert Storrs (1840), Oliver Wendell Holmes (1842) e Ignaz Semmelweis (1848) sobre la transmisión de la fiebre puerperal; los de P.L. Panum (1846) sobre la contagiosidad del sarampión, y los de William Budd (1857) sobre la transmisión de la fiebre tifoidea. Estos trabajos, incluido el de Snow, muestran el enorme esfuerzo intelectual que estos investigadores debieron hacer para documentar mediante la pura observación propuestas sobre la capacidad transmisora, los mecanismos de contagio y la infectividad de agentes patógenos sobre los que aún no podía demostrarse una existencia real. El enorme valor de estos trabajos se encuentra en el hecho de que los agentes infecciosos responsables de cada una de estas enfermedades se descubrieron entre veinte y treinta años más tarde, en el mejor de los casos (López-Moreno et al., 2000).

El inicio de la metodología epidemiológica parte del análisis de datos sobre un asunto en particular generalmente asociado a una enfermedad. No obstante, en la actualidad, la epidemiología es una disciplina dinámica y su concepto no solo se restringe a las enfermedades microbianas, sino que abarca un espectro más amplio que incluye la salud y sus desórdenes, problemas nutricionales y fenómenos producto del comportamiento humano como los accidentes. Para muchos especialistas, es una manera de pensar, de técnicas para explorar las causas y orientar la prevención frente a cualquier causa de morbilidad y mortalidad, sin que necesariamente sea una enfermedad transmisible.

Puede entonces hacerse un primer enunciado consistente en que la metodología epidemiológica de análisis de datos incorpora el estudio de las condiciones o características de los individuos y las de su entorno, pues en ese conjunto de características o factores, podría estar la respuesta al porqué unos individuos se han enfermado y otros no, y cómo pueden mejorarse las obras civiles de su entorno para que no alteren su salud. A partir de esto conviene apreciar ahora la evolución histórica de las redes de distribución de agua.

2.1.2 Las redes de distribución de agua

La existencia del agua en la actividad humana se remonta a los albores del tiempo; podría decirse que el ser humano es “hijo” del agua pues es un hecho que la composición de un individuo es de un 70% de este elemento. Los primeros indicios del uso del agua por parte del hombre se han establecido cerca del 12.000 a. C., pues estudios arqueológicos han podido determinar que algunos grupos de cazadores – recolectores volvían cíclicamente a los valles fértiles de entonces.

De igual forma, se ha podido establecer que muchos grupos que habían hecho su transición de nómadas a sedentarios, se enfrentaron a un primer problema de escasez de agua por cambios climáticos y eventos geológicos alrededor del 7.000 a. C. (Sala

et al., 2018). Esto llevó a aquellos grupos a inventar *sistemas de riego* muy primitivos y de alcance muy limitado. La evolución de las comunidades comenzó en este instante temporal, pues requería de la unión de varios individuos en una labor que les daría beneficios en forma equitativa. El éxito aparente que se observó con los sistemas de riego llevó a tener los primitivos desarrollos hidráulicos respecto al uso del agua para consumo humano tal y como se ha podido apreciar en hallazgos datados en el 4.000 a. C. respecto a la potabilización del agua: hervir el agua, ponerla al sol, filtrarla a través de arena. En la medida que las comunidades se fueron agrupando y comenzó el desarrollo de las primeras ciudades de la antigüedad, también se iniciaron las soluciones para el suministro de agua a los pobladores; esto se registra en el primer sistema de “acueducto” en la ciudad sumeria de Nippur por el 3.750 a. C. (*ilustración 12*) en los inicios de la Edad de Bronce.

Este sistema arcaico consistía en una serie de canales conectados en forma permanente a los ríos circundantes, el Éufrates y el Tigris, dada su posición estratégica en medio de ellos.

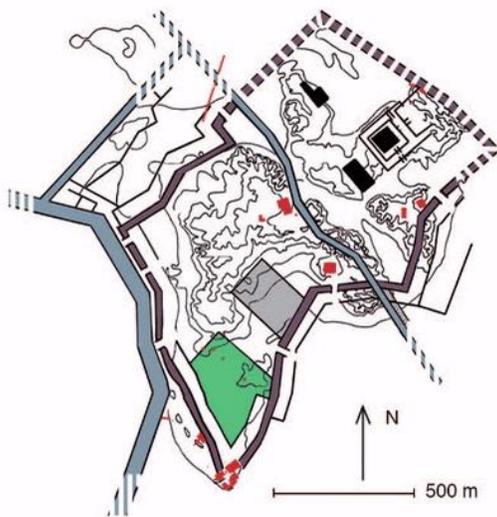


Ilustración 12. Mapa de Nippur. (Pedersén et al., 2010)

Al observar el plano de la ciudad de Nippur se pueden apreciar los mencionados canales de suministro, que para este caso están en color gris azulado, las murallas de la ciudad en color marrón, los jardines en color verde, los templos en color negro y los administrativos en rojo. Resulta claro que el ingenio humano siempre se desarrolla en función de las dificultades que debe afrontar de tiempo en tiempo; este sistema de suministro simple, debió satisfacer a la población de la ciudad que, según los registros de las tablillas encontradas, ascendía a unos 9.000 o 10.000 habitantes. Era una metrópoli de la época.

Pero los problemas surgen al crecer la población y más adelante en el tiempo, por el año 3.000 a. C., se presentarían las primeras grandes ciudades con un completo sistema de suministro de agua y paralelamente, de canalización de aguas residuales.

Tan imponentes ciudades son Mohenjo-Daro y Harappa en el fértil Valle del Indo, que actualmente corresponden a Pakistán. Estas ciudades, además de contar con un completo sistema de abastecimiento de agua, también poseían infraestructuras elaboradas para baños públicos, servicio de agua caliente y desagües independientes en sus viviendas. Estas ciudades, según los escritos encontrados, llegaron a albergar una población estable de 40.000 habitantes, algo sorprendente, pues mantener una población de esa magnitud en esa época implicaba una infraestructura y una organización muy definida. De hecho, fueron las primeras ciudades que urbanísticamente se erigieron mediante diseños de cuadrícula como las actuales ciudades donde vivimos, y sorprende saber que cada calle de Mohenjo-Daro o Harappa, tuviese su propio alcantarillado y sistemas de drenaje que eran, según las opiniones de algunos estudiosos del tema, más avanzados que los más recientes sistemas que se puedan encontrar en algunas poblaciones actuales de la India o del mismo Pakistán.

En efecto, las excavaciones arqueológicas han sacado a la luz edificaciones que *“tenían baño privado y retrete (cuarto refinado y acondicionado) que estaban comunicados por medio de sumideros (zanja o depósito que recibe y almacena temporalmente aguas en el punto inferior de un sistema de circulación) y tuberías de barro a un sistema de cloacas (tubo subterráneo o canal abierto en un sistema de alcantarillas para trasladar el agua sucia a la zona adecuada) que desembocan en enormes pozos sépticos. Las alcantarillas disponían de aberturas a intervalos para su limpieza”* (Fundación AQUAE, 2021). La ingeniería civil y urbanística de la época era, con todo, demasiado sofisticada, equiparable quizás con las obras europeas de comienzos del siglo XX.

Avanzando en el tiempo, se puede apreciar cómo los sistemas de redes de suministro se van convirtiendo en obras cada vez más grandes y complejas, tal y como se puede apreciar en el antiguo Egipto por el año 2.770 a. C., época en que se construyó el primer dique o presa de que se tenga conocimiento. Esta infraestructura se denominó Sad El-Kafara, que significa *presa de los infieles* en el lenguaje árabe y se ubica a 30 Km de la actual ciudad de El Cairo. Su construcción obedecía a la necesidad de controlar las inundaciones que ocasionaban las crecidas del río Nilo, con todo lo que esa situación conllevaba para la época. Tenía una altura de 14 metros y una longitud entre 90 y 100 metros; el grosor del muro era de unos 98 metros en su base y unos 56 metros en su cresta. Los trabajos duraron entre 10 y 12 años, según los escritos encontrados, pero al parecer una inundación anticipada del río la destruyó parcialmente a causa de la inexistencia de un aliviadero, lo cual condujo a abandonar el proyecto.

Los ingenieros egipcios desarrollaron con mucha lógica la parte superior, pero la inferior no era consistente para la magnitud de la obra que, de haber sido completada, hubiese sido capaz de almacenar aproximadamente entre 465.000 m³ y 624.000 m³ de agua. Este fracaso condujo a postergar obras similares por un espacio de ocho siglos. De haber tenido éxito, la presa podría abastecer a buena parte de las poblaciones vecinas de manera controlada, no solo para las redes de suministro, sino también para las redes de riego. Fue el primer proyecto pensado para un área mucho más grande que las ciudades de la época y para una población cercana a los 100.000 habitantes, incluyendo los animales domésticos.

Entrando al año 700 a. C. hace su aparición en Asiria el rey Senaquerib, hijo del gran Sargón II. Su reinado se caracterizó por continuas batallas heredadas con sus vecinos, pero su verdadero legado lo constituyó el auge e impulso a la arquitectura y a las obras civiles; sus logros fueron la reconstrucción colosal de la ciudad de Nínive, la gran capital de Asiria con templos, jardines, palacios y fortificaciones y lo más importante el denominado “acueducto de Jerwan”, gigantesca obra para suministrar agua a toda

la inmensa ciudad. La historiadora Gwendolyn Leick señala: “*Nínive, con su población heterogénea de personas de todo el Imperio Asirio, era una de las ciudades más hermosas del Cercano Oriente, con sus jardines, templos y palacios espléndidos*” (Leick, 2010, p.132) y además cita a Nínive como poseedora de una serie de canales y acueductos cuidadosamente planeados y ejecutados para asegurar un suministro constante de agua, no sólo para el consumo humano, sino también para mantener irrigados los parques y jardines públicos; una característica de la vida urbana que no todas las ciudades han atendido con tanto cuidado y planificación. Estudios recientes afirman que los famosos Jardines Colgantes de Babilonia en realidad estaban ubicados en Nínive y fueron construidos bajo el reinado de Senaquerib (Leick, 2010).

La red de suministro de agua de Nínive se alimentaba de los ríos Tigris y Khosr, pero al igual que el Nilo en Egipto, estos ríos tenían temporadas de sequías y se bajaba mucho su caudal, por lo que se buscaron manantiales y arroyos de las colinas al norte de la ciudad y fueron conducidos por una infraestructura de aproximadamente 10 Km en canal y piedra maciza hacia el río Khosr; esta era la magnitud del mencionado acueducto de Jerwan. Y como la ciudad de Nínive, el acueducto fue colosal, pues debía dar abastecimiento a una población cercana a los 600.000 habitantes.

Otro ingenio que surge en esta época en Nínive es la denominada “bomba de agua” que permitiría llevar el sistema de irrigación de los jardines a los niveles superiores de las construcciones, con lo cual no resulta descabellado indicar que una de las siete maravillas del mundo antiguo estaba en Nínive: los jardines colgantes de Babilonia.

Avanzando un poco más en el tiempo se llega a Roma y a la construcción de su primer acueducto, el *Aqua Appia* por el 312 a. C. Esta obra se constituye en el primer ejercicio de los ingenieros romanos que buscaba suministrar a la población de la ciudad un volumen cercano a los 75.000 m³ de agua en un recorrido aproximado de 16,4 km; lo sorprendente de esta estructura es que buena parte de su recorrido es subterráneo (*ilustración 13*) con una caída de unos 10 metros, situación que lo protegía de los continuos ataques que sufría la ciudad por las denominadas guerras

samnitas. La iniciativa de los dirigentes romanos, especialmente de los dos censores que gobernaban la ciudad (lo que hoy serían alcaldes) Gaius Plautius y Appius Cladius Grassus, era mejorar notablemente el saneamiento de la ciudad y aprovechar la infraestructura de plomería y alcantarillado interior, que convertiría a Roma en el líder mundial en redes de suministro y que otorgaría amplios beneficios en la salud de la población que solo serían superados en la actualidad. Pese a que este primer desarrollo de la ingeniería fue primitivo y poco sofisticado en comparación a los diez acueductos que tenía la ciudad, se mantuvo en servicio hasta los tiempos del emperador Augusto César (27 a. C.) mediante mantenimiento regular, renovaciones e incluso de una ampliación en la cobertura del mismo.



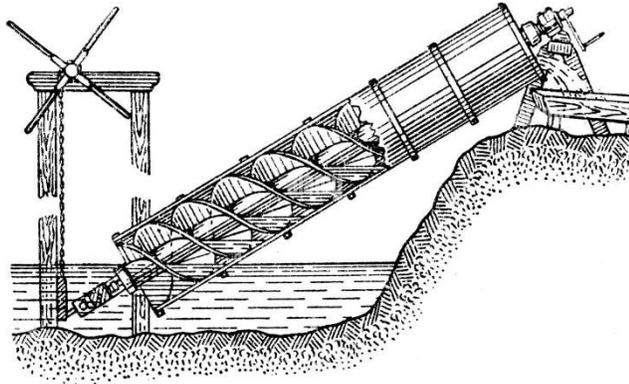
Ilustración 13 Una sección de probablemente 500 metros de largo del Aqua Appia en los barrios de Centocelle. Tomado de <http://www.romanaqueducts.info/aquasite/romappia/foto2.html>

la fundación de la ciudad, los romanos estaban satisfechos con el uso de aguas tales como las extraídas del Tíber, de pozos o de manantiales. Pero ahora corren por la ciudad el acueducto Apio, ...”

Sobre este mismo periodo de tiempo, en el año 300 a. C. surgen dos elementos muy importantes para una red de distribución de agua: el tornillo de Arquímedes (*ilustración 14*) y la rueda hidráulica de Filón de Bizancio. El tornillo de Arquímedes es una máquina helicoidal usada para elevar algún elemento a un nivel superior, se usó para extraer material excavado, cereales y por supuesto agua.

Este hecho histórico de las obras civiles romanas en el Aqua Appia fueron detallados por parte de Sextus Julius Frontinus, comisionado de agua, unos cien años después de su finalización; manifiesta en su libro, *Los Acueductos de Roma*, que “... durante cuatrocientos cuarenta y un años desde

El físico siciliano Arquímedes fue su inventor, pero dicha invención partió de los estudios que hiciera sobre la nombrada bomba de agua que mandó elaborar Senaquerib para el riego de sus jardines en Nínive.



El siguiente elemento es la rueda hidráulica, que se constituyó en el primer motor de movimiento en la

antigüedad; su nombre original era la **noria** y Filón de Bizancio la inventó con el fin de bombear agua.

Los ingenieros romanos se interesaron en este invento y lo mejoraron con la denominada rueda hidráulica horizontal que fue conocida después como el “molino romano” y el molino de rueda vertical de paletas o “molino de Vitruvio” nombrado así en honor al ingeniero Vitruvio que estudió y mejoró este tipo de maquinaria. Es de apreciar que el tipo de molino descrito por Vitruvio fue el más común y el de más largo uso por muchos siglos tanto en Europa como en Norte América.

El avance de los acueductos en Roma iba de la mano del aumento de la población, algo que en su momento no podía determinarse con precisión ya que las autoridades romanas censaban solo a los hombres para fines políticos y militares, nunca a las mujeres o los niños. Según la documentación existente, la población de Roma fluctuaba entre 452.000 y 486.000 ciudadanos (varones) lo cual permite hacer una estimación de cerca de 950.000 habitantes.

Esta población era la denominada población permanente, pero Roma contaba con un número significativo de población itinerante, con lo cual bien se podía afirmar que Roma fue una ciudad que alcanzaba a albergar más de un millón de habitantes entre el siglo II a. C. y el siglo I d. C. Satisfacer a una creciente población, implicaba una



Ilustración 14 Tornillo de Arquímedes. Tomado de <https://c8.alamy.com/compes/dyp2pw/tornillo-de-arquimedes-ilustracion-de-la-enciclopedia-sovietica-1926-dyp2pw.jpg>

serie de obras civiles muy exigente, de ahí que Roma fue la primera ciudad del mundo que tuvo diez acueductos en funcionamiento simultáneo.

Si la ciudad de Roma albergaba esta población, en la extensión total del Imperio se estimaba que tenía cerca de cuarenta y cinco millones de habitantes, un 20% de la población mundial en esa época y eso llevó a los ingenieros a construir una infraestructura de 507 kilómetros de acueductos de los cuales unos 435 kilómetros eran acueductos subterráneos, unos 15 kilómetros eran acueductos de superficie y unos 59 kilómetros eran soportados por construcciones arqueadas (Rodà de Llanza, 2014).

Las redes de suministro romanas se veían beneficiadas por el auge de las cañerías tanto para suministro como para recolección de aguas residuales. Su punto de expansión se vio hacia el año 100 a. C. momento en que además se implementaba la reciente creación del grifo.

Un hecho de salubridad que se desprende de estos acontecimientos en la elaboración de las redes de suministro, es que el agua de los ríos se contaminó con plomo en unas cantidades muy superiores a los presentados en las fuentes de captación de los acueductos. Según esos estudios realizados en muestras tomadas cerca de la desembocadura del río Tíber, más exactamente en el sitio del puerto de Trajano, arrojan como resultados que el agua de las cañerías en la Antigua Roma contenía altos niveles de plomo. Los manantiales no presentaron índices peligrosos de metales pesados. En ningún momento se habla de si esa contaminación pudo ocasionar problemas de salud a la población de la ciudad, pero un análisis de la documentación conservada y posterior al Imperio Romano, permite concluir que el plomo fue una de las varias razones que produjeron el colapso de la civilización romana (Chappuis, 2014).

Avanzando en el tiempo se pueden apreciar otros avances en las redes de suministro de distintos sitios del planeta, como en Mesoamérica con los desarrollos de la cultura maya o en el sudeste asiático con los impresionantes desarrollos de Angkor Wat en

Camboya (Kampuchea). Las obras civiles en estos sitios se hicieron entre el año 1.000 y 1.100 de nuestra era. Sobre los Mayas cabe decir que destacaron por sus sorprendentes avances en escritura, arte, arquitectura, matemáticas y astronomía. Respecto a las complejas obras para redes de suministro salta en primer lugar las estructuras construidas en la ciudad de Palenque (tierras altas de Chiapas en México). Esta ciudad se ubica en una estrecha meseta de piedra caliza a una altura de 150 metros sobre las llanuras de Tabasco.

Es una alta escarpa que, además de brindar una excelente posición defensiva en las constantes guerras de la época, otorgaba una fuente de manantiales naturales, 56 en total, que generaban nueve cursos de agua separados, de ahí justamente el nombre de Palenque (Lakam-há: que en maya significa tierra donde las aguas fluyen) que se encuentra en su libro épico el Popol-Vuh. Esta situación resultaba muy atractiva para sus residentes pues la ciudad siempre contaba con agua aún en las épocas de sequía estacional. Sin embargo, demasiadas corrientes de agua dentro de un área limitada donde había casas y templos llevaban a inundaciones frecuentes de la plaza y zonas residenciales aún en la temporada de sequía anual. Esto hizo que los ingenieros mayas construyeran un sistema único de control de agua, canalizando el agua bajo las plazas. Fueron los primeros trabajos registrados en que las redes de suministro abordaban problemas no solo de cobertura, sino de sobreoferta.

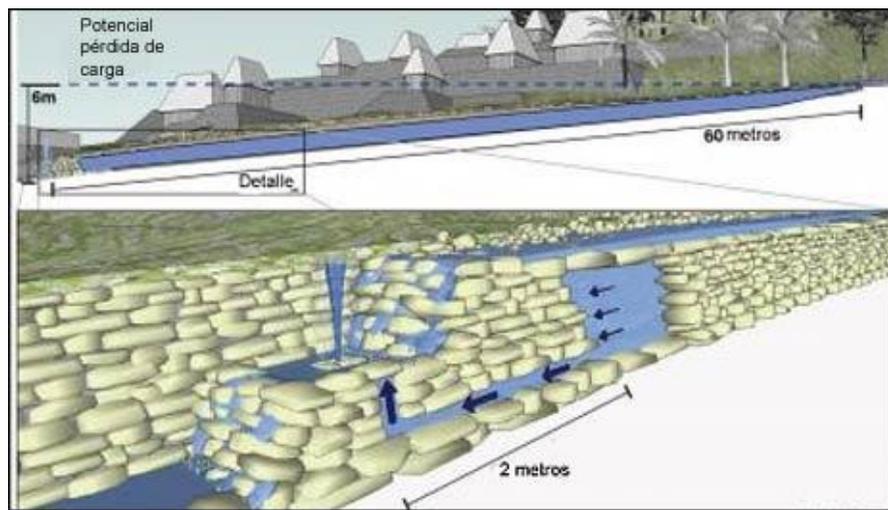


Ilustración 15. Sistemas de agua en Palenque. Tomado de <https://haraldviking.blogspot.com/2012/11/obras-hidraulicas-y-acueductos-mayas.html>

La característica del control del agua en la mayoría de los sitios mayas era una constante, principalmente para tener agua almacenada en la temporada de sequía para sus pobladores; pero el caso de Palenque era contrario; en él se aprovechaba el agua con elaborados acueductos subterráneos que guiaban la corriente por debajo de sus edificaciones; esto significaría que los ingenieros mayas lograron generar agua a presión antes de la llegada de los españoles (*ilustración 15*).

A pesar de todo ello, las principales ciudades mayas colapsaron en parte por su sistema político de ciudades-estado y en parte por severas sequías continuas y prolongadas que arruinaron las cosechas; siendo estos sitios abandonados.

El caso de Angkor Wat es distinto, pues las obras civiles que desarrollaron los ingenieros de la cultura Khmer fueron sorprendentemente colosales para la época. Angkor Wat es un complejo de templos erigidos en el 1.100 por el rey Suryavaman II y se encontraba a 1,7 km de su capital Angkor Thom, la cual fue considerada por los primeros viajeros occidentales que arribaron en el año 1.600 “tan fantástica como la Atlántida de Platón”. La construcción de Angkor Wat tardó aproximadamente 30 años y contaba con un área de 1.500 m², era una construcción rodeada por un enorme foso de 195 m de ancho.



Ilustración 16. Templo de Angkor Wat. Crédito: Charles J Sharp/Wikimedia Commons

Esta construcción tiene la característica de ser toda en piedra; el foso, tal y como se aprecia en la imagen (*ilustración 16*), es perfectamente cuadrado y es precisamente su hechura artificial lo que evita que el complejo se hunda ya que el agua se filtra bajo la superficie endureciendo la tierra y la roca, lo cual era una forma natural de hacer concreto. Los ingenieros Khmer tuvieron que realizar cálculos muy precisos, pues un error hubiese inundado el complejo y lo hundiría. La complejidad de la red de suministro en esta estructura comenzaba por las denominadas “rocas huecas” que iban dentro de las construcciones llevando las aguas de lluvias a los diferentes puntos del complejo.

No utilizaban cañerías como se conocía en Roma, pero horadaban la roca artificialmente para hacer este tipo de canalizaciones. Al igual que la capital Angkor Thom, los ingenieros Khmer planificaban sus obras en cuadrados monumentales. La ciudad de Angkor Thom, tenía 9 km² de superficie. La red de suministro en este espacio ocupaba una serie de intrincados canales superficiales y subterráneos que abastecían satisfactoriamente a una población cercana a 150.000 habitantes. Los canales subterráneos tenían el mismo origen que las “rocas huecas” de Angkor Wat. Extrañamente, tanto Angkor Thom, como Angkor Wat fueron abandonados hacia el 1.610 (Bonet, 2008).

Curiosamente después de tan grandes logros de ingeniería, los herederos de los Khmer volvieron a sus ancestrales sistemas de distribución; un hecho muy similar al de los constructores de las pirámides que llegaron a un muy alto nivel y sus herederos “olvidaron” como hacer esas construcciones. Avanzando un poco más en el tiempo vemos la evolución en la gestión del agua nos lleva al periodo comprendido entre 1800 y 1900 en el que se hacen populares las obras civiles relacionadas con la construcción de presas en todo el mundo conocido. Es importante apreciar que, a partir de los descubrimientos e inventos del pasado, en esta época se perfeccionaron aquellos de modo que el uso del agua no solo se hace para abastecer a una población del vital líquido, sino que también su aplicación a la producción de energía eléctrica, por el

accionar del agua en turbinas enormes, se vuelve una constante. Hay entonces dos tipos de presas, las que generan electricidad y las que alimentan uno o más acueductos y sistemas de riego.

El objetivo se centra ahora en mejorar la distribución del agua no solo para consumo humano, sino para producción agrícola e industrial; no obstante, la calidad del agua se afecta por los desarrollos de productos químicos para la industrial y la agricultura, pues deterioran el medio ambiente con las aguas residuales alterando así el ciclo del agua. Basta con recordar que esto ya se observó en tiempos del esplendor de Roma con el aumento de metales pesados en puntos de la desembocadura del principal río de la ciudad.

En los últimos cincuenta años, el trabajo generalizado transcurre por dos vías: la mejora en la calidad del agua suministrada, incluyendo también su cobertura, y la percepción global en el cuidado del agua dándole una naturaleza jurídica a la misma e incluyéndola dentro de sus propias constituciones políticas. Entre un 35% y un 40% de los países han dado este paso jurídico y político.

2.2 ESTADO DEL ARTE

2.2.1 El papel de la Epidemiología y sus herramientas de análisis de datos

Como se ha podido apreciar, la historia de la epidemiología se ha sucedido en forma paralela al desarrollo social de las poblaciones y éstas al ritmo de los avances en la gestión del agua. Puede decirse que un punto de partida de tipo académico se inicia en 1580 cuando se documentaron en forma precisa y detallada, los contagios masivos de sarampión, difteria y la peste bubónica. Lo más destacable en esa documentación consistía en la descripción de las características de cada contagio y especialmente, su propagación.

Precisamente, estudiar el modo de propagación de las enfermedades fue determinante para desarrollar unas metodologías que permitieran recolectar datos sobre las causas o circunstancias, que daban origen a esos contagios. El perfeccionamiento de estas metodologías y la búsqueda continua de un orden lógico para las clasificaciones de las enfermedades llevaron al nacimiento de la estadística sanitaria en 1.660, momento en el que se comienzan a realizarse los respectivos análisis de datos relacionados con los patrones de decesos en Londres. Surgen así los primeros modelos de análisis de datos epidemiológicos.

De forma más exclusiva, se puede apreciar una diferenciación en los datos que entraban en un estudio, independientemente del tipo que fuera. Se habla entonces de los datos *cuantitativos* y de los datos *cualitativos*. Durante cerca de trescientos años (1.660 a 1.950), todos los modelos partían y llegaban a datos cuantitativos; técnicamente el uso de este tipo de datos generaba conclusiones confiables y, ciertamente, válidas.

Mientras que con los datos de tipo cualitativo ocurría lo contrario. En 1.979, Miles señalaba que un analista cualitativo tenía pocas herramientas para cuidarse de un autoengaño o de obtener conclusiones poco fiables; igualmente, los métodos de análisis rara vez describían con detalle cómo un investigador obtuvo sus conclusiones finales a partir de, por ejemplo, 3.600 páginas de notas de campo (Huberman & Miles, 2000).

Actualmente la situación se ha equilibrado considerablemente, pues existen un gran número de publicaciones y de conferencias relacionadas con lo cualitativo, así como grupos de trabajo especializados en áreas de estudio y, por supuesto, software para computadoras. Puede afirmarse que ahora, el trabajo de análisis de datos se comparte en forma explícita entre diversos sectores del conocimiento.

Un ejemplo de esta forma de compartir el análisis de datos se aprecia en que se buscan nuevas estrategias y/o metodologías basadas en la integración de los Sistemas de Información Geográfica (SIGs) como forma de georeferenciación espacial y visualización de las variables analizadas, junto con métodos de aprendizaje automático (Machine Learning) que permitan entender características puntuales, variables influyentes y dinámicas de los sistemas de abastecimiento de agua potable (Navarrete-López et al., 2017). En este ejemplo, el objetivo consiste en permitir que los gestores de abastecimiento localicen fallos en la red y puedan contar con una herramienta para la toma de decisiones ante esos fallos.

La forma en que se realiza la investigación en epidemiología siempre combina el conocimiento de la ciencia biológica con la ciencia social y, por lo mismo, técnicas cuantitativas como cualitativas para el análisis de datos; siendo estas últimas de relevancia por ser más flexibles y por tener en cuenta enfoques socio-ambientales que brindan respuestas a cuestionamientos en situaciones complejas.

Al hacer una revisión de las publicaciones que tratan específicamente sobre los métodos epidemiológicos para análisis de datos, se encuentran algunos elementos interesantes de abordar.

Pero antes de entrar en el tema se describirá brevemente el trabajo realizado: Se consultaron dos bases de datos de referencias *Scopus* y *Web of Science*, se hizo el análisis de clústeres y gráficos de producción para entender el campo de interacción de los mencionados métodos epidemiológicos.

La idea fue tomar el término de búsqueda literal: “*epidemiology*” AND “*data análisis*” OR “*complex networks*” y observar su comportamiento en clústeres. La primera aproximación se hizo a través de la herramienta online carrot2.org y arrojó un resumen que se puede apreciar en la siguiente ilustración (17).

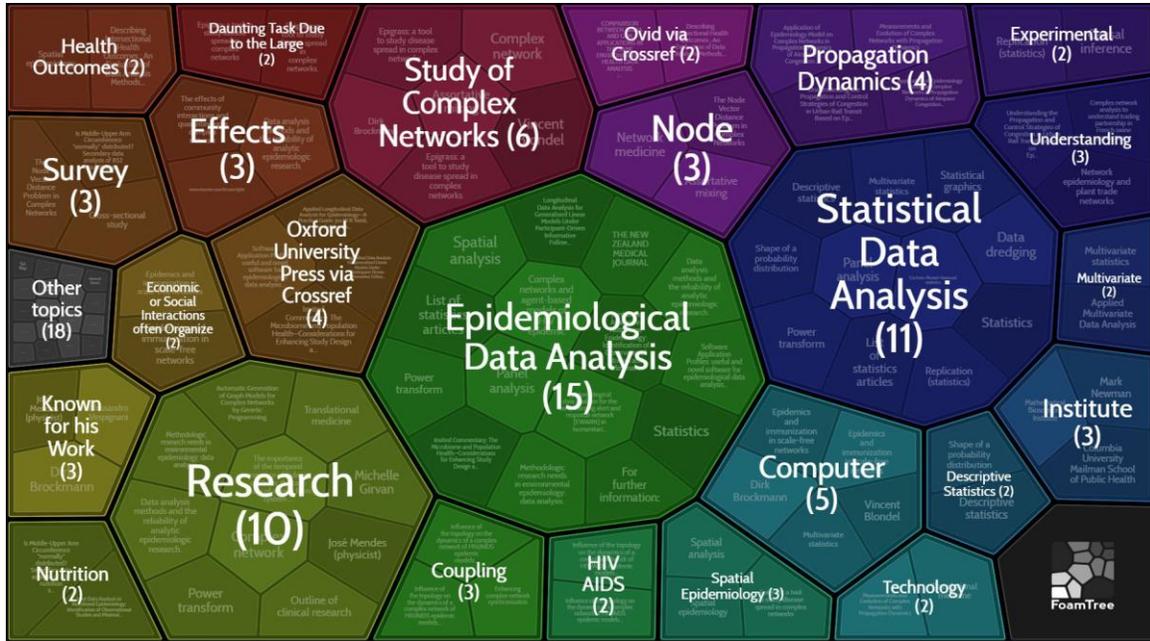


Ilustración 17. Métodos epidemiológicos para análisis de datos. <https://carrot2.org/>

Sin embargo, esta aproximación maneja unas 79 publicaciones. Una vez que se obtienen los resultados en las dos bases de datos mencionadas (*Scopus* y *WoS*), la consulta arroja unas **288.097 publicaciones** con las cuales ya se puede analizar mucho mejor las interacciones entre los temas que trabajan los *métodos epidemiológicos para análisis de datos*.

La tabla 1 resume en qué campos la interacción de las herramientas epidemiológicas son más utilizadas.

Si bien los métodos epidemiológicos para análisis de datos tienen su fuerte en categorías relacionadas con la salud y especialmente con la salud pública; también se están implementando en publicaciones de otras áreas, aunque con menor impacto.

CATEGORÍAS	PORCENTAJE DE PARTICIPACIÓN	PUBLICACIONES
Engineering Electrical Electronic	7,03	19343
Biochemistry Molecular Biology	5,46	15019
Multidisciplinary Sciences	5,30	14580
Chemistry Multidisciplinary	4,59	12620
Computer Science Information Systems	4,08	11233
Neurosciences	4,05	11129
Computer Science Artificial Intelligence	3,90	10718
Chemistry Inorganic Nuclear	3,86	10615
Public Environmental Occupational Health	3,49	9609
Cell Biology	3,42	9396
Materials Science Multidisciplinary	3,25	8947
Telecommunications	3,14	8646
Chemistry Physical	2,79	7673
Environmental Sciences	2,75	7563
Oncology	2,75	7560
Physics Multidisciplinary	2,71	7457
Crystallography	2,65	7298
Computer Science Interdisciplinary Applications	2,62	7206
Mathematical Computational Biology	2,35	6469
Medicine General Internal	2,32	6384
Genetics Heredity	2,17	5976
Biochemical Research Methods	2,08	5710
Infectious Diseases	2,03	5573
Biotechnology Applied Microbiology	2,02	5558
Computer Science Theory Methods	1,90	5219

Tabla 1. Categorías en las que se aplican métodos epidemiológicos de análisis de datos

Retomando el número total de las publicaciones y aplicando la herramienta *VOSviewer*, se pudo tener una visualización de mapa de interacciones (*ilustración 18*) mediante técnicas de diseño y de agrupamiento *VOS*. La primera hace referencia a la *fuerza de la asociación*, la cual normaliza la fuerza de los vínculos entre elementos del mapa y la segunda realiza la normalización del mismo modo que en la técnica de disposición y modularidad técnica de agrupamiento.

En esta ilustración se puede apreciar la fuerza de asociación existente entre los elementos relevantes de las publicaciones referentes a los métodos epidemiológicos para análisis de datos; independientemente del área que se desarrolle, los elementos que destacan son análisis, resultado final, análisis de supervivencia, meta análisis, revisión y evaluación. Aparecen dos elementos como son China e Irán y resultan relevantes sobre todo en los años 2020 y 2021 debido a los eventos de la pandemia de Covid-19.

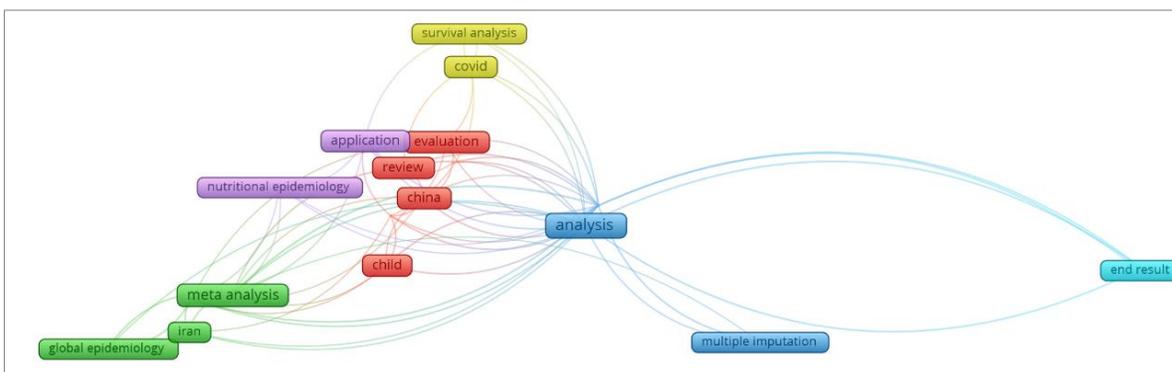


Ilustración 18. Fuerza de asociación de elementos relevantes para métodos epidemiológicos de análisis de datos. VOSviewer

En la siguiente ilustración (19) se muestra el comportamiento de la producción de estas publicaciones tomando solo cuatro fuentes como son Plos One, Statistics in

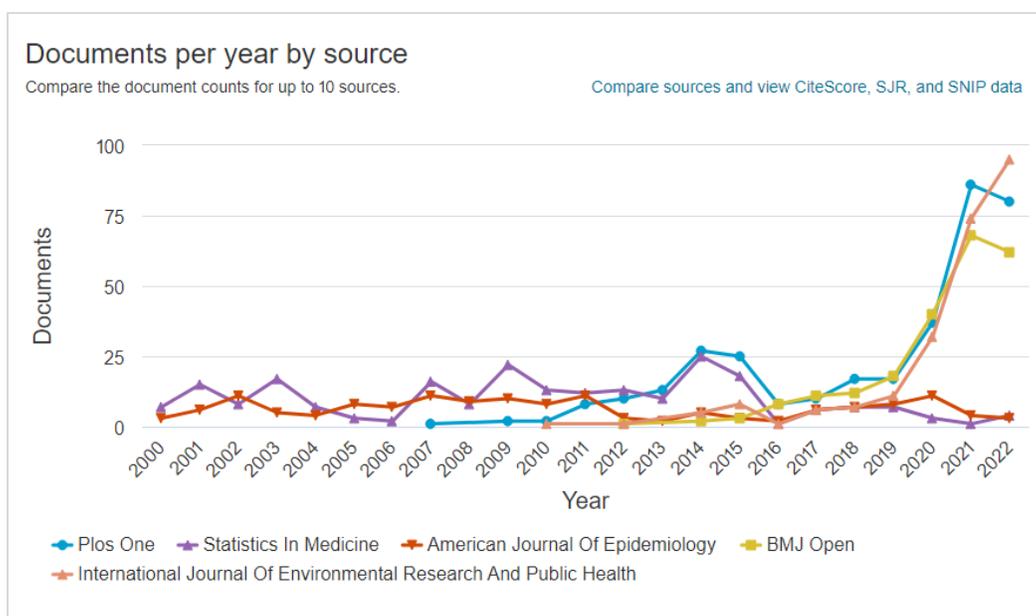


Ilustración 19. Comportamiento de producción para métodos de análisis de datos. Fuente: Scopus

Medicine, American Journal of Epidemiology, BMJ Open e International Journal of Environmental Reserch and Public Health.

Resulta claro que el papel de la epidemiología en la dinámica de las demás áreas de conocimiento, gracias al desarrollo de métodos de análisis de datos de tipo cualitativo, se ha incrementado en los últimos veinte años, permitiendo encontrar soluciones a diferentes problemáticas que la sociedad tiene con su entorno. Precisamente, la mayor participación de los métodos epidemiológicos se tiene en la salud pública y ambiental, lo cual permite hacer una conexión entre la epidemiología y la gestión de agua urbana. Se busca establecer ese nexo entre epidemiología e hidráulica, para evidenciar posibles aspectos en común y herramientas que se puedan explorar, estudiar y adaptar a la gestión de los sistemas de distribución de agua urbana, bajo un enfoque de análisis inteligente de datos y tradicionalmente usadas en el área de la epidemiología.

Hay ejemplos de cómo se ha venido adaptando este concepto en sistemas de distribución de agua urbana y dentro del contexto de Data Science.

Se presentan diferentes metodologías de análisis de datos usadas en Data Mining que han sido tradicionalmente aplicadas a análisis basados en el área epidemiológica, y que se abordarán para la gestión de los sistemas de distribución de agua urbana (Navarrete López).

2.2.1.1. Métodos heredados de la epidemiología

Como ya se indicó, la investigación en epidemiología siempre combina los conocimientos de la ciencia biológica con la ciencia social; sin embargo, los métodos de análisis de datos presentan variaciones en virtud de los campos de acción en que se desenvuelve la epidemiología como tal, a saber: la epidemiología descriptiva, la epidemiología analítica y la epidemiología experimental.

Estos campos se aplican en gran medida dentro del desarrollo clínico y epidemiológico de los eventos sanitarios. Para el caso de la gestión del recurso hídrico, la aplicación del campo de la epidemiología descriptiva es el que mejor se adapta, pues realiza estudios sobre los comportamientos de un evento dentro de una comunidad. La epidemiología descriptiva propone tres elementos concretos: frecuencia, magnitud y enfermedad. En el área de la gestión del recurso hídrico estos elementos se pueden asimilar así: la enfermedad con los fallos en el suministro de agua pública a causa de la mala adaptación a su ambiente interno y/o externo; la frecuencia con el número de casos de un fallo presentados en un tiempo y lugar determinado; y, por último, la magnitud con el número elevado de fallos.

Lo anterior refuerza la necesidad de promover los estudios epidemiológicos de corte descriptivo, como diseño que permite determinar “qué está ocurriendo” con la salud y la enfermedad en poblaciones específicas según Arias et al. (2020) y en nuestro caso con los fallos y el suministro de agua.

La epidemiología descriptiva se sirve de varios métodos de análisis de datos, como los métodos estadísticos, evolutivos, predictivos, etc. A manera de resumen, en la siguiente ilustración (20) se muestran los principales métodos heredados de la epidemiología que son aplicables a la gestión del recurso hídrico.

Estos métodos cuentan con diversas herramientas que ya han sido implementadas en diversos procesos en la gestión del agua. Por ejemplo, las ecuaciones diferenciales se han usado ampliamente para el establecimiento de métodos de modelamiento matemático; los modelos estocásticos y de optimización se han usado para apoyo de la administración y planificación de oferta y demanda de agua. Por otra parte, las series temporales, que se construyen a partir de la toma de datos a mediano o largo plazo, tienen por objeto identificar patrones o tendencias en las cadenas de información y los modelos poblacionales y los modelos evolutivos desarrollan soluciones más robustas ante modelos de simulación complejos.



Ilustración 20. Métodos heredados de Epidemiología

Para finalizar, los modelos de supervivencia se han aplicado para la estimación de la vida útil de los componentes del sistema de redes o para la estimación de fallos; mientras que las redes complejas se han utilizado en el estudio de la estructura de las redes y de las interacciones entre sus componentes.

2.2.1.2. La pandemia de COVID-19

A lo largo de su historia, la epidemiología generalmente se ha focalizado en una población determinada o en un espacio geográfico específico. Este patrón ha sido constante salvo contadas excepciones que se enumeran a continuación:

1. La *peste negra* en el siglo XIV, se estimó que fallecieron unos 25 millones de personas en Europa.
2. La *gripe española* entre 1.918 y 1.919, se estimó que fallecieron unos 50 millones en el mundo.
3. La *gripe asiática* en 1.957, se estima que falleció un millón de personas.
4. La *gripe de Hong Kong* entre 1.968 y 1.969, la H3N2 se cobró la vida de unas 40 mil personas.
5. El *VIH-SIDA* en 1.980, se ha cobrado la vida de 25 millones de personas en todo el mundo.
6. La *viruela* a lo largo de la historia humana se ha cobrado la vida de millones de personas; no obstante, la OMS ya ha declarado que esta ha sido erradicada en todo el mundo.
7. El *cólera* a lo largo de la historia humana, que fue muy fuerte en el siglo XIX donde fallecieron entre 30 y 40 mil personas; pero ya ha sido erradicada en la mayoría de los países industrializados, quedando algunos remanentes en zonas marginadas de África o Sudamérica.
8. La influenza AH1N1 entre 2.009 y 2.010, que fue una pandemia que afectó a 35 países del mundo, de los que se reportaron fallecimientos en 26 de ellos; cerca de 20 mil personas.

Este recuento sirve de preámbulo al último evento sanitario global originado por el virus SARS-CoV-2 y conocido como la pandemia de COVID-19, el cual surgió en noviembre de 2019 y aún persiste a lo largo del mundo con sus variantes. Se ha calculado un número de fallecidos cercano a los 6,4 millones de personas. Esta pandemia, a diferencia de las anteriores, ha traído un cambio radical en los comportamientos sociales, políticos, médicos, medioambientales, entre otros.

Una vez más, la epidemiología resulta de vital importancia para comprender y contener la COVID-19. El aporte en este caso se inició con la aplicación de estudios de cohorte, estudios de casos y control y estudios transversales. De acuerdo con cada

región se han utilizado diversos enfoques que permiten gestionar la información epidemiológica sobre el COVID-19 como, por ejemplo, los grafos de conocimiento (Delgado Fernández et al., 2021), o los modelos estadísticos para las predicciones de COVID-19 (Escobar et al., 2020).

Todo lo relacionado con la pandemia ha sido ampliamente documentado, tal y como se puede apreciar en la siguiente ilustración (21), la cual muestra el volumen de publicaciones durante la aparición del COVID y la fecha actual.

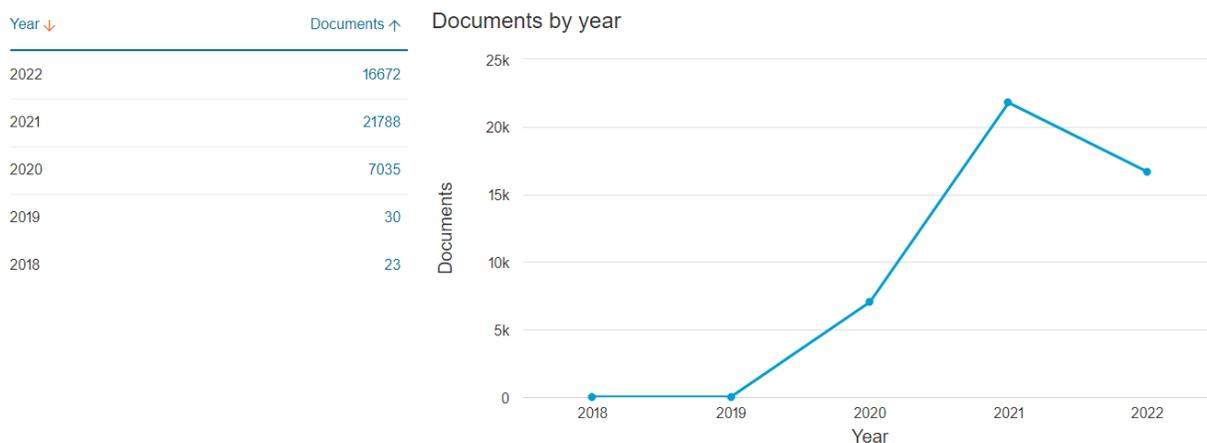


Ilustración 21. Volumen de publicaciones COVID. Fuente: Scopus

Como puede apreciarse, la literatura existente sobre este evento es abundante. En la siguiente ilustración (22) se puede apreciar en forma resumida, qué países son los que más publican respecto al COVID.

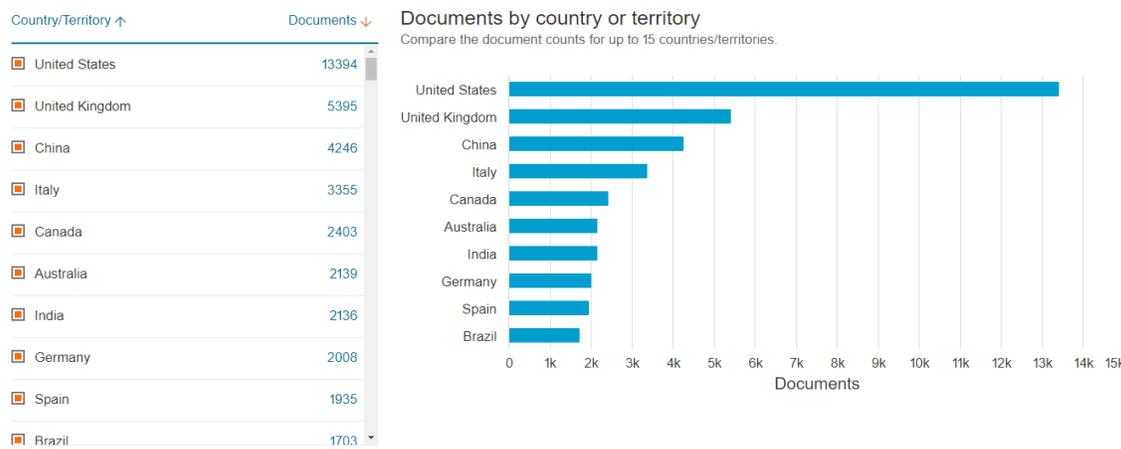


Ilustración 22. Países que más publican sobre COVID. Fuente: Scopus

Buena parte de la literatura sobre el COVID-19 relacionada con los métodos de análisis de datos epidemiológicos hacen referencia a los modelos matemáticos SIR, SIER y SIRD. El primer modelo fue planteado en 1.927 por W. O. Kermack y A. G. McKendrick y muestra los estados de los individuos en un espacio de tiempo frente a un contagio: (s)uceptibles, (i)nfectados y (r)emovidos. Este modelo evoluciona al SIER que muestra un estado adicional: (s)uceptibles, (i)nfectados, (e)xpuestos y (r)ecuperados; simultáneamente se maneja el modelo SIRD, en el que se contiene como variable los (d)ecesos (Polo et al., 2020).

La ausencia de cualquier conocimiento previo sobre el SARS-CoV-2, el virus que causa la enfermedad del COVID-19, necesariamente retrotrae los métodos de análisis que desde hace más de un siglo han sido eficientes en el estudio de este tipo de eventos.

Después de comenzar los estudios mediante observación directa, se registran los parámetros de comportamiento del virus, lo cual permite utilizar posteriormente los métodos propios de la epidemiología descriptiva que han servido para determinar los modelos de evolución-contagio entre la población y su predicción, siendo esto práctico pues se ha podido, en cierta forma, predecir los denominados picos de contagio.

La aplicación de estos modelos y de los métodos de análisis de datos sobre este evento sanitario del COVID-19, puede ser útil para los estudios de la gestión de fallos en los sistemas de distribución de agua pública, en especial el uso del modelo SIER.

2.2.1.3. Epidemiología de la energía

En la introducción a este trabajo, se indicó que el enfoque sistémico de la “Epidemiología de la energía” (Hamilton et al., 2013) se centra en la investigación empírica, proporcionando un marco metodológico que involucra profesionales en diversas áreas: físicos, ingenieros, sociólogos y economistas, en un trabajo

interdisciplinario que permite observar y describir patrones de demanda de energía y demás asociaciones entre los factores que la determinan.

En el documento “Uso de métodos epidemiológicos en la investigación de energía y edificios para alcanzar los objetivos de emisiones de carbono” (Hamilton et al., 2017) se indica que, a medida que los gobiernos se desplazan hacia intervenciones sectoriales a gran escala, se necesita una base de investigación y evidencia para apoyar el desarrollo, la implementación y la evaluación continua de la política de demanda de energía. La reducción de la demanda de energía de los edificios es ampliamente reconocida como un componente clave de los gases de efecto invernadero.

Los datos de requisitos previos de las existencias de edificios necesarios para respaldar este cambio esencial en el rendimiento energético de los edificios no están necesariamente disponibles o son inaccesibles, o incompletos. Se recopila más información sobre el uso de energía de los edificios a través de sensores de alta frecuencia y la forma del edificio. Los métodos de análisis se aplican a los innumerables y diversos subsectores de la “población” del parque inmobiliario debe ser proporcional a la heterogeneidad del parque inmobiliario.

Construir la epidemiología de la energía plantea que el desarrollo y la aplicación de un enfoque epidemiológico para investigar la demanda de energía puede mejorar la comprensión de los factores interrelacionados para la orientación y evaluación de políticas y proporcionar información sobre los mecanismos que influyen en la demanda de energía.

2.2.2 El papel del agua urbana y la gestión del recurso hídrico

El anterior acápite muestra cómo los métodos de la epidemiología tienen gran utilidad en las diferentes áreas de conocimiento y que cada día que pasa adquieren más relevancia e importancia, especialmente los métodos que incorporan técnicas cualitativas. Centrándose ahora en un campo específico como lo es el del agua urbana

y su gestión, es conveniente ver qué elementos de esas metodologías pueden adaptarse y cuales adoptarse.

En el libro “Desafíos del Agua Urbana en Las Américas: Perspectivas de las Academias de Ciencias” escrito por (Vammen, 2015) se hacen unos cuestionamientos que vale la pena citar en este estudio:

- ¿Se pueden solucionar los problemas de abastecimiento de agua y saneamiento urbano mediante una mejor gestión de los mismos?
- ¿Es posible dar solución a los retos de mejora de saneamiento y gestión de aguas residuales?
- ¿Puede mejorarse la atención que se presta actualmente a los problemas de salud y enfermedades transmitidas por el agua en las zonas urbanas?
- ¿Cuáles son los desafíos de adaptación al cambio climático relacionados con el agua en las zonas urbanas y cómo pueden solucionarse?
- ¿Cuáles son los modelos y conceptos a seguir que contribuyan a mejorar la gestión del agua en las zonas urbanas?

Dar respuesta a estos cuestionamientos, o hacer una aproximación a sus posibles soluciones, es importante pues actualmente cerca del 60% de la población mundial vive en las ciudades y ese número sigue en aumento. En la introducción plasmada en este trabajo se afirma que el abastecimiento de agua y, por lo mismo, el acceso a la misma debe cumplir con condiciones de calidad, cantidad y continuidad; el uso global del agua se ha multiplicado seis veces más en el último siglo y la proyección es que siga aumentando por causa del crecimiento demográfico y el cambio en los patrones de consumo. A este ritmo, y si la situación no cambia, el mundo se enfrentará a un déficit hídrico cercano al 40% para el 2030 (Hídricos, 2020).

Esa afirmación sobre el incremento del uso del agua frente al crecimiento demográfico se hace más palpable al repasar la historia y ver los espacios de tiempo tan amplios que han existido; basta con mirar la evolución del agua urbana en la ciudad de Nippur para abastecer a una población que estaba entre 9.000 y 10.000 habitantes hacia el año

3.700 a. C. y cómo tres siglos después aparecen ingenios hidráulicos en Mohenjo-Daro para abastecer a una población cercana a los 40.000 habitantes. El esplendor de las construcciones civiles para abastecer agua urbana no solo a sus habitantes sino también para elementos ornamentales, se presentaría en Nínive cerca de dos mil años después en el 700 a. C. para abastecer a una población cercana a los 600.000 habitantes. Y después de transcurridos dos siglos y medio, tenemos a Roma y sus diez acueductos que por cerca de cuatro siglos dieron sustento a una población muy cercana al millón de habitantes. Pero ahora el crecimiento demográfico se aprecia como una progresión geométrica en tiempos muy cortos; por ejemplo, la población de la ciudad de Londres hacia el año 1850 era cercana a los tres millones de habitantes y para este 2022 su población ya ronda los 24 millones de habitantes. Esto significa un incremento en 172 años de poco más del 80%. Otro ejemplo es China, país que para el año 1850 se estimaba tenía una población total cercana a los 430 millones de habitantes y para el 2022 su población ronda los 1.400 millones de habitantes. En una palabra: sorprendente.

La explosión demográfica conlleva grandes retos y uno de ellos consiste en el adecuado abastecimiento de agua para la población lo cual genera los interrogantes que antes se enunciaron. Desde la perspectiva de este estudio, puede contestarse con un rotundo “sí” a los tres primeros interrogantes planteados, pues es un hecho que el acceso a los servicios de abastecimiento y de saneamiento han logrado una cobertura cercana al 98% de esa población mundial que vive en las ciudades.

Ahora bien, el aumento en la cobertura del suministro de agua y en el servicio de saneamiento básico ha contribuido en gran medida a que se reduzcan los brotes de enfermedades transmitidas por este elemento vital (a través de vectores o bacterias), adicionando a ello las mejoras en la continuidad del servicio de abastecimiento y en un dedicado plan de mantenimiento y/o renovación de las redes de suministro.

En el artículo "Las maneras de fluir del agua visible e invisible en las ciudades. Implicaciones de la presencia del agua en el futuro urbano" los autores (Sánchez

Llorens & Fernández Alonso, 2017) muestran el papel del agua urbana y cómo “la planificación de los recursos hídricos y el diseño de las infraestructuras relacionadas con el agua han estado siempre presentes en nuestras ciudades, ya que el abastecimiento fue uno de los primeros problemas a resolver. Atravesar el espacio ocupado por el agua y retener sus movimientos inciertos también han moldeado nuestros paisajes urbanos”. Y, posteriormente, agregan que “la simultaneidad de ejemplos puede descubrir, deductivamente, una visión global de la importancia del agua en la historia urbana de las ciudades donde el crecimiento urbano está vinculado a las infraestructuras de agua”. Tal y como se pudo apreciar en la reseña histórica sobre las redes de suministro, esta afirmación sobre la importancia del agua deja en un punto preponderante el papel desempeñado por el agua urbana en la historia misma de las urbes y de sus habitantes.

Realizando una revisión de las publicaciones que tratan específicamente sobre la gestión del agua urbana en redes de abastecimiento, o simplemente la gestión del recurso agua, ocurre un caso similar al de los métodos epidemiológicos para análisis de datos, se encuentran algunos elementos interesantes de abordar. Utilizando el

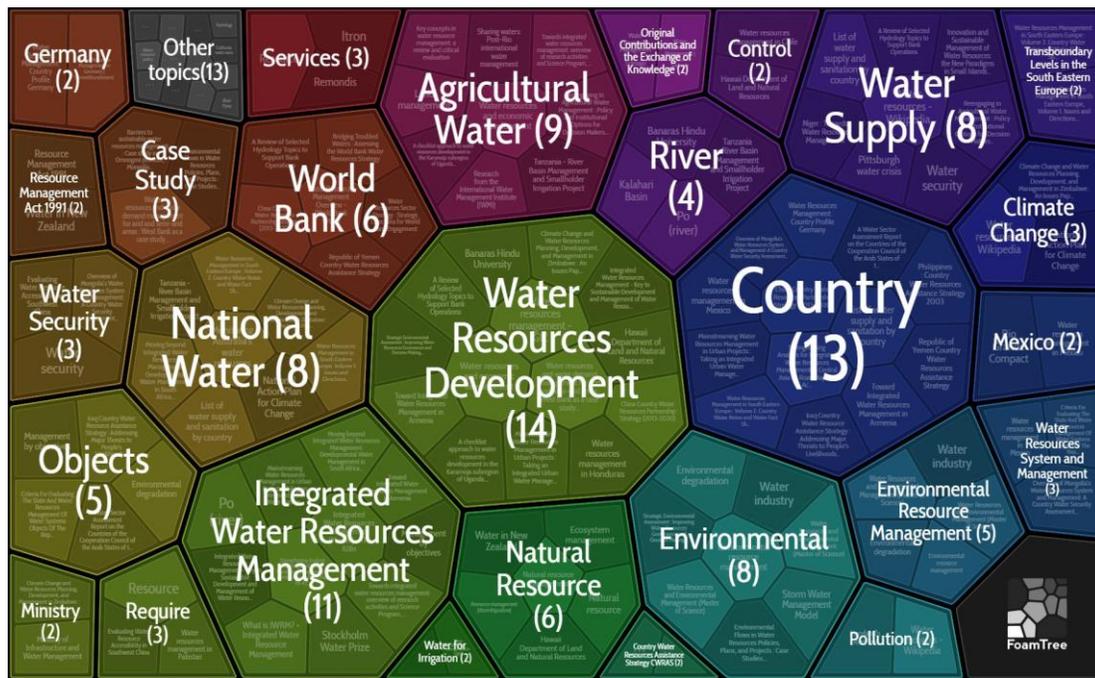


Ilustración 23. Gestión del recurso hídrico. <https://carrot2.org>

mismo procedimiento, se usa el término de búsqueda: “*Water resources management*” y se observa su comportamiento en clústeres.

Nuevamente se hace una primera aproximación con la herramienta online carrot2.org y arroja el siguiente resumen (*ilustración 23*).

Nuevamente, esta aproximación maneja unas 80 publicaciones. Una vez que se obtienen los resultados en las dos bases de datos ya mencionadas (Scopus y WoS), la consulta arroja unas **106.440 publicaciones** con las cuales ya se puede analizar mucho mejor las interacciones.

En la siguiente tabla (2) se puede apreciar en qué campos la interacción de la gestión del recurso hídrico es más utilizada.

CATEGORÍAS	PORCENTAJE DE PARTICIPACIÓN	PUBLICACIONES
Environmental Sciences	37,44	27483
Water Resources	31,20	22900
Geosciences Multidisciplinary	10,94	8027
Engineering Civil	8,50	6238
Engineering Environmental	8,30	6089
Ecology	7,86	5772
Environmental Studies	6,84	5019
Green Sustainable Science Technology	5,38	3949
Agronomy	4,96	3643
Meteorology Atmospheric Sciences	4,46	3272
Marine Freshwater Biology	4,24	3111
Soil Science	2,85	2088
Multidisciplinary Sciences	2,71	1987
Energy Fuels	2,68	1970
Biodiversity Conservation	2,49	1826
Fisheries	2,17	1596
Forestry	2,12	1556
Limnology	2,06	1509
Public Environmental Occupational Health	1,07	785
Biotechnology Applied Microbiology	0,98	716

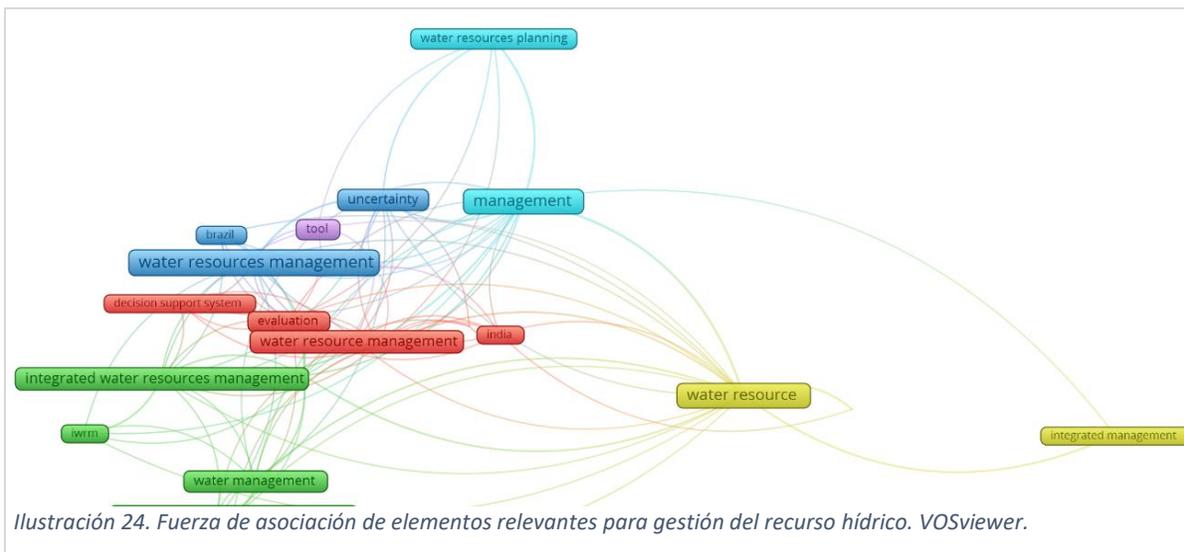
Tabla 2. Categorías en las que se aplica la gestión del recurso hídrico

Es preciso hacer una salvedad, pues tanto en esta tabla 2 como en la tabla 1, los datos presentados son una pequeña muestra resumen del total de las publicaciones. ¿Qué se puede apreciar entonces en esta tabla 2? Que las publicaciones sobre la gestión del recurso hídrico se hallan presentes en una buena cantidad de categorías, incluyendo aquellas del ámbito de la salud pública y la salud particular.

Si bien su impacto no es tan grande, si es importante apreciar una coincidencia interesante; mientras que los métodos epidemiológicos para análisis de datos tienen una participación del 5,3% en la categoría de *Multidisciplinary Sciences*, la participación de la gestión del recurso hídrico en la misma categoría es de 2,71%.

Esto se puede entender como la coexistencia tanto de la epidemiología como de la gestión del agua, en buena parte de las publicaciones en esta categoría, y así hay más ejemplos de coexistencia, a saber: *Public Environmental Occupational Health, Surgery, Toxicology, Biotechnology Applied Microbiology, Computer Science Interdisciplinary Applications, Chemistry Medicinal, Zoology, Mathematics Applied y Agronomy* entre otras.

Volviendo al número total de publicaciones sobre la gestión del recurso hídrico (106.440), se procede a la aplicación de la herramienta VOSviewer a fin de visualizar un mapa de interacciones tal y como se aprecia en la ilustración 24. Acá se puede observar la fuerza de la asociación entre los elementos relevantes de las publicaciones



sobre la gestión del recurso hídrico, tales elementos son recurso hídrico, gestión del recurso hídrico, gestión integrada y gestión integrada de recurso hídrico.

Tal y como se hizo en el acápite anterior, el comportamiento de la producción de publicaciones relacionadas con la gestión del recurso hídrico se muestra en la siguiente ilustración (25), tomando cinco fuentes: Water Swizerland, Science of the Total Environment, Journal of Hydrology, Water Resources Management y Water Science and Technology

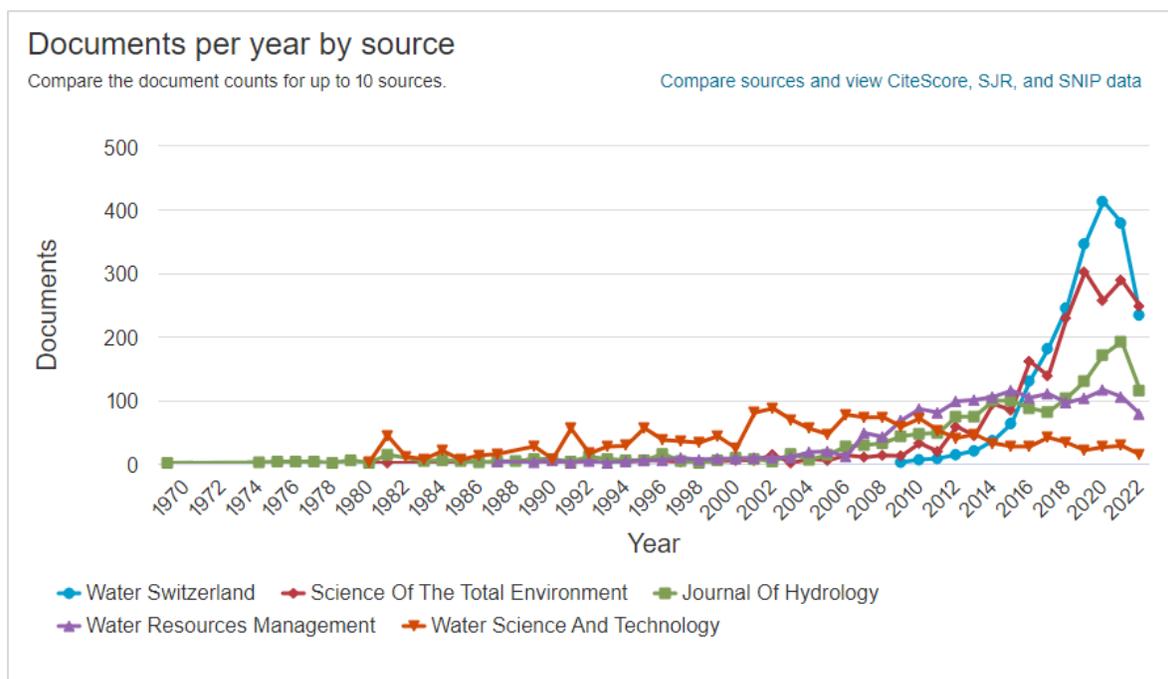


Ilustración 25. Comportamiento de producción para la gestión del recurso hídrico. Fuente Scopus

El papel del agua urbana y la gestión del recurso hídrico a través de redes de suministro se desarrolla veloz y enormemente en virtud del crecimiento demográfico.

Si bien las publicaciones en los dos temas tratados hasta ahora (métodos epidemiológicos para análisis de datos y la gestión del recurso hídrico), se han incrementado desde el año 2.008, en gran medida por el aumento en el acceso abierto a tales documentos, resulta claro que la producción relacionada con el agua se ha multiplicado por cuatro frente a la producción relacionada con la epidemiología. Y no

debería sorprender este resultado, pues como se ha podido apreciar a lo largo de la historia, el papel del agua urbana es vital para el crecimiento de una ciudad y para la identificación social de su población; lamentablemente, el contexto actual de la globalización y su siempre opresora artificialidad, ha llevado a que los habitantes de una ciudad pierdan su conexión con la identidad histórica de la misma, así como con el medio que le rodea. La adecuada gestión del recurso hídrico mediante las redes de distribución, tanto para abastecimiento como para ornamentación, permite volver a naturalizar las ciudades dentro del concepto moderno de las *Smart Cities*.

2.2.2.1. El agua urbana dentro de los procesos epidemiológicos

Como se pudo ver en el acápite anterior, existe una coexistencia de la epidemiología y el agua en las publicaciones revisadas hasta ahora. Conviene apreciar en detalle un ejemplo de lo desarrollado hasta ahora, para ello se hará un análisis del artículo “Chlorination disinfection by products in municipal drinking water – A review” publicado por Mazhar, Mohd Aamir; Khan, Nadeem A.; Ahmed, Sirajuddin; Khan, Afzal Husain; Hussain, Azhar; Rahisuddin; Changani, Fazlollah; Youse, Mahmood; Ahmadi, Shahin y Vambol, Viola. Este documento tiene la característica de ser uno de los más citados en el año 2.020; un breve resumen del mismo establece que *durante mucho tiempo, la desinfección del agua se realiza para inactivar o controlar las poblaciones de bacterias en el agua tratada y la cloración simple es el método más común para desinfectar el agua. Es casi seguro que durante más de una década la cloración se ha utilizado de forma eficaz para controlar la contaminación transmitida por el agua. Se forma una variedad de subproductos de desinfección (DBP) como trihalometanos (THM), ácidos haloacéticos (HAA) y acetonitrilos halogenados (HAN), etc., cuando las materias orgánicas naturales (NOM) presentes en el agua cruda reacciona con el cloro, lo que puede tener efectos adversos para la salud de los seres humanos. Estos compuestos deterioran la calidad del agua y existen confirmaciones epidemiológicas de estrecha relación entre su aparición y resultados*

desfavorables, especialmente los crecimientos malignos de órganos fundamentales en las personas.

Las categorías de esta publicación se relacionan en los siguientes clústeres (ilustración 26)



Ilustración 26. Clústeres de Chlorination disinfection by products in municipal drinking water – A review (WoS)

La publicación está referida en siete categorías: Environmental Sciences, Chemistry Analytical, Engineering Environmental, Food Science Technology, Green Sustainable Science Technology, Meteorology Atmospheric Sciences, Microbiology. La participación en la primera categoría es de un 40,0%, en las otras se distribuye uniformemente con un 10,0% en cada una.

En un aparte del artículo, los autores exponen un problema social y técnico a solucionar: *“El cloro se introdujo por primera vez en Chicago, Estados Unidos, en 1.908 en un suministro de agua urbano para disminuir el recuento de bacterias. ... La cloración del agua potable sigue siendo un logro notable en la prevención de enfermedades del siglo XX. Sin embargo, la existencia de DBP clorados en el agua del grifo es motivo de gran preocupación, ya que se sospecha que causan cáncer”*.

Y este problema se agudiza no solo por la ingesta del agua, sino también por inhalación y por contacto dérmico. Los procesos epidemiológicos que se usaron fueron técnicas cuantitativas, propias de la epidemiología tradicional; no obstante, se pueden aplicar elementos cualitativos que permitan encontrar una solución a dicha problemática. Este ejemplo es útil para mostrar uno de los múltiples nexos que se pueden generar entre la epidemiología y la administración del agua urbana.

2.2.3 La gestión de fallos en redes de distribución de agua

Las redes de distribución de agua son un elemento vital dentro de la infraestructura de los actuales acueductos, máxime al tener tanta complejidad en el abastecimiento para las ciudades modernas cuya población crece exponencialmente en cortos periodos de tiempo. La gestión técnica eficaz del agua potable exige el manejo de una cantidad importante de información tanto física como espacial, haciendo referencia la primera a la tubería en sí y la segunda a su ubicación; esta gestión de redes de agua urbana se desarrolla mediante la unión de las nuevas tecnologías de tratamiento de la información con nuevas técnicas para la construcción de modelos de redes, principalmente derivadas de sistemas de información geográfica SIG, técnicas que permitan analizar datos y elaborar desde consultas sencillas hasta complejos modelos (Caballero-Núñez, 2017).

Martínez-Codina (2015) indica que un hecho inevitable en las redes de distribución de agua es su deterioro, causado principalmente por desgaste de materiales o por variaciones en las presiones servidas. Como consecuencia, se suceden los agrietamientos o roturas de las tuberías, los cuales llevan asociados enormes costos sociales, económicos y medioambientales. Pero, además, puede presentar un costo en salud pública, pues la intrusión patógena es un fenómeno que se presenta en los sistemas de distribución de agua potable, provocando la disminución de la calidad del agua distribuida cuando los fluidos (contaminados) circundantes a las conducciones entran en las mismas a través de defectos en condiciones de depresión. La

contaminación en condiciones más desfavorables puede provocar enfermedades epidémicas debido al consumo de agua contaminada (Mora-Rodríguez, 2013).

Conocer los procedimientos que operan en una red de distribución se constituye en un elemento fundamental para atacar el problema ante un fallo físico que pueda provocar la intrusión de componentes contaminantes a la red. Es claro que los fallos, aunque llegasen a producir una rotura que automáticamente suspende el servicio de agua, en ocasiones se presentan por agrietamientos leves y estos son los que hay que ponerles más atención ya que pueden ocurrir filtraciones hacia afuera de la red, pero también intrusiones en la misma; lo anterior puede ocurrir sin necesidad de suspender el servicio, permitiendo así su propagación a sectores más amplios de la misma red lo cual llevaría al colapso de la misma.

Puede decirse que desde la década de los 90 o incluso más atrás, la literatura técnica internacional ha venido reflejando la intención de introducir explícitamente la estadística en el diseño y la operación de las redes de suministro de agua potable. Dados los fenómenos del crecimiento de la población servida por la red, el fallo mecánico de sus componentes, el envejecimiento de las tuberías (que implica la reducción de su capacidad de conducción) y otros, se ha venido trabajando en el cálculo y la asignación cuantitativa de valores probabilísticos a estos fenómenos, con el fin de llegar a cuantificar la confiabilidad o garantía de funcionamiento de la red (Rodríguez, 2003); y actualmente se busca agregar el elemento cualitativo en esos cálculos.

A continuación, y tal como se presentó en los acápites anteriores, se realiza una revisión de las publicaciones que versan sobre la gestión de fallos en las redes de distribución de agua. Usando el término de búsqueda "*Fault management*" AND "*water distribution networks*" se puede observar su comportamiento en clústeres mediante una aproximación con la herramienta online carrot2.org y se obtiene la siguiente ilustración (27).

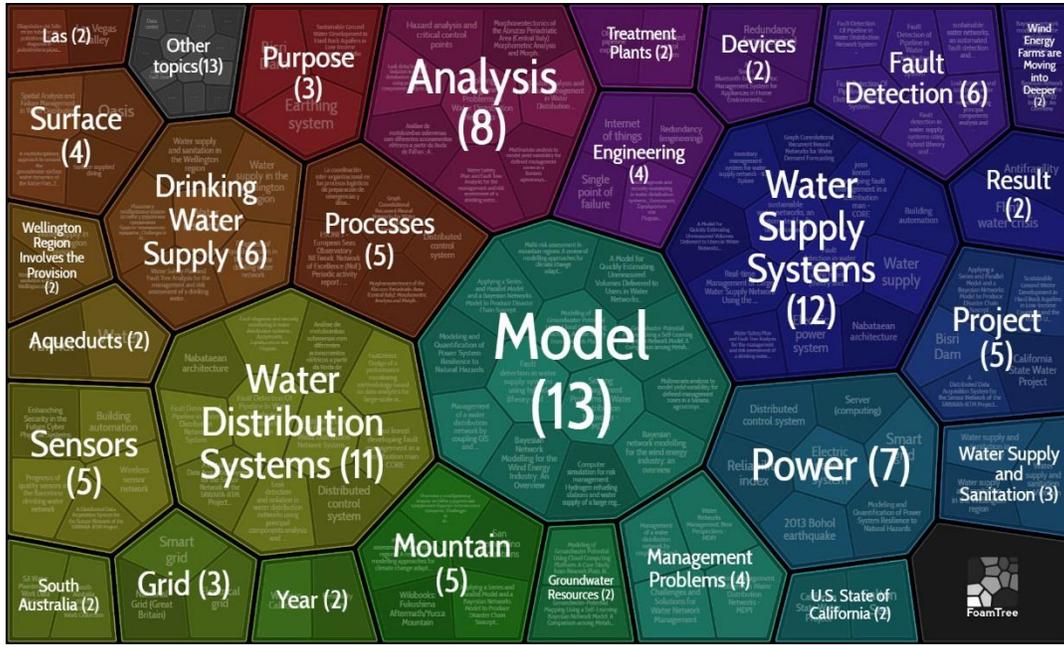


Ilustración 27. Gestión de fallos. <https://carrot2.org>

Una vez más, esta aproximación resume unas 87 publicaciones. De nuevo al obtener los resultados en las dos bases de datos ya trabajadas, la consulta arroja unas **1.451 publicaciones** para las cuales se analizan sus interacciones.

La siguiente tabla (3) resumen presenta los clústeres de gestión de fallos en redes de distribución:

CATEGORÍAS	PORCENTAJE DE PARTICIPACIÓN	PUBLICACIONES
Water Resources	20,40	185
Geosciences Multidisciplinary	15,44	140
Engineering Civil	15,33	139
Environmental Sciences	13,56	123
Engineering Geological	8,60	78
Engineering Environmental	4,96	45
Construction Building Technology	3,42	31
Physics Applied	2,32	21
Engineering Industrial	2,21	20
Ecology	2,10	19
Computer Science Interdisciplinary Applications	1,99	18
Chemistry Multidisciplinary	1,76	16
Multidisciplinary Sciences	1,54	14

Environmental Studies	1,21	11
Biotechnology Applied Microbiology	0,55	5
Computer Science Hardware Architecture	0,55	5
Engineering Biomedical	0,55	5
Statistics Probability	0,55	5
Public Environmental Occupational Health	0,33	3
Toxicology	0,22	2

Tabla 3. Categorías en las que se aplica la gestión de fallos

Las publicaciones que incorporan la gestión de fallos están orientadas más a las categorías de geociencias, ingeniería civil, recursos hídricos y ciencias ambientales, alcanzando así una participación del 64,73% mientras que, en las demás categorías, su participación es mucho menor. Volviendo a apreciar la categoría de *Multidisciplinary Sciences* se puede ver la participación de los tres temas tratados hasta ahora, epidemiología (5,3%), gestión de agua (2,71%) y gestión de fallos (1,54%); esto confirma la existencia de un nexo entre estos temas en buena parte de las publicaciones para esta categoría. Y se repite también en otras categorías, como: *Public Environmental Occupational Health, Toxicology, Biotechnology Applied Microbiology, Computer Science Interdisciplinary Applications*, entre otras.

Aplicando ahora la herramienta VOSviewer al número total de publicaciones en este tema (1.451) se puede visualizar un mapa de interacciones como se aprecia en la siguiente ilustración (28).

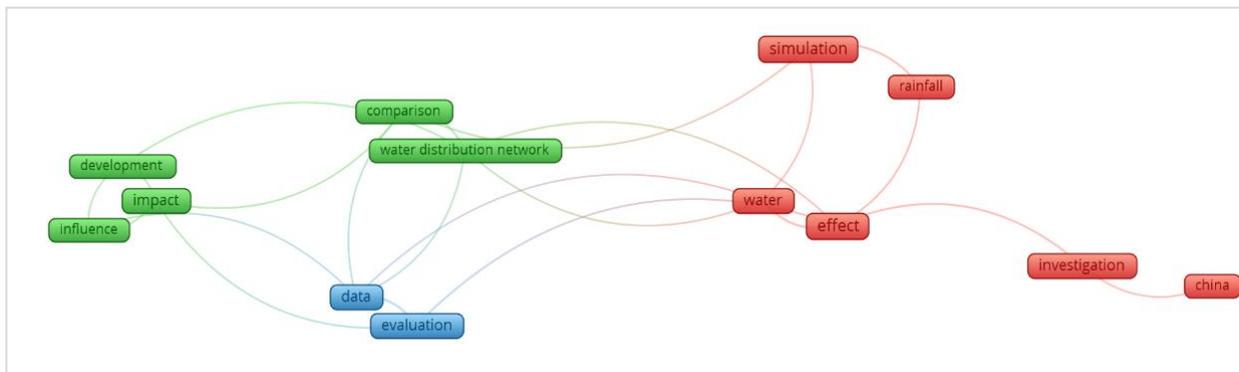


Ilustración 28. Fuerza de asociación de elementos relevantes para gestión de fallos. VOSviewer.

La fuerza de la asociación entre los elementos relevantes de las publicaciones sobre gestión de fallos, se aprecia en redes de distribución de agua, impacto, datos de simulación y evaluación. El comportamiento de la producción de publicaciones relacionadas con la gestión de fallos se muestra en la siguiente ilustración (29), tomando cinco fuentes: *Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering OMAE*, *Applied Mechanics and Materials*, *Advanced Materials Research*, *Nace International Corrosion Conference* y *American Society of Mechanical Engineers Pressure Vessels and Piping Division Publication PVP*

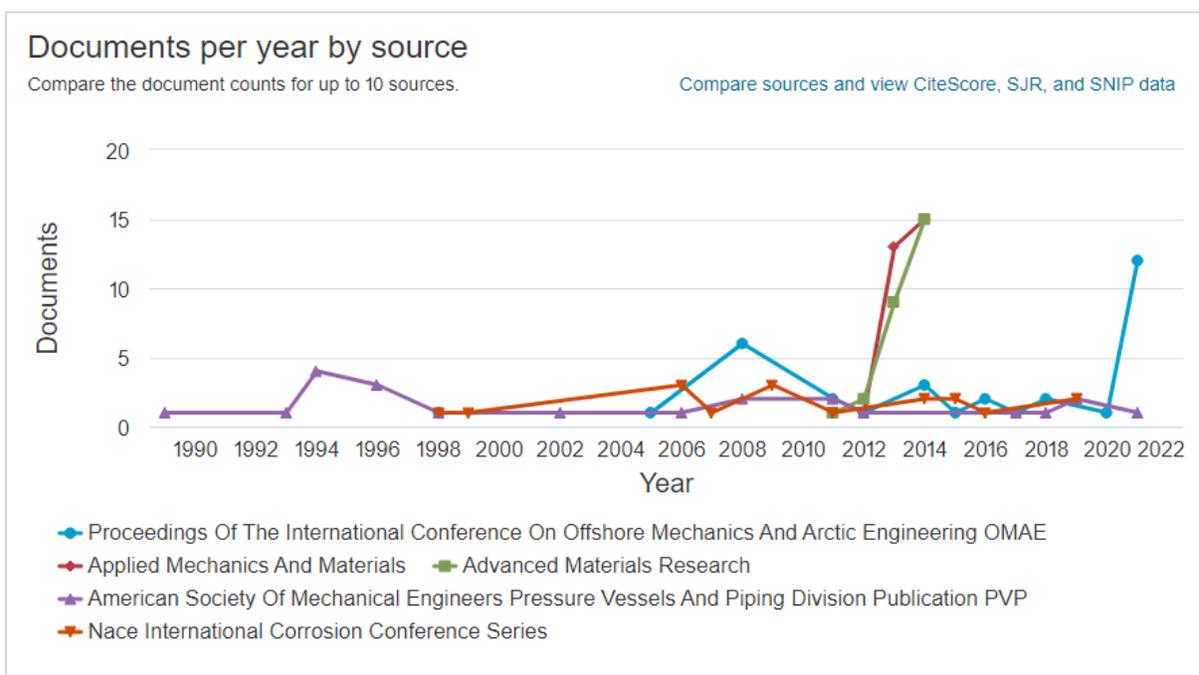


Ilustración 29. Comportamiento de producción para gestión de fallos. Fuente Scopus

Si bien la gestión de fallos en redes de abastecimiento de agua urbana es un tema importante para los gestores, la producción de publicaciones sobre el tema no refleja un incremento significativo frente a los otros dos temas tratados: la epidemiología y la gestión del recurso hídrico. Pero bien puede argumentarse que la gestión de fallos en redes de abastecimiento es parte del proceso mayor de gestión del recurso hídrico. Destacan dos momentos, el primero entre el 2.012 y el 2.014, cuando se puede ver un leve incremento de la producción respecto de investigaciones en materiales avanzados

y en aplicación de materiales. El otro momento se presenta en el 2.021 con otro leve incremento de la producción relacionados más con estructuras, seguridad y confiabilidad de materiales y tuberías para desarrollos marinos y submarinos. Esto lo que muestra es que la producción en el tema de gestión de fallos en redes de distribución, se centra más en el diseño y búsqueda de materiales óptimos, pero muy poco sobre herramientas para predicción de los fallos.

GESTIÓN DE DEMANDA

“En este capítulo se trata en más detalle uno de los componentes de la administración del recurso hídrico como lo es la gestión de la demanda, partiendo del ¿por qué es importante? Se definen los patrones de consumo de agua y se analizan varios de los modelos para la estimación de la demanda, destacando el modelo SAX y su eficiencia en varios entornos de aplicación, incluyendo el análisis de la demanda de agua urbana.”

3.1 DEMANDA DE AGUA URBANA

El concepto de agua urbana generalmente incluye suministro de agua y saneamiento. Bajo esta perspectiva, el saneamiento involucra la recolección y tratamiento de efluentes domésticos e industriales, no incluye drenaje urbano, manejo de residuos sólidos, pues aún persiste una visión desfasada de la gestión urbana del agua en la ciudad. El agua urbana involucra componentes que permiten un desarrollo ambiental sostenible y utilizan los conceptos de gestión integrada de los recursos hídricos (GIRH), necesarios para la planificación, ejecución y mantenimiento de la infraestructura de la ciudad (Tucci, 2008).

Como ya se indicó, el acceso al servicio de agua potable es vital e indispensable y así está considerado en las legislaciones de muchos países, en donde debe cumplir con condiciones de calidad, cantidad y continuidad.

Conocer la demanda de agua requerida por un sistema de distribución de agua (WDS) hace posible un mejor control del sistema (Di Nardo et al., 2015) (Scarpa et al., 2016).

Esto naturalmente considera el comportamiento de los diferentes tipos de consumidores dentro de cada área de medición del distrito (DMA) (ilustración 30) (Panagopoulos et al., 2012).

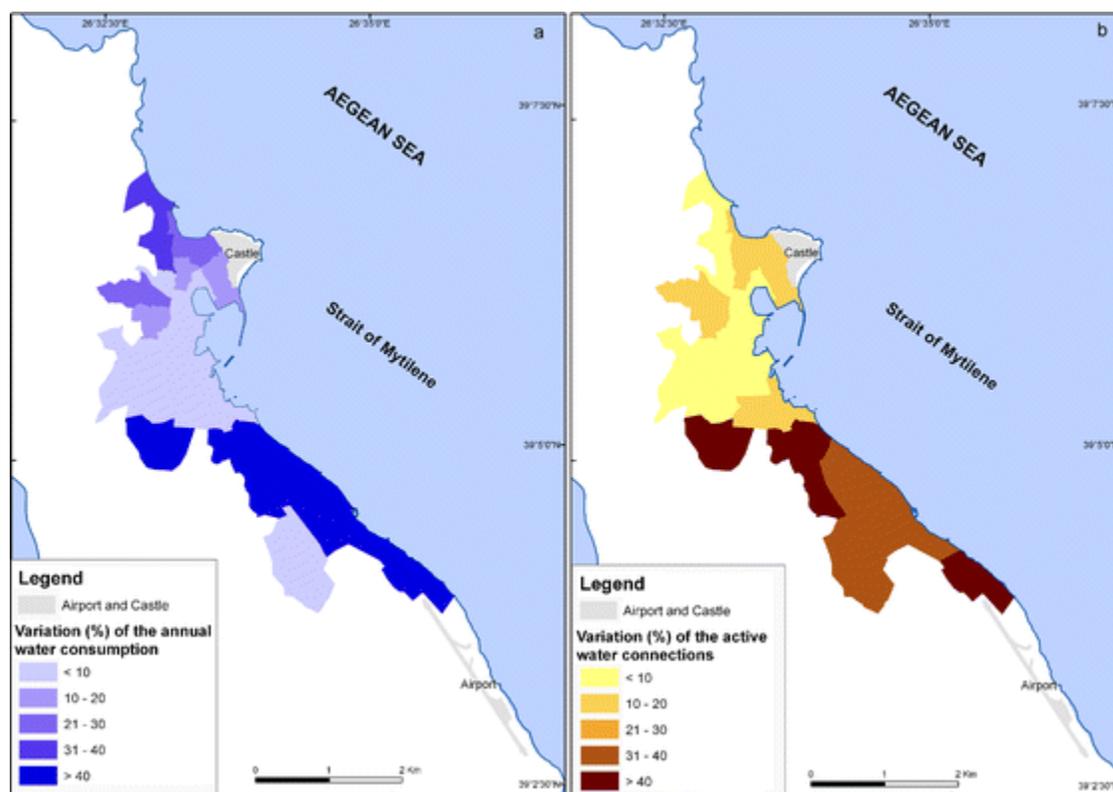


Ilustración 30. Ejemplo de las variaciones (%) para los años 2000 y 2007 en cada distrito de Mitilene de la tasa anual de consumo de agua (a) y el número de conexiones de agua activas (b) (Panagopoulos et al., 2012)

Además, teniendo en cuenta la amplia variabilidad y distribución espacial de las áreas WDS, el análisis puede ampliarse aún más hacia un análisis integral que tenga en cuenta otros servicios públicos urbanos como las redes de gas o electricidad. Aun así, considerar las relaciones espaciales en el análisis de la demanda de agua beneficia la provisión de balances hídricos más precisos, ya que el análisis puede enfocarse ahora en cada uno de los DMA que abarcan el WDS (Herrera, 2011). Al mismo tiempo, facilita la gestión de activos y el control de los componentes del sistema y la localización y cuantificación de fugas (Buchberger et al., 2004).

El rendimiento de los activos de abastecimiento de agua, que debe satisfacer los comportamientos de demanda de los usuarios, puede aprovechar los patrones de demanda de agua de clientes con características similares o que se encuentran en condiciones ambientales similares (Candelieri, 2017)

3.2 GESTIÓN DE DEMANDA DE AGUA

Para abordar este punto, conviene hacerse una pregunta: ¿por qué es importante la gestión de la demanda de agua? Una respuesta adecuada a esta inquietud consiste en que, si se aumenta el suministro de agua a una locación específica en forma eficiente, puede redundar en el desarrollo de esa locación; se podría decir que una buena gestión de la demanda de agua es un instrumento a favor del desarrollo sostenible (López, 2002).

Hoy en día, alrededor del 55 % de la población mundial, unos 4.200 millones de habitantes, vive en ciudades. Se cree que esta tendencia continuará. En el año 2.050, la población urbana se duplicará, y casi 7 de cada 10 personas vivirán en ciudades.

La velocidad y la intensidad de estos crecimientos presentan nuevos desafíos para hacer frente a la planeación o reorganización de sistemas de infraestructuras de redes complejas como lo son las redes de abastecimiento de agua, la infraestructura eléctrica o de transporte y sus múltiples interacciones. La gestión y el desarrollo sincronizados generan ciudades más robustas y resilientes capaces de hacer frente a las crisis o retos que se puedan llegar a presentar no solo a nivel de cobertura de servicios, sino también a nivel económico, social, de salud pública, por ejemplo.

Frente a la construcción de nuevas infraestructuras de oferta y en el continuo incremento de la presión sobre los ecosistemas naturales, la nueva cultura del agua propugna por la optimización del uso de las infraestructuras, la introducción de nuevas tecnologías de control, regulación, utilización y reutilización.

Se trata pues de conseguir con la buena gestión del agua, por una parte, un uso técnicamente eficiente y, por otra, un uso eficiente social y económicamente; perspectivas estas que se enmarcan en la consecución y conservación del estado del bienestar (López, 2002).

Entre los instrumentos de gestión de la demanda de agua, los abastecimientos urbanos tienen un peso específico importante. A las medidas de ahorro, propugnadas ya por Postel, (1993), se les agregan las de optimización del uso del agua así:

- Mejoras de redes de transporte y distribución
- Mejora de la fontanería domestica
- Reordenación de riegos urbanos
- Xerojardinería y acolchados
- Mejora tecnológica de los sistemas de riego urbano y baldeo de viales.
- Reutilización de agua
- Desdoblamiento de redes de distribución y alcantarillado según calidades del agua.
- Campañas de información y concienciación ciudadana (cultura del agua)

Es así como el ejercicio planificador de la gestión de redes de abastecimiento de agua enmarcado en los nuevos planteamientos trata de desvelar no sólo la relación demanda global-recursos disponibles o potenciales, sino también conocer la demanda con detalle, saber cuál será el volumen esperado, concretar su composición y ponerla en relación con tendencias y hábitos de la población y posibles medidas diseñadas para su modulación (Pedregal Mateos, 2004).

3.3 PATRONES DE CONSUMO PARA LA DEMANDA DE AGUA

El objetivo de las redes de abastecimiento es satisfacer las demandas de agua requeridas por los usuarios del servicio. La respuesta del sistema va a estar condicionada fundamentalmente por dichas demandas, que pueden ser vistas como las cargas que debe soportar una estructura. Sin embargo, en el caso de las demandas, estas no permanecen constantes a lo largo del tiempo, sino que varían aleatoriamente siguiendo unos determinados patrones temporales (García Carrasco, 2019).

Las fluctuaciones en el consumo de agua urbana pueden obedecer a variaciones temporales de dichos consumos; dichas variaciones pueden estar representadas como variables periódicas las cuales siguen comportamientos similares o repetitivos según ciertos ciclos o zonas de distribución. Los patrones de consumo de agua pueden observarse en ciclos de tiempo relativamente cortos (24 horas), semanales, mensuales o anuales. Por ejemplo, en un ciclo de 24 horas en un hogar, se observan incrementos del consumo a medida que se empiezan las actividades diarias, durante el día el consumo disminuye cuando hay un desplazamiento a las actividades laborales y en la tarde/noche nuevamente se incrementa dicho consumo al llegar y permanecer en casa (*ilustración 31*).

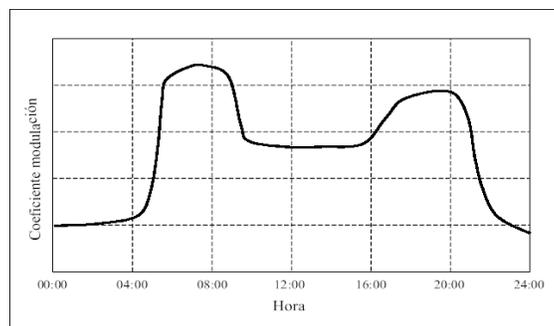


Ilustración 31. Curva de modulación para un área residencial. (García Carrasco, 2019)

Los patrones de consumo de los días laborables difieren de los correspondientes a los fines de semana. Las diferencias estacionales del consumo de agua están relacionadas con los cambios climáticos, tales como la temperatura y la precipitación, así como el cambio de hábitos de los usuarios, de forma que durante los meses de verano aumentan las actividades recreativas y agrícolas en países estacionales (García Carrasco, 2019).

La previsión de la demanda de agua es de suma importancia para la planificación y operación del suministro de agua, para garantizar un servicio continuo a niveles adecuados de cantidad y presión. Por lo tanto, los modelos precisos de demanda de agua son clave para abordar la gestión inteligente de activos físicos tales como bombas, válvulas y medidores de agua (Creaco et al., 2018) (Fecarotta et al., 2018) (Nguyen et al., 2018). Además, gracias a una comprensión más profunda de la demanda de agua, es posible proponer mejor la operación casi en tiempo real del suministro de agua urbana. Al igual que otros problemas de pronóstico, el pronóstico de la demanda de agua es esencialmente un problema de regresión y ha sido abordado por modelos de regresión tanto simples como multivariados (Adamowski et al., 2010) y por modelos clásicos de series de tiempo (Caiado, 2010). Estos últimos modelos suelen considerar variables de tiempo y calendario, ya que los usuarios demandan agua generalmente a las mismas horas. Sin embargo, se comportan de manera diferente los fines de semana o días festivos que entre semana. La aplicación de técnicas innovadoras para la previsión a corto plazo ha cobrado fuerza en los últimos años superando la precisión de los modelos clásicos de previsión a corto plazo (Herrera et al., 2010). Esto se ha debido principalmente al crecimiento de la capacidad informática y al desarrollo de herramientas eficientes de aprendizaje automático, dentro de las cuales se pueden destacar las redes neuronales artificiales (ANN) como un enfoque importante para pronosticar la demanda de agua (Msiza et al., 2008). El uso de máquinas de aprendizaje extremas (Tiwari et al., 2016), sistemas de inferencia neurodifusos adaptativos (Vijayalaksmi et al., 2015), y las técnicas híbridas de optimización y aprendizaje automático (Ji et al., 2014) pueden considerarse

actualmente investigación de vanguardia en la demanda de agua. Sin embargo, a medida que evolucionan las nuevas técnicas de previsión de la demanda, los datos disponibles también crecen exponencialmente. Los nuevos desafíos a abordar son, entonces, cómo manejar bases de datos de series temporales de gran tamaño y cómo obtener una respuesta casi en tiempo real de los modelos predictivos, manteniendo los niveles más altos de precisión. Los algoritmos híbridos pueden mejorar la precisión de las herramientas de pronóstico, haciendo coincidir mejor los efectos climáticos y sociales con las tendencias de las series temporales (Cadenas et al., 2016) (Zhou et al., 2016). En esta dirección, varios autores han propuesto métodos híbridos destacando el uso de modelos ANN y ARIMA (Herrera et al., 2011) (Khashei et al., 2011) (Zhang, 2003) o el uso de métodos de aprendizaje automático vinculados a varios modelos de transformación, como wavelets (Campisi-Pinto et al., 2012) o series de Fourier (Brentan et al., 2017).

3.4 MODELOS PARA LA ESTIMACIÓN DE LA DEMANDA DE AGUA

Desde su inicio, los modelos conocidos de cálculo de redes de distribución de agua fueron creados en países desarrollados donde el servicio del agua es continuo y la demanda de agua se satisface todo el tiempo. Cuando un usuario necesita agua, sólo basta con que abra la llave y reciba el líquido que necesita. Es por eso que tales modelos son normalmente "modelos dependientes de la demanda", es decir, que la demanda de agua es fija y el modelo obtiene las presiones en la red en función de la demanda dada (Cabrera-Bejar et al., 2012). La mayoría de modelos de redes de distribución conocidos, tales como EPANET (Rossman, 2000), Infoworks WS, ScadRed (Tzatchkov et al., 1996) y otros son de ese tipo.

Por el contrario, las redes de distribución de agua con servicio intermitente operan por un tiempo limitado, sólo cuando se proveen de agua. La cantidad de agua que el usuario puede captar depende de la presión disponible en su conexión y la duración del servicio y no siempre la demanda del usuario se satisface completamente, es decir, la operación de dichas redes es dependiente del suministro o "dependiente de la presión" y debe ser modelada de esta manera. (Cabrera-Bejar et al., 2012). El comportamiento hidráulico de un modelo dependiente del suministro o dependiente de la presión es muy diferente del comportamiento de un modelo dependiente de la demanda (Batish, 2003) (Tzatchkov et al., 2008).

Los modelos para el dimensionado de los sistemas hidráulicos, tanto de distribución como de evacuación de agua, deben tener como principal objetivo proporcionar soluciones que respondan de forma adecuada y eficiente a las situaciones reales y operativas, garantizando una adecuada prestación del servicio. Es por tal motivo que actualmente crece el interés en realizar un preciso ajuste de dichos modelos, con el fin de alcanzar un mayor nivel de detalle de las condiciones reales de consumo; a través del uso de técnicas y recursos de modelado más ajustados a la realidad del proceso (Caicedo Londoño, 2012).

3.4.1 Desarrollo de modelos de previsión para la demanda.

Con el transcurrir del tiempo se han utilizado metodologías diversas en los modelos para la previsión de la demanda de agua urbana, la cual involucra diferentes variables que pueden ser controlables, así como otras variables que son no controlables. Dentro de las primeras se encuentran algunas como las restricciones del sistema y el precio; en las segundas, variables dinámicas no controlables como los diversos hábitos de los consumidores, variables socioeconómicas y aquellas relacionadas con factores asociados al clima: periodos de humedad o sequía etc.

Las diferentes propuestas que se han realizado para la previsión de la demanda incluyen métodos y enfoques que varían desde la extrapolación histórica más simple hasta los modelos analíticos sofisticados. Si bien existen muchos factores que afectan el rendimiento de los modelos de predicción, se ha encontrado que la fidelidad del pronóstico de la demanda de agua urbana está significativamente influenciada por la periodicidad de la demanda, el horizonte de pronóstico, el método de pronóstico, la especificación del modelo y algunas características específicas del estudio, como el tamaño de la muestra, el año de publicación y el nivel de desarrollo del país en el que se realizó el estudio. Sebri, M. (2016).

En su gran mayoría, los modelos desarrollados para pronosticar la demanda requieren datos históricos de la misma y asumen que el comportamiento que se ha observado, se reproducirá en un futuro (*ilustración 32*) (Adamowski, 2008) (Bakker et al., 2013). Por tal razón, el modelado y pronóstico de series temporales de la demanda de agua se apoya desde el comienzo en el análisis de Fourier, pues se pueden identificar los componentes periódicos de la demanda (a nivel estacional, semanal, horario etc.) y sus resultados permiten encontrar los armónicos predominantes (Gortaire et al., 2016).

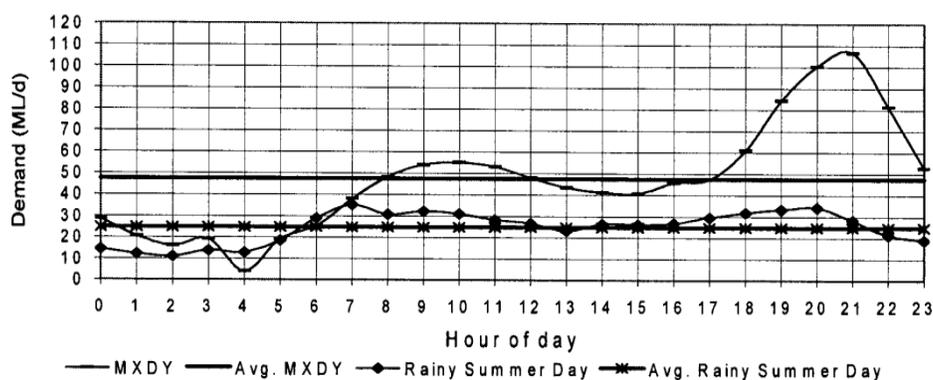


Ilustración 32. Uso típico diario de agua durante el verano en la zona del centro oeste de Ottawa, (MXDY día de mayor demanda para todo el registro de 1994–2002; promedio MXDY promedio demanda máxima para todo el registro. (Adamowski, 2008)

Los modelos de predicción para estimar la demanda pueden ser univariados, donde solo los registros de los sistemas de distribución se consideran entradas para los

modelos (Alvisi et al., 2007) (Alvisi & Franchini, 2014) (Romano & Kapelan, 2014). Además, pueden ser multivariantes, donde algunos de los factores influyentes se consideran entradas del modelo para predecir la demanda (Zhou et al., 2002) (Herrera et al., 2010b) (Adamowski et al., 2012) (Odan & Reis, 2012) (Tian et al., 2016).

Tanto los modelos univariados, como los multivariados, se pueden basar en varios enfoques comúnmente aplicados como lo son *los modelos basados en patrones* (Gagliardi et al., 2017) (Pacchin et al., 2017) (Pacchin et al., 2019), *los de análisis de regresión* (Adamowski & Karapataki, 2010) (Bakker et al., 2014), *los modelos clásicos de series de tiempo* (Guarnaccia et al., 2020) (Viccione et al., 2020) (Mu et al., 2020), *los métodos de inteligencia artificial (IA)* como redes neuronales artificiales (ANN) (Gagliardi et al., 2017) (Pacchin et al., 2019) (Pesantez et al., 2020), *los de bosque aleatorio* (Herrera et al., 2010a) (Mu et al., 2020) (Nasser et al., 2020) (Pesantez et al., 2020), *los de máquinas de vectores de soporte (SVM)* (Candelieri et al., 2019) (Pesantez et al., 2020), *los de lógica difusa y los modelos neuro-difusos* (Vijayalaksmi & Babu, 2015) (Jithish & Sankaran, 2017). Más recientemente, también se han propuesto *los modelos híbridos* que combinan dos de los enfoques citados anteriormente (Quevedo et al., 2010) (Alvisi & Franchini, 2014) (Suhartono et al., 2018).

Investigadores como House-Peters & Chang, (2011), Donkor et al., (2014), Ghalehkhondabi et al., (2017) y de Souza Groppo et al., (2019) afirman que se han aplicado diferentes métodos y modelos en estudios anteriores para predecir la demanda de agua, incluidos los modelos tradicionales, los de IA y los híbridos de IA.

Modelos tradicionales, como los de análisis de series temporales y regresión (Gato et al., 2007a, 2007b), se emplearon por primera vez en la simulación de la demanda de agua. Sin embargo, los enfoques tradicionales carecían de precisión al momento de pronosticar la demanda, lo que podía causar problemas importantes en la operación y gestión del sistema de suministro. Además, el crecimiento del impacto por el cambio climático y la urbanización, causan incertidumbre, de modo que la predicción y el

pronóstico sean más complejos; por lo que ello motivó a los investigadores a desarrollar aún más sus modelos (Zubaidi et al., 2018), incluido el uso de técnicas de IA.

Los modelos de pronóstico de series de tiempo para la demanda de agua, se pueden dividir en modelos deterministas y estocásticos. Los modelos deterministas se utilizan para las series de tiempo mediante cadenas de patrones estacionales, semanales y diarios teniendo en cuenta su naturaleza física. Los modelos estocásticos generalmente adoptan un enfoque numérico ya que comúnmente se expresan utilizando modelos estadísticos y probabilísticos que se basan en datos históricos (Box & Jenkins, 1976; Chau, 2006).

En el diseño de series de tiempo para demanda de agua, los modelos estocásticos más comunes son: el modelo de ruido aleatorio puro (o blanco), el modelo autorregresivo (AR), el modelo de promedio móvil (MA), el modelo de promedio móvil autorregresivo (ARMA), el modelo de promedio móvil integrado autorregresivo (ARIMA) y el promedio móvil integrado autorregresivo estacional (SARIMA) (Box & Jenkins, 1976) (Chau, 2006) (Rojas, 2013). Estos modelos, también conocidos como modelos estadísticos tradicionales, son lineales porque las predicciones de los valores futuros están limitadas a ser funciones lineales de observaciones pasadas. Debido a su relativa simplicidad en la comprensión y la implementación, los modelos lineales han sido el principal enfoque de investigación y las herramientas aplicadas ampliamente usadas durante muchos años (Machiwal & Jha, 2012).

Aunque los modelos ARIMA son bastante flexibles, su principal limitación es la forma lineal preasumida del modelo; suponiendo estructuras de correlación lineal entre los valores, no puede capturar patrones no lineales. Como alternativa a esta restricción de los modelos ARIMA fueron propuestas de manera favorable las redes neuronales artificiales (ANN) para tratar con modelados no lineal principalmente. Los resultados obtenidos al aplicar ANN han sido superiores a todo tipo de modelo

convencional en muchos escenarios Mouatadid & Adamowski, (2017) y Guo et al., (2018).

Con el transcurrir del tiempo se han probado diversos modelos de demanda de agua usando las metodologías más tradicionales basadas en las series temporales y otras, integrando modelos clásicos con metodologías de aprendizaje automático (ML), dando origen a nuevos modelos que han favorecido el desarrollo de nuevas tecnologías a nivel científico.

En el año 2016 surge una nueva representación matemática de series temporales llamada *Matrix Profile*. Esta es una estructura de datos y algoritmos asociados que ayuda a resolver el problema dual de detección de anomalías y descubrimiento de patrones (Foundation, 2020), basándose en el concepto de unión de similitud que consiste en comparar subsecuencias de la serie temporal con el resto de subsecuencias de la serie (Barreno Recio, 2021). A continuación, se puede apreciar (*ilustración 33*) un ejemplo que complementa la definición de una Matrix Profile.



Ilustración 33. Cálculo de un perfil de distancia a partir de la segunda ventana. (<https://towardsdatascience.com/introduction-to-matrix-profiles-5568f3375d90>)

Definición de Matrix Profile: Una representación de la serie temporal ST' consiste en la transformación de la serie temporal $ST = \{y_1, y_2, \dots, y_n\}$ en otro vector de dimensión inferior $ST' = \{y'_1, y'_2, \dots, y'_p\}$ donde $p < n$. Si dos series ST_a y ST_b son similares entonces sus representaciones ST'_a y ST'_b también deben ser similares a ellas.

Actualmente una de las representaciones o abstracciones de alto nivel de los datos de series temporales sin procesar, es el algoritmo Symbolic Aggregate approxImation (SAX). El SAX permite una discretización de las series temporales, reduciendo la dimensionalidad de la serie. De acá que los modelos predictivos basados en el análisis de datos de epidemiología, son a menudo los más utilizados.

3.4.1.1. Interoperabilidad y explicabilidad de los modelos

Los modelos creados usando metodologías de ML han logrado extraer información útil y valiosa de grandes cantidades de datos para construir sistemas computacionales inteligentes que toman decisiones sobre problemas específicos en varios campos como finanzas, banca, medicina, educación (Pintelas et al., 2020) y también en el área de la ingeniería. Uno de los grandes retos actuales en las investigaciones que involucran modelos de ML es la de encontrar modelos que además de ser eficientes y precisos, sean fácilmente interpretables.

La interpretabilidad y la explicabilidad constituyen factores clave que son de gran importancia en los modelos prácticos de ML. La interpretabilidad es la capacidad de comprender y observar el mecanismo de un modelo o el comportamiento de predicción en función de su estímulo de entrada, mientras que la explicabilidad, es la capacidad de demostrar y explicar en términos comprensibles para un ser humano, la decisión de los modelos sobre los problemas de predicción (Arrieta et al., 2020).

Según su nivel de explicabilidad y precisión los modelos de ML podrían estar clasificados en tres categorías: modelos de caja blanca, caja negra o caja gris.

- Caja blanca: es un modelo cuya lógica interna, funcionamiento y pasos de programación, son transparentes y, por lo tanto, su proceso de toma de decisiones es interpretable (Pintelas et al., 2020).
- Caja negra: suele ser un modelo de ML más preciso cuyo funcionamiento interno no se conoce y es difícil de interpretar, lo que significa que un probador de software, por ejemplo, solo podría conocer las entradas esperadas y las salidas correspondientes de este modelo (Pintelas et al., 2020).
- Caja gris: es una combinación de modelos de caja negra y caja blanca (Bohlin, 2006). El objetivo principal de un modelo de Caja Gris es el desarrollo de un conjunto de modelos de Caja Negra y Caja Blanca, con el fin de combinar y adquirir los beneficios de ambos, construyendo un modelo compuesto global más eficiente (Pintelas et al., 2020).

Teniendo en cuenta el desarrollo de algunos de los modelos de ML y en función de su nivel de explicabilidad se podría tener la siguiente clasificación (*ilustración 34*):

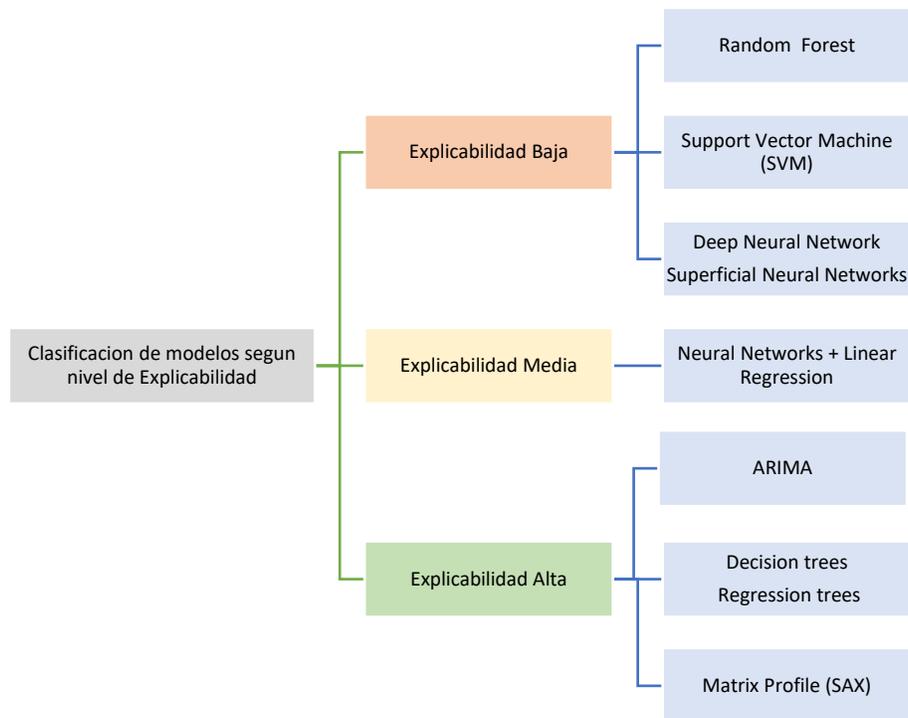


Ilustración 34. Clasificación de explicabilidad de modelos

Basado en las principales técnicas de interpretación y en los tipos de resultados, se ha realizado un importante trabajo de investigación en los últimos años para desarrollar métodos y algoritmos sofisticados, para interpretar modelos de aprendizaje automático. Estos algoritmos siempre están limitados por el compromiso entre explicabilidad y precisión, lo que significa que para lograr una alta precisión a menudo se sacrifica la explicabilidad, aunque el objetivo principal de todos estos algoritmos es el desarrollo de modelos precisos e interpretables (Pintelas et al., 2020).

Por lo anteriormente comentado en función de la relevancia en la explicabilidad y simplicidad de los modelos en este trabajo de investigación se planteó trabajar con el algoritmo SAX (Symbolic Aggregate aproXimation) con el cual se logra hacer la discretización e identificación de patrones en series temporales de redes de abastecimiento de agua urbana.

3.4.2 Análisis de datos basados en epidemiología para la demanda de agua

Los modelos predictivos basados en el análisis de datos de epidemiología a menudo incluyen la evaluación de múltiples relaciones entre dos o más series temporales. Este es el caso de comprender cómo cambian los datos a lo largo del tiempo y cómo aplicar adecuadamente los métodos de pronóstico (Navarrete-López et al., 2019). Cuando se trabaja con múltiples modelos de series temporales, es habitual establecer asociaciones entre la exposición a una amenaza y el resultado de salud. Por ejemplo, a partir del estudio de las asociaciones entre la exposición a elementos como la contaminación del aire, las variables climáticas o el polen, se pueden derivar resultados importantes con respecto a las formas de mostrar los síntomas y las consecuencias de la enfermedad (Cai et al., 2009; Jackson et al., 2017). En el abastecimiento de agua, asociaciones similares muestran, por ejemplo, cómo las altas temperaturas inciden en producir progresivamente mayores niveles de demanda de

agua para una determinada población (Brentan et al., 2017). Como consecuencia, es útil investigar las relaciones causa-efecto entre las variables exógenas y los niveles de demanda de agua.

Las bases de datos para llevar a cabo estos estudios están naturalmente disponibles a intervalos de tiempo regulares. Esto explica por qué el análisis de series temporales juega un papel tan importante en la epidemiología. El objetivo principal es explicar los cambios en la principal variable dependiente del tiempo del proceso a través de cambios en los niveles de las otras variables bajo análisis (Navarrete-López et al., 2019). Esto hace que la regresión de Poisson y los modelos ARIMA hayan sido ampliamente utilizados dentro de esta área de investigación (Bhaskaran et al., 2013).

Otros modelos más complejos como los ARIMA multivariante y modelos dinámicos que buscan encontrar relaciones entre las series temporales relativas a la exposición y el resultado, también se han utilizado (Clayton et al., 2013). También es habitual trabajar con funciones de transferencia y modelos de intervención, cuyo uso y rendimiento está frecuentemente limitado por la complejidad de dichos modelos (Helfenstein, 1991).

Los enfoques de análisis de datos epidemiológicos (EDAA: Epidemiological Data Analysis Approaches) facilitan el enfoque de análisis de demanda de agua múltiple al considerar las diversas áreas de medición de distrito (DMA) en las que a menudo se divide un WDS (Herrera, 2011). Dado que los métodos de epidemiología modelan las relaciones de causa y efecto, los EDAA se amplían fácilmente para investigar el impacto en el consumo de agua relacionado con los efectos de eventos que van desde operaciones de válvulas hasta políticas relacionadas con condiciones climáticas extremas. Los modelos de pronóstico EDAA permiten desagregar los efectos mencionados anteriormente para las DMA individuales (Izquierdo et al., 2010). Esto se hace mediante un análisis de datos a nivel de paciente que permite desarrollar un modelo predictivo diferente por DMA. El enfoque se completa naturalmente teniendo en cuenta las correlaciones existentes entre DMA de un mismo WDS (Herrera et al.,

2016). Los EDAA proporcionan, de esta manera, un sentido espacial de análisis al enfoque temporal común al abordar la modelización de la demanda de agua a nivel urbano (Navarrete-López et al., 2019).

3.5 ALGORITMO SAX

El algoritmo Symbolic Aggregate aproXimation (SAX) es un algoritmo de discretización de series temporales. SAX se desarrolló inicialmente para reducir la información densa y extensa de una serie temporal en cadenas de caracteres creadas para ser más breves que la serie temporal original (Navarrete-López et al., 2019). SAX también se ha encontrado útil para varias tareas de minería de datos (*ilustración 35*), en particular, indexación (Toshniwal, 2009), agrupamiento (Aghabozorgi et al., 2014; Martínez-Álvarez, 2011) y clasificación (Yuan et al., 2015). La vocación principal de los métodos basados en SAX es proporcionar una vía adecuada para los modelos de reconocimiento de patrones asociados con series temporales (Rajaraman et al., 2011).

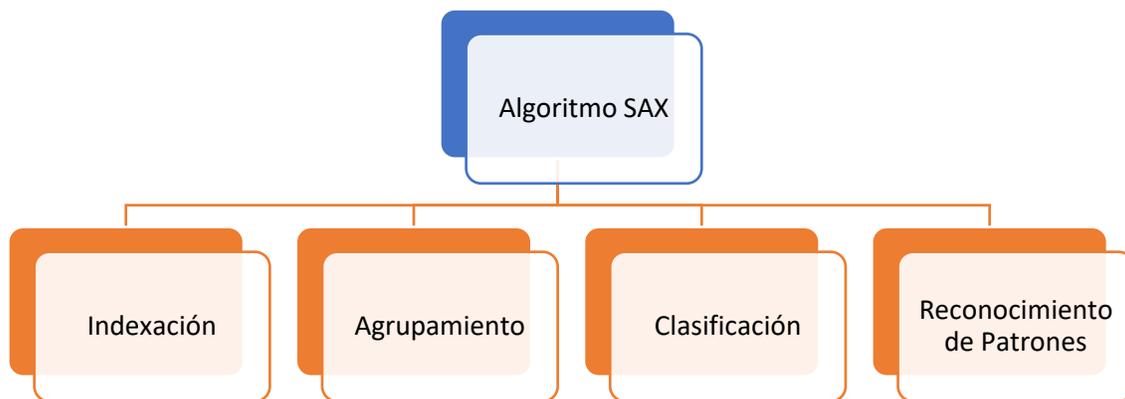


Ilustración 35. Aplicaciones del algoritmo SAX

Una propiedad muy deseable de cualquier representación de serie temporal es la representación simbólica; casi todos los conjuntos de datos de series temporales son de muy alta dimensión. Este es un hecho desafiante porque todos los algoritmos de indexación y minería de datos no triviales se degradan exponencialmente con la

dimensionalidad. Por ejemplo, por encima de 16 a 20 dimensiones, las estructuras de índice se degradan a un escaneo secuencial (Hellerstein et al., 1997).

Una forma de reducción de la dimensionalidad en series de tiempo es la discretización de series temporales (*ilustración 36*), la cual consiste en convertir valores continuos en valores discretos (Rechy-Ramírez et al., 2011). Por lo general, este proceso necesita definir intervalos o cortes de valores en el espacio continuo de cada variable (Márquez Grajales, 2019).

Dichos intervalos son secuencias o segmentos de la serie de tiempo. Al conjunto de estos se le conoce como esquema de discretización y se expresa como $D = \{[d_0, d_1]_1, [d_1, d_2]_2, \dots, [d_{n-2}, d_{n-1}]_{g-1}, [d_{n-1}, d_n]_g\}$ donde d_0 y d_n son los valores máximos y mínimos del componente temporal de la serie (Márquez Grajales, 2019). Cada intervalo en D representa un segmento donde el conjunto de valores de la serie contenidos dentro de este son transformados en valores discretos. Al número total de segmentos g se le conoce como grado de discretización (Rechy-Ramírez et al., 2011).

Los valores que puede tomar la serie discreta son obtenidos mediante diversos

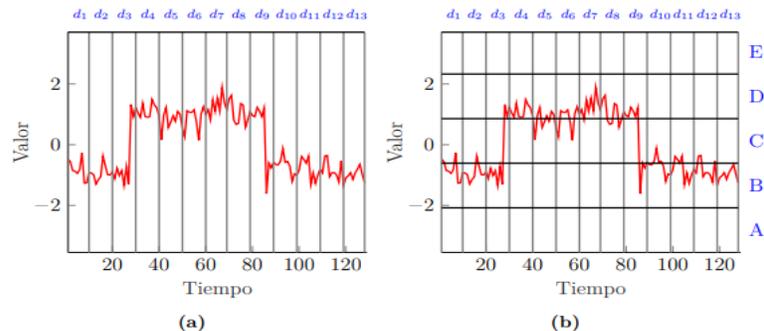


Ilustración 36. (a) Esquema de discretización del componente temporal. (b) Discretización simbólica de series de tiempo. Tomado de: (Márquez Grajales, 2019).

métodos propuestos en la literatura especializada (Bondu et al., 2013); (Bai et al., 2013; Lkhagva et al., 2006). Una de las formas más utilizadas es calcular el promedio de los valores encontrados en cada intervalo de tiempo $x_{[d_i, d_{i+1}]}^{(-i)}$, obteniendo una serie temporal reducida con valores continuos $(ST)^- = \{x^{(-1)}, x^{(-2)}, \dots, x^{(-s)}\}$; a este método se le conoce como Piecewise Aggregate Approximation (PAA). Para

convertir estos promedios en valores discretos, se utilizan representaciones simbólicas donde la serie de tiempo discreta es conocida como strings o cadenas (Márquez Grajales, 2019).

Hay una enorme riqueza de algoritmos existentes y estructuras de datos que permiten la manipulación eficiente de dichas cadenas. Algunos ejemplos de “herramientas” que son no definidas para secuencias de valores reales, pero están definidas para enfoques simbólicos incluyen hashing, árboles de sufijos (*ilustración 37*), modelos de Markov (*ilustración 38*), árboles de decisión (Lin et al., 2007).

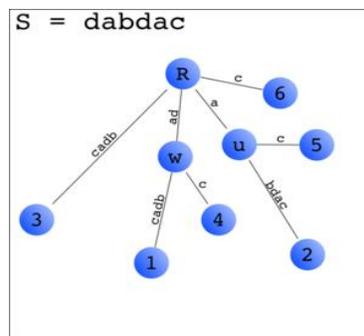


Ilustración 37. Ejemplo árbol de sufijos. Tomado de: <https://kripkit.com/arbore-de-sufijos/>

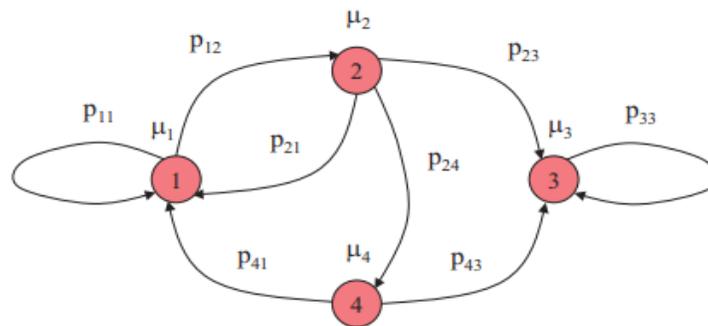


Ilustración 38. Cadena de Markov en tiempo continuo. Tomado de: <https://complex-systems-ai.com/es/proceso-de-markov/cadenas-de-markov-en-tiempo-continuo/>

La salida SAX es una cadena de caracteres o "palabra" que se aproxima a la serie temporal (McLaurin et al., 2014). Por lo tanto, es esencial trabajar con métodos apropiados para capturar patrones formados por palabras que caracterizan y representan los datos de la serie temporal. Los árboles de sufijos (Lin et al., 2007; Rasheed et al., 2010) se han utilizado ampliamente para este propósito y, por lo tanto,

proporcionan una forma de organizar subcadenas de palabras SAX, lo que facilita los procesos de extracción de patrones. Los árboles de sufijos tienen varias aplicaciones (Lin et al., 2007; Rasheed et al., 2010) debido a su capacidad para proporcionar soluciones de tiempo lineal a configuraciones de cadenas complejas (Keogh et al., 2002).

3.5.1 Enfoque simbólico del algoritmo SAX.

SAX permite que una serie de tiempo de longitud arbitraria n se reduzca a una cadena de longitud arbitraria w , ($w < n$, típicamente $w \ll n$). El tamaño del alfabeto también es un número entero arbitrario a , donde $a > 2$. El procedimiento de discretización que propone el algoritmo SAX, es único en el sentido de que utiliza una representación intermedia entre la serie temporal sin procesar y las cadenas simbólicas. Primero se transforman los datos en la representación de aproximación agregada por partes (PAA) y luego se simbolizan las representaciones PAA en una cadena discreta (Lin et al., 2007).

Este enfoque usado por el algoritmo SAX tiene dos ventajas importantes:

Reducción de dimensionalidad: se puede utilizar el poder de reducción de dimensionalidad bien definido y bien documentado de PAA (Keogh et al., 2001); (Yi et al., 2000), y la reducción se traslada automáticamente a la representación simbólica (Lin et al., 2007).

Límite inferior: una medida de distancia entre dos cadenas simbólicas limita la distancia real con la serie de tiempo original. La medida de distancia simbólica limita inferiormente la medida de distancia PAA (Lin et al., 2007).

3.5.1.1. Reducción de dimensionalidad vía PAA.

Una serie temporal C de longitud n se puede representar en un espacio de dimensión w – mediante un vector $\bar{C} = \bar{c}_1, \dots, \bar{c}_w$. El i -ésimo elemento de \bar{C} se calcula mediante la siguiente ecuación:

$$\bar{c}_i = \frac{w}{n} \sum_{j=\frac{n}{w}(i-1)+1}^{\frac{n}{w}i} c_j. \quad (1)$$

Para reducir la serie temporal de n dimensiones a w dimensiones, los datos se dividen en w "marcos" de igual tamaño. Se calcula el valor medio de los datos que caen dentro de un marco y un vector de estos valores se convierte en la representación de datos reducidos. La representación se puede visualizar como un intento de aproximar la serie de tiempo original con una combinación lineal de funciones de base de caja (Lin et al., 2007) (*ilustración 39*).

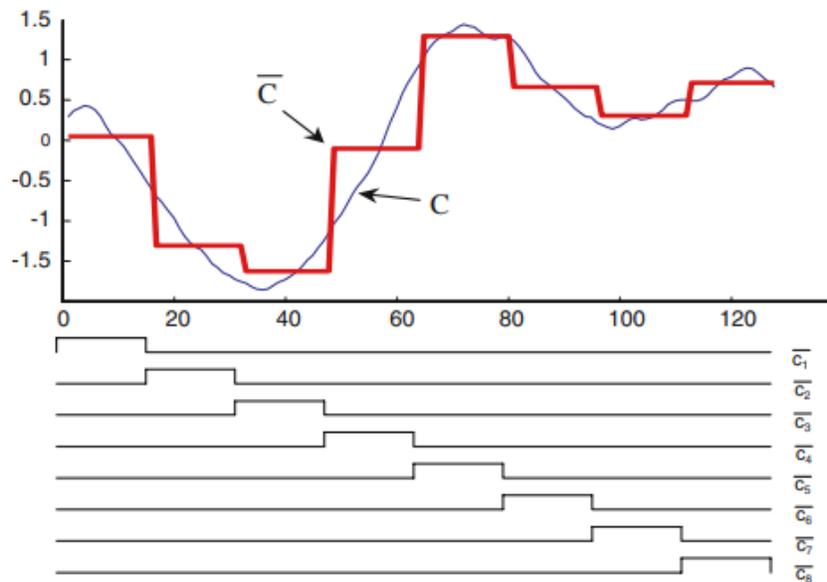


Ilustración 39. La representación PAA se puede visualizar como un intento de modelar una serie de tiempo con una combinación lineal de funciones de base de caja. En este caso, una secuencia de longitud 128 se reduce a ocho dimensiones. (Lin et al., 2007)

3.5.1.2. Discretización.

Una vez se ha transformado una base de datos de series temporales en PAA, es posible aplicar una transformación adicional para obtener una representación discreta (*ilustración 40*). Es deseable tener una técnica de discretización que produzca símbolos con equiprobabilidad (Apostolico et al., 2002); (Lonardi, 2001). Dado que las series de tiempo normalizadas tienen una distribución altamente gaussiana, es posible determinar los "puntos de ruptura" que producirán áreas de igual tamaño bajo la curva gaussiana (Larsen et al., 2005).

Definición 1. Puntos de ruptura: los puntos de ruptura son una lista ordenada de números $B = \beta_1, \dots, \beta_{a-1}$ tal que el área bajo una curva gaussiana $N(0, 1)$ de β_i hasta $\beta_{i+1} = \frac{1}{a}$ (β_0 y β_a se definen como $-\infty$ y ∞ , respectivamente).

Estos puntos de corte pueden determinarse buscándolos en una tabla estadística. Una vez que se han obtenido los puntos de corte, se puede discretizar una serie de tiempo de la siguiente manera. Primero obtener un PAA de la serie temporal. Todos los coeficientes PAA que están por debajo del punto de corte más pequeño se asignan al símbolo "*a*", todos los coeficientes mayores o iguales que el punto de corte más pequeño y menores que el segundo punto de corte más pequeño se asignan al símbolo "*b*". Se llama palabra a la concatenación de símbolos que representan una subsecuencia (Lin et al., 2007).

Definición 2. Palabra: Una subsecuencia C de longitud n se puede representar de la siguiente manera como una *palabra* $\hat{C} = \hat{c}_1, \dots, \hat{c}_w$. Sea α_i el i -ésimo elemento del alfabeto, es decir, $\alpha_1 = \mathbf{a}$ y $\alpha_2 = \mathbf{b}$. Entonces el mapeo de una aproximación PAA, \bar{C} a una palabra \hat{C} se obtiene de la siguiente manera:

$$\hat{c}_i = \alpha_j \text{ si } \beta_{j-1} \leq \bar{c}_i < \beta_j. \quad (2)$$

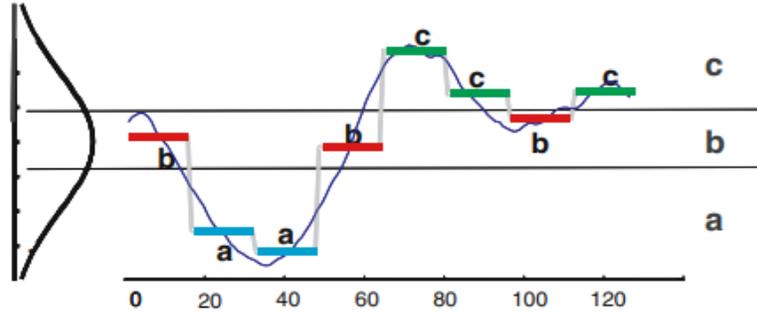


Ilustración 40. Una serie de tiempo se discretiza obteniendo primero una aproximación PAA y luego usando puntos de corte predeterminados para mapear los coeficientes PAA en símbolos SAX. En el ejemplo anterior, con $n = 128$, $w = 8$ y $a = 3$, la serie temporal se asigna a la palabra **baabccbc**. (Lin et al., 2007)

La ilustración 40 presenta un ejemplo con un diccionario $\Sigma = \{a, b, c\}$ de 3 letras.

3.5.1.3. Medida de distancia.

Una característica importante de SAX es tener asociada una métrica, una distancia entre cada dos series de tiempo codificadas en SAX, que es un límite inferior de la distancia entre series de tiempo cuando se calcula en el espacio original. La distancia tradicionalmente usada entre dos palabras SAX es la llamada MINDIST, y consiste en buscar las distancias entre cada par de símbolos o letras (Navarrete-López et al., 2019). Dadas dos series de tiempo Q y C de la misma longitud l , la ecuación 3 muestra cómo se define MINDIST:

$$\text{MINDIST}(\hat{Q}, \hat{C}) = \sqrt{\frac{l}{w}} \sqrt{\sum_{i=1}^w (\text{dist}(\hat{q}_i, \hat{c}_i))^2}. \quad (3)$$

donde w es el número de segmentos en que se dividen ambas series temporales y $\hat{Q} = \hat{q}_1, \dots, \hat{q}_w$ y $\hat{C} = \hat{c}_1, \dots, \hat{c}_w$ son las palabras en las que ambas series temporales, Q y C , se codifican respectivamente utilizando la Ecuación (1). La función **dist()** se puede implementar utilizando una tabla de búsqueda que contiene distancias calculadas previamente entre las letras del alfabeto SAX en uso $\Sigma = \{\alpha_1, \dots, \alpha_m\}$. Esas son distancias mínimas entre letras por lo que MINDIST también

devuelve la distancia mínima entre la serie temporal de los valores originales a través de sus correspondientes palabras (Navarrete-López et al., 2019).

Específicamente, la distancia entre dos símbolos $\alpha_i, \alpha_j \in \Sigma$ es calculada por la ecuación (4)

$$\text{dist}(\alpha_i, \alpha_j) = \begin{cases} 0, & \text{si } |i - j| \leq 1 \\ \beta_{\max(i,j)-1} - \beta_{\min(i,j)}, & \text{de otra manera} \end{cases} \quad (4)$$

dónde β es el valor de la Normal estandarizada en el correspondiente punto de ruptura de la configuración PAA.

3.5.1.4. Reconocimiento de patrones en palabras SAX.

Una tabla que contenga una entrada para cada palabra SAX puede ser una herramienta útil para contar el número de ocurrencias de ciertas subcadenas. Después de crear y guardar una tabla adecuada, los elementos del mismo tamaño de la cadena SAX para los que se encuentran los patrones se pueden identificar fácilmente. Se observa que cualquier consulta en la tabla toma tiempo constante (Keogh et al., 2002). Una forma conveniente de organizar un directorio de subcadenas a partir de una palabra es la técnica llamada árbol de sufijos (Gusfield, 1997). El árbol de sufijos es un tipo de búsqueda digital que representa un conjunto de cadenas sobre un alfabeto finito. Los árboles de sufijos son ideales para extraer patrones periódicos (Rasheed et al., 2010) como los que se encuentran comúnmente en la demanda de agua urbana. Además, la técnica del árbol de sufijos permite encontrar anomalías o patrones sorprendentes, también, que pueden entenderse como consecuencias de maniobras de operación en el suministro de agua (abrir/cerrar válvulas) o la presencia de eventos disruptivos (explosión de tuberías, por ejemplo) (Navarrete-López et al., 2019).

Una definición formal del árbol de sufijos, T , asociado con una cadena x de n caracteres extraídos de un alfabeto ordenado Σ es el siguiente. La cadena x se anota

primero como $x[1:n]$, y cualquier subcadena de la forma $x[i:n]$, será llamada un sufijo de x . Sea $\$$ un carácter especial, que no coincida con ningún carácter en Σ . El carácter $\$$ es un marcador final derecho y su propósito es separar en T el sufijo $x[1:n]$ del sufijo $x[j:n]$, para todo $i > j$, siempre que el primero sea prefijo del segundo. Esto asegura que ningún sufijo sea un prefijo de otro sufijo, y que habrá n nodos hoja, uno para cada uno de los sufijos de T . En consecuencia, cada hoja de T está etiquetada con un entero distinto j tal que el camino desde la raíz hasta la hoja j corresponde al sufijo $x[j:n]$ (Grossi et al., 1993).

La Tabla 4 proporciona un resumen de los pasos principales para el proceso de creación del árbol de sufijos.

Sea T un árbol correspondiente a una cadena x de longitud n :
<ol style="list-style-type: none"> 1. T tiene exactamente n hojas, numeradas del 1 al n 2. Cada borde tiene una etiqueta, que es una subcadena de x 3. Cada nodo interno tiene al menos 2 hijos 4. Etiquetas de dos aristas a partir de un interior el nodo no comienza con el mismo carácter 5. La etiqueta del camino desde la raíz hasta una hoja. Numerado i es el sufijo de x a partir de posición i, es decir, $x[i, \dots, n]$

Tabla 4 Visión general para la construcción de árbol de sufijos

Hay representaciones compactas para los árboles de sufijos que permiten codificar una gran cantidad de subcadenas diferentes, manejando como máximo $2n$ nodos (McCreight, 1976). Esto es importante ya que este hecho permite indexar palabras grandes a través de varias operaciones de orden $\varphi(n)$.

La decisión de tomar el método del árbol de sufijos para la extracción de patrones de SAX proviene de su eficiencia computacional, la representatividad de los patrones que logra encontrar y su gran aplicabilidad. Esto ha llevado a esta técnica a

desempeñar un papel central también en una práctica más compleja de extracción de patrones de cadenas, como es el caso del genoma humano (Välimäki et al., 2007).

Como ejemplo de cómo funciona el proceso del árbol de sufijos, la palabra *cabdcabb* se construyó usando el alfabeto propuesto $\Sigma = \{a, b, c, d\}$. La ilustración 41 muestra el proceso de extracción de patrones para este ejemplo.

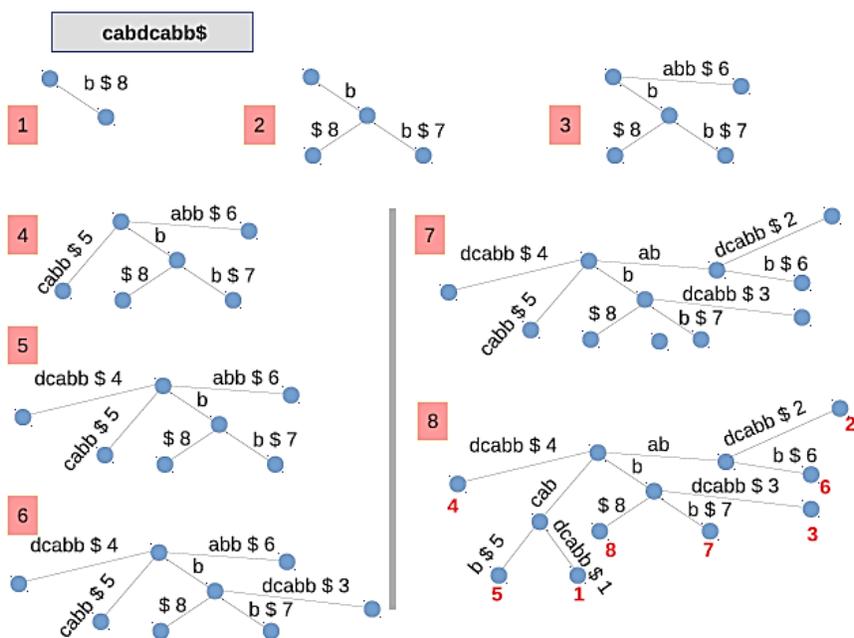


Ilustración 41. Ejemplo de proceso de extracciones en árbol de sufijos.

En el ejemplo, la primera acción es tomar el sufijo más pequeño *b \$*, comenzando a la derecha de la palabra. Después de crear una raíz, se agrega una primera rama relacionada con *b \$* a la raíz y se indexa por 8, que es la posición original del sufijo en la palabra. Pasando a la posición 7 de la palabra se encuentra el sufijo *bb \$*; como ya existe una rama *b*, se agrega otra rama a partir de la primera (ver paso 2 de la *ilustración 41*) y se parte la anterior dando lugar a otra con la marca final *\$*. El siguiente sufijo es *abb \$*. No hay rama que represente *a*. Se crea una nueva rama a partir de la raíz para representar todo el sufijo, realizando el índice 6 (ver paso 3 de la

ilustración 41). El proceso se itera siguiendo el enfoque teórico explicado anteriormente hasta conseguir indexar convenientemente cada sufijo en T (paso 8 de la *ilustración 41*) (Navarrete-López et al., 2019).

3.6 ARBOLES DE SUFIJOS Y ALGORITMO SAX EN DIVERSOS CONTEXTOS

Las aplicaciones de los árboles de sufijos han demostrado ser útiles en diversos campos de acción. Dada su naturaleza para el manejo de cadenas de caracteres son eficientes y ampliamente usados en la identificación de patrones y secuencias en áreas como la bioinformática y más puntualmente en la secuenciación de grandes colecciones de cadenas de ADN. En la secuenciación de cadenas la implementación clásica se vuelve impracticable para grandes volúmenes de información dado que ocupan demasiado espacio, que siempre muchas veces mayor que el texto mismo. Luego, no pueden ser almacenados en memoria principal, lo que en la práctica significa un aumento importante del tiempo de respuesta (Abeliuk Kimelman, 2012).

En el campo actual de la tecnología informática el algoritmo SAX está siendo estudiado para mejorar aplicaciones que permitan el aumento de potencia de procesamiento de información en computadores. El trabajo propuesto por Dumss et al., (2020) presenta un enfoque utilizando Aproximación agregada simbólica (SAX) combinada con datos recuperados de sistemas de información industrial (IIS) como planificación de recursos empresariales para enriquecer conjuntos de datos con contexto adicional y, al mismo tiempo, reducir el tamaño de la serie temporal necesaria para guardar o transmitir los datos representativos. El enfoque presentado en la investigación de Dumss usando SAX combinado con información adicional, permitió la detección rápida de tendencias no deseadas y sus causas.

El algoritmo SAX también ha sido implementado más recientemente para analizar datos de series de tiempo en el campo de la astronomía, donde se producen grandes volúmenes de información recopilada a partir de observaciones terrestres y espaciales.

En astronomía, existe una población de asteroides entre las órbitas de Marte y Júpiter. Un mecanismo que contribuye a que los asteroides del cinturón principal entren en la región cercana a la Tierra podrían ser las colisiones de alta velocidad entre asteroides (Menichella et al., 1996). Mientras que, en la región cercana a la Tierra, definida como una región que satisface la desigualdad q (distancia del perihelio) $< 1,3$ au y Q (distancia del afelio) $> 0,98$ au, los asteroides pueden experimentar encuentros cercanos con ciertos planetas (Mercurio, Venus, la Tierra y Marte) y también la luna (Riza et al., 2022). Como resultado de estos encuentros cercanos, los asteroides pueden fragmentarse debido a las fuertes fuerzas de las mareas o, si sobreviven, sus órbitas pueden cambiar drásticamente en poco tiempo. Los cambios drásticos en las órbitas pueden cambiar el destino futuro de los asteroides, ya sea que sigan orbitando alrededor del sol, colisionen con objetos masivos o sean expulsados del sistema solar. Según (Gallardo, 2006), es laborioso identificar cuál de los cientos de resonancias de movimiento medio (MMR) que teóricamente existen cerca del semieje mayor de la órbita es la que afecta el movimiento del asteroide.

Puntualmente en el trabajo desarrollado por (Riza et al., 2022) se propone un nuevo modelo computacional integrando los algoritmos SAX y de proyección aleatoria para cambiar la serie de tiempo a una representación en cadena y posteriormente detectar motivos o patrones presentes en los datos, analizando las MMR.

3.7 ALGORITMO SAX PARA EL ANÁLISIS DE LA DEMANDA DE AGUA

El pronóstico epidemiológico a menudo funciona con un análisis espacial para complementar el análisis de series temporales. Se puede afirmar que, en base a las similitudes que se pueden encontrar entre varias DMA, es posible mejorar el conocimiento previo de la demanda de agua para una DMA objetivo específico (Navarrete-López et al., 2019). Esto se logra al considerar la información que se encuentra en aquellos DMA que exhiben patrones de demanda similares a la DMA objetivo. Estas similitudes se basan en los métodos descritos anteriormente con respecto a las distancias SAX y los árboles de sufijos. Existen varias formas alternativas de abordar las distancias o (des)similitudes entre series temporales (Aghabozorgi et al., 2015). En muchos casos, estas alternativas se basan en algoritmos desarrollados para el agrupamiento de series temporales completas tomando las características principales de los algoritmos de agrupamiento estático. Luego, el proceso continúa modificando la definición de similitud a una nueva apropiada, o aplicando una transformación a la serie de tiempo para obtener características estáticas (Liao, 2005). El método en el que se basa este documento tiene la ventaja de funcionar bien con grandes bases de datos de series temporales, ya que la dimensión del problema se reduce naturalmente con los métodos relacionados con SAX. El proceso general se muestra en el diagrama de flujo siguiente (*ilustración 42*), donde DMA_i está destinado a mejorar la precisión predictiva en su demanda.

El proceso de predicción de mejora que se muestra, comienza con la selección de un DMA como objetivo. Luego, cada serie temporal de demanda de agua se transforma en SAX, así para todos los DMA. Se evalúan todas las distancias por pares entre el objetivo y los otros DMA. Se selecciona el DMA más cercano al objetivo para mejorar la información del modelo predictivo sobre el DMA objetivo. Este proceso tiene una clara ventaja sobre un único modelo predictivo y, además, muestra una forma de aplicar el conocimiento producido por SAX (Navarrete-López et al., 2019).

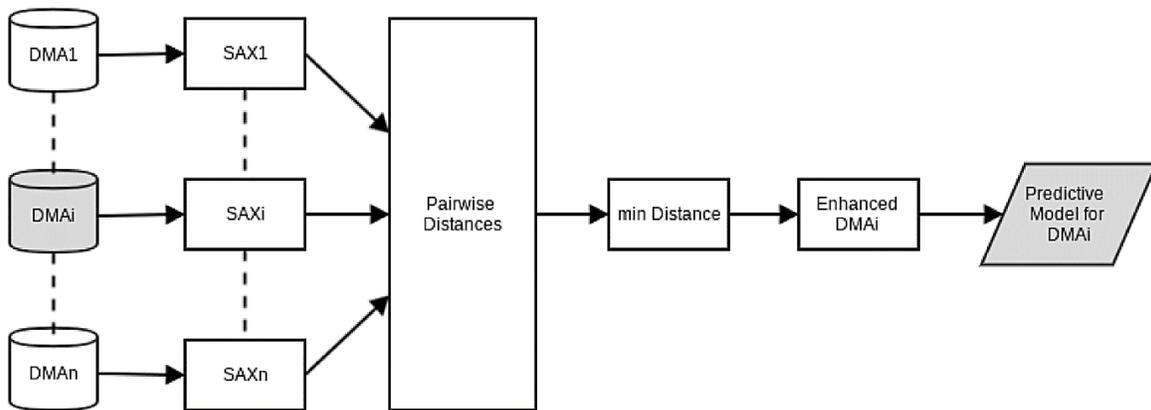


Ilustración 42. Diagrama de flujo del proceso general para mejorar los modelos predictivos de demanda de DMA. (Navarrete-López et al., 2019)

Este modelo SAX será aplicado más adelante en el desarrollo del capítulo cinco, el cual tratará el caso de estudio sobre gestión de demanda en la red de suministro de agua urbana de la ciudad de Franca, Brasil.

CAPÍTULO 4

GESTIÓN DE FALLOS

“El capítulo presente trata puntualmente el otro componente básico de la administración del recurso hídrico como lo es la gestión de fallos. Comienza con la exposición de algunos conceptos sobre fallos, luego muestra las principales causas de fallos y por qué resulta importante procurar su anticipación. Se aprecian algunos métodos de la epidemiología para la estimación de fallos y se cierra con la aplicabilidad de los sistemas multiagente y el modelo SIR .”

4.1 CONCEPTO DE FALLOS EN REDES DE ABASTECIMIENTO DE AGUA

Para tratar el concepto de *fallos en redes de abastecimiento de agua*, conviene apreciar lo que es una red de abastecimiento de agua. Una completa definición de este tipo de red se aprecia en el trabajo de Vargas J. & Huyhua H. (2020) en el que una red de abastecimiento de agua consiste en un conjunto de obras necesarias para captar, conducir, tratar, almacenar y distribuir el agua desde fuentes naturales ya sean subterráneas o superficiales, hasta las viviendas de los habitantes que serán favorecidos con dicho sistema

Por consiguiente, un diseño adecuado de este tipo de red mejora sustancialmente la calidad de vida de una población y su consecuente desarrollo. Por esta razón un sistema de abastecimiento de agua potable debe cumplir con normas y regulaciones vigentes para garantizar su correcto funcionamiento, Vargas J. et al., (2020).

En concordancia con la anterior definición, resulta evidente pensar en que una red de distribución de agua es un sistema complejo que se encuentra sometido a cambios frecuentes, a una evolución constante (*ilustración 43*). Y dicha evolución se presenta

no solo en el aspecto físico de la red y sus componentes, sino también en los aspectos lógicos y tecnológicos; esto facilita la operación de la misma, pero al mismo tiempo la vuelve más compleja. De ahí que la labor de gestionar una red de distribución abarca varios componentes tanto de infraestructura, como de tecnología, que de no administrarlos adecuadamente pueden llevar a un primer escenario de alto impacto en las compañías que prestan el servicio.

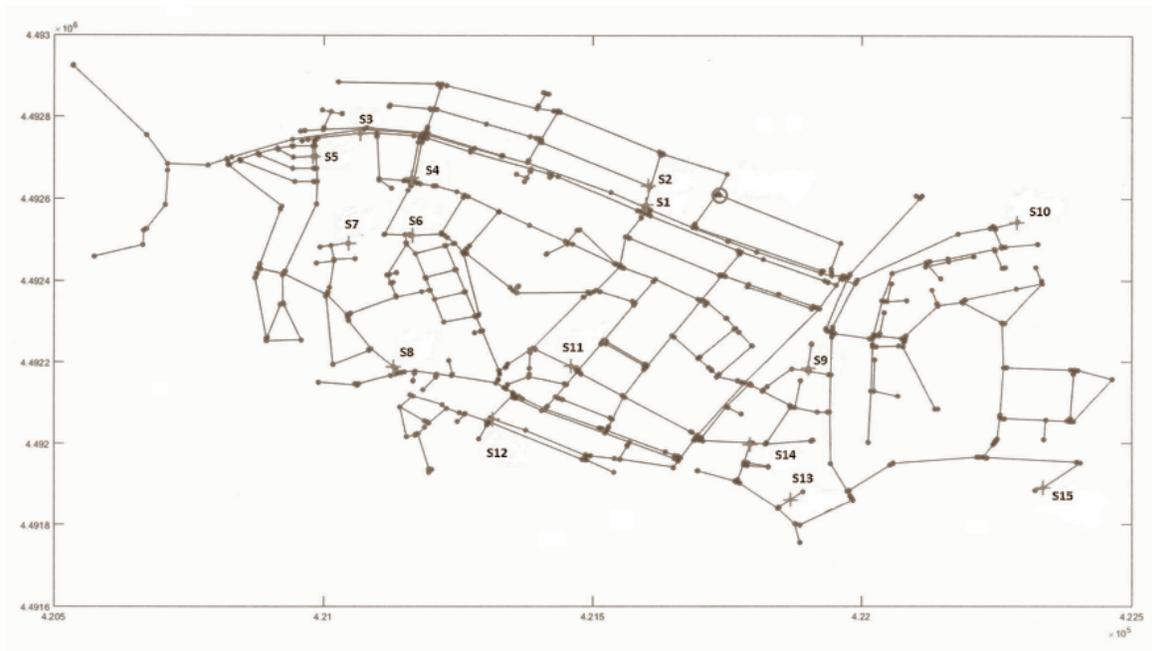


Ilustración 43. Ejemplo de Red de Distribución de agua en Madrid en un sector de 1000 nodos. Pereira Martín, M., Blesa, J., Puig, V., & Muñoz De la Peña, D. (2019).

Partiendo de lo anterior, surge la noción de la *gestión de fallos* como un componente necesario e imprescindible en la administración de una red de distribución, por lo que la gestión de fallos, a grandes rasgos, se ocupa de descubrir, aislar y corregir los problemas que se presenten y que puedan cerrar la prestación del servicio; los fallos deben ser aislados con la mayor rapidez posible para proteger al resto de la red, por ejemplo, desconectando la porción afectada de la red (Du Toit et al., 2015).

Como ya se indicó anteriormente en la introducción, existen diferentes escenarios que pueden afectar a las redes de suministro de agua; sin lugar a dudas, el paso inevitable

del tiempo lleva a que la infraestructura envejezca y la red sufrirá procesos de deterioro lo cual, en efecto, incidirá en el transporte y distribución del agua. Las fisuras que se puedan presentar no solo generan una pérdida en el volumen del agua, sino también una pérdida de energía y la consecuente pérdida económica. Puede decirse que la cantidad de agua perdida oscila entre de un 20% a un 30% de la producción total (Daoudi et al., 2005).

Además del paso del tiempo, otros factores que también inciden en la presencia de fallos dentro de una red de distribución de agua son: malas conexiones entre tuberías, corrosión interna o externa de la tubería, daño mecánico causado por una carga excesiva en la tubería, movimiento de tierra, altas presiones en la red, daños en las tuberías por excavaciones, temperaturas extremas, defectos en las tuberías y baja calidad de la mano de obra (Puust et al., 2010).

4.2 PRINCIPALES CAUSAS DE FALLOS

Es indudable que la presencia de fallos en una red de distribución se causa por diversos factores, los cuales pueden originarse en forma separada e individual o, en conjunto y simultáneamente. Además, mientras más extensa y compleja sea la red, más compleja será la gestión de fallos de la misma. Pero independientemente de la complejidad de la red, se han considerado tres aspectos principales como causas de fallos:

- a) Un aspecto relacionado con las propiedades estructurales de las tuberías, del material, de la interacción entre el medio y la tubería y, por último, la calidad de la instalación.
- b) Otro aspecto es el relacionado con las cargas internas en función de la presión de operación y cargas externas, sobrecargas del suelo, tráfico pesado, cargas de heladas y daños accidentales o intencionales.
- c) Por último, un aspecto relacionado con el deterioro del material debido a cambios internos y/o externos y, por ambientes químicos. (Rajani et al., 2001)

Si bien los aspectos (a) y (c) tienen un impacto grande como causales de fallos, se puede pensar en que dichas causas son “previsibles” por así decirlo; pues tarde o temprano el desgaste del material se hace evidente y es necesaria su renovación. La interacción entre el medio y la tubería, así como los cambios internos y/o externos pueden anticiparse de modo que pueda mitigarse su impacto, aunque nunca se deben descartar los imprevistos en estos factores.

Sobre el aspecto (b) se tiene que los factores que lo componen presentan la mayor incidencia en los fallos, a pesar de ser un aspecto altamente “controlable” tanto por instrumentos específicos como por observación directa.

Al abordar en detalle las principales causas de fallos se puede establecer entonces, que las redes de distribución de agua potable son sistemas basados en una extensa y compleja infraestructura que se expande y desarrolla progresivamente en el tiempo, a la vez que se encuentra expuesta a un proceso de deterioro continuo. Espín (2015)

La operación y mantenimiento de este tipo de infraestructuras requiere de una gestión eficaz y la *gestión de activos* es actualmente el proceso de solución más práctico. Las empresas de abastecimiento de agua poseen dos tipos principales de activos: activos dispersos geográficamente, como los sistemas de distribución de agua, y grupos concentrados de activos, como las plantas de tratamiento (McKibben et al., 2002).

4.2.1 Síntesis del proceso de gestión de activos.

Una de las cuestiones clave para la implementación de este modelo de gestión consiste en establecer el conjunto de herramientas tecnológicas (tanto hardware como software) necesarias para satisfacer las necesidades de recopilación, almacenamiento y procesamiento de la información requerida, así como para integrar dichos sistemas con el fin de apoyar la interoperabilidad y el intercambio eficaz de dicha información. Espín (2015)

Conviene precisar algo antes, muchas veces los operadores de las redes de abastecimiento tienden a usar de forma indistinta las expresiones gestión de activos y gestión de mantenimiento; pero en realidad son procesos comerciales muy diferentes. La gestión de activos se orienta a obtener el máximo rendimiento de los bienes o recursos en propiedad y, la gestión de mantenimiento se enfoca en reducir el costo de mantenimiento mientras se extiende la vida útil del activo. Un sistema eficaz de gestión de activos debe incluir un sistema eficaz de gestión de mantenimiento. McKibben, J., et al. (2002).

Precisamente, McKibben, J., & Davis, D. establecen como funciones de la gestión de activos, entre otras, las siguientes:

- Diseño efectivo de activos.
- Gestión eficaz del mantenimiento.
- Rehabilitación o reacondicionamiento y sustitución de activos efectivos.
- Supervisión eficaz del estado de los activos.
- Planificación financiera eficaz.

Se destaca la *supervisión eficaz del estado de los activos*, pues las condiciones actuales de los activos físicos junto con las tendencias en el deterioro de éstos, son medidas importantes que se utilizan para planificar su rehabilitación o su reemplazo. Monitorear el estado de los activos es una actividad que se puede desarrollar a partir de la información operativa, como lo es la medición del tiempo de ejecución, la de temperatura y las de vibración. Las redes subterráneas de servicios públicos representan un problema único ya que estos activos son difíciles de inspeccionar. La inspección periódica de los activos subterráneos es un elemento importante de un sistema efectivo de monitoreo. (McKibben, J., et al.,2002). Es indiscutible que en las redes subterráneas se presente la mayor incidencia de fallos.

Como ya se enunció en la introducción, hay tres herramientas principales: los modelos GIS, los modelos hidráulicos y SCADA, cada una de estas herramientas tiene sus propias funciones, las cuales se resumen en la siguiente tabla.

HERRAMIENTA	FUNCIONES
Sistema de supervisión, control y adquisición de datos (SCADA)	<ul style="list-style-type: none"> - Registro y monitorización del estado de los componentes. - Posibilidad de actuación sobre los elementos en tiempo real. - Almacenamiento de las señales registradas.
Sistema de Información Geográfica (SIG)	<ul style="list-style-type: none"> - Proporcionar la ubicación de los activos. - Consulta de las características de los elementos y del resto de información asociada al SIG y elaboración de informes. - Comprobación de polígonos de corte y posibilidades de suministro alternativo ante una avería. - Apoyo en la elaboración de modelos de simulación hidráulica.
Modelo de Simulación Hidráulica	<ul style="list-style-type: none"> - Realizar estimaciones de la capacidad de los sistemas con valores de consumo reales o previstos. - Comprobación de la fiabilidad y calidad del suministro. - Detección de anomalías, roturas y resto de disfunciones. - Entrenamiento de operadores ante situaciones de alerta.
Modelos de ayuda a la decisión	<ul style="list-style-type: none"> - Evaluación de la red según diferentes criterios: funcionamiento de la red, deterioro de los elementos, etc.

Tabla 5. Principales herramientas involucradas en la gestión de activos. Fuente: Espín, 2015

Valga decir que hay más herramientas como la gestión de mantenimiento asistido por ordenador (GMAO), sistema de información financiera (SIF) o sistemas de gestión de bases de datos (SGBD). Pero para la temática en desarrollo solo se hará una breve profundización de la herramienta SCADA.

La estructura de SCADA se fundamenta en tres componentes: componente de detección de variables y acción de comando, componente de transmisión de señales y, por último, componente de control y automatización.

Para Espín Leal, P. (2015), la interoperabilidad de estos componentes permite en las redes de distribución, entre otras, las siguientes acciones:

- Determinar los consumos por zonas.
- Informar en tiempo real el estado del sistema.
- Actuaciones automáticas.
- Detección rápida de averías.
- Historial de funcionamiento de la red.
- Posibilidad de supervisión y mejora del funcionamiento del sistema.
- Posibilidad de controlar la red de modo optimizado.

La eficacia de estas acciones depende, en buena medida, de la instrumentación de campo: los sensores. Con estos instrumentos se puede recaudar la información de las condiciones de la red a través de la medición de los principales parámetros que inciden en la operación de la red. Esos parámetros son: *caudal, presión y nivel*. (Espín Leal, P., 2015).

El monitoreo continuo de estos parámetros es lo que permite elaborar un adecuado funcionamiento de la red al contar con una cuidadosa gestión de fallos. No obstante, la mayoría de fallos en las redes de suministro se originan a partir de excesos o cargas inadecuadas en estos tres parámetros.

La siguiente imagen (*ilustración 44*) proporcionada por Huang, Y. et al., (2021) resume, en buena medida, lo expuesto hasta el momento sobre las principales causas de fallos.

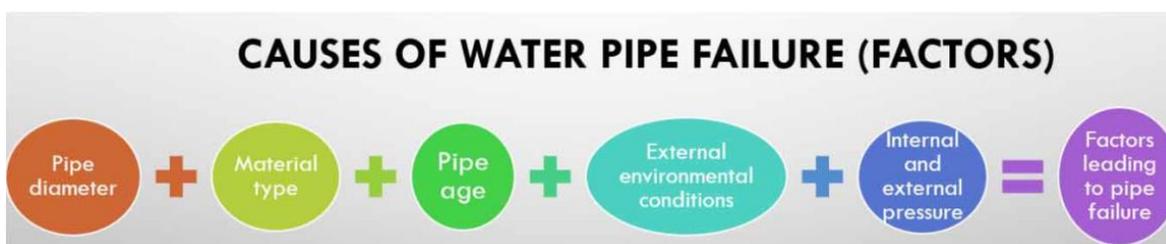


Ilustración 44. Principales causas de fallos. Huang, Y.; Rameezdeen, R.; Chow, C.; Gorjian, N.; Li, Y. (2021)..

4.3 LA IMPORTANCIA DE LA PREDICCIÓN DE FALLOS EN LOS SISTEMAS DE ABASTECIMIENTO DE AGUA

En el contexto mundial actual, la predicción de fallos en una red de abastecimiento de agua adquiere una importancia muy relevante, ya que desempeña un rol crucial para satisfacer la creciente demanda de agua potable en las grandes ciudades.

Como ya se pudo apreciar, las interrupciones del servicio de suministro de agua pueden ocurrir debido a instancias de fallas de componentes, como en tuberías y válvulas, o en sobrecargas, como en caudales o variaciones de presiones. Por lo general, este tipo de fallos provoca la necesidad de excluir la parte de la red de

abastecimiento que está asociada con una disminución en la cantidad de agua suministrada o con una falta total del suministro. Las interrupciones en el suministro de agua o, el suministro de agua con parámetros inadecuados (por ejemplo, a una presión demasiado baja) dan lugar a reclamos por parte de los usuarios finales (Pietrucha-Urbanik, K. 2015).

En virtud de lo anterior, muchas compañías dedicadas al abastecimiento de agua urbana toman acciones para reducir el número de fallas en la red, por ejemplo:

- Renovación técnica de tuberías (renovaciones, mantenimiento y diagnóstico),
- Sustitución de tuberías y accesorios,
- Mejora en la detección de lugares de fugas,
- Limitación de la presión de la red al valor más bajo permitido,
- Adecuada operación, diseño y ejecución del sistema de abastecimiento de agua.

El aumento de la confiabilidad del suministro de agua se ve afectado por la modernización y operación racional de los sistemas de suministro de agua, lo que se relaciona, entre otros, con la estabilización de la presión en la red y el servicio de emergencia eficiente. Sin duda, un aspecto importante que influye en la confiabilidad del suministro de agua potable es el uso de métodos activos de gestión de redes de suministro de agua que permiten reducir la duración de las interrupciones no planificadas en el suministro de agua Pietrucha-Urbanik, K. (2015).

La predicción de fallos mediante un adecuado monitoreo de la red incide notoriamente en la reducción de las interrupciones del suministro de agua y, por lo mismo, en los posibles efectos nocivos en la salud pública. Igualmente, el aspecto económico de la gestión de fallos, que determina directamente el tomar u omitir procedimientos que reduzcan el riesgo, se convierte en una importancia estratégica para la compañía.

En concordancia con lo anterior, la importancia de la predicción de fallos es significativa ya que envuelve temáticas operativas, económicas y de salud. Asimismo, gana importancia en la medida que las redes de abastecimiento de agua se vuelven

más complejas y aún más, cuando el análisis de fallos permite establecer las causas fundamentales de esos fallos en función de varios modelos de información útil (Cai et al. 2016).

4.3.1 La incertidumbre en la predicción de fallos.

Al realizar una revisión de los estudios centrados en los fallos en una red de abastecimiento de agua, se puede apreciar que la *incertidumbre* es un tema de preocupación en el análisis de fallas; la inexactitud en la medición y en el registro de conjuntos de datos de fallas, las variaciones en los juicios de los expertos debido a diferentes antecedentes de experiencia y la credibilidad relacionada, son las razones de tal incertidumbre (Kabir G. et al., 2015).

El estudio de Kabir G. et al. (2015) desarrolla un marco de predicción de fallas de tuberías principales de agua basado en la actualización del desempeño del modelo de riesgo proporcional bayesiano de Weibull (BWPHM). La aplicabilidad del marco propuesto se ilustra con la predicción de fallas de modelos de tuberías de hierro fundido y hierro dúctil de la red de distribución de agua de la ciudad de Calgary, Canadá. Con la actualización, los modelos bayesianos han mejorado efectivamente la predicción de fallas en la tubería principal de agua cada vez que se dispone de nuevos datos o nueva información (*ilustración 45*). El marco propuesto puede evaluar el desempeño del modelo a la luz de información incierta y en evolución.

Generalmente, el *análisis de supervivencia* ha sido el modelo estadístico más utilizado para fallos en tuberías de suministro de agua, pues permite ocuparse del deterioro y la falla a lo largo del tiempo e involucra el modelado del tiempo transcurrido entre un evento inicial y un evento terminal.

Este modelo estadístico incorpora el hecho de que mientras algunas tuberías se rompen, otras no y esta información tiene un fuerte impacto en el análisis de fallas de tuberías. Los modelos utilizan covariables (es decir, diámetro, longitud, resistividad

del suelo) para diferenciar las distribuciones de fallos de tuberías sin dividir los datos de fallas, lo que brinda una mejor comprensión de cómo las covariables influyen en la falla global de la tubería (Kabir G. et al., 2015).

Las *incertidumbres* se convierten en una parte integral de los modelos de predicción de fallos debido a la integración o fusión, de datos/información de diferentes fuentes, a la participación del juicio humano (experto) para la interpretación de datos y a las observaciones, la información parcial e incompleta.

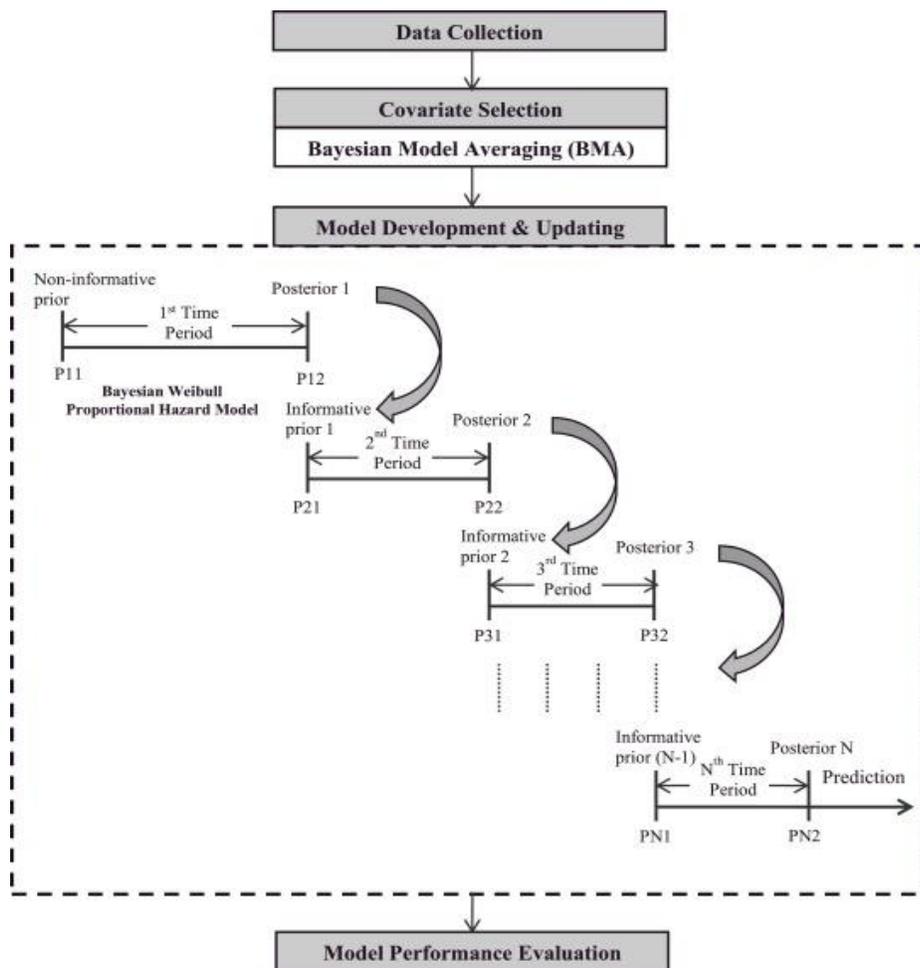


Ilustración 45. Ejemplo de actualización del modelo bayesiano Fuente: Kabir, G. (2015)

4.3.2 Tasa de fallos y frecuencia en la predicción de fallos.

Como ya se ha visto, buena parte de los estudios sobre fallos en redes de abastecimiento se enfocan en un componente de la red como lo pueden ser las tuberías o las válvulas; pero omiten el hecho que este tipo de redes conforman un sistema complejo con diferentes subsistemas y componentes, sin tener en cuenta todos los elementos del mismo simultáneamente (Aghapoor Khameneh, P., et al., 2020).

Bajo ese entendido, la idea es que los incidentes de fallos de los accesorios individuales y las conexiones domiciliarias de una red de abastecimiento se incluyen en el análisis matemático para que se refleje la condición real de la red (Aghapoor Khameneh et al., 2020). Lo anterior constituye otro agregado a la importancia en la predicción de fallos, pues involucra a todos los elementos que conforman la red de abastecimiento de agua.

Sin embargo, poder acceder a suficientes cantidades de datos en un sistema tan complejo es difícil y, por lo tanto, es una estimación precisa de las tasas de fallos. En términos técnicos, la *tasa de fallos* es un parámetro principal en el análisis de fallos de cualquier sistema complejo y las características mediante las cuales se puede calcular dicha tasa son las propiedades físicas de los componentes: el diámetro, la edad y el material. (Aghapoor Khameneh et al., 2020).

No siempre se puede tener la disponibilidad de esos datos de la totalidad de la red, Por lo que utilizar la frecuencia de ocurrencia de fallos se convierte en base para un análisis cuantitativo, siendo esto otro de los puntos que reiteran la importancia en la predicción de fallos.

4.4 ALGUNOS MÉTODOS FAVORABLES EN LA ESTIMACIÓN DE FALLOS

Existe una certeza en torno a la completa comprensión de los procesos de fallos en las tuberías subterráneas de transporte de agua, es una tarea compleja. Ante esa complejidad, se han desarrollado modelos estadísticos que permitan hacer una predicción de fallos próxima al estado real de la red. Algo que predomina, en mayor o menor medida, es que los métodos estadísticos de predicción se ven afectados por la cantidad y la calidad de los datos que se puedan recoger y, además, por las técnicas estadísticas adoptadas. Esto convierte el desarrollo de modelos estadísticos para fallos en un verdadero desafío, más cuando esos fallos emanan de múltiples factores en forma individual o conjunta (Kabir et al., 2015).

4.4.1 El Análisis de Supervivencia.

Este modelo estadístico es el más utilizado para *gestión de fallos*. En él se describe un conjunto de métodos estadísticos, en los que resalta el momento de ocurrencia de un evento específico. Este modelo permite estimar, por ejemplo, la esperanza de vida de los pacientes o el período de funcionamiento de dispositivos. El análisis de supervivencia se aplicó inicialmente en biología y medicina, pero su uso se ha extendido a la economía, la sociología y las ciencias sociales, así como en la ingeniería.

Diferentes investigadores aplicaron diferentes métodos de análisis de supervivencia como el estimador de Kaplan-Meier (Christodoulou, 2011), proceso de Poisson homogéneo o regresión de Poisson (Asnaashari, 2009), proceso de Poisson no homogéneo (NHPP) (Rogers, 2011), NHPP inflado a cero (Rajani et al., 2012), modelo exponencial/Weibull (Dridi et al., 2005), modelo exponencial multivariable

(Mailhot et al. 2000), modelo de riesgos proporcionales de Cox (Cox-PHM) y el modelo de riesgos proporcionales de Weibull (WPHM) (Kimutai et al., 2015).

Analizando un poco la función de riesgo proporcional, la probabilidad de que ocurra un evento no deseado determina la función de riesgo en la siguiente forma:

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P[t \leq T < t + \Delta t | X \geq t]}{\Delta t} \quad (5)$$

Como la función de riesgo no acepta valores negativos, se puede escribir en forma exponencial:

$$h(t, X) = h_0(t)e^{\beta X} \quad (6)$$

en donde $h_0(t)$ es el nivel de riesgo de referencia, β es un vector de coeficientes de regresión y X es el vector de variables explicativas. Esto permite establecer la función de supervivencia, dada por la siguiente forma:

$$\hat{S}(t, X) = \hat{S}_0(t) \exp(X\beta) \quad (7)$$

en la que $\hat{S}_0(t)$ es la función de supervivencia de referencia.

El criterio principal para evaluar el estado de las tuberías de agua es el índice de tasa de falla $-\lambda$. El estimador del índice de fallas por año para un tipo particular de tuberías de agua (conexiones principales, de distribución y de suministro de agua), se determinó a partir de la fórmula:

$$\lambda = \frac{k}{L\Delta t} \quad (8)$$

donde k es el número total de fallas en el tipo de red, L es la longitud del tipo de red dado [km] y Δt es la unidad de tiempo igual a un año (Pietrucha-Urbanik, K. 2015).

Como ya se indicó, el modelo de riesgos proporcionales ha servido para presentar un modelo de predicción de fallos relativamente más simple frente a los usados actualmente con mayor nivel de complejidad. El trabajo de Kimutai et al. (2015) enseña que el modelo de riesgos proporcionales de Cox (Cox-PHM) sirve para evaluar los impactos de diferentes covariables en las tasas de riesgo. El modelo asume que un vector de covariables $z = [z_0, z_1, z_2, \dots, z_q]$ actúa multiplicativamente sobre la función de riesgo de línea base $h_0(t)$ y es equivalente a la tasa de riesgo total cuando las covariables no tienen influencia en la falla de la tubería. La función de riesgo de línea base representa el proceso de envejecimiento, como el efecto de la corrosión interna y externa, que ocurre no solo en función del tiempo, sino también de otras variables estresantes como la resistividad del suelo, la corrosividad del suelo o el contenido de humedad.

Para terminar, del trabajo de Pietrucha-Urbanik, K. (2015) “*Failure Prediction in Water Supply System – Current Issues*” se tiene que un elemento importante que influye en la vida útil del sistema de suministro de agua: es monitorear las fallas y detectar estos elementos que causan las fallas con mayor frecuencia. Al seleccionar los elementos que causan la falla con mayor frecuencia, se debe recordar que la “vida útil” de algunos elementos puede ser mucho más corta o más larga de lo que se supone.

El tiempo de funcionamiento de la red de suministro de agua debe determinarse no solo sobre la base de los datos teóricos establecidos por los diseñadores, sino también sobre los datos operativos tomados del monitoreo de la red. Tal procedimiento está justificado porque no es común que la vida útil real de un elemento técnico se

superponga con un tiempo de trabajo teórico. También se debe tener en cuenta que hay casos en los que el tiempo real de operación del objeto es mucho más largo que el teórico. Para todo el sistema de suministro de agua, como el final de la vida útil debe considerarse un momento en que la tasa de fallas es tan alta que no es económico reparar las fallas (Pietrucha-Urbanik, 2015).

4.4.2 Redes Neuronales.

Como se ha indicado a lo largo de este trabajo, las condiciones actuales en el suministro de agua han entrado a un estado de operación delicada; esto en virtud que la creciente demanda de agua que se deriva del aumento de la población mundial y de la continua urbanización, está estresando las redes de abastecimiento de agua hasta sus límites. Una estimación un poco pesimista, indica que para el año 2025 cerca de 1800 millones de personas experimentarán escasez absoluta de agua y dos tercios del mundo vivirán en condiciones de escasez de agua. UN-Water, FAO (2007).

Ante este panorama poco alentador, surgen estudios nuevos como el de Aramane et al. (2019) en el que hace uso de las denominadas redes neuronales, ya que han demostrado ser un enfoque adecuado para la detección de fugas de agua, pues tienen el mejor y más extenso alcance del problema debido a su naturaleza computacional. Otro punto a favor del uso de las redes neuronales consiste en que no tienen fallas básicas, a diferencia de los métodos alternativos, como la detección de fugas mediante sensores acústicos que no pueden diferenciar entre picos de flujo y fugas (Aramane et al., 2019).

El uso de las redes neuronales se complementa con el uso de otras herramientas como pueden ser las de código abierto, por ejemplo, Neuroph Studio. Esta herramienta es un marco de red neuronal escrito en Java. Se puede utilizar para crear y entrenar redes neuronales en programas Java. Neuroph proporciona una biblioteca de clases Java, así como una herramienta GUI para crear rápidamente componentes de red neuronal Java.

El modelo que consiste en una red neuronal de perceptrón multicapa identifica fugas simultáneas en múltiples regiones con éxito (Aramane et al., 2019).

La aplicación de la inteligencia artificial (IA) para monitorear las redes de abastecimiento implica utilizar la red neuronal de base radial para el modelado.

Es un perceptrón multicapa, que considera el valor medio de los datos y la distancia entre los datos al establecer los valores de peso de la red neuronal. Los datos de entrenamiento se obtienen utilizando la herramienta de simulación hidráulica. (Salam et al., 2014).

Asimismo, la aplicación de los métodos SVM (máquinas de vectores de soporte) y

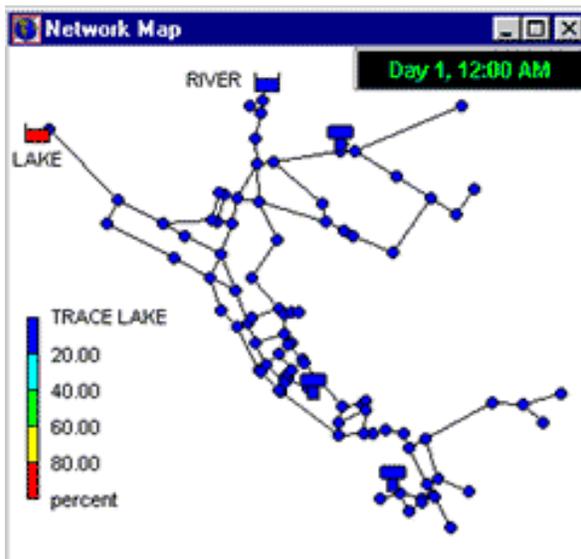


Ilustración 46. Maqueta EPANET de un sistema de agua.
Fuente: www.epa.gov

ELM (máquinas de aprendizaje extremo) parte del modelado con EPANET, que es una aplicación de software utilizada para modelar las redes de abastecimiento de agua potable. Esta herramienta permite realizar simulaciones de largo plazo del comportamiento hidráulico y de la calidad del agua dentro de las redes de tuberías presurizadas, que consisten en tuberías, nodos (uniones), bombas, válvulas y tanques de almacenamiento.

Se puede usar para rastrear el flujo de agua en cada tubería, la presión en cada nodo, la altura del agua en cada tanque, una concentración química, la edad del agua y el origen del rastreo en toda la red durante un período de simulación (ilustración 46) (Janke et al., 2021).

El método EML complementa la red neuronal de modo que la tubería de agua se trata como un conjunto de regiones que comienzan y terminan en sensores de presión. Estas lecturas del sensor de presión sirven como entrada al sistema. El sistema consta de

una red neuronal perceptrón multicapa recurrente. Esta red identifica regiones de la tubería con fugas cuando se alimenta con lecturas de presión de entrada (Aramane et al., 2019).

La aplicación de la técnica de Redes Neuronales Artificiales (ANN) tiene la ventaja de ser muy robusta y permite modelar sistemas lineales y no lineales complejos sin hacer suposiciones implícitas. Una red neuronal artificial está compuesta por una red de neuronas y sigue el ejemplo de sus contrapartes biológicas, de manera que las neuronas que son capaces de aprender pueden ser entrenadas para encontrar soluciones, reconocer patrones, clasificar datos e incluso pronosticar eventos futuros.

Las ANN han encontrado una amplia aplicación en la simulación de relaciones muy complejas y, como tales, encontraron una amplia aplicación en el modelado de problemas de gestión de recursos hídricos, incluida la detección de fugas, la optimización de la red de distribución de agua, el reemplazo y la rehabilitación de tuberías de agua, la previsión de la demanda de agua y el control de la presión (Bakker et al., 2003)

Las ANN se pueden entrenar para pronosticar la dinámica del flujo en una red de distribución de agua. Dicha dinámica de flujo se puede comparar con las demandas de agua en un área de medición de distrito en particular. La diferencia entre el consumo real de agua (consumo medido) y la demanda de agua simulada para un área con medidor de distrito representa la fuga de agua en la red de distribución de agua del área con medidor de distrito. Se descubrió que una ANN podría entrenarse y usarse para pronosticar el flujo con hasta un 99 % de confianza. Por lo tanto, la técnica ANN es un enfoque flexible y eficiente para la detección de fugas en las redes de distribución de agua. (Makaya et al., 2015).

Un ejemplo de esto se tiene con la red neuronal de percepción multicapa (MLP), que es una red en la que hay una capa de entrada que consta de nodos que simplemente aceptan los valores de entrada. Las salidas de las neuronas en una capa son entradas

para las neuronas en la siguiente capa. La última capa se llama capa de salida. Las capas entre las capas de entrada y salida se conocen como capas ocultas.

Con la excepción de las neuronas de entrada, hay dos pasos separados involucrados en la retransformación de entradas a salidas para las neuronas en una red (Parida et al, 2006).

En el primer paso cada neurona en una capa recibe una suma de activaciones ponderadas de todas las neuronas en la capa anterior de entradas a la red. Un término constante, generalmente denominado valor umbral de neurona (*bias* en inglés), se agrega a esta suma para producir la entrada neta, Y_{net} , a la neurona:

$$Y_{net} = \sum_{i=1}^N Y_i W_i + W_0 \quad (9)$$

donde N es el número total de neuronas en la capa anterior o matriz de entrada, Y_i es la entrada de neurona recibida de la i -ésima neurona en la capa anterior o matriz de entrada, W_i es el peso o la fuerza de la conexión de la neurona a una i -ésima neurona en la capa anterior o matriz de entrada, y W_0 es el valor de sesgo/umbral de la neurona.

Para una salida lineal, un término de sesgo es equivalente a una intersección en un modelo de regresión. Un término de sesgo se puede tratar como un peso de conexión de una unidad especial con un valor constante negativo. Dependiendo de la terminología, ya sea que se agreguen o resten sesgos o umbrales, el rendimiento de la red generalmente no se verá afectado según la función de transferencia (Makaya et al., 2015).

El segundo paso implica la transformación de la entrada neta, Y_{net} , en salida, Y_{out}

$$Y_{out} = f(Y_{net}) = f(\sum_{i=1}^N Y_i W_i + W_0) \quad (10)$$

Donde $f(\cdot)$ denota la función de transferencia de neurona seleccionada.

No existe una regla general que permita determinar el número de capas y neuronas ocultas en dichas capas y, por lo tanto, el diseño y las comparaciones de diferentes arquitecturas de red a partir de los mismos conjuntos de datos, parece la mejor opción para obtener una estructura de red óptima. La estrategia para seleccionar el número óptimo de capas ocultas es variando el número de neuronas, capas ocultas y funciones de entrenamiento. (Makaya et al., 2015).

Aunque existen avances significativos en el modelado del flujo de los sistemas de distribución de agua y los métodos de detección de fugas, se precisa aún mejorar todavía más la eficiencia y la confiabilidad de estos métodos en aras de modelar sistemas complejos con dinámicas de flujo complejas. Las ANN son por ahora, el enfoque más conveniente para dichos sistemas.

Generalmente, los métodos que se basan en el aprendizaje automático y que se utilizan en el diagnóstico de fallos, utilizan las mediciones del sistema como características para detectar y aislar fallas en cada componente. Sin embargo, en redes industriales como las redes eléctricas, no es posible detectar y aislar fallas simplemente monitoreando componentes individuales. Por ejemplo, el bajo voltaje en una estación eléctrica puede deberse a una falla en la estación o a una alta demanda de las estaciones vecinas. Por lo tanto, en muchos casos, la relación entre las mediciones de los componentes es mucho más informativa en comparación con las mediciones de los componentes individuales para la detección y el aislamiento de fallas (Khorasgani et al., 2019)

Para este tipo de redes complejas, los modelos del sistema no son fáciles de desarrollar y continúan actualizándose durante el ciclo de vida del sistema; por lo tanto, no siempre se dispone de modelos fiables de estos, a menudo incompletos y plagados de incertidumbres en el seguimiento del comportamiento del mismo.

Un enfoque alternativo consiste en aplicar una solución basada en datos que, además de las mediciones del sistema, también pueda usar la relación entre los componentes para el diagnóstico de fallas (Khorasgani, 2018).

A diferencia de las redes neuronales clásicas como el Multilayer Perceptron (MLP), las redes neuronales convolucionales (CNN) pueden extraer relaciones locales entre los datos de entrada mediante el uso de filtros convolucionales localizados. Estos filtros extraen características idénticas en diferentes ubicaciones espaciales dando lugar a avances en las tareas de clasificación, por ejemplo, de imágenes y videos (Karpathy et al. 2014) Análogamente, extraer relaciones locales entre variables industriales en diferentes componentes puede mejorar la detección y el aislamiento de fallas.

Sin embargo, las CNN funcionan muy bien para estructuras de datos regulares, y no pueden lograr el mismo rendimiento para dominios con estructura irregular, como las redes industriales. Ahora se han introducido las redes neuronales convolucionales gráficas (GCNN) para generalizar CNN a algunos dominios irregulares o más generalmente no euclidianos, como moléculas químicas y redes sociales (Khorasgani et al., 2019)

En el caso de una red de abastecimiento, para detectar y aislar un componente defectuoso, resulta insuficiente monitorear las mediciones en el componente. Un comportamiento inesperado de una medición de serie temporal podría deberse a una falla en el componente o una falla en un componente relacionado. Por lo tanto, usar solo las mediciones locales para la detección y el aislamiento de fallas puede disminuir las tasas de detección y aumentar las falsas alarmas (Khorasgani et al., 2019)

Muchos de los algoritmos de Machine Learning (ML) están desarrollados para operar en un espacio vectorial. En espacios vectoriales, podemos usar métricas de distancia como la distancia euclidiana para cuantificar la diferencia entre los puntos de datos. Los grafos representan una estructura de datos mucho más compleja.

Para sistemas con pocos nodos y conexiones limitadas entre vecinos de orden superior, usar grafos puede ser una solución sencilla. Sin embargo, esto puede conducir a la pérdida de información y a una complejidad computacional adicional. Un enfoque más eficiente es adaptar los algoritmos de ML al dominio gráfico. Las redes neuronales convolucionales (CNN) utilizan filtros convolucionales para extraer características de las imágenes. De manera similar, Graph CNNs (GCNN) usa Graph Fourier Transform (GFT) para extraer características de los gráficos. (Khorasgani et al., 2019).

GCNN aplica GFT para transferir las señales al dominio gráfico de Fourier utilizando la matriz U , que es la matriz considerada de vectores propios de L ; y ésta es la matriz laplaciana. Luego aplica un filtro en el dominio de Fourier del gráfico para extraer las características del gráfico. Un filtro localizado k -hop utiliza cualquier par de nodos con la distancia de ruta más corta menor que k para extraer las características. Finalmente, aplica Inverse-GFT para transferir características al dominio del tiempo. Las características extraídas en cada capa del gráfico se pueden presentar como

$$\text{Inverse_GFT}(g_\theta Z) = U g_\theta(\Lambda) U^T X, \quad (11)$$

donde el filtro del grafo $g_\theta(\Lambda) \in R^{N \times N}$ es una matriz diagonal con parámetros θ

$$g_\theta(\Lambda) = \begin{bmatrix} g_\theta(\lambda_1) & 0 & 0 & \dots & 0 \\ 0 & g_\theta(\lambda_2) & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & g_\theta(\lambda_N) \end{bmatrix}, \quad (12)$$

donde λ_i es el valor propio i -ésimo de L .

En la siguiente ilustración (47) se puede apreciar un modelo que representa una capa convolucional de grafos en GCNNU para un grafo con 4 nodos ($N = 4$). X_n representa señales gráficas en el nodo n ; g_l representa un filtro para ser aprendido y λ_i es el valor propio i -ésimo.

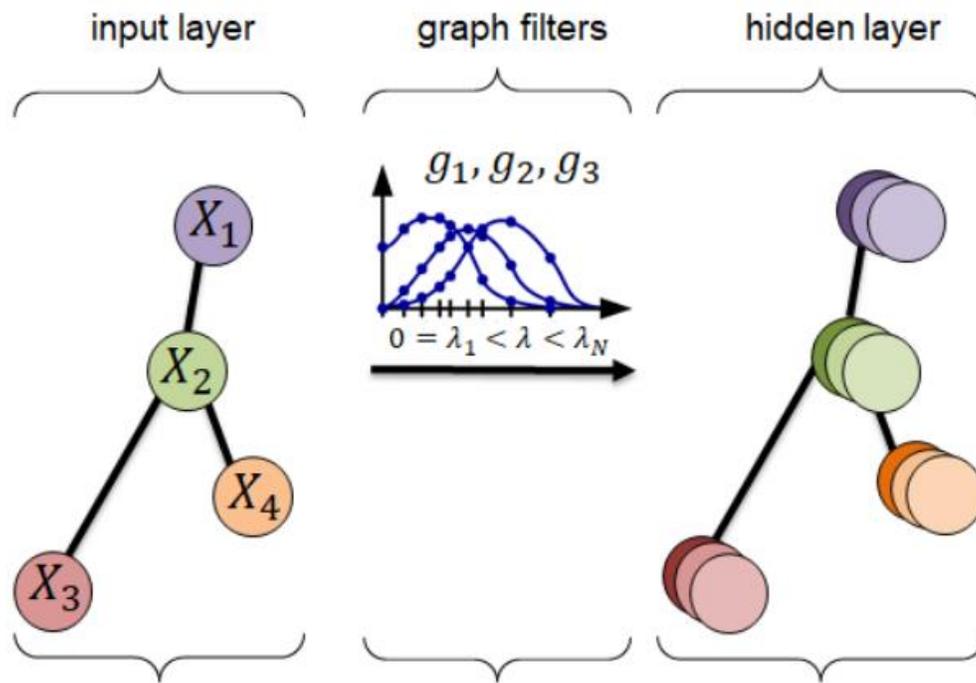


Ilustración 47. Una capa convolucional de gráfo para un gráfo con 4 nodos ($N=4$). Fuente: (Khorasgani et al., 2019)

La descomposición propia de la matriz es computacionalmente costosa. La aproximación polinomial de Chebyshev se puede utilizar para extraer características gráficas y aprender los parámetros del filtro sin descomposición propia (Defferrard et al., 2016)

Este tipo de solución con redes neuronales convolucionales gráficas, mejora significativamente la detección y el aislamiento de fallas al tener en cuenta la conexión entre los componentes sin usar el modelo físico del sistema. La conexión puede ser una conexión física, como tuberías, o puede representar correlaciones entre los componentes (Khorasgani et al., 2019).

4.5 FALLOS EN CASCADAS

Dada la complejidad de las redes de suministro de agua por los diversos factores que intervienen en ella, así como por su exposición a los frecuentes cambios tanto del medio físico, como del medio tecnológico; resulta necesario abordar el tema de los denominados *fallos en cascada*. De acuerdo con el trabajo de Bialek et al. (2016) titulado “Benchmarking and validation of cascading failure analysis tools”, un fallo en cascada se define como la pérdida sucesiva no controlada de elementos del sistema provocada por un evento inicial o disparador (*ilustración 49*). Normalmente, son eventos de baja probabilidad y alto impacto, lo que los convierte en una amenaza destacable para la estabilidad y seguridad de la red.

En la actualidad se viene desarrollando el concepto de las “Smart cities” cuya finalidad es lograr una eficiencia general mediante la integración de infraestructuras y servicios. Font et al. (2014) indica que las infraestructuras que se consideran críticas en las “Smart cities” son las energéticas, las de telecomunicaciones y las de distribución de agua entre otras. Se consideran críticas ya que al integrarlas permiten una oferta de servicios con calidad, oportunidad y eficiencia; pero a la vez, al aumentar su complejidad y su dependencia, se genera el riesgo de vulnerabilidad o de fallos en una de esas infraestructuras que pueda extenderse a otras y así, progresivamente, provocar un fallo en cascada.

Un ejemplo de ese tipo de *dependencia* lo constituye el aumento en el uso de las tecnologías de información en las infraestructuras (*ilustración 48*), pues generalmente necesitan de la información transmitida por otra infraestructura a través de la infraestructura de telecomunicaciones. Este tipo de dependencia va en aumento debido a la extensa utilización de los sistemas de control industrial para la Supervisión, Control y Adquisición de Datos (SCADA). (Font et al., 2014).

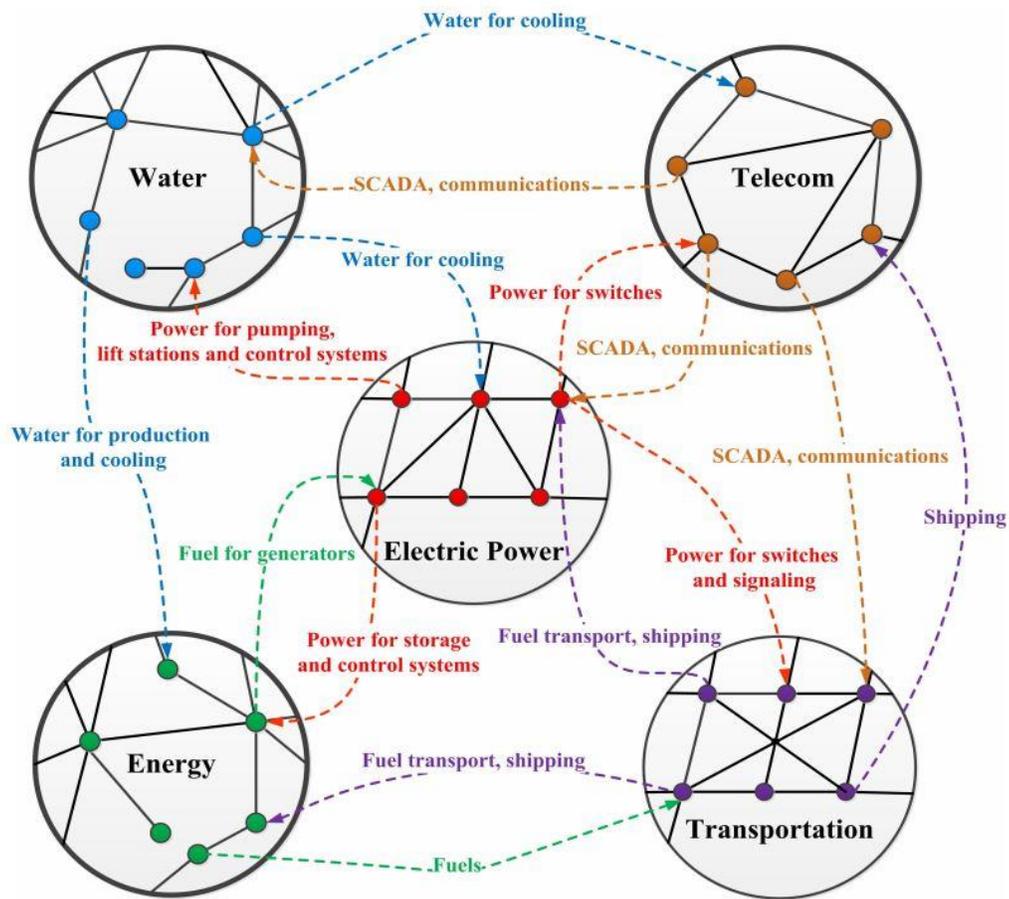


Ilustración 48. Relación de interdependencia entre diferentes infraestructuras. Fuente: Wikipedia. CC BY 4.0

Lo anterior se ve amplificado por el hecho, cada vez más común, de la ubicación de los centros de control que gestionan las infraestructuras; los cuales tienden a estar cada vez más alejados de las infraestructuras que controlan, con lo que la dependencia con las TIC se resalta aún más. Es acá donde surge un nuevo aspecto a considerar en las redes de distribución de agua y que agrupa un nuevo tipo de fallos, que sin tener en cuenta los problemas derivados de las numerosas interconexiones y dependencias inherentes, puede afectar la red: los problemas de seguridad informática clásicos, como lo son malware, rootkits, ataques de denegación de servicio distribuidos (DDoS), etc. (Font et al., 2014).

4.5.1 Los desafíos de la simulación de fallos en cascada.

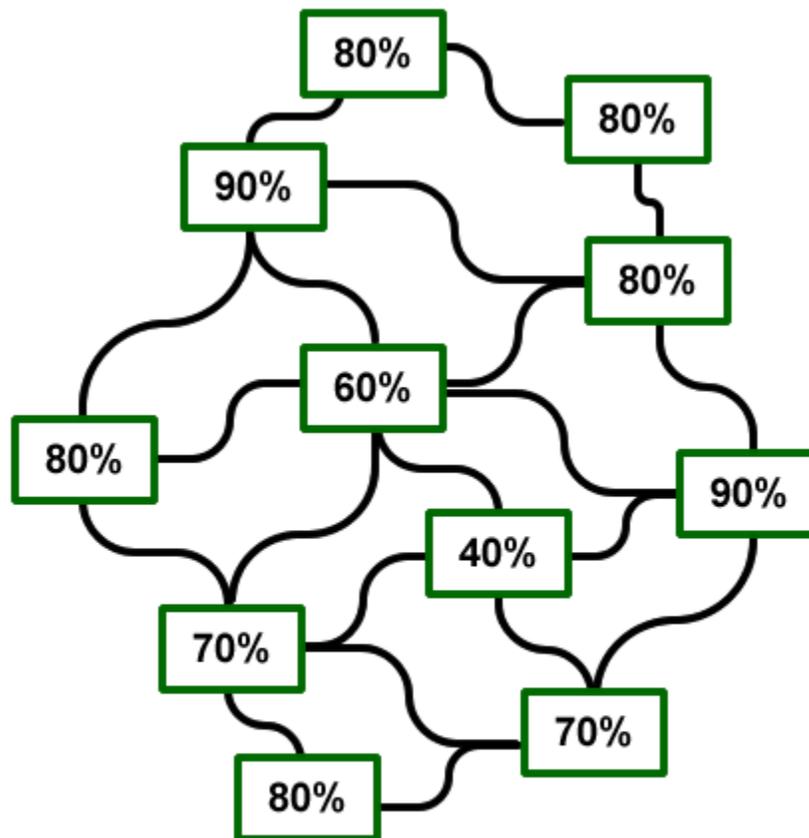
Un punto a revisar en torno a los fallos en cascada y su análisis, consiste en que la mayoría de los enfoques implican el uso de una herramienta de simulación de cortes en cascada. Desarrollar y, en últimas, validar estas herramientas es un desafío sustancial debido al conjunto numeroso y diverso de mecanismos por los cuales se propagan todas las cascadas reales. Para cada mecanismo adicional de cascada incluido en un modelo, es necesario hacer suposiciones sobre cómo reaccionará un sistema ante condiciones operativas extremas, raramente observadas (Bialek et al., 2016).

Observando los desarrollos hechos para los casos en redes de energía, los mecanismos potenciales que podrían modelarse incluyen una variedad de fenómenos tradicionales de inestabilidad y protección, incluidas sobrecargas en cascada interrumpidas por relés, colapso de voltaje, inestabilidad dinámica y oscilaciones entre áreas (Kundur et al., 2007)

Adicionalmente, las fallas en cascada también pueden involucrar una variedad de factores de comunicación, de control, económicos y sociales. Estos factores ocurren en diversas escalas de tiempo, lo que complica aún más el proceso de modelado. Y valga agregar, que los operadores humanos juegan un papel particularmente importante y aún más difícil de modelar. De una parte, el conocimiento inadecuado de la situación por parte de los operadores, puede ser un factor importante en varios sucesos de fallos; pero, de otra parte, las acciones del operador también pueden reducir el riesgo en un sistema; las acciones de mitigación apropiadas pueden detener la propagación de fallos en cascada. Modelar las acciones del operador es un desafío sustancial (Bialek et al., 2016).

Otros factores clave que dificultan el modelado lo constituyen las incertidumbres del estado del sistema y la naturaleza estocástica tanto de los eventos desencadenantes (*exógenos*) que conducen al inicio de un evento de fallo (día, hora, clima, etc.) como

de los eventos interdependientes (*endógenos*) que implica control, dinámica y protección a través de los cuales se propaga el fallo. Aunque hay avances actualmente en el modelado de algunos de los mecanismos potenciales (tanto de activación como de propagación, deterministas y estocásticos), se desconoce en gran medida la importancia relativa de estos (Bialek et al., 2016).



Network running normally

Ilustración 49. Ejemplo de fallos en cascada en una red. Fuente: Wikipedia CC BY-SA 3.0

4.5.2 El Análisis del árbol de fallos.

En el análisis de *árbol de fallos* convencional, las probabilidades de falla de los componentes de un sistema se tratan como valores exactos al estimar la probabilidad

de falla del evento principal. Para muchos sistemas, a menudo es difícil evaluar las probabilidades de falla de los componentes de ocurrencias pasadas porque los entornos de los sistemas cambian. Además, podría ser necesario considerar la posible falla de los componentes, incluso si nunca antes han fallado (Tanaka et al., 1983).

El análisis de árbol de fallos ha demostrado ser una herramienta analítica útil para el análisis de confiabilidad y seguridad de sistemas complejos (*ilustración 50*) (Barlow et al., 1973). Esta herramienta parte de la previa selección de un "suceso no deseado o evento que se pretende evitar", sea éste un accidente de gran magnitud (explosión, fuga, derrame, etc.) o sea un suceso de menor importancia (fallo de un sistema de cierre, etc.) para averiguar en ambos casos los orígenes de los mismos (Piqué et al., 1992).

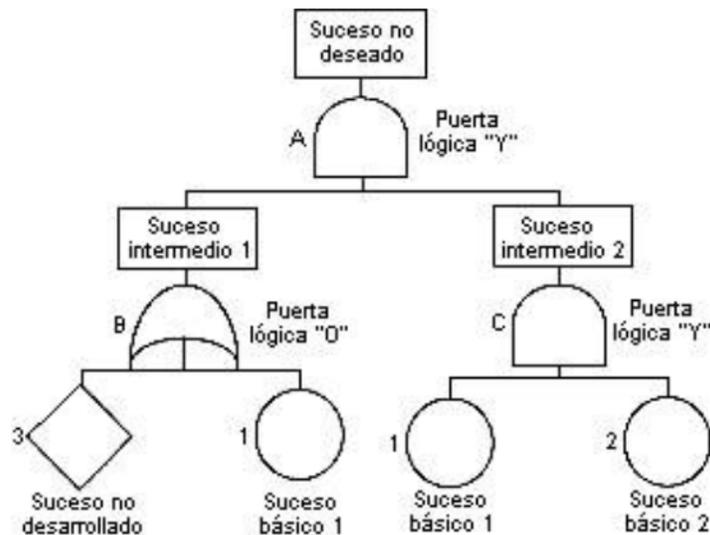


Ilustración 50. Representación gráfica de un árbol de fallos. Fuente: NTP: 333 Análisis probabilístico de riesgos

Para estudiar los fallos en cascada se han planteado algunos modelos como OPA (Dobson et al., 2001), CASCADE (Dobson et al., 2003) y TRELSS (Hardiman et al., 2003), entre otros. Estas metodologías diseñan un estado cuasi estacionario (QSS) para simular el comportamiento del sistema durante las contingencias y evaluar así el

riesgo de las fallas en cascada. En la siguiente ilustración (51) se resume el enfoque desarrollado por estos modelos (Rivera Muñoz, 2021).

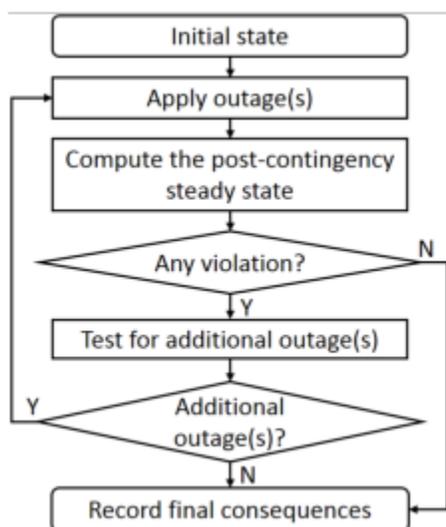


Ilustración 51. Metodología típica en fallos en cascada. (Henneaux, P. et al 2018)

Si bien el método de *árbol de fallas* permite cuantificar el número de fallos usando una aproximación “*top-down*” para determinar las condiciones de fallo del sistema, el método de *análisis del árbol de fallas* usa una aproximación “*bottom-up*” para identificar los modos de fallo y los efectos en la seguridad del sistema (Barbosa Cubillos, 2008).

4.5.3 Modelo de IO de inoperabilidad dinámica: efectos de cascada.

Como ya se ha indicado en este trabajo, los fallos en una red de distribución de agua tienen su causa en diferentes factores externos e internos. Acorde con la tendencia global de lucha contra el cambio climático, el uso eficiente del agua potable se convierte entonces en la mejor estrategia de adaptación a ese cambio. La disponibilidad de agua puede estar sujeta a interrupciones a corto plazo debido a

periodos de sequía, mientras que las interrupciones a largo plazo, que ocurren durante la vida útil de las plantas industriales, pueden causar riesgos operativos futuros (Zhang et al., 2022).

Dada la actual interdependencia de infraestructuras, el aumento de los vínculos entre éstas puede hacer que todo un sistema sea vulnerable a fallas en cascada en caso de pérdida de flujo de agua en algunas infraestructuras. El posible impacto indirecto de la escasez de agua en tales sistemas integrados puede no ser evidente sin el uso de modelos apropiados. El modelo de entrada-salida (IOM, también conocido como modelo de Leontiev) fue pensado principalmente para hacer estudios en economía, pero su uso se ha extendido a otras áreas como la energía, el suministro de agua y otros. La idea consiste en aplicar IOM para apreciar la propagación de inoperabilidad dentro de una red interdependiente y evaluar su vulnerabilidad (Zhang et al., 2022).

El IOM considera una red que consta de muchas subredes y se utiliza para analizar la interdependencia de las infraestructuras. Suponiendo que una red tenga n sectores, IOM se puede representar mediante el siguiente sistema de ecuaciones lineales:

$$x = Ax + c \quad (13)$$

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1j} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ij} \end{bmatrix} \quad (14)$$

$$a_{ij} = \frac{x_{ij}}{x_j} \quad (15)$$

donde:

x es un vector $n \times 1$, que denota la producción total de los n sectores;

c es un vector $n \times 1$, que denota la demanda externa de los n sectores;

A es una matriz $n \times n$, que denota las interdependencias entre los sectores;

x_j es la producción total del j -ésimo sector;

x_{ij} es la cantidad de entrada del i -ésimo sector al j -ésimo sector.

Un sistema interdependiente o parte de él, puede verse afectado por un desastre repentino o por una falla interna, lo cual puede causar un efecto de propagación del fallo a nivel global y es ahí donde la inoperabilidad se utiliza para analizar y cuantificar ese efecto. La clasificación de los sectores afectados según su vulnerabilidad a los fallos originados en un sector primario, sirve de elemento fundamental para la gestión del riesgo. Este análisis permite comparar la aplicabilidad del modelo de entrada-salida de inoperabilidad (IIM) que es promovido por la demanda contra el que es promovido por la oferta. El modelo de reducción de la demanda presentado por Santos y Haimes (2004) es el siguiente:

$$\delta x = A\delta x + \delta c \leftrightarrow (\hat{x} - \tilde{x}) = A(\hat{x} - \tilde{x}) + (\hat{c} - \tilde{c}) \quad (16)$$

donde, δx es la diferencia entre el total normal \hat{x} y el degradado \tilde{x} ;

salidas; δc es la diferencia entre la demanda final planificada \hat{c} y la degradada \tilde{c} .

Para cualquier sector i del sistema, la expresión de inoperatividad es la siguiente:

$$q_i = \frac{\hat{x}_i - \tilde{x}_i}{\hat{x}_i}, \forall i \quad (17)$$

donde, q_i denota la inoperancia del sector i . Al cuantificar el riesgo, es una métrica adimensional que va desde 0 (estado normal de funcionamiento del sistema o sector) hasta 1 (estado de falla completa del sistema o sector) (Zhang et al., 2022).

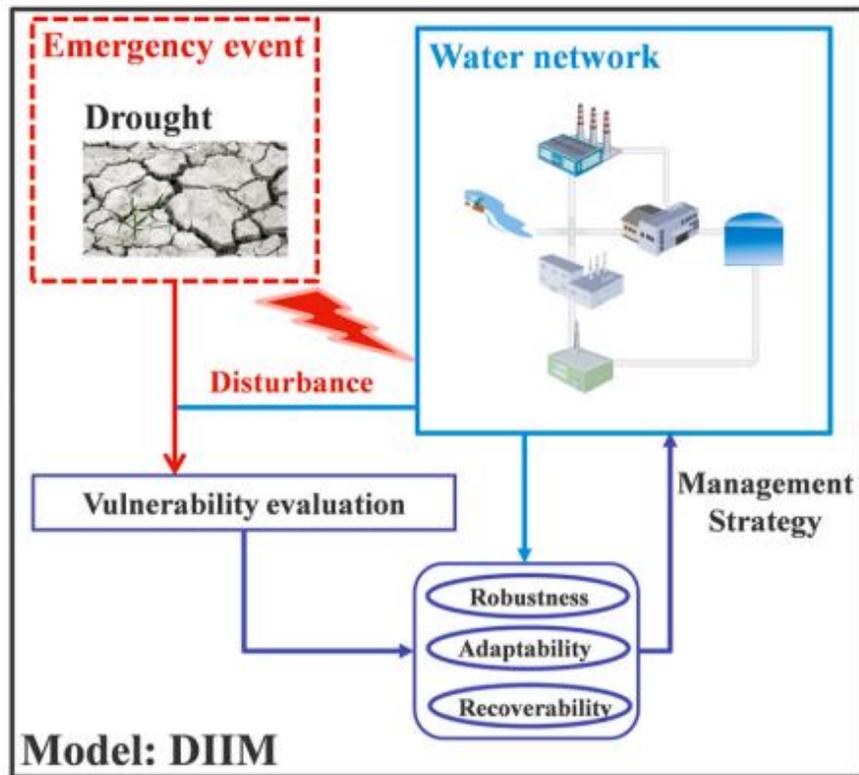


Ilustración 52. Diagrama del modelo dinámico de entrada-salida. Zhang, Y. et al., (2022).

La evaluación de la fragilidad del sistema integra los indicadores de vulnerabilidad en el modelo dinámico de entrada-salida (DIIM) (ilustración 52), lo cual permite analizar la propagación de fallas en las redes de agua mediante la simulación de los efectos en cascada de los fallos. Para Zhang, Y. et al., (2022) la vulnerabilidad se define como el estado de respuesta de un sistema o sector, a un fallo. Incluye la capacidad de resistir la perturbación, la máxima inoperabilidad que se puede lograr y la capacidad de recuperarse.

Cuanto mayor sea el coeficiente de resiliencia, más rápida será la recuperación, pero aquí el coeficiente de resiliencia solo considera el efecto del sector en sí; es decir, suponiendo que el componente del sector i no cumple la función esperada, se examina el coeficiente de resiliencia del sector mediante la reparación (fallo de la máquina) para volver al nivel deseado (Zhang et al., 2022).

4.6 SISTEMA MULTIAGENTE PARA LA SIMULACIÓN DE EPIDEMIAS

Apreciando la reseña histórica en este trabajo, puede decirse que la humanidad ha sufrido enfermedades infecciosas a lo largo de toda su historia, aún a día de hoy. Las últimas que se han evidenciado son las epidemias de SARS, MERS, AH1N1 y el SARS Cov-2, que han obligado a los gobiernos a tomar medidas restrictivas mientras se encuentran soluciones para mitigar la afectación; pero no resulta fácil conocer el impacto que esas acciones frente a un “sistema complejo y dinámico como lo es la propagación de una enfermedad en una población” (Vázquez, 2022).

Anteriormente, los modelos epidemiológicos consideraban a la población como homogénea, de modo que todos los individuos podían contagiar a cualquier otro (Bailey, 1975). En la medida que pasa el tiempo, se aprecia la necesidad de incorporar diversas características a estos modelos para permitirles ser más acertados, como la estructura de contactos que existe en la población, características específicas de los individuos como su estado socio-económico, edad, sexo, entre otros. Una forma de lograr esto fue modelar a la población mediante el uso de redes (Vázquez, 2022).

Usar modelos que involucren mucho detalle en el comportamiento de los individuos y del agente infeccioso, permite lograr un mayor poder de predicción; pero debido a la complejidad, también se requiere de un mayor uso computacional. Un modelo determinista no puede capturar las incertidumbres asociadas al mundo real como cuando un individuo infectado contagia a otro siempre con alguna probabilidad, y no con certeza absoluta (Vázquez, 2022).

Lo anterior lleva a adoptar modelos estocásticos que busquen obtener unos resultados más reales; es acá donde las técnicas de simulación multiagente permiten analizar y cuantificar los efectos de las distintas medidas de contención ante la propagación de

enfermedades, para que la comprensión de estas diferencias conlleve a una toma de decisiones facilitada y basada en evidencia científica (Vázquez, 2022).

Representar las conexiones e interacciones entre nuestros agentes por medio de una red, permite el uso de técnicas matemáticas bien estudiadas de teoría de grafos y sistemas de propagación en redes. De modo que al combinar la estructura de contacto poblacional usando redes con las características que se le asignan a un agente, se busca modelar de una manera eficiente, y de cierta forma más realista, la propagación de una epidemia (Vázquez, 2022).

4.6.1 Concepto de sistemas multiagente y de redes complejas.

El concepto de sistemas multiagente es el siguiente: un conjunto de entidades heterogéneas, llamadas agentes, que interactúan entre sí acorde a un sistema de reglas. Las interacciones permiten apreciar los fenómenos a nivel del sistema. Según Mei S et al. (2010) los agentes tienen tres características principales:

- Autonomía: el agente actúa conforme a su base de creencias.
- Visión local: ningún agente tiene una visión global del sistema o la complejidad del mismo es más grande que el entendimiento del agente.
- Descentralización: No hay control del agente.

Asimismo, los modelos basados en agentes se pueden definir acorde con ciertas características principales (Macal C., 2016):

- Individual. Aquel en el que los agentes en el modelo son representados individualmente y tiene diversas características.
- Autónomo. Aquel en el que los agentes individuales tienen comportamientos internos que les permiten ser autónomos, con capacidad para sentir cualquier condición que ocurre en el modelo en cualquier tiempo y actuar apropiadamente acorde con su comportamiento.

- Interactivo. Aquel en el que los agentes autónomos interactúan con otros agentes y con el entorno.
- Adaptativo. Aquel en el cual los agentes autónomos que interactúan cambian sus comportamientos durante la simulación, conforme los agentes aprenden, se encuentran con situaciones nuevas, o como población ajustan su composición para incluir grandes proporciones de agentes sean los que se adaptan satisfactoriamente.

El concepto de redes complejas puede ser descrito como un grafo compuesto de un conjunto de vértices y un conjunto de aristas. Existen diversos tipos como, libre de escala, mundo pequeño y redes aleatorias (*ilustración 53*).

- Libre de escala. Es un tipo de red con una distribución de grados que sigue una ley de potencia de la forma $P(k) \sim k^{-\alpha}$, donde k es el grado, $P(k)$ es la probabilidad de que un nodo tenga grado k , y α es el exponente de la ley de potencia (Vázquez, 2022).
- Mundo pequeño. Son redes en donde la mayor parte de pares de vértices están conectados a través de un camino corto (Vázquez, 2022).
- Redes aleatorias. Son la forma más simple de un grafo complejo, en el cuál, cada par de nodos están conectados aleatoriamente con una probabilidad independiente p . La distribución de los grados de este tipo de grafos sigue una distribución de Poisson (Vázquez, 2022).

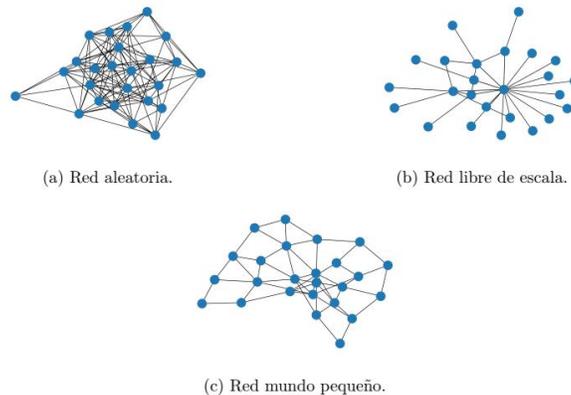


Ilustración 53. Ejemplos de redes complejas. Vázquez E. (2022)

4.6.2 Redes complejas multiagente.

El concepto de una red compleja multiagente es el mismo de la red compleja, solo que ahora, los vértices se reemplazan por agentes. De igual forma, las aristas representan los contactos, interacciones o relaciones entre los agentes y estas pueden modificarse con el tiempo. Esta funcionalidad de modificarse en el tiempo hace que la red compleja sea dinámica, por lo que los nodos agentes cambian entre estados y las aristas varían entre si existen o no, o en el peso que se les pueda asignar (Vázquez, 2022).

Este tipo de red compleja multiagente se define así:

$$G(t) = (V(t), E(t)) \quad (18)$$

donde si $u \in V(t)$, se le llama nodo agente. Cada arista $a \in E(t)$ conecta a dos nodos agentes formada por la interacción de estos dos (Vázquez, 2022).

4.6.3 Modelos epidemiológicos.

En la simulación de epidemias se trabaja con modelos epidemiológicos de compartimentos, a tiempo continuo, modelando la población subyacente con un grafo $G = (V, E)$. En estos se tiene una cantidad finita de compartimentos C_1, \dots, C_q , y en el tiempo t , cada nodo $u \in V$ es asignado a un C_i para un solo i (Khashei & Bijari, 2010). A continuación, se mencionan algunos de los modelos más conocidos (Newman, 2018).

4.6.3.1. El modelo SI.

Es el modelo más simple ya que se consideran solamente dos estados S (susceptible) y I (infectado); en el caso de que un individuo sea susceptible significa que es alguien que aún no tiene la enfermedad, pero podría tenerla si entra en contacto con otro que sí la tenga. Un individuo infectado es aquel que tiene la enfermedad y que puede transmitirla si entra en contacto con algún individuo que sea susceptible (*ilustración 54*).

4.6.3.2. El modelo SIS.

A diferencia del modelo anterior, en modelo SIS se puede tener una reinfección de la enfermedad, es decir el individuo infeccioso pasa de nuevo a ser susceptible después de un tiempo, no adquiere una inmunidad a la infección por haberla tenido (*ilustración 54*).

4.6.3.3. El modelo SIR.

En este modelo se trabaja con otro estado R, al que se le denomina recuperado, el cual representa que un individuo infeccioso adquiere inmunidad a la enfermedad después haberse recuperado (*ilustración 54*).

4.6.3.4. El modelo SIRS.

Existen modelos más complejos, en los cuales se trabaja con inmunidad temporal. Un individuo infeccioso pasa a recuperado después de un tiempo determinado y está en dicho estado solo un tiempo para pasar nuevamente a ser susceptible y poder infectarse nuevamente de la enfermedad (*ilustración 54*).

4.6.3.5. El modelo SEIR.

Para este modelo se trabaja con un estado extra que se representa con E y significa que un individuo susceptible estuvo expuesto a la enfermedad y ésta aún no se ha desarrollado al punto de que pueda transmitirla a otros individuos (*ilustración 54*).

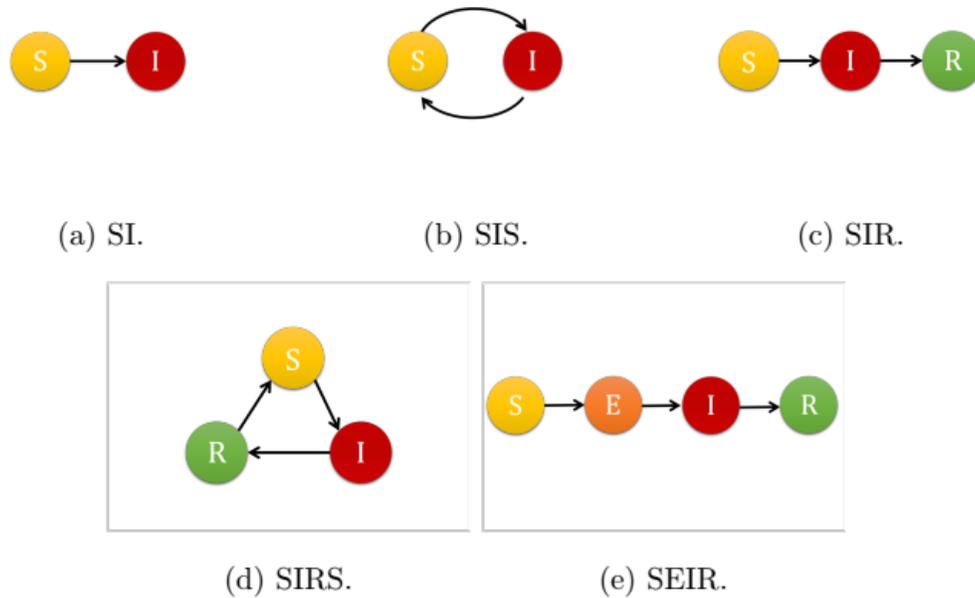


Ilustración 54. Modelos compartimentales. Vázquez E. (2022)

4.6.4 Ejemplo de simulación con sistema multiagente.

El siguiente ejemplo, tomado de *“Integration of small world networks with multi-agent systems for simulating epidemic spatiotemporal transmission”* por Liu, T., Li, X., & Liu, X. (2010); propone un modelo integrado basado en una red mundial pequeña (SWN) y un sistema multiagente (MAS) para simular la transmisión espaciotemporal epidémica. En este modelo, MAS representa el proceso de interacciones espaciotemporales entre individuos, y SWN describe la red de relaciones sociales entre agentes.

El modelo se compone de definiciones de atributos de agentes, reglas de movimiento de agentes, vecindades, construcción de redes de relaciones sociales entre agentes y reglas de transición de estado. Los efectos de descomposición de la "memoria" de la infección, la distancia y la relación social entre los agentes se introducen en el modelo, que no están disponibles en los modelos tradicionales (Liu et al., 2010)

El modelo propuesto se usa para simular el proceso de transmisión de la gripe en la ciudad de Guangzhou basado en la plataforma de software Swarm. El modelo de integración tiene un mejor rendimiento que el modelo SEIR tradicional y el modelo epidémico basado en MAS puro. Este modelo se ha aplicado a la simulación de la transmisión de epidemias en un entorno geográfico real. La simulación puede proporcionar información útil para la comprensión, predicción y control de la transmisión de epidemias (Liu et al., 2010).

En la actualidad, los métodos de modelado de epidemias incluyen principalmente ecuaciones dinámicas de sistemas, modelos de autómatas celulares, sistemas multiagente y redes complejas. El modelo SIR (S-susceptible, I-infectious, R-removed) basado en ecuaciones diferenciales fue formulado originalmente por Kermack al. (1927); a partir de entonces, se propusieron algunos derivados, como el modelo SEIR (E-exposed) y el modelo SIRS.

Sobre la base de la simulación del proceso estocástico de infección, inicio y propagación de cada paciente con SARS, Shi Y. (2003) construyó un modelo dinámico de sistema de propagación del SARS. Wang H. et al., (2003) presentó un modelo dinámico mejorado para evaluar el potencial epidémico del SARS y la eficacia de las medidas de control. Sin embargo, las ecuaciones dinámicas del sistema epidémico asumen que las poblaciones se mezclan por completo y desprecian los efectos espaciales de propagación de las epidemias (Liu et al., 2010).

Componente: *el modelo MAS*. La influenza (gripe) es una infección viral altamente contagiosa de las vías respiratorias. La gripe se propaga a través de gotitas en el aire

infectadas con virus que se tosen o estornudan en el aire. Casi todas las personas son susceptibles al virus de la influenza. La persona que tiene gripe infectada obtendrá inmunidad después de la recuperación y mantendrá una duración corta, alrededor de 8 a 12 meses. Es importante tratar la gripe con seriedad porque puede provocar neumonía y otras complicaciones potencialmente mortales, especialmente en bebés, ancianos y personas con problemas de salud a largo plazo, en otoño e invierno debido a la alta incidencia de gripe (Liu et al., 2010).

El ciclo de infección de la gripe se puede dividir en cuatro etapas: período susceptible, período expuesto, período infeccioso y período de eliminación. Las poblaciones se pueden dividir correspondientemente en cuatro categorías: susceptibles (S), expuestas (E), infecciosas (I) y eliminadas (R). $S(t)$, $E(t)$, $I(t)$ y $R(t)$ son la cantidad de S, E, I, R en el paso de tiempo t respectivamente. La población es igual a la suma de $S(t)$, $E(t)$, $I(t)$ y $R(t)$ y puede considerarse aproximadamente como una constante (Liu et al., 2010).

El sistema dinámico epidémico SEIR se puede expresar como

$$\frac{dS}{dt} = -\beta \frac{S(t)I(t)}{N}, \quad (19)$$

$$\frac{dE}{dt} = \beta \frac{S(t)I(t)}{N} - \sigma E(t), \quad (20)$$

$$\frac{dI}{dt} = \sigma E(t) - \eta I(t), \quad (21)$$

$$\frac{dR}{dt} = \eta I(t), \quad (22)$$

$$S(t) + E(t) + I(t) + R(t) = N, \quad (23)$$

donde β denota la tasa de infección, es decir, la tasa de contacto adecuada, σ define la tasa de exposición que se vuelve infecciosa, η define la tasa de infección que se elimina. Por lo tanto, la duración promedio de la exposición viene dada por $1/\sigma$ y la duración promedio de la infectividad viene dada por $1/\eta$ (Liu et al., 2010).

Los atributos de un agente incluyen principalmente días expuestos, días infectivos, dosis de infección, grupo de edad, residencia, etc. Los días expuestos indican la duración que un agente E ha estado en el período expuesto. Los días infectivos indican la duración que un agente I ha estado en el período infectivo. Los modelos de umbral epidémico indican que los agentes S no se infectan inmediatamente cuando entran en contacto con los agentes I, pero después de recibir una cierta cantidad de "toxinas", es decir, dosis de infección, se infectarán. Por lo tanto, hasta cierto punto, el valor del atributo de dosis de infección es crucial para decidir si los agentes S se infectarán o no (Liu et al., 2010).

El modelo SWN refleja una característica importante de las relaciones sociales, es decir, la mayoría de los amigos viven en la misma calle o sirven a la misma empresa. La longitud de ruta promedio se define como la distancia promedio de cualquier par de nodos. El coeficiente de agrupamiento de la red es igual al coeficiente de agrupamiento promedio de todos los nodos. Dos características críticas de la red de mundo pequeño son la longitud de ruta promedio más corta y el coeficiente de agrupamiento más alto. El grado del nodo es el número de aristas que se conectan a él (Liu et al., 2010).

Ahora bien, el modelo MAS puede usarse para describir la estructura de SWN, mientras que los nodos en SWN pueden considerarse como los agentes en MAS. Los enlaces entre nodos se pueden utilizar para representar la relación social de los agentes. Por lo tanto, se puede construir un nuevo modelo basado en SWN y MAS. En el modelo propuesto, la relación social puede ser considerada en la simulación de forma explícita. El método de aleatorización agregando borde para construir SWN ha sido ampliamente adoptado. Las relaciones sociales del agente se pueden clasificar en

diferentes tipos, como familia, compañero de escuela, colega, vecino y amigo (Liu et al., 2010).

Componente: reglas de transición de estado del agente. Los estados de los agentes pueden ser modificables. Si los agentes S contactan a los agentes I con frecuencia, estarán sujetos a infección. Después de ser infectados, entran en período expuesto sin infectividad. Tan pronto como entran en el período infeccioso, tienen la capacidad de infectar a otros agentes S. Tendrán la inmunidad después de recuperarse de la epidemia y no podrán volver a contagiarse (Liu et al., 2010).

La regla de transición de estado es el núcleo del modelo. El parámetro de probabilidad de infección generalmente se aplica para decidir si la infección ocurrirá o no en la mayoría de las investigaciones, mientras que el parámetro permanece sin cambios en todo el proceso de simulación. Sin embargo, se deben tener en cuenta las influencias temporales y espaciales. Si contacta con el agente I recientemente, el agente S se infectará con alta probabilidad. Por otro lado, cuanto más corta es la distancia entre el agente S y el agente I, mayor es la probabilidad de infección. Si los agentes S se infectarán y cuándo se debe a combinaciones de factores temporales y espaciales (Liu et al., 2010).

Cálculo de la dosis total de infección que recibe el agente S de los agentes I. La dosis de infección $d_{ik}(t')$ que el agente S recibe del agente I en el paso de tiempo t' se puede expresar como:

$$d_{ik}(t') = \begin{cases} d, & w = 1 \\ d/R_s & w = 0 \end{cases} \quad \alpha \leq p \quad (24)$$

donde d es la dosis de infección, R_s representa la distancia entre el agente S y el agente I, α es una variable aleatoria, p denota la probabilidad de recibir dosis de infección de agentes S cuando contraen agentes I.

En el paso de tiempo t' , las dosis totales $D_i(t')$ del agente S son iguales a la suma de las dosis que este agente S recibió del agente I con el que entra en contacto. Las dosis totales del agente S se pueden representar de la siguiente manera:

$$D_i(t') = \sum_{k=1}^K d_{ik}(t') \quad (25)$$

donde K es el número de I agentes que contactan a este S agente en el paso de tiempo t' .

Cálculo de la dosis acumulada respectiva que recibió cada agente S. Dodds, P. et al (2004) exploró un modelo generalizado para introducir la memoria de las exposiciones posteriores a una influenza contagiosa. En este ejemplo, se pueden ver los efectos de descomposición de la memoria de infección. La dosis acumulada se puede dar como la siguiente ecuación:

$$D_i(t) = \sum_{t'=t-T+1}^t \gamma^{t-t'} D_i(t'). \quad (26)$$

Cada agente mantiene una memoria de las dosis recibidas de I agentes durante los T pasos de tiempo anteriores. El parámetro γ se introduce como tasa de disminución de la dosis de infección. Cuando $\gamma = 1$, la disminución de la dosis de infección no se considerará en la memoria del período T .

Comparación de dosis acumuladas de agente S con el umbral de dosis. Si la dosis acumulada que recibió el agente S supera un umbral de dosis dado, el agente se infectará; de lo contrario, no.

$$D_i(t) = \begin{cases} \geq d^*, \text{ infected,} \\ < d^*, \text{ not infected} \end{cases} \quad (27)$$

donde d^* es el umbral de dosis, que decide si el agente S será infectado o no.

La estructura del modelo de transmisión espaciotemporal epidémica basado en SWN y MAS se puede mostrar en la siguiente ilustración (55).

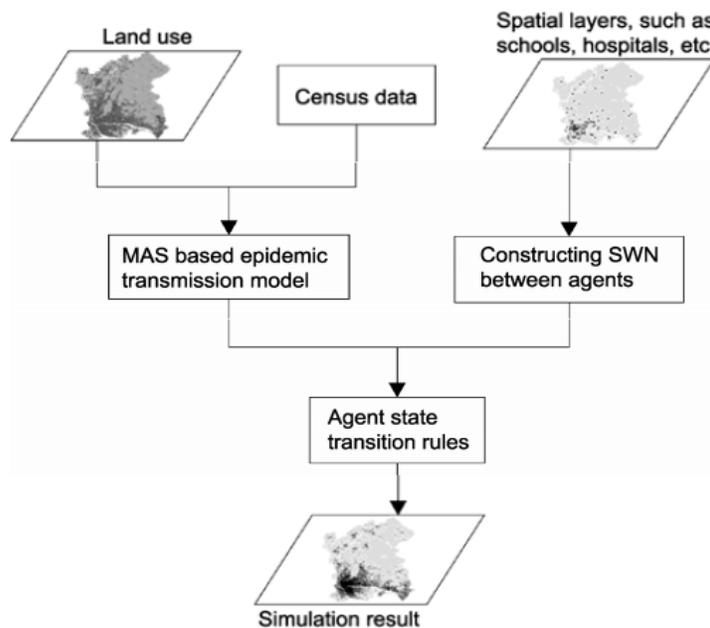


Ilustración 55. Transmisión espaciotemporal epidémica basada en modelos SWN y MAS. Liu, T. et al., (2010)

4.7 APLICACIÓN DE LOS SISTEMAS MULTIAGENTE PARA LA ESTIMACIÓN DE FALLOS EN CASCADA

Uno de los escenarios más críticos para los sistemas de distribución de agua (WDS) ocurre cuando la disfunción de un solo activo impacta en el desempeño de otros activos como consecuencia de condiciones físicas e hidráulicas alteradas. El modelado de propagación de fallas es de gran importancia para la operación y gestión

de infraestructuras de agua urbana. Por ejemplo, un evento de explosión de una tubería en cascada puede interrumpir el suministro regular en un área de la red de agua e incluso colapsar todo el sistema. Esta dinámica de fallas comparte características clave con la forma en que los contagios se propagan a través de una población de individuos.

La hipótesis principal se basa en que el estallido de una sola tubería puede propagarse de forma doble: induciendo el cierre de las tuberías vecinas (y elevando los niveles de riesgo de otras), y provocando variaciones bruscas en los niveles de presión nodal. Estas consecuencias podrían tener un efecto en cascada por la ocurrencia de más roturas de tuberías y así continuar el proceso. Los modelos de epidemiología compartimental específicamente diseñados son capaces de replicar escenarios de propagación de fallas y ayudar a descubrir tuberías y áreas de WDS de alto riesgo de interrupción del suministro.

Como se ha apreciado hasta el momento, la propagación de fallos sobre los activos de un sistema de distribución de agua (WDS) es uno de los principales temas de investigación en la actualidad. Resolver esto es una tarea difícil ya que la gran cantidad de combinaciones de ocurrencias de fallas múltiples y la dimensión natural e interconectividad de un WDS hacen de este un problema de alta complejidad. A diferencia de los problemas operativos relacionados con fallas de tubería única (impacto limitado en el rendimiento hidráulico del WDS), las ocurrencias de propagación de fallas pueden interrumpir áreas sensibles del WDS e incluso colapsar todo el sistema.

Los escenarios críticos asociados a la propagación de fallas son aquellos provenientes de la ocurrencia de eventos naturales extremos como deslizamientos (Choi et al., 2018), periodos fríos (Karamouz et al., 2010) o periodos de sequía (Head, 2014); o ataques intencionales como vandalismo, ataques terroristas (Jeong et al., 2006) o ciberataques (Taormina et al., 2018). En todos los casos, la falla o interrupción de un solo activo o una pequeña porción del WDS puede causar el cierre de áreas más

grandes para solucionar el problema inicial. Esto puede provocar más condiciones de estrés en otras partes del sistema, variando las presiones de servicio, provocando eventos transitorios y, en última instancia, afectando otros activos de WDS. Aun así, los nuevos activos afectados pueden tener incidencia en el suministro de agua y el rendimiento hidráulico en otras partes de la red de agua.

El objetivo consiste en tomar prestados modelos de epidemiología para localizar activos sensibles propensos a desencadenar fallas en cascada a través de un WDS. Esto ayudará a las empresas de servicios públicos de agua en cualquier proceso de toma de decisiones para administrar y controlar los problemas de flujo de la red, como las presiones del servicio, la calidad del agua y la contaminación.

El trabajo de Shuang et al. (2014) puede ser considerado como el principal antecedente de lo aquí presentado. En ese trabajo, los autores realizaron una evaluación de vulnerabilidad de un WDS cuyos nodos fallan en cascada. Los mismos autores (Shuang et al., 2017) investigaron la confiabilidad del sistema y el tiempo de propagación de fallas para los WDS que experimentan fallas en cascada (Crucitti et al., 2004). Partiendo de la hipótesis de que la rotura de una sola tubería provocaría el cierre de las tuberías vecinas para su reparación. Esto tendrá un doble efecto: variaciones indeseables en la presión de los nodos de consumo del entorno, y aumento de los niveles de riesgo de rotura de tuberías en la red.

La calidad del servicio (QoS) para los clientes se ve afectada de la misma manera que una enfermedad que se propaga a través de una red. Las tuberías de la red de agua (y otros activos) también pueden “infectarse”, aumentando su estrés debido a presiones internas anormales y, en consecuencia, aumentando la probabilidad de rotura. Este escenario podría tener una propagación fatal en cascada a través del WDS.

Se han investigado modelos epidemiológicos para replicar fallas en cascada en redes (Madar et al., 2004), (Ajelli et al., 2010). Manzano et al., (2014) en su trabajo “*Epidemic and Cascading Survivability of Complex Networks*” proponen una medida sobre la llamada supervivencia epidémica que evalúa las vulnerabilidades de los

nodos de red bajo un proceso epidémico. También se investigó la capacidad de supervivencia en cascada, midiendo qué tan propenso es un nodo a propagar fallas en cascada.

Si bien los modelos epidemiológicos se han aplicado ampliamente a las líneas aéreas, el transporte y las redes eléctricas, se ha realizado poco trabajo más allá de lo mencionado anteriormente por (Bardet & Little, 2014) y (Navarrete-López et al., 2019) en WDS.

La mayor parte del trabajo realizado hasta ahora sobre fallos en cascada en redes complejas se enfoca en la propagación de fallas de nodo; pero el tema es amplio ya que la propagación de baja presión se analiza a través de nodos WDS. Vale la pena entonces destacar el importante papel de las tuberías de red (enlaces WDS) como vectores epidémicos. Las tuberías transmiten la enfermedad (baja presión) mientras que también son objeto de más interrupciones (estallido de tuberías) en sí mismas. El modelado de propagación de fallas de nodos y tuberías es clave para la evaluación de la confiabilidad del WDS y los planes de rehabilitación (Winkler et al., 2018).

Las tuberías son un elemento clave para la distribución de agua y se ha dedicado bastante investigación a sus fallas y a su importancia relativa individual entre el conjunto de todas las tuberías de la red (Izquierdo et al., 2008). El problema de las fallas en cascada también se extiende a otros activos de WDS, como válvulas y bombas (Berardi et al., 2014).

4.7.1 Riesgo y confiabilidad de los sistemas de distribución de agua.

Para profundizar en este aspecto, es necesario introducir los conceptos de riesgo y confiabilidad. Ambos están directamente relacionados con la propagación de fallos y los procesos en cascada en las redes de servicios públicos. En el caso de los WDS, la

complejidad y dimensión de las infraestructuras físicas y digitales asociadas al abastecimiento de agua generan nuevas fuentes de vulnerabilidades y fallas en la interconectividad urbana que, en última instancia, ponen en riesgo el bienestar y la salud de los ciudadanos.

El riesgo para una infraestructura pública es la posibilidad de tener consecuencias indeseables o pérdidas de cualquier tipo como consecuencia de la utilización de sus servicios. En términos generales, el riesgo es una estimación de la exposición del sistema a un evento disruptivo particular y potencial (Hopkins, 2012). El riesgo se estima por la probabilidad de ocurrencia de la pérdida multiplicada por el impacto de sus consecuencias. Las pérdidas se deben a fallas o interrupciones de servicios, estructuras organizacionales e instalaciones involucradas en el desempeño de la infraestructura.

Desde un punto de vista físico, las infraestructuras urbanas pueden entenderse como equipamientos en red en los que fluye un determinado recurso como el gas, la electricidad, el agua, los vehículos, entre otros. En general, el recurso se mueve a través de la red de un nodo a otro. Esta influencia tanto en los enlaces como en los nodos que experimentan una carga a medida que el recurso pasa activamente.

En un sistema que opera sin estrés adicional o sin los efectos de alguna perturbación (llamado estado normal de la red), los ajustes automáticos de los flujos de enlace y las cargas de los nodos pueden ocurrir como consecuencia de variar el balance de entrada/salida, entre otras condiciones.

Sin embargo, tanto la tasa de flujo como el estado de la topología de la red pueden cambiar. En el contexto de la evaluación y gestión de riesgos de seguridad de infraestructuras críticas, también deben tenerse en cuenta factores como la resiliencia y la vulnerabilidad (Khakzad & Reniers, 2015). En este contexto, la evaluación de la resiliencia y vulnerabilidad de una infraestructura crítica (por ejemplo, un sistema de distribución de agua) se vuelve crucial no solo para la red en sí, sino también por la

propagación de posibles efectos negativos mediante el proceso en cascada en otras infraestructuras relacionadas (Soldi et al., 2015).

Una sola falla en un elemento de la infraestructura puede propagarse a través de su propia infraestructura y afectar también a otros sistemas relacionados, propagando así esa falla en cascada. En este sentido, el proceso en cascada se puede clasificar en el orden de su interacción de dos maneras: 1) interdependencias de los elementos del sistema propio, y 2) interdependencias con otras infraestructuras.

En primer lugar, el impacto de la ocurrencia de cualquier evento de interrupción en una red podría originar el evento en cascada; es decir, una falla individual desencadena eventos adversos sucesivos en otros vértices de la red para finalmente extenderse a las zonas de menor resiliencia del sistema. En segundo lugar, la generalización de los procesos de propagación de fallos de la infraestructura primaria aumenta la complejidad de las infraestructuras secundarias afectadas y a su vez la vulnerabilidad a posibles fallos de funcionamiento de estos sistemas secundarios a gran escala (Pescaroli et al., 2016).

En caso de interrupción o disfunción de una tubería principal de agua, el flujo que pasa a través de ella debe redistribuirse a través de otras tuberías en la red (Ash & Newth, 2007; Zio & Sansavini, 2011). Esta redistribución (también denominada evento desencadenante) (Simonsen et al., 2008), puede provocar que otros activos de la red y tuberías sobrepasen su capacidad provocando más eventos de rotura en el caso de tuberías, o un mal control del suministro de agua. De esta manera, la cantidad de activos afectados por la única interrupción inicial se propaga por toda la red y finalmente colapsa toda la red.

Las fallas en cascada se pueden describir por su frecuencia (Plietzsch et al., 2016), (Schäfer et al., 2018), intensidad (Pahwa et al., 2014), tiempo de recuperación relacionado (Hong et al., 2017) y extensión del daño (Liu et al., 2014). Sun et al., (2017) introdujo una serie de decisiones con el objetivo de evitar o, al menos, minimizar el impacto de estos problemas en el rendimiento y el servicio de la red, y

mejorar la resiliencia. Para ello, entre los procedimientos más utilizados se destaca: Reducir su riesgo aumentando la conectividad entre redes de infraestructuras críticas (Korkali et al., 2017), potenciar la robustez de las estructuras asociadas reforzando la conectividad de los nodos más influyentes en la conectividad de la red (Gong et al., 2017), desarrollar estrategias basadas en redes complejas para la recuperación incluso de un colapso total de la infraestructura (Di Muro et al., 2016), (Afrin & Yodo, 2019) y crear capacidades de autorreparación a través de topologías de anillo de red que explotan enlaces redundantes diseñados para obtener una recuperación automática del sistema (Angskun et al., 2010), (Quattrociocchi et al., 2014).

4.7.2 Epidemiología para modelar fallos en sistemas de distribución.

En este apartado se presenta el proceso de la red de difusión y, específicamente, los modelos de epidemiología adaptados para ejecutar su transmisión infecciosa en una red compleja. En aras de la simplicidad en un proceso de incorporación adicional de la propagación de fallos en WDS y la simulación hidráulica, se presentan los llamados modelos compartimentales. Estos modelos simplifican modelos matemáticos complejos y se enfocan en el análisis y la replicación de la dinámica de transmisión de enfermedades y cómo la población en riesgo evoluciona en el tiempo.

4.7.2.1. Modelos compartimentales en epidemiología.

Los modelos compartimentales son un procedimiento para el modelado de enfermedades basado en la división de una población en compartimentos o grupos de individuos de características similares (Brauer, 2008), (Blackwood & Childs, 2018).

Luego, en función de las tasas de flujo entre compartimentos, es posible modelar la propagación del contagio además de determinar el estado infeccioso de cada individuo

y rastrear la evolución de los compartimentos a lo largo del tiempo (Dodds & Watts, 2004). El principal modelo que se emplea en este trabajo es el SIR, que ya se enunció en el apartado 4.6.3. Este modelo comprende tres compartimentos: susceptible - infectado - recuperado (SIR) (Kermack & McKendrick, 1927) . Cada individuo tiene asociado uno de estos tres estados en un momento particular, t , del proceso de transmisión de la enfermedad. En cada paso de tiempo t , cada individuo i entra en contacto con otro individuo j , vecinos en la red compleja. Si i es susceptible y j está infectado entonces, con probabilidad p , i recibe una dosis positiva $d_i(t)$, tomada aleatoriamente de alguna distribución $f(d)$; de lo contrario, $d_i(t) = 0$. Cada individuo mantiene una memoria de las dosis recibidas en los T pasos de tiempo anteriores. Los individuos susceptibles se infectan si reciben una dosis acumulada por encima de un determinado umbral d_i^* . Una vez infectados, los individuos pueden recuperarse con probabilidad r .

Para la especificación completa del modelo, la información sobre el cambio de un compartimento a otro debe venir con la tasa de transición correspondiente. Por ejemplo, el cambio de S a I viene con una probabilidad de contagio β . Para representar mejor el proceso que ocurre en un WDS, la probabilidad β debe ponderarse por la conectividad entre los individuos en el estado S e I y el flujo que pasa. Entre I y R, la tasa de transición es γ (tasa de recuperación). Si la duración de la infección se denota D , entonces $\gamma = 1 / D$, ya que un individuo experimenta una recuperación en D unidades de tiempo.

4.7.2.2. Epidemiología en redes.

Los modelos de epidemiología y difusión proponen un enfoque de vanguardia en el análisis de la dinámica de propagación en redes complejas (Pastor-Satorras et al., 2015). La teoría del campo medio basada en el individuo (IBMF) hace que el proceso

markoviano sea inherente a las tasas de transición entre los compartimentos (S, I, R, entre otros) de cualquier red.

Esto se aborda bajo la hipótesis de la independencia de tales estados entre un individuo y su vecindario. Los modelos IBMF se basan en las propiedades espectrales de la matriz de adyacencia asociada con la red (Di Nardo et al., 2018) junto con los resultados de la simulación para la versión estática de la red. El campo medio basado en grados (DBMF) proporciona un marco dinámico para los modelos de epidemiología en una red.

Precisamente, el tipo de modelo basado en agentes SIR se aplicará más adelante en el desarrollo del capítulo seis, dentro de la herramienta especial denominada NetLogo, para tratar el caso de estudio sobre gestión de fallos en la red de suministro de agua urbana de la ciudad de Coro, Venezuela.

CAPÍTULO 5

CASO DE ESTUDIO PARA GESTIÓN DE DEMANDA

“En este capítulo se desarrolla un caso de estudio en el que se aplican los procesos de gestión de demanda ya tratados en un objetivo concreto: la ciudad de Franca perteneciente al estado de São Paulo en Brasil. Comienza con una breve descripción del objetivo para contextualizar el estudio, y luego se aplican las metodologías de análisis usadas a redes de abastecimiento de agua para comprobar su efectividad, concretamente el modelo SAX. Se termina con la discusión de los resultados.”

5.1 LA CIUDAD DE FRANCA, SAO PAULO, BRASIL

El primer caso de estudio para este trabajo se centra en la ciudad de **Franca**, ubicada en la parte noreste del estado (São Paulo), distante 401 km de la capital del estado y 676 km de la ciudad de Brasilia. Cubre un área total de 605,679 km² de los cuales 86,92 km² comprenden el área urbana. De acuerdo con los datos del censo de 2021, la población de la ciudad era de 358.539 habitantes (IBGE, 2022). Se encuentra a 1.040 metros sobre el nivel del mar (*ilustración 56*). Ha sido conocida con diversos nombres como *Vila Franca do Rei* o el de *Vila Franca do Imperador*; recibió el estatus de ciudad en el año 1824. La ciudad tiene una de las mayores refinerías de azúcar en Brasil; posee cultivos de café, arroz, maíz, soja y feijão (frijoles). Franca se destaca industrialmente por ser el mayor productor de calzado de Brasil, de ahí que reciba el epíteto de “capital nacional del calzado”, además de la fabricación de muebles, vehículos, productos químicos, productos farmacéuticos, entre otros. Las mercancías viajan por ferrocarril y carretera a la ciudad de São Paulo y a otros puntos del estado

como Ribeirão Preto (80 km al sur), y al estado vecino de Minas Gerais (Enciclopedia Británica, 2015). Franca es el centro industrial y económico de un aglomerado urbano que comprende 19 municipios con un total de 672.053 habitantes, que se encuentran consecutivamente entre las principales ciudades brasileñas con el mejor saneamiento urbano (IBGE, 2022).



Ilustración 56. Vista aérea de Franca. Wender Henrique - Own work, CC BY-SA 4.0

En efecto, la ciudad de Franca ha liderado por siete años las estadísticas nacionales, estando acompañada en el 2022 por las ciudades de Santos (SP), Uberlândia (MG), São José dos Pinhais (PR) y São Paulo (SP). Entre los criterios para formar la clasificación se encuentran las tasas de suministro de agua a la población, donde Franca alcanza al 100% de las casas, además del tratamiento del agua y las inversiones (Goulart, 2022).

Al respecto, el IBGE - Instituto Brasileiro de Geografia e Estatística – ha desarrollado una investigación sobre el abastecimiento de agua (*ilustración 57*) y alcantarillado (*ilustración 58*), que arrojó para el año 2017 los siguientes resultados:

Ano: 2017		Franca			
<ul style="list-style-type: none"> ▼ ÁREA DO SETOR DE SANEAMENTO ▼ ABASTECIMENTO DE ÁGUA ▼ Número de economias abastecidas 				139.797	unidades
<ul style="list-style-type: none"> <ul style="list-style-type: none"> NÚMERO DE ECONOMIAS ATIVAS ABASTECIDAS RESIDENCIAIS 		124.596	unidades		
<ul style="list-style-type: none"> ▼ Características do Abastecimento de Água > COM REDE DE DISTRIBUIÇÃO DE ÁGUA 				Sim	unidades
<ul style="list-style-type: none"> <ul style="list-style-type: none"> Extensão total da rede de distribuição de água 		1.516	km		
<ul style="list-style-type: none"> > Volume de água tratada distribuída por dia 		72.706	m³		
<ul style="list-style-type: none"> > Volume de água consumido por dia 		55.668	m³		
<ul style="list-style-type: none"> Índice de perdas calculado 		23,4	%		

Ilustración 57. IBGE - Pesquisa Nacional de Saneamento Básico - Abastecimento de água e Esgotamento sanitário 2017 a

Ano: 2017		Franca			
<ul style="list-style-type: none"> ▼ ESGOTAMENTO SANITÁRIO POR REDE COLETORA ▼ Número de economias esgotadas 				138.002	unidades
<ul style="list-style-type: none"> <ul style="list-style-type: none"> NÚMERO DE ECONOMIAS ATIVAS ESGOTADAS RESIDENCIAIS 		123.110	unidades		
<ul style="list-style-type: none"> ▼ Características do Esgotamento Sanitário > COM REDE DE ESGOTAMENTO SANITÁRIO 				Sim	unidades
<ul style="list-style-type: none"> <ul style="list-style-type: none"> Extensão total da rede coletora de esgoto 		1.250	km		
<ul style="list-style-type: none"> > Volume de esgoto tratado por dia 		43.961	m³		

Ilustración 58. IBGE - Pesquisa Nacional de Saneamento Básico - Abastecimento de água e Esgotamento sanitário 2017 b

En la siguiente ilustración (59) se puede apreciar el plano del perímetro urbano de Franca, en línea violeta, y la línea de expansión a 5 años, en línea roja.

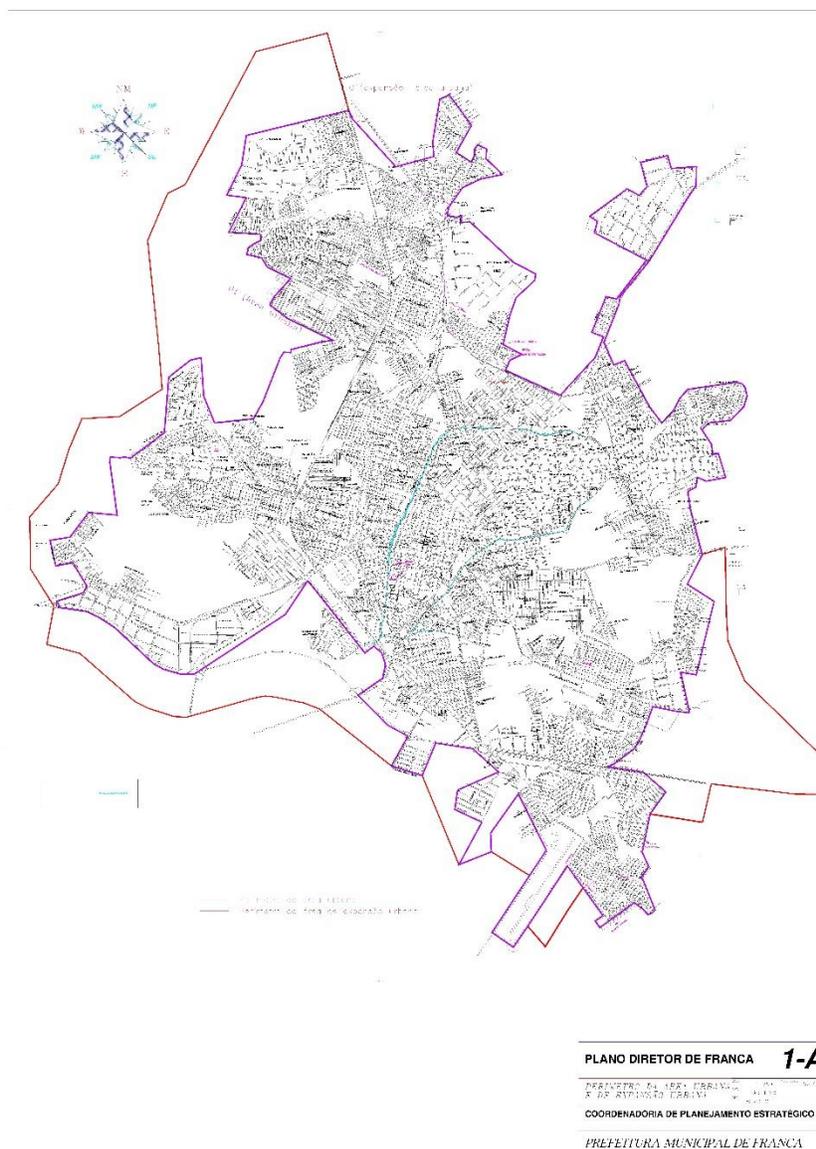


Ilustración 59. Perímetro Urbano de Franca.
https://www.franca.sp.gov.br/index.php?option=com_phocadownload&view=file&id=8134:mapa-1-a-perimetro-da-area-urbana-e-expansao-urbana-conforme-lc-50-2003-extensao-jpg&Itemid=258

Se puede apreciar que, dentro del perímetro urbano, existen áreas despobladas (muy pocas) y que el límite de expansión urbana agrega una cantidad significativa de terreno para construcción, por lo que el análisis para la efectividad del sistema de abastecimiento de agua urbana debe ser muy preciso.

En la ciudad de Franca, el operador de abastecimiento de agua es Sabesp - Companhia de Saneamento Básico do Estado de São Paulo - quien asumió los servicios de agua y alcantarillado en marzo de 1977. El municipio es abastecido por la Estación Potabilizadora del Norte con una capacidad de 878 litros por segundo (SABESP, 2023)

Las aguas residuales son procesadas en nueve sistemas - Franca, Luiza, Paulistano I, Paulistano II, City Petrópolis, Aeroporto, Palestina, São Francisco y Morada do Verde - con una capacidad total de 862,8 litros por segundo (SABESP, 2023).

5.2 DESCRIPCIÓN DEL CASO DE ESTUDIO

La metodología presentada se aplica a cuatro áreas de medición de distrito (DMA) del sistema de distribución de agua (WDS) de la ciudad de Franca (Navarrete-López et al., 2019).

El estudio de estos cuatro DMA tiene en cuenta la distribución espacial y las propiedades hidráulicas de la demanda de agua en todo el WDS. En la ilustración 60 se muestra cómo Franca WDS se divide en varias DMA. Dicha partición se realizó considerando las condiciones de presión de las áreas principales. Los bloques grises en esa ilustración (60) corresponden a las DMA involucradas en el caso de estudio de este trabajo. En este WDS, el agua se suministra desde una sola planta de tratamiento de agua a todas las DMA mediante una tubería principal. Los tanques más importantes para la distribución de agua se destacan en la figura y son los responsables de controlar y regularizar el suministro para las DMA que integran todo el WDS (Navarrete-López et al., 2019).

Tres de las cuatro DMA elegidas para el análisis son de tamaño similar en cuanto al número de nodos de demanda. Específicamente, el número de conexiones domiciliarias para SA-3 es 2.168, para SL es 2506 y para SA-ZA es 2728. Estas tres

DMA son distritos urbanos con pequeñas empresas, una configuración común para las áreas residenciales de Brasil. La cuarta DMA involucrada en el estudio de caso es más grande que las otras tres. Se llama ETA, tiene 10.439 conexiones domiciliarias y también abastece a un centro penitenciario; por lo tanto, es un sector de particular importancia para la gestión del agua de la ciudad. La serie temporal de cada DMA objeto de estudio recoge datos de consumo de demanda de agua de 4.000 h medidos en litros por segundo. Así, los datos del estudio de caso abarcan un período aproximado de cinco meses (Navarrete-López et al., 2019).

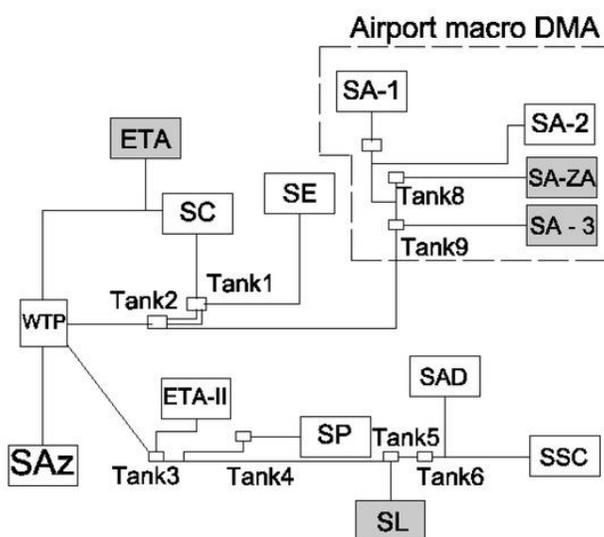


Ilustración 60. Diseño DMA del WDS de Franca. Navarrete-López et al., (2019)

5.3 RESULTADOS

La dimensión de la serie temporal correspondiente a los cuatro DMA del caso de estudio se reduce utilizando SAX. La ilustración 61 muestra el rendimiento del proceso de ajuste de PAA. La selección de reducción considera: (a) la precisión del PAA, en términos de la suma de cuadrados—SS—computada como la distancia entre el promedio aritmético para un PAA y los valores originales de la serie de tiempo; y (b) la reducción del número, representada por el número de segmentos. Bajo ambos criterios, se toma 400 como el número de segmentos que forman la partición PAA,

con un diccionario o alfabeto de cuatro letras $\{a, b, c, d\}$. Las letras están relacionadas con los distintos niveles de valores de demanda de agua en la serie temporal; donde a es el nivel más bajo y d es el más alto (Navarrete-López et al., 2019).

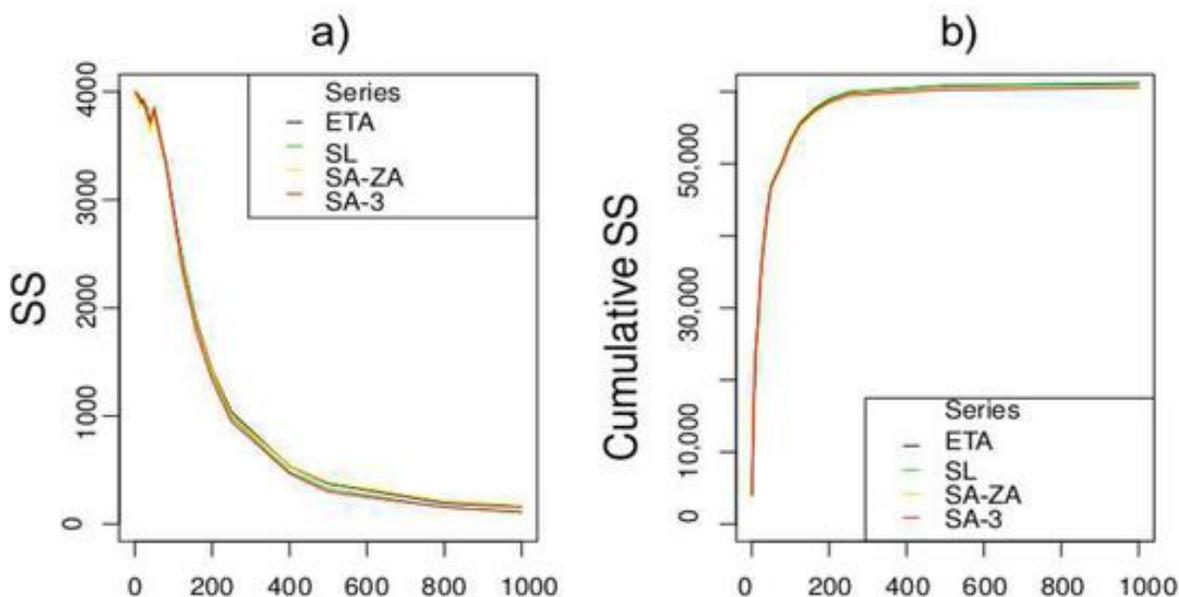


Ilustración 61. Número de segmentos para la configuración PAA; (a) SS vs. número de segmentos; (b) SS acumulada vs. número de segmentos. Navarrete-López et al., (2019)

La precisión de los diversos PAA se explica mejor en la ilustración 62. Allí se presentan varios niveles de PAA de reconstrucción de la demanda de agua para el caso del distrito SL: la ilustración 62 a–d muestra el PAA para las reconstrucciones de series de tiempo en 50, 100, 200 y 400 segmentos, respectivamente. Este es un buen ejemplo representativo de cómo el nivel de precisión de un PAA depende del número de segmentos en los que se divide la serie temporal, que va desde la reconstrucción de series temporales de baja fidelidad para un PAA con 50 segmentos (Ilustración 62- a) hasta el PAA de superior fidelidad al utilizar 400 segmentos (Ilustración 62- d) (Navarrete-López et al., 2019).

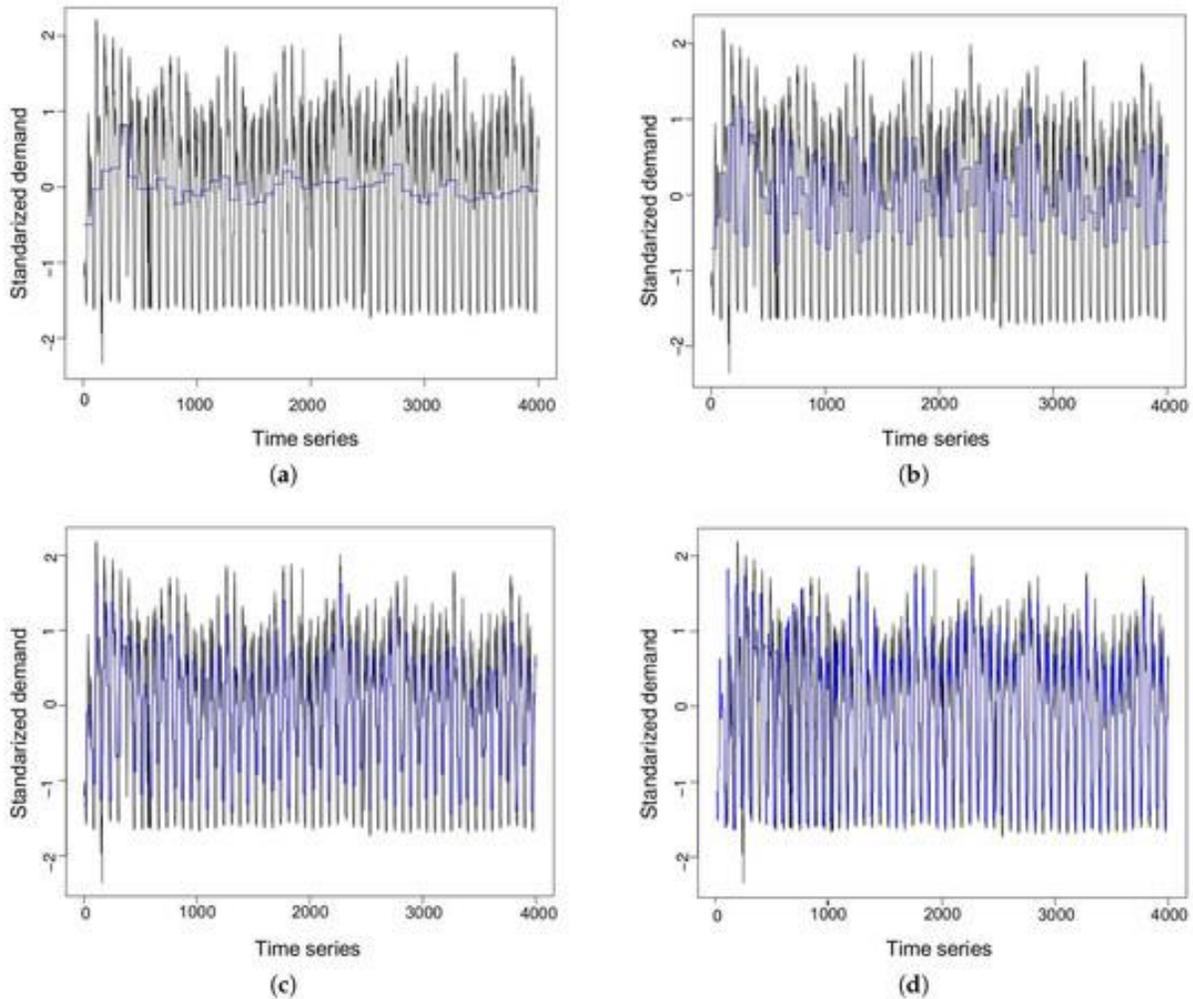


Ilustración 62. Sensibilidad de las reducciones de PAA. Caso de DMA de SL. (a) SL DMA: valores originales de demanda de agua y PAA calculados a través de 50 segmentos; (b) SL DMA: valores originales de demanda de agua y PAA calculados a través de 100 segmentos; (c) SL DMA: valores originales de demanda de agua y PAA calculados a través de 200 segmentos; (d) SL DMA: valores originales de demanda de agua y PAA calculados a través de 400 segmentos Navarrete-López et al., (2019)

La Tabla 6 muestra las distancias MINDIST entre los cuatro DMA cuando se utilizan 400 segmentos para calcular el PAA, ya que se trata de varios segmentos adecuados para explicar la mayor parte de la variabilidad encontrada en la serie temporal en estudio, según las ilustraciones 61 y 62. SL, SA -ZA y SA-3 son los DMA que están más cerca según la distancia MINDIST. Además, cabe señalar que la palabra SAX para la serie temporal de demanda de agua correspondiente a SA-ZA se puede obtener directamente combinando los otros dos DMA (Navarrete-López et al., 2019).

	ETA	SL	SA-ZA	SA-3
ETA	0	-	-	-
SL	4,77	0	-	-
SA-ZA	5,22	0,00	0	-
SA-3	4,27	2,13	2,13	0

Tabla 6. MINDIST para los 4 DMA del caso de estudio.

La subcadena **aaddcdbaadddddbaacdddca** es el patrón SAX más largo común a todas las series encontradas a través de árboles de sufijos aplicados a las palabras SAX. Considerando que cada letra representa valores para 10 registros de consumo de agua, este patrón muestra una tendencia similar para las 230 h de la serie de tiempo original (es decir, más de 9 días consecutivos con demandas de agua de nivel similar en todas las DMA). También se han encontrado otros patrones más cortos en las palabras SAX del caso de estudio, y cada uno se encuentra al menos dos veces dentro de cada serie. Patrones similares a los encontrados en SL se encuentran tanto en SA-ZA como en SA-3 (Navarrete-López et al., 2019).

Una de las principales hipótesis de este trabajo fue la de combinar estos patrones de similitud entre DMAs con los obtenidos a partir de las MINDIST. Esto ayudaría a respaldar otros métodos predictivos basados en datos históricos disponibles mediante el uso de información sobre comportamientos de consumo de agua similares en otras áreas (Navarrete-López et al., 2019).

Nuestra afirmación es que esto mejoraría la precisión de los modelos predictivos y, además, la gestión de condiciones de respuesta de demanda similares en diferentes DMA facilitarían la detección de lecturas anormales en los medidores de flujo o incluso la presencia de usos o fugas de agua inesperados. Esta idea (como se describe en la *ilustración 42*) se propone para mejorar el modelo predictivo para el DMA SL utilizando algún DMA de cierre MINDIST. En este caso se elige SA-3, en lugar de SA-ZA (que está a distancia nula), por razones de diversidad. Tenemos en cuenta que el uso de SA-ZA habría llevado al uso de características demasiado similares (casi

idénticas) y esto no habría enriquecido el proceso. Por el contrario, SA-3 introduce una mayor variabilidad en el proceso (Navarrete-López et al., 2019).

La ANN aquí considerada para trabajar junto con SAX tiene una capa oculta, y las neuronas usan funciones de activación logarítmicamente lineales. Se probaron varias configuraciones mediante la combinación de los parámetros típicos de ANN: 3, 6 y 9 nodos para la capa oculta; y 0,01 y 0,1 para el parámetro de tasa de aprendizaje. Los 2250 valores por DMA de la serie de datos se han reducido a 500, 400 y 200 valores por serie temporal para volver a entrenar la ANN en diferentes muestras. Esto permite trabajar con los valores promedio de precisión de los modelos en lugar de usar un solo valor. Las variables de entrada para la ANN son en todos los casos medidas de demanda de agua a 20 min, 40 min y 1 semana antes del tiempo actual t (salida). Las bases de datos se han dividido en particiones de 70% para entrenamiento y 30% para validación (Navarrete-López et al., 2019).

Los resultados obtenidos de los errores para el DMA SL y la combinación DMA SL + SA-3 muestran un comportamiento estable, con los errores de predicción esperados, es decir, los valores obtenidos en la predicción de las redes que combinan los dos DMA muestran valores consistentemente más bajos en todos los casos. Se realizó una prueba t de Student para comparar las medias de error de cada predicción: SL y SL + SA-3 obteniendo un valor de $p = 0,0033$, que muestra diferencias significativas entre las medias de los dos grupos. El error promedio global usando el único DMA SL es 1.0106 mientras que para la combinación de ambos DMA es 0.8779. La tabla 7 resume estos resultados; t es el estadístico de prueba y df los grados de libertad. El intervalo de confianza indica que el 95% de la muestra se encuentra entre sus límites (Navarrete-López et al., 2019).

Mean Error SL	Mean Error SL + SA-3	t	df	p -value	Confidence Interval
1,011	0,878	2,955	476	0,003	0,044-0,221

Tabla 7. Error medio de modelos predictivos utilizando DMA SL y DMA combinación SL + SA-3. Comparación por la prueba t de Student para igualdad de medias

La ilustración 63 muestra la distribución de errores por cada configuración de las ANN. La configuración 1 está formada por 4 bloques de 500 valores de demanda de agua, la configuración 2 está formada por 8 bloques de 500 valores, la configuración 3 está formada por 20 bloques de 200 valores y la configuración 4 está formada por 8 bloques de 400 valores. El resto de parámetros son los mismos para las cuatro configuraciones. El error promedio mínimo se encuentra para la Configuración 3 (Navarrete-López et al., 2019).

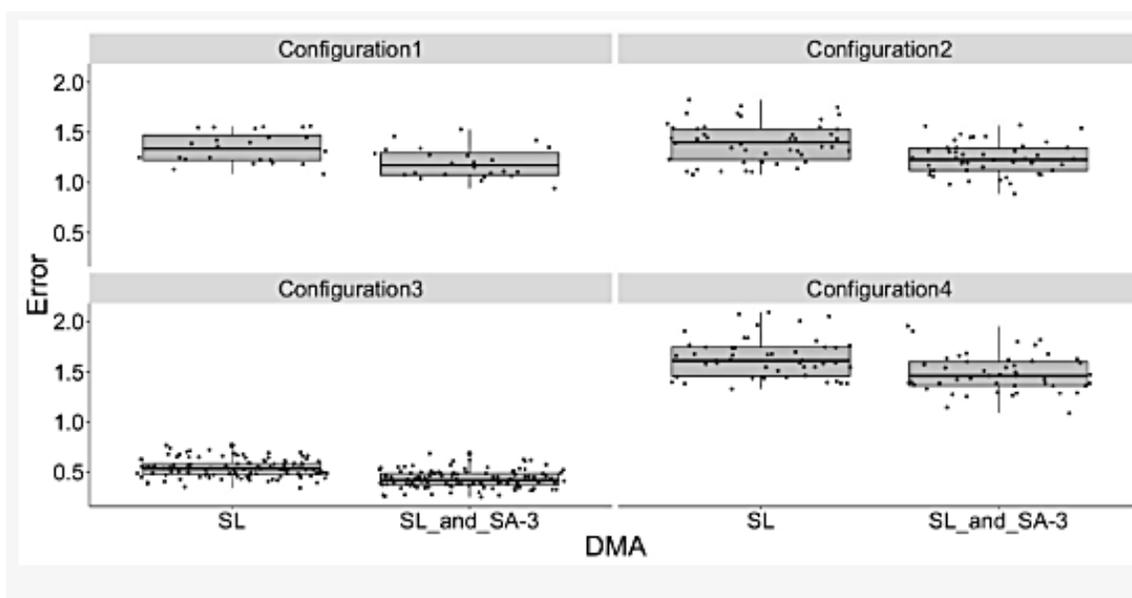


Ilustración 63.. Box-plots de los errores asociados por configuración de ANN (Navarrete-López et al., 2019)

Los errores calculados en la *ilustración 63* se producen al ejecutar de forma iterativa la ANN con datos nuevos para fines de capacitación y validación en cada momento. Repitiendo varias veces el mismo modelo con diferentes datos, es posible llegar a cálculos más robustos. El modelo predictivo final se selecciona en promedio y no está sesgado por el efecto de su dependencia de una muestra única. La tabla 8 muestra los resultados para la configuración superior de ANN (configuración 3) (Navarrete-López et al., 2019).

Tasa de aprendizaje	Nodos en capa oculta	Error medio SL	Error medio SL + SA-3
0,01	3	0,549	0,455
0,01	6	0,533	0,410
0,01	9	0,526	0,425
0,1	3	0,567	0,488
0,1	6	0,529	0,401
0,1	9	0,517	0,399

Tabla 8. Error medio de predicción (configuración superior de ANN) para DMA SL y para la combinación de DMA SL + SA-3.

Dos pensamientos principales surgen del análisis de las distribuciones de error que se muestran en la ilustración 63. En primer lugar, el modelo mejorado proporciona resultados más precisos en todos los casos. Esto significa que la estrategia de usar información de varias DMA para el modelo predictivo en una sola DMA es adecuada. En segundo lugar, existe un límite superior para la longitud óptima de la información utilizada para realizar el modelo predictivo de la demanda de agua. Esto sucede porque el error mínimo se encuentra utilizando modelos predictivos de 200 valores (Navarrete-López et al., 2019).

CAPÍTULO 6

CASO DE ESTUDIO PARA GESTIÓN DE FALLOS

“En este capítulo se desarrolla un caso de estudio en el que se aplican los procesos de gestión de fallos ya tratados con un objetivo concreto: la ciudad de Coro perteneciente al estado de Falcon en Venezuela. Comienza con una breve descripción de la herramienta a usar para este caso y posteriormente se describe el objetivo para contextualizar el caso. Luego se aplican las metodologías de análisis usadas a redes de abastecimiento de agua para comprobar su efectividad. Se termina con la discusión de los resultados.”

6.1 LA CIUDAD DE CORO, ESTADO DE FALCON, VENEZUELA

El segundo caso de estudio para este trabajo se centra en la ciudad de **Coro**, capital del municipio de Miranda en el Estado de Falcón al noroeste de Venezuela. Se encuentra a unos 320 km al oeste-noroeste de Caracas, en el extremo sur del istmo que une la Península de Paraguaná a tierra firme. Tiene una altitud media aproximada de 17 metros sobre el nivel del mar, con un clima semiárido peninsular. Posee construcciones en tierra únicas en toda la región del Caribe y por ser el único ejemplo subsistente de una fusión lograda de las técnicas y estilos arquitectónicos autóctonos, mudéjares españoles y holandeses, la ciudad de Coro fue declarada como Patrimonio de la Humanidad por la Unesco en 1993 (Enciclopedia Británica, 2014).

Para el año 2020, la ciudad tenía una población aproximada de 265.569 habitantes en su área urbana y en su área total alcanzaba los 313.428 habitantes según cálculos del INE (Instituto Nacional de Estadística de Venezuela); su crecimiento demográfico se estima en un 2,2% anual. La economía de la ciudad gira en torno al turismo. La red de abastecimiento de la ciudad, se puede apreciar en la siguiente ilustración (64).

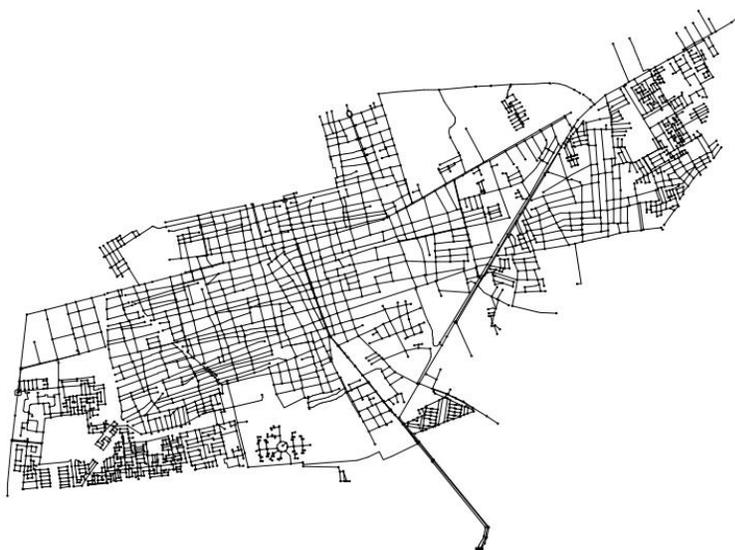


Ilustración 64. Red de abastecimiento de Coro (Martins, 2020)

La administración de esta ciudad carece de un sitio oficial de información que sea accesible completamente, lo cual impide obtener datos precisos y más cercanos a la cotidianidad de la población, sus costumbres de uso, sus infraestructuras y demás condiciones de habitabilidad (<http://alcaldiamiranda.com/index.php>). Carecer de información complementaria precisa permite inferir que, ante una situación de contingencia que afecte a la mayoría de la ciudad, los tiempos de respuesta serían lentos y por lo mismo, puede llegarse a un colapso. Precisamente, ese es el problema que se aborda: *los tiempos de respuesta*. Ante un fallo en la red de abastecimiento de agua urbana y su propagación en cascada, el modelado basado en agentes permite establecer el comportamiento de los tiempos de respuesta óptimos para evitar el colapso de la red.

6.2 LA HERRAMIENTA DE MODELADO NETLOGO

Antes de continuar este caso de estudio, conviene apreciar una breve descripción de la herramienta que se utilizó en su análisis: la aplicación NetLogo. NetLogo es un lenguaje de programación multiagente y un entorno de modelado para simular fenómenos naturales y sociales. Es particularmente adecuado para modelar sistemas complejos que evolucionan con el tiempo. Los modeladores pueden dar instrucciones a cientos o miles de "agentes" independientes, todos operando al mismo tiempo. Esto hace posible explorar las conexiones entre los comportamientos de nivel micro de los individuos y los patrones de nivel macro que surgen de sus interacciones (Wilensky et al., 2004).

NetLogo permite a los usuarios abrir simulaciones y “jugar” con ellas, explorando su comportamiento bajo diversas condiciones (*Ilustración 65*). También es un entorno de creación que es lo suficientemente simple como para permitir a los estudiantes e investigadores crear sus propios modelos, incluso si no son programadores profesionales (Wilensky et al., 2004).

Es una aplicación independiente escrita en Java, por lo que puede ejecutarse en todas las principales plataformas informáticas. Es un software gratuito: cualquiera puede descargarlo de forma gratuita y crear modelos sin restricciones. Viene con una extensa documentación y tutoriales y una gran colección de modelos de muestra (Wilensky et al., 2004).

Como lenguaje, NetLogo es miembro de la familia Lisp que admite agentes y concurrencia. Los agentes móviles llamados "tortugas" se mueven sobre una cuadrícula de "parches", que también son agentes programables. Todos los agentes pueden interactuar entre sí y realizar múltiples tareas al mismo tiempo (Wilensky et al., 2004).

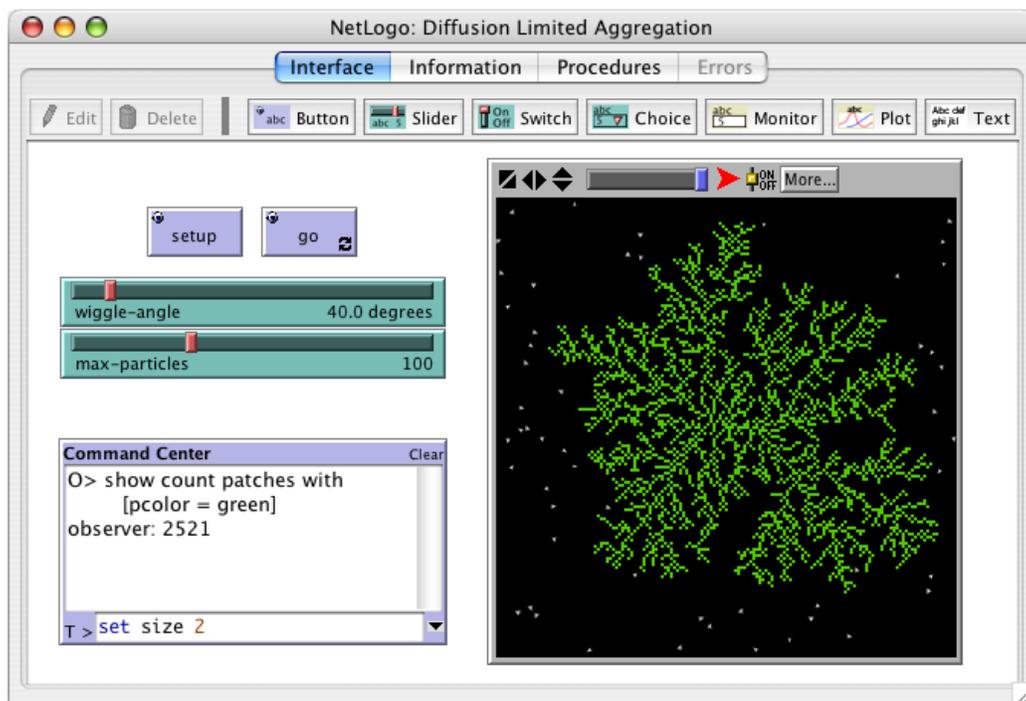


Ilustración 65. Ejemplo de interfaz de Netlogo (Wilensky et al., 2004)

6.3 MODELACIÓN BASADA EN AGENTES CON NETLOGO

Siempre que haya necesidad de estudiar un sistema real compuesto de elementos interconectados, donde cada uno de estos tiene su propia dinámica, a menudo es imposible prever el surgimiento de una dinámica global para el sistema. Por lo tanto, para que el comportamiento de un sistema esté descrito apropiadamente es necesario conocer, además de sus elementos, las interacciones o relaciones entre ellos y, además, sus estados (valores instantáneos de todos los elementos), así como sus transiciones (cambios dinámicos entre esos estados). En resumen, debe describirse la estructura y la dinámica de un sistema (Caparrini, 2015).

Para este caso de estudio, lo que está en juego es un sistema complejo, cualquier tipo de modificación, incluso si ella fuera marginal, en términos de uno o varios de sus

elementos puede conducir a un cambio dramático en la operación general del sistema. Es claro que estos fenómenos pueden ser observados y entendidos solo a través de la construcción de un modelo, por lo que el conocimiento de herramientas como NetLogo para el modelamiento de sistemas complejos se convertirá en una necesidad a futuro y el modelado basado en agentes es una de las maneras más exitosas para modelar este tipo de sistemas.

6.4 APLICACIÓN DEL MODELADO DE NETLOGO A LA RED DE CORO

Con lo expuesto hasta ahora, se procede a estudiar esta red con de la herramienta de NetLogo para poder hacer la simulación de fallos en cascada. Para ello se utilizará la versión 6.3 de la herramienta que fue publicada el 29 de septiembre del 2022 y que posee las siguientes características: Extension API version: 6.1 / OpenJDK 64-Bit Server VM 17.0.3 (BellSoft; 17.0.3+7-LTS) / sistema operativo: Windows 10 10.0 (amd64 processor) / Scala version 2.12.16 / Java heap: used = 43 MB, max = 989 MB / JOGL: (3D View sin inicializar) / OpenGL Graphics: (3D View sin inicializar).

La interfaz para el modelo basado en agentes que se desarrolló en NetLogo, se puede apreciar en la siguiente ilustración (66):

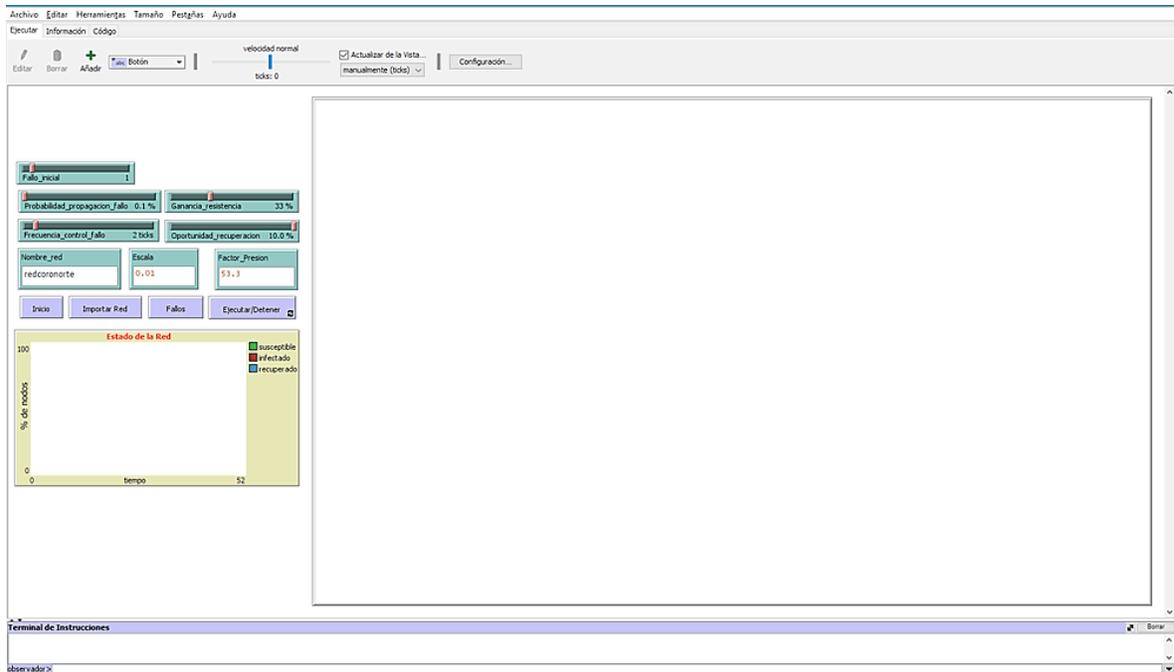


Ilustración 66. Interfaz de trabajo para el modelo MBA

La interfaz consta de tres pestañas principales que se corresponden con las acciones a desarrollar: pestaña de ejecutar, pestaña de información y pestaña de código.

La primera permite ver las instrucciones básicas que trae NetLogo para cualquier tipo de modelado (ilustración 67).

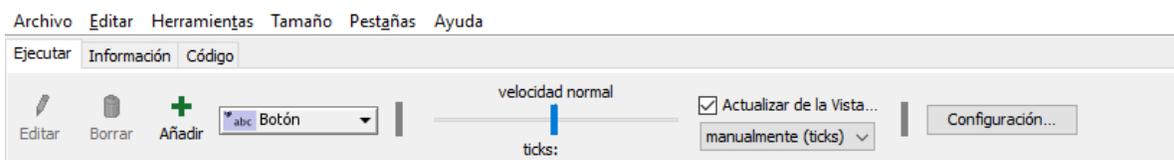


Ilustración 67. Pestaña 1 de la interfaz

La segunda pestaña incorpora información de base que facilita el diseño y construcción del modelo (ilustración 68).

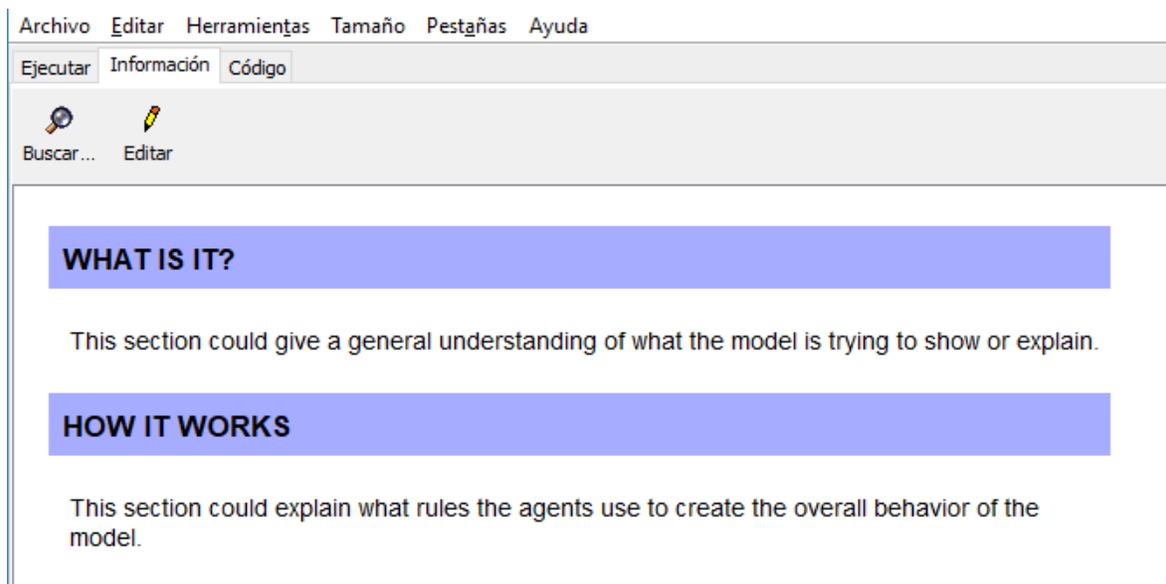


Ilustración 68. Pestaña 2 de la interfaz

La tercera pestaña incorpora funcionalidades para construir, editar y depurar el código de NetLogo para el modelo a trabajar (ilustración 69).

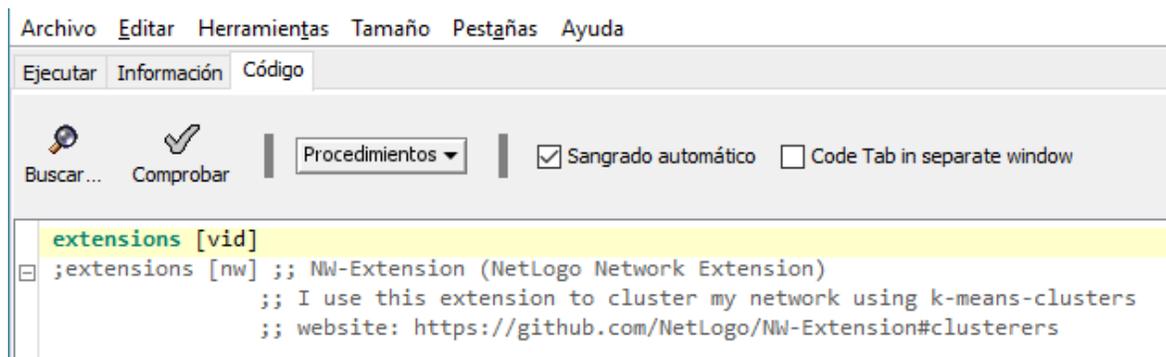


Ilustración 69. Pestaña 3 de la interfaz

En la ilustración (70) se puede apreciar el área de trabajo de NetLogo de autoría propia, donde se diferencian claramente dos secciones, la de la derecha en negro es el área donde se incorporan los datos de la red a trabajar; y a la izquierda los comandos

creados para ejecutar el modelo que se ha creado con la herramienta NetLogo para redes de agua urbana.

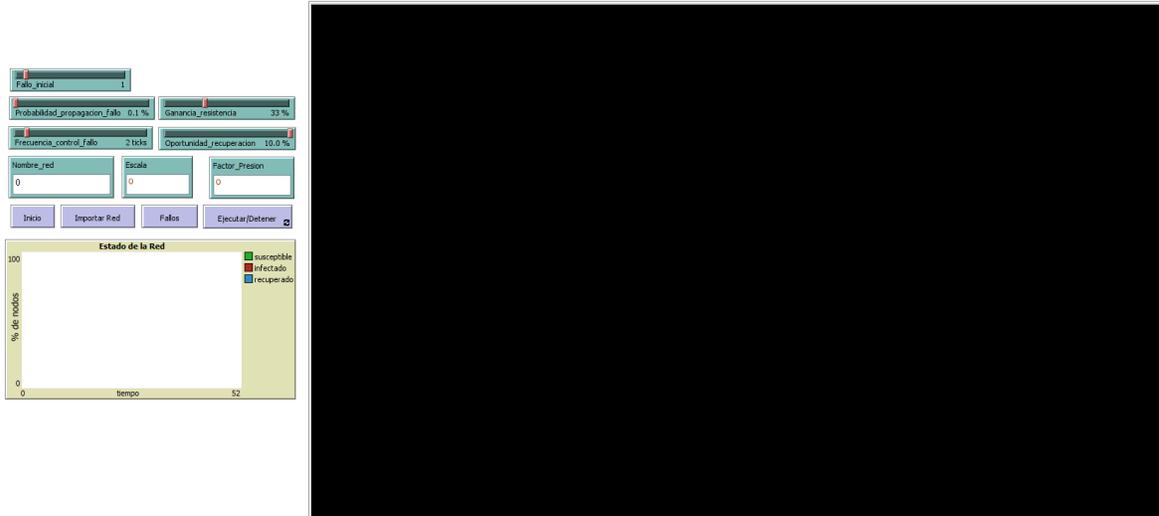


Ilustración 70. Area de trabajo de la interfaz

Entrando en detalle, en la sección de comandos (*ilustración 71*) se pueden apreciar ocho (8) cuadros de configuración para el experimento, destacados en color verde, donde se parametrizan:

1. Número de ocurrencias del fallo inicial: número de nodos donde se puede presentar el fallo inicial; se encuentra en el rango de 1 a 10.
2. Porcentaje de probabilidad de propagación del fallo: mide la severidad del fallo en términos de su propagación a otros nodos; se encuentra en el rango de 0% a 10%.
3. Porcentaje de ganancia de resistencia: mide el efecto que genera en la red los nodos resistentes; su rango va de 0% a 100%.
4. Número de frecuencia de control de fallo: número de ticks (momentos) en que se hace control mediante el modelo al fallo inicial y su propagación; su rango está entre 1 y 20 ticks.

5. Porcentaje de oportunidad de recuperación: mide la posibilidad de obtener nodos recuperados en un tick (momento); su rango se encuentra entre el 1% y el 10%.
6. Nombre de la red: indica el archivo del que se incorporan los datos de la red a analizar con el modelo.
7. Escala: indica la escala de la red que se puede apreciar en el área de trabajo a la derecha de estos comandos. Maneja un zoom (acercamiento/alejamiento) entre 0 y 100 unidades.
8. Factor de presión: indica el límite máximo de la presión en los nodos y tuberías en las que la red tiene su desempeño normal; cualquier valor por arriba de ese límite aumenta la probabilidad de fallos y los consecuentes fallos en cascada.

En la misma ilustración (71) se pueden apreciar cuatro botones de control para la ejecución del modelo, destacados en color morado, cuyas acciones son:



Ilustración 71. Comandos de ejecución para el modelo

1. Botón de inicio: limpia el área de trabajo de la red y prepara el modelo a su estado inicial para incorporar los datos de la red a analizar.
2. Botón de importar red: trae e incorpora todos los datos de la red desde un archivo adecuado. Dependiendo del tamaño, puede demorar un poco.
3. Botón de fallos: incorpora el número de nodos con fallo, según los parámetros indicados al modelo.
4. Botón de ejecutar/detener: inicia la simulación del modelo según los parámetros configurados y la detiene manualmente, si los ticks (momentos) aún no finalizan la simulación.

Para terminar, se encuentra una ventana de seguimiento en el que se muestran de manera simultánea los comportamientos de los agentes (nodos) durante el tiempo. Contrasta el tiempo en ticks (momentos), que varían según los parámetros configurados, con el porcentaje de agentes afectados. Estos agentes se agrupan en tres conjuntos: susceptibles, infectados y recuperados (notación tomada del modelo aplicado SIR), donde los susceptibles son los nodos que tienen probabilidad de fallar, los infectados son los nodos con fallos y que tienen oportunidad de recuperarse, y los recuperados son los nodos que tienen resistencia al fallo.

El código de programación original que se desarrolló (de autoría propia) para el modelo SIR en el lenguaje de NetLogo, se puede apreciar en el anexo 1 de este trabajo.

Al aplicar los datos de la red de suministro de agua urbana de la ciudad de Coro (Venezuela), se tiene que el total de nodos para ejecutar la simulación es de 2.887 y su ejecución se puede apreciar en el video que la misma herramienta de NetLogo produce y que, con una edición propia, se puede visualizar en el siguiente enlace:

<https://youtu.be/a6o1UZ05BwM>

6.5 ANÁLISIS DE RESULTADOS Y DISCUSIÓN

Una vez estructurada la red para analizarla con el modelo SIR en la herramienta de NetLogo, se procede a realizar cinco simulaciones con diez momentos cada una. Estos momentos se corresponden con los ticks: 1, 10, 30, 80, 180, 240, 300, 360, 410, 450 los cuales fueron tomados así porque son los momentos que mejor reflejan los cambios del comportamiento de los agentes dentro del modelo.

Cada uno de esos momentos arroja una tabla que se presenta en el anexo 2 del documento y que muestran los resultados por cada simulación, por cada momento y por cada estado, entendiendo esto último como estado susceptible, infectado y recuperado.

Tomando entonces resultados por cada estado, los obtenidos para los nodos susceptibles son los siguientes:

SUSCEPTIBLES (TRUE)										
S	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_10
1	2680	96	55	20	4	2	7	4	4	4
2	2742	102	57	38	7	3	4	6	4	5
3	2759	92	65	23	6	5	3	1	6	4
4	2719	84	62	24	10	3	5	3	5	6
5	2751	79	60	32	7	4	5	5	6	NA

Tabla 9. Resultados simulación: estado Susceptibles (true)

En la tabla (9), la **S** significa el orden de la simulación y la **M_** significa el orden de los momentos capturados. Tomando estas cifras, se realiza un gráfico, el cual se aprecia en la ilustración (72) que muestra el comportamiento de los nodos susceptibles a lo largo de cada simulación con una fuerte tendencia a disminuir mientras pasa el tiempo. Esto se evidencia en el momento 2 de cada simulación.

Este resultado se debe a que el porcentaje de probabilidad de propagación del fallo utiliza un valor del 0,1% de modo que la severidad del fallo en cascada no incida drásticamente en la simulación; si fuese más alto, el efecto será casi que inmediato entre el tick 0 y el tick 1, colapsando totalmente la red.

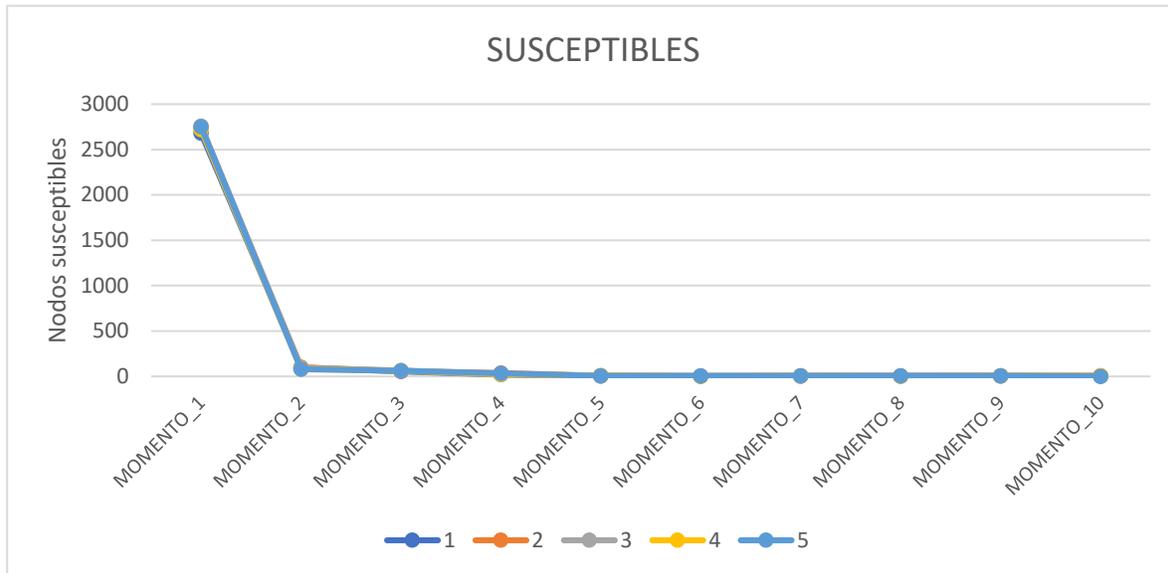


Ilustración 72. Comportamiento de nodos susceptibles

El hecho de que a partir del momento 5 se mantiene más o menos estable el comportamiento de los nodos susceptibles, indica que el porcentaje de ganancia de resistencia genera un efecto estabilizador en la totalidad de la red al ir aumentando los nodos recuperados; si ese porcentaje fuera inferior, la red se estabilizaría a muy largo plazo; mientras que, al aumentar, la red se estabiliza más pronto, impidiendo la propagación de los fallos en cascada. Este comportamiento es muy similar al que se pueda obtener con otras herramientas epidemiológicas computacionales en casos con nodos más reducidos.

Tomando entonces resultados obtenidos para los nodos infectado se tiene:

INFECTADOS (TRUE)										
S	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_10
1	203	2395	1695	746	152	64	15	6	2	0
2	142	2373	1700	717	132	44	15	6	2	0
3	126	2399	1740	756	139	59	25	15	5	3
4	167	2421	1721	759	157	58	21	11	4	0
5	135	2393	1714	768	137	51	11	2	0	NA

Tabla 10 Resultados simulación: estado Infectados (true)

Con las cifras obtenidas en la tabla (10), se realiza un gráfico, el cual se aprecia en la ilustración (73) que muestra el comportamiento de los nodos infectados a lo largo de cada simulación con un rápido crecimiento entre los momentos 2 y 3 para luego tener una tendencia a disminuir mientras pasa el tiempo. Esto se evidencia en el momento 10 de cada simulación, pues el fallo deja de existir en ese instante.

Al igual que en el estado de los susceptibles, el porcentaje de probabilidad de propagación del fallo con un valor del 0,1% evita que la severidad del fallo no colapse la red. Conviene apreciar que la frecuencia de control del fallo, tanto en el estado de los susceptibles como en el de los infectados y, posteriormente en el de los recuperados, permite un margen de maniobra que facilita la recuperación de los nodos en un tiempo breve, manteniendo alertas en la tasa de recuperación.

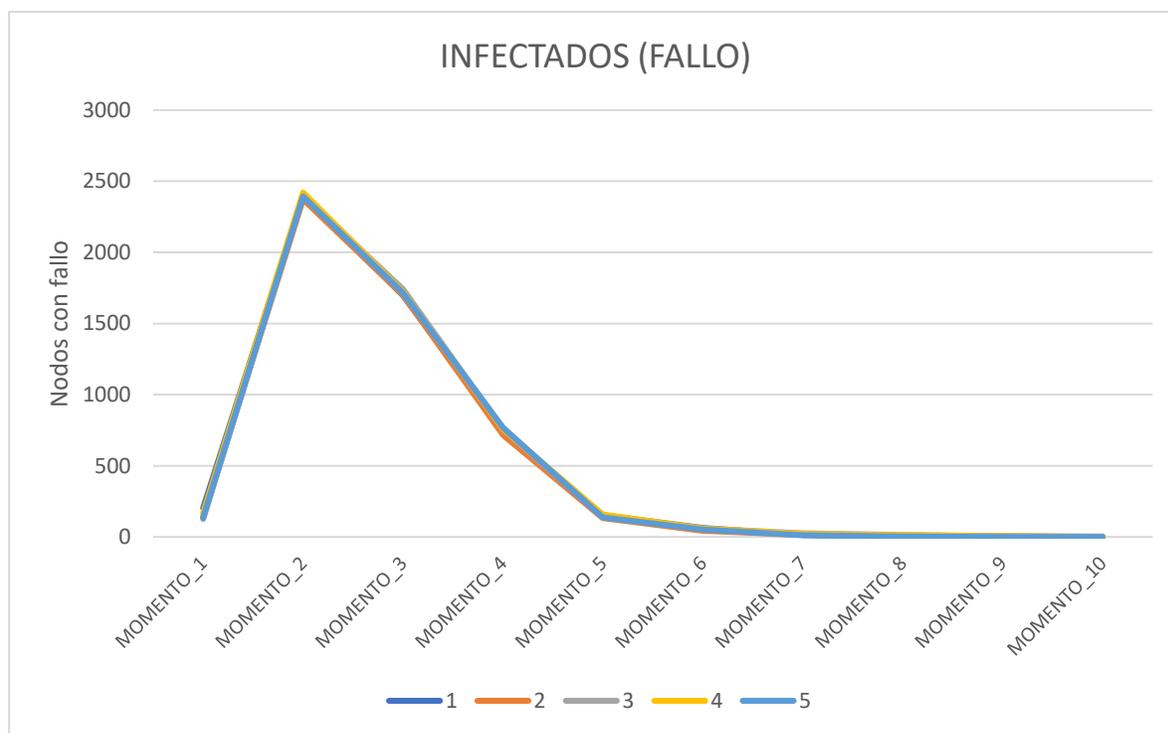


Ilustración 73. Comportamiento de los nodos infectados (con fallo)

Es de tener en cuenta que la cantidad total de nodos de la red es de 2.887 y que, en todas las simulaciones, el fallo en cascada nunca abarcó a todos los nodos, algo normal para este tipo de redes de suministro de agua urbana.

Ahora, al ver los resultados obtenidos para los nodos recuperados se tiene:

RECUPERADOS (TRUE)										
S	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_10
1	3	395	1136	2120	2730	2820	2864	2876	2880	2882
2	2	411	1129	2131	2747	2839	2867	2874	2880	2881
3	1	395	1081	2107	2741	2822	2858	2870	2875	2879
4	0	381	1103	2103	2719	2825	2860	2872	2877	2880
5	0	414	1112	2086	2743	2831	2870	2879	2880	

Tabla 11. Resultados simulación: estado Recuperados (true)

Estas cifras obtenidas en la tabla (11), muestran como los nodos se han recuperado luego de la infección, teniendo presente que, en el momento de máxima infección, había también un número importante de resistentes que llevaron al incremento de los recuperados en forma constante. Esos resistentes eran nodos que no se volvieron susceptibles y se mantuvieron a lo largo de cada simulación más o menos estables. Esto se aprecia mejor en la ilustración (74).

Estos resultados permiten ver la incidencia que el porcentaje de la oportunidad de recuperación tiene sobre el comportamiento de los nodos de la red; mientras sea alto, habrá más resistentes que susceptibles y por lo mismo, el fallo en cascada se puede controlar en un menor tiempo posible; por el contrario, si el porcentaje es mínimo, la factibilidad de colapso de la red se incrementa y, por lo tanto, los tiempos de recuperación serán más extensos.

Otro punto a tener en cuenta en este estado consiste en apreciar que, conociendo el número total de nodos de la red (2.887), los nodos susceptibles no superan de 2.750 y que los nodos con fallos por mucho llegan a unos 2.400; la cantidad de nodos recuperados es bien alta (2.880) debido a la resistencia que algunos nodos generan, y puede decirse, además, que la misma red de abastecimiento ocasiona.

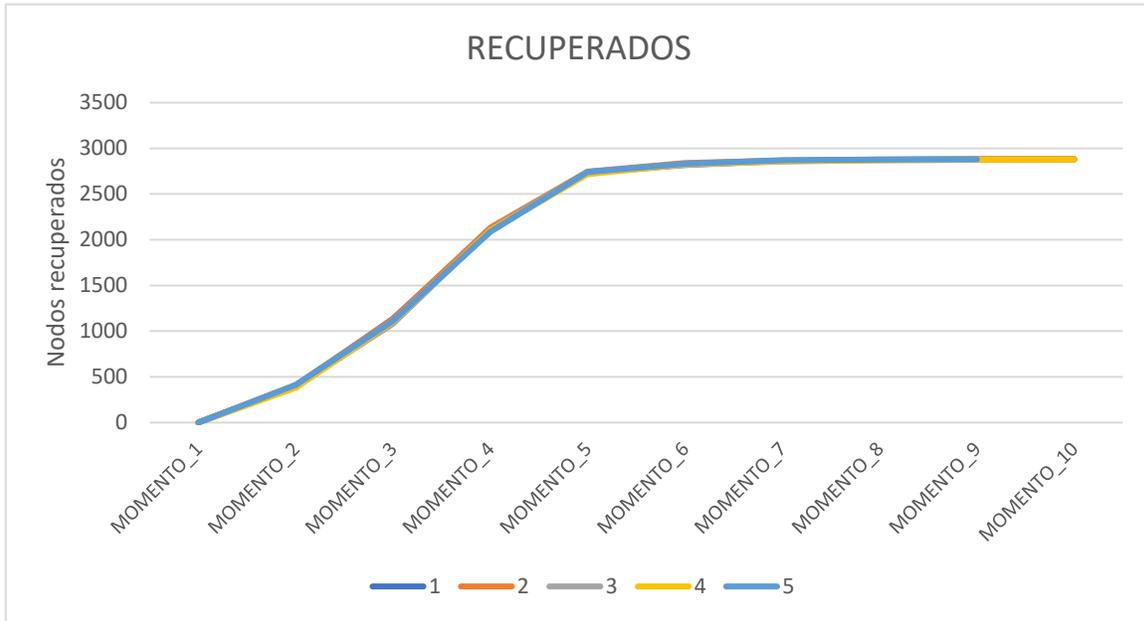


Ilustración 74. Comportamiento de los nodos recuperados (con fallo)

Apreciando y combinando las ilustraciones (73) y (74) se puede encontrar que momentos pueden ser considerados importantes para este análisis.

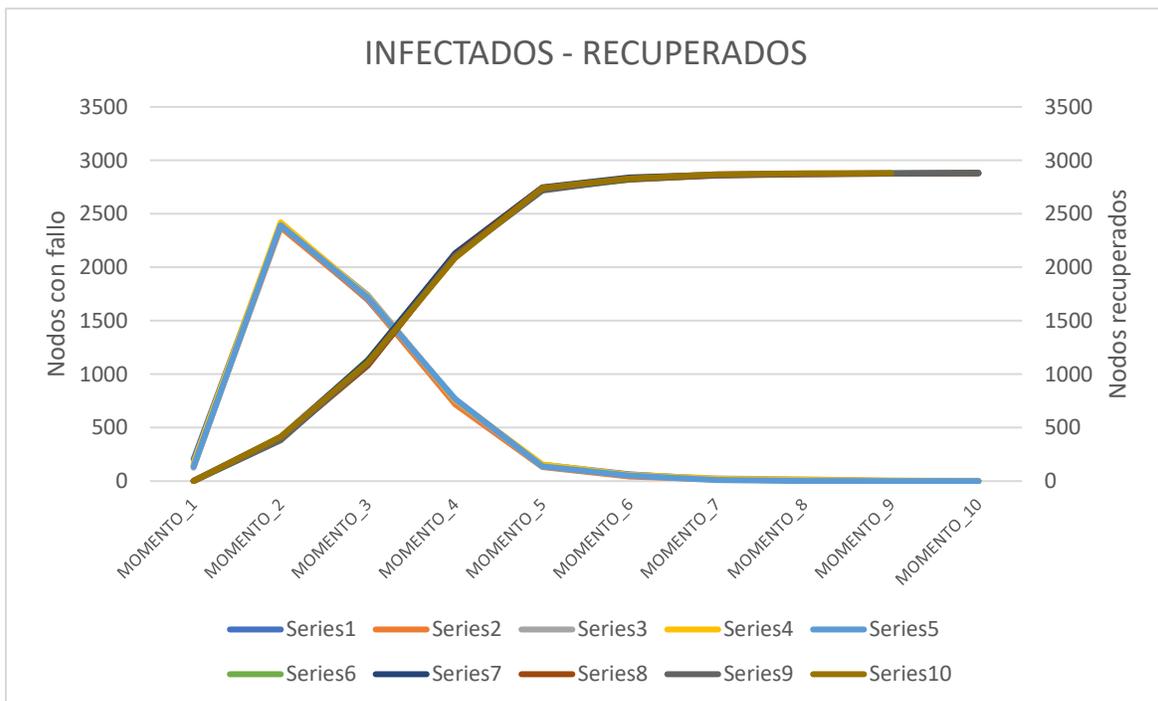


Ilustración 75. Combinación entre infectados y recuperados (true)

La ilustración (75) muestra una combinación de los estados de infección y recuperación, tal y como pueden apreciarse en la ventana de seguimiento que presenta NetLogo. Es de recordar que en las herramientas epidemiológicas computacionales pueden apreciarse gráficas muy similares. Para este caso, los momentos 2, 3 y 5 merecen una atención especial, pues muestran los momentos importantes de las simulaciones y que, bajo la configuración establecida, pueden pensarse como generales para la red de Coro ante un evento de fallo en cascada.

El momento 2 muestra un crecimiento significativo de los fallos sin que colapse la red y, simultáneamente, un impulso importante en la recuperación. Al finalizar el momento 3, la recuperación es tan fuerte que los fallos decrecen y llegan a un “punto de equilibrio” por así decirlo, haciendo significativa la tasa de recuperación como factor determinante de ese punto. A partir del momento 5, los fallos decrecen progresivamente a cero, mientras que las recuperaciones alcanzan cotas altas afectando incluso a aquellos nodos susceptibles que no fallaron y pasan al estado de recuperación.

Un elemento adicional que se puede inferir de la ilustración (75) consiste en el comportamiento global de la red (con 2.887 nodos), para determinar los mapas de calor que facilitan la predicción de aquellas áreas en que se pueden presentar facilidades para la propagación de fallos en cascada o de mayor vulnerabilidad. La aplicación del modelo SIR, permite obtener las siguientes gráficas de resumen (*ilustraciones 76, 77 y 78*).

Para construir los mapas de calor de la red de Coro, se realizaron interacciones entre el programa NetLogo y Python. Se tomaron los datos de la red de Coro con sus correspondientes nodos y se realizó la simulación de presiones usando el paquete WNTR (Water Network Tool for Resilience) de Python, la salida de la red se llevó a Netlogo y se inició la simulación deteniéndose en el primer tick de tiempo, esta salida se llevó nuevamente a Python cerrando las tuberías de los nodos que resultaron infectados en el primer tick y simulando nuevamente las presiones de los nodos que

permanecieron sin infectar. Los resultados nuevamente fueron llevados a NetLogo y se continuó con el segundo tick de tiempo, lo que generó nuevos nodos infectados, adicionales a los que ya estaban del primer tick. Nuevamente, la base de datos fue llevada a Python simulando las presiones para los nodos que permanecían sin infectar. Al finalizar el segundo tick y simulación de presiones, los estados de los nodos y las presiones se plasmaron en los mapas de calor correspondientes, permitiendo identificar las zonas más vulnerables a fallos de la red.

Todas las simulaciones marcan una constante a partir de los momentos indicados en la ilustración (79), por eso solo se hace énfasis en estos ticks (momentos).

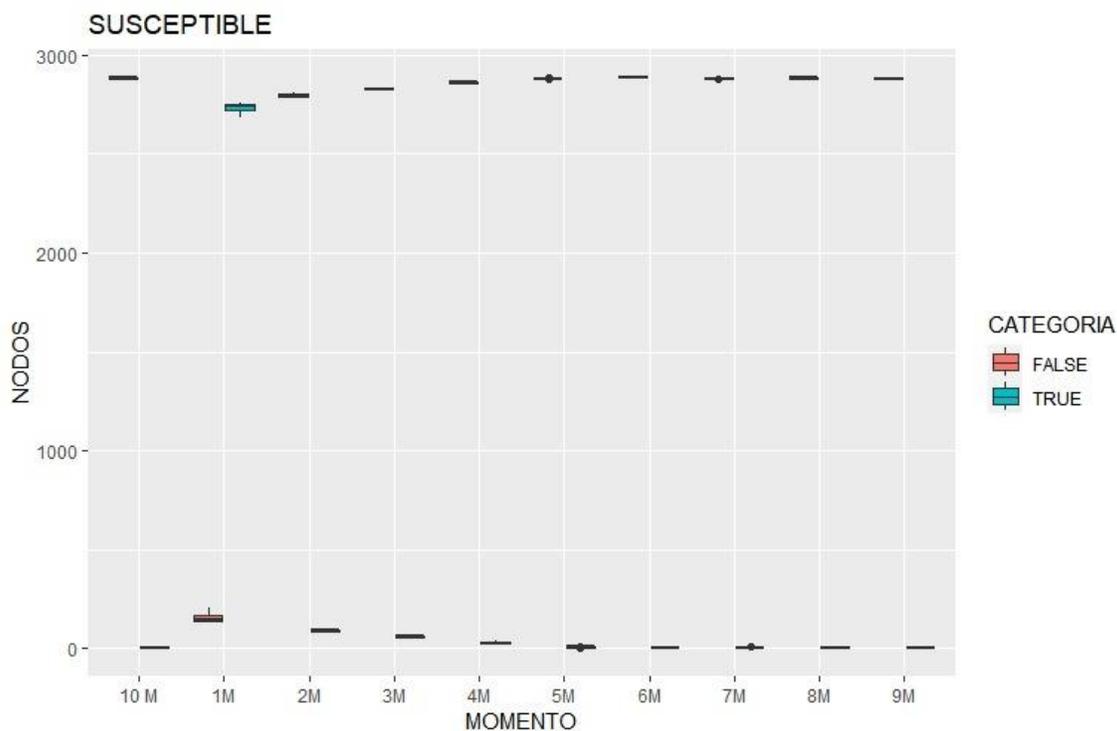


Ilustración 76. Resumen de nodos susceptibles

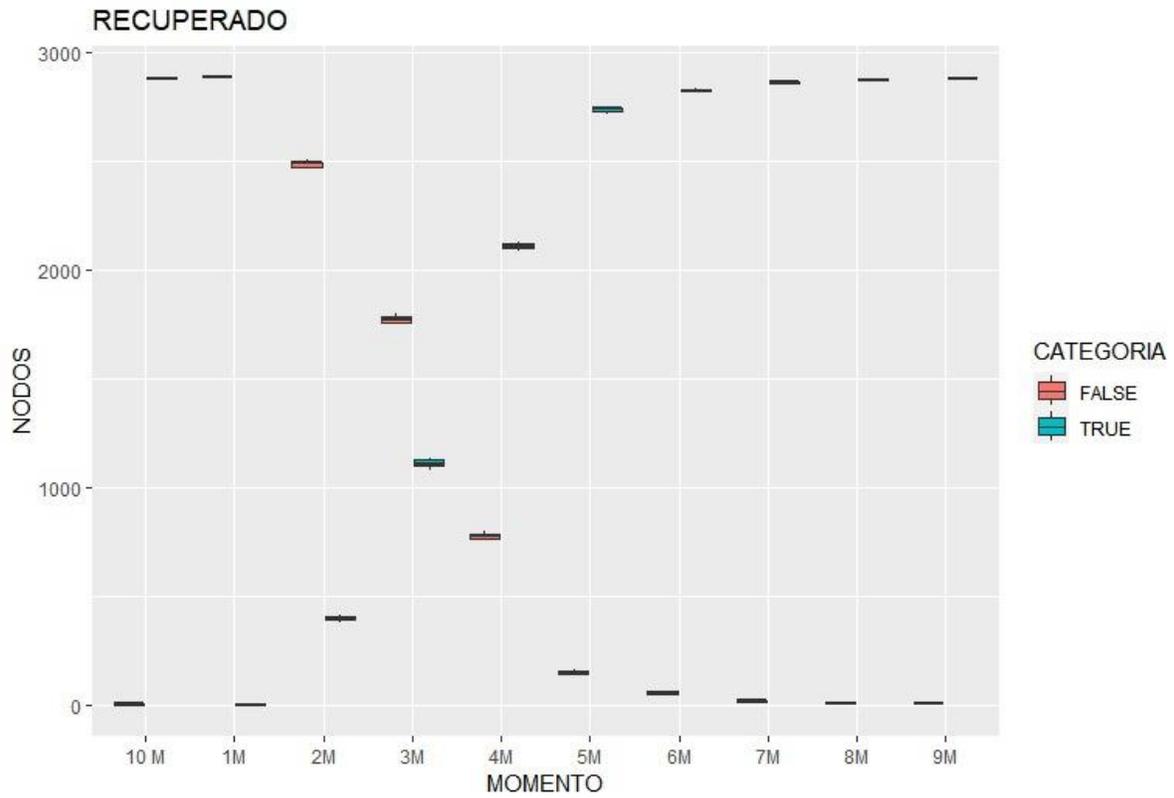


Ilustración 77. Resumen de nodos recuperados

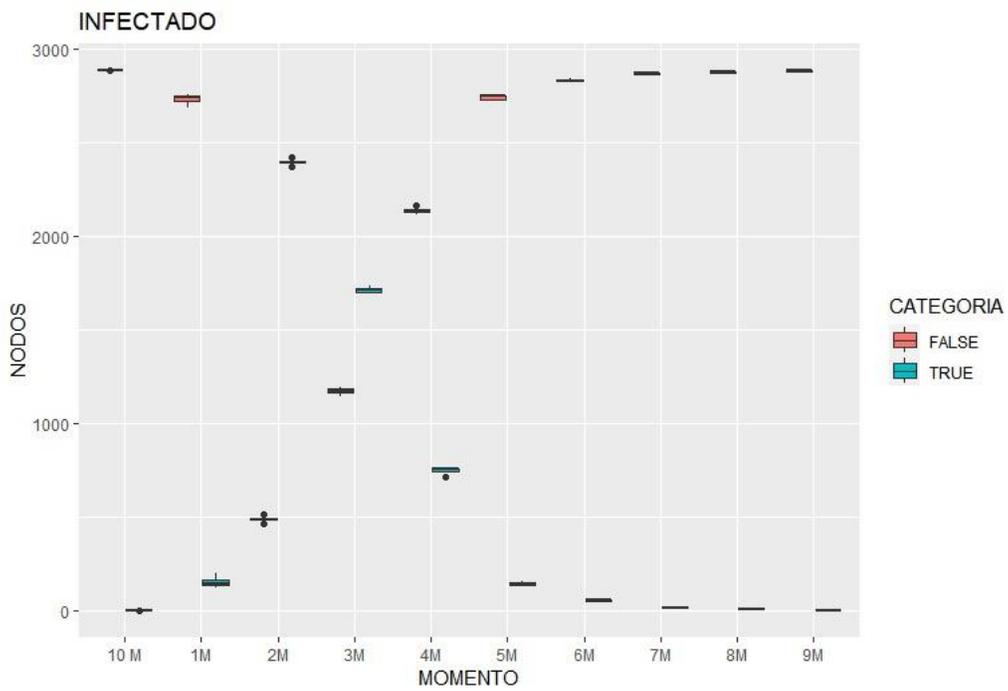


Ilustración 78. Resumen de nodos infectados (fallos)

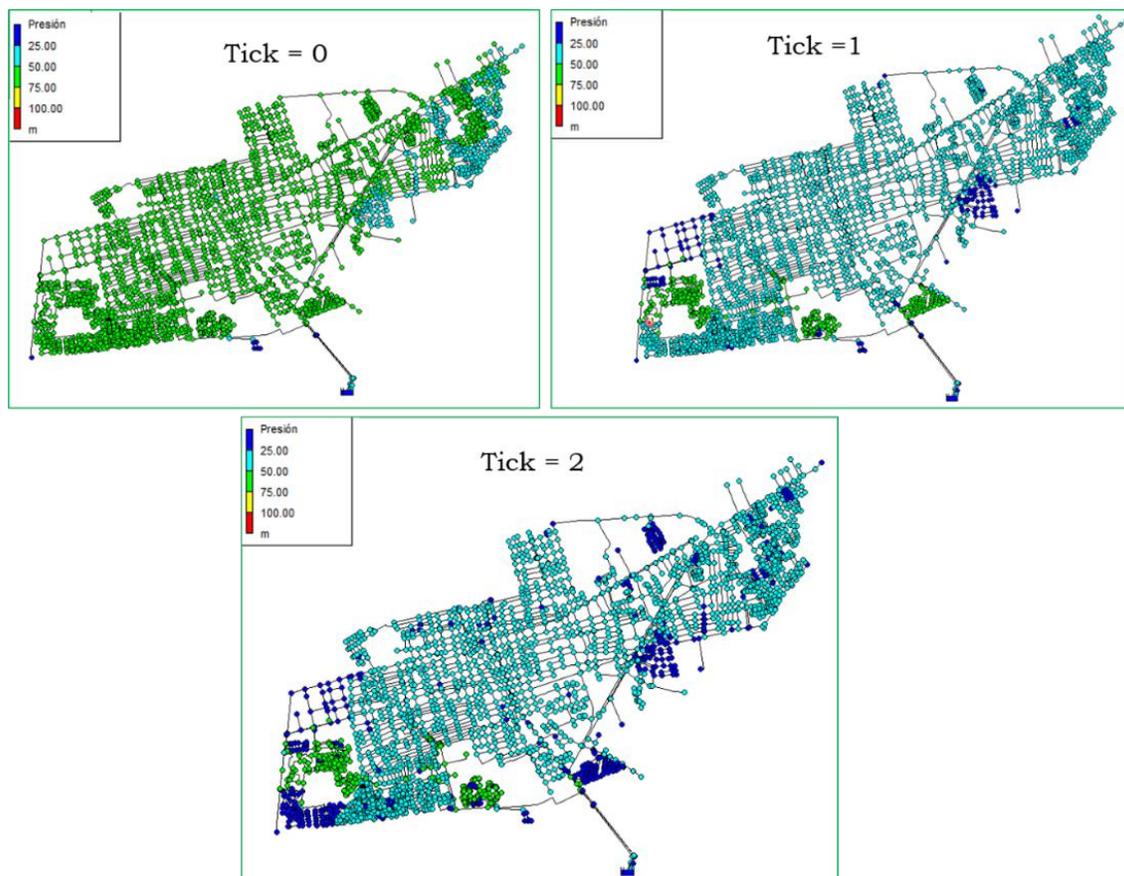


Ilustración 79. Mapas de calor de la Red de Coro

CAPÍTULO 7

CONCLUSIONES

“Este es el último capítulo del trabajo; en él se desarrolla las conclusiones que se presentan en dos apartados definidos: el primero contiene una breve discusión que incluye la idea de la metodología preventiva y el segundo contiene las conclusiones, así como unas líneas futuras de investigación. Comienza con unas consideraciones históricas y de metodología para que con los resultados obtenidos se puedan evidenciar las conclusiones. Se termina con el planteamiento de unas líneas futuras de investigación que completarían lo acá tratado.”

7.1 METODOLOGÍA PREVENTIVA

7.1.1 Consideraciones previas.

El desarrollo de esta tesis ha permitido apreciar cómo el elemento agua ha gozado de una enorme importancia en la civilización desde sus comienzos y cómo ha contribuido a la creación y crecimiento de los asentamientos urbanos desde hace miles de años. El agua potable no solo es un recurso vital para la población, sino que es esencial para la viabilidad económica de las ciudades; basta con ver el auge que tuvieron algunos centros urbanos como la ciudad de Menfis en el antiguo Egipto o la ciudad de Nippur en Sumeria.

Podría decirse que la existencia de esos asentamientos se logró en el momento en que el género humano consiguió disponer de un suministro de agua más o menos estable; además, el desarrollo de esas sociedades necesariamente se encuentra atada a las mejoras en el control del agua.

Ese control ha evolucionado desde simples sistemas de riego a elaboradas redes de suministro de agua potable, las cuales poseen características especiales únicas como lo es su complejidad, su dificultad para monitoreo o el alto riesgo al que están expuestas y que puede afectar la salud de los ciudadanos. En efecto, la calidad del agua se puede deteriorar por múltiples causas; de ahí que surja la idea de que el agua es un elemento creador de vida, pero a la vez, puede convertirse en un elemento muy peligroso para la misma.

Este pensamiento llevó a realizar diversos estudios que, desde 1850, han buscado soluciones o aproximaciones que permitan determinar cómo la distribución de agua puede afectar la salud de una comunidad, tal y como ocurrió con los brotes de cólera en Inglaterra por esas fechas. A día de hoy, esto se mantiene vigente. El interés sigue puesto, en gran medida, en la posible intrusión de contaminantes dentro de una red de suministro, bien sea por un incidente casual (rotura de tuberías) o por un ataque premeditado. Esto plantea la difícil tarea de mantener un monitoreo continuo de la calidad del agua a lo largo de la red de suministro.

Ahora bien, no solo se trata de monitorear la red en sí misma, sino de la gestión de la red de suministro, lo cual implica una gestión de demanda y una gestión de fallos que, a pesar de ser completamente independientes (una surge de un componente demográfico y la otra de un componente técnico operativo), en su conjunto dependen mutuamente para su buena realización y esto incide no solo en la oportunidad para el suministro de agua, sino también en su calidad.

Para que la calidad del agua no se deteriore y, por lo mismo, se ocasionen enfermedades contagiosas en la población que puedan llegar a situaciones epidémicas de consideración, se han desarrollado múltiples estudios con modelos matemáticos para encontrar metodologías efectivas que permitan mejorar y optimizar el control de una red de suministro de agua.

7.1.2 Orientación metodológica.

Revisando el estado del arte se recuerda que la literatura relacionada con la gestión del recurso hídrico tiene, dentro de las publicaciones de ciencias medioambientales, una participación del 37,44% en comparación con la participación que la literatura sobre epidemiología tiene en esa área (2,75%). Igualmente, cabe ver la participación que en las publicaciones de ciencias multidisciplinarias tiene la gestión del recurso hídrico (2,71%) frente a las que tiene la epidemiología en esa área (5,30%).

Las anteriores cifras ¿qué significan? Lo primero que salta a la vista es que los estudios sobre el recurso hídrico tienen un alto impacto en el manejo y cuidado del medio ambiente, mientras que en su carácter multidisciplinario es lo contrario. De otra parte, los estudios epidemiológicos tienen una baja participación frente a esas dos áreas. Lo anterior se puede entender como la aplicación metodológica de la técnica del recurso hídrico orientada al cuidado del medio ambiente y, por lo mismo, a la salud del ser humano previniendo posibles contagios. Gráficamente se puede apreciar así (*ilustración 80*):

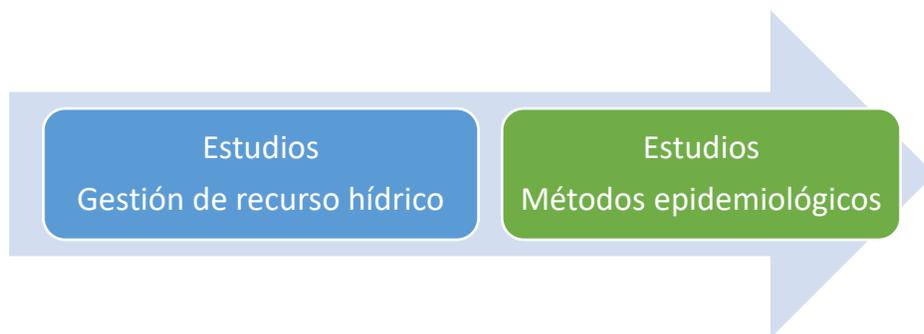


Ilustración 80. Orientación metodológica

Este trabajo propone un marco novedoso basado en la adaptación de metodologías de análisis de datos de epidemiología para hacer frente a los desafíos de la gestión de servicios públicos de agua (*ilustración 81*).

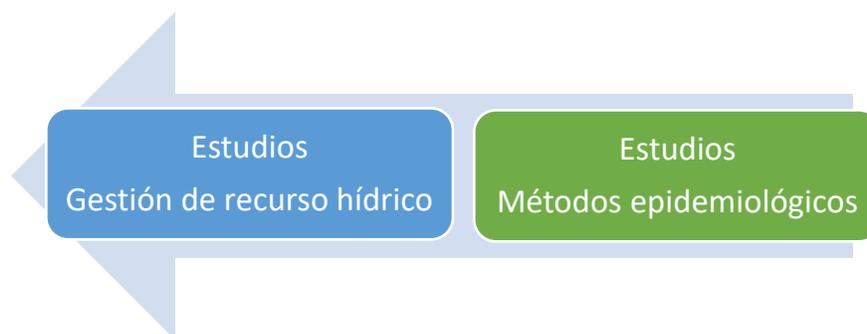


Ilustración 81. Propuesta metodológica

Para mostrar la utilidad de la propuesta, se desarrolla una metodología innovadora para potenciar los modelos predictivos de demanda de agua. Esta metodología utiliza la distancia SAX como base para encontrar DMA similares a la DMA objetivo, en la que se desarrolla el modelo predictivo.

La segunda propuesta de este trabajo es utilizar SAX para el modelado de pronóstico basado en epidemiología. En particular, SAX funciona bien como soporte para encontrar motivos y patrones sorprendentes en series temporales de demanda de agua y para abordar análisis de visualización. Todas estas características son adecuadas para trabajar bien en cuestiones de epidemiología. Sin embargo, esto también es importante en el caso de la demanda de agua, ya que SAX puede:

- extraer patrones de consumo regulares, útiles para planificar diversas operaciones en la gestión del agua potable y planes de rehabilitación posteriores;
- apoyar la detección de anomalías, un problema capaz de señalar directamente las causas de los transitorios de presión en las tuberías, o para localizar las fuentes de contaminación y su impacto en el tratamiento de efluentes;

- trabajar con series de tiempo largas y ser un marco preciso para calcular similitudes. En las aplicaciones de demanda de agua, esto permite abordar las relaciones entre los patrones de consumo en diferentes áreas sectoriales del sistema de distribución de agua. Como consecuencia, también se pueden rastrear los efluentes con mayor precisión.

El enfoque SAX para series temporales de demanda de agua se combina con árboles de sufijos, que también utilizan los investigadores del genoma humano en algoritmos para descubrir las estructuras más importantes ocultas en las secuencias biológicas. Esto es posible dada la alta capacidad de los árboles de sufijos para almacenar y comprimir una gran cantidad de datos simbólicos en tiempo lineal. Esta función es susceptible de ser utilizada para aplicaciones casi en tiempo real en series temporales más cortas, como puede ser el caso de la demanda de agua. En la propuesta actual, los árboles de sufijos respaldan la extracción de patrones de consumo y la comparación de la demanda de agua en varias DMA.

Con información de varias DMA, el análisis de este trabajo muestra cómo se pueden mejorar los modelos predictivos agregando información de bases de datos de series temporales con características similares a la DMA de destino. En aras de la claridad, se ha propuesto el uso de una sola ANN como predictor. Sin embargo, cualquier modelo predictivo puede ser adecuado para integrarse en el proceso.

SAX se puede expandir aún más para usar también cuantiles o regresiones locales para cada una de sus particiones, en lugar de usar directamente el promedio. Esto mejoraría los resultados actuales, brindando mejores capacidades de modelado para los picos de consumo de agua y para encontrar patrones de demanda aún más precisos. Dado que SAX es también un método para reducir la dimensión de la base de datos, trabajar con SAX permite extender la metodología propuesta a modelos de series temporales de gran tamaño.

7.2 CONCLUSIONES Y LÍNEAS FUTURAS

La propuesta de este trabajo es adaptar un enfoque de análisis de datos epidemiológicos como un enfoque novedoso para la hidráulica urbana. Entre la amplia gama de aplicaciones que los EDAA abren para seguir investigando, este trabajo aborda con éxito una metodología de pronóstico basada en la epidemiología para la predicción de la demanda de agua. Esto sugiere la idoneidad de las EDAA para abordar cuestiones clave de gestión del suministro de agua. Aun así, las EDAA necesitan más investigación en hidráulica urbana como un marco prometedor para tareas relacionadas con la operación y gestión de redes de agua potable.

Precisamente, el marco de gestión tiene que ver con la adecuada toma de decisiones, hecho que se puede corroborar respecto al uso de las EDAA con un ejemplo suscitado durante la época de pandemia. Los modelos epidemiológicos fueron utilizados como una de las herramientas más eficientes para asesorar a los diferentes gobiernos durante la propagación del SARS-CoV2. Aunque resulta muy positivo su uso predictivo, queda un sinsabor al intentar establecer el verdadero alcance que tiene ese tipo de modelos para evaluar los diferentes escenarios de política pública. En forma similar ocurre con las organizaciones encargadas del suministro de agua urbana, cuyas políticas no siempre dependen de los resultados de unos modelos epidemiológicos.

Volviendo al ejemplo de la pandemia, si la política pública para ese caso se fundamentara únicamente en modelos epidemiológicos, se ignoraría toda una serie de variables que esos modelos no están diseñados para analizar. Lo mismo ocurre con las organizaciones que prestan el servicio de agua urbana. Es por esto que seguir ampliando la capacidad de los modelos SAX y de los modelos basados en agentes, se constituye en un objetivo para que desarrollos posteriores puedan incluir todas esas variables y llevar a una predicción más cercana a la realidad, más precisa. El marco metodológico de la epidemiología computacional para la detección de fugas en redes de agua, podría incluso predecir las fugas antes de que ¡aparezcan!

Asimismo, el desarrollo y mejoramiento de la herramienta NetLogo le ha permitido convertirse en un elemento importante a la hora de desarrollar y ejecutar modelos basados en agentes. En el caso de estudio tratado en el capítulo 6 se pudo apreciar el comportamiento del modelo al aplicarse en una red grande con 2.887 nodos, concretamente el modelo SIR.

Las metodologías epidemiológicas que se derivan de este trabajo y que hacen referencia a la gestión de demanda y la gestión de fallos, en redes de abastecimiento de agua urbana, pueden ser heredables a otras infraestructuras brindando nuevas herramientas de análisis a redes complejas como son las redes de energía, telecomunicaciones, transporte y energías limpias, teniendo en cuenta las particularidades que conllevan cada una de ellas.

Para terminar, las metodologías trabajadas presentan enfoques simples que se orientan a la labor de toma de decisiones, identificando modelos con un nivel de interpretabilidad y explicabilidad adecuados sobre los problemas de predicción.

Las líneas futuras de investigación se centran en la unicidad. Cada área correspondiente al sistema denominado *Smart cities*, tiene un desarrollo independiente; por un lado, está el desarrollo de las redes de suministro de energía eléctrica, las de suministro de agua urbana, las de transporte, las de telecomunicaciones y por último las de energías limpias. Todas ellas manejan sus propios algoritmos con el fin de gestionar eficientemente su demanda, sus fugas y los fallos, constituyéndose en subsistemas autónomos.

El avance tecnológico implica una interacción entre esos subsistemas, lo cual permitiría apreciar un desarrollo de un modelo multiagente a gran escala que logre permitir una predicción de fallos en cascada que pueda llegar a colapsar una ciudad en su totalidad o incluso una región de varias ciudades interconectadas.

REFERENCIAS

- Abeliuk Kimelman, A. J. (2012). Árboles de Sufijo Comprimidos para Textos Altamente Repetitivos.
- Adamowski, J., Fung Chan, H., Prasher, S. O., Ozga-Zielinski, B., & Sliusarieva, A. (2012). Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada. *Water Resources Research*, 48(1).
- Adamowski, J., & Karapataki, C. (2010). Comparison of multivariate regression and artificial neural networks for peak urban water-demand forecasting: evaluation of different ANN learning algorithms. *Journal of Hydrologic Engineering*, 15(10), 729-743.
- Adamowski, J. F. (2008). Peak daily water demand forecast modeling using artificial neural networks. *Journal of Water Resources Planning and Management*, 134(2), 119-128.
- Afrin, T., & Yodo, N. (2019). A concise survey of advancements in recovery strategies for resilient complex networks. *Journal of Complex Networks*, 7(3), 393-420
- Aghabozorgi, S., Shirkhorshidi, A. S., & Wah, T. Y. (2015). Time-series clustering—a decade review. *Information systems*, 53, 16-38.
- Aghabozorgi, S., & Wah, T. Y. (2014). Clustering of large time series datasets. *Intelligent Data Analysis*, 18(5), 793-817.
- Aghapoor Khameneh, P., Miri Lavasani, S.M., Nabizadeh Nodehi, R. et al. Water distribution network failure analysis under uncertainty. *Int. J. Environ. Sci. Technol.* 17, 421–432 (2020). <https://doi-org.ezproxy.unal.edu.co/10.1007/s13762-019-02362-y>
- Ahedo García, V. A. U., Álvarez, M., Arinyo i Prats, A., Barceló, J. A., Bocanegra Barbecho, L., Bogdánovic, I., Castillo Bernal, M. F. d. (2020). *Terra Incognita: Libro blanco sobre transdisciplinariedad y nuevas formas de investigación en el Sistema Español de Ciencia y Tecnología.*
- Ajelli, M., Gonçalves, B., Balcan, D., Colizza, V., Hu, H., Ramasco, J. J., . . . Vespignani, A. (2010). Comparing large-scale computational approaches to epidemic modeling: agent-based versus structured metapopulation models. *BMC infectious diseases*, 10(1), 1-13.
- Ali, A. M., Shafiee, M. E., & Berglund, E. Z. (2017). Agent-based modeling to simulate the dynamics of urban water supply: Climate, population growth, and water shortages. *Sustainable Cities and Society*, 28, 420-434.
- Allen, L. J. (2008). An introduction to stochastic epidemic models. *Mathematical epidemiology*, 81-130.
- Alvisi, S., & Franchini, M. (2014). Assessment of the predictive uncertainty within the framework of water demand forecasting by using the model conditional processor. *Procedia Engineering*, 89, 893-900.
- Alvisi, S., Franchini, M., & Marinelli, A. (2007). A short-term, pattern-based model for water-demand forecasting. *Journal of hydroinformatics*, 9(1), 39-50.
- Angskun, T., Fagg, G., Bosilca, G., Pješivac-Grbović, J., & Dongarra, J. (2010). Self-healing network for scalable fault-tolerant runtime environments. *Future Generation Computer Systems*, 26(3), 479-485.
- Apostolico, A., Bock, M. E., & Lonardi, S. (2002). Monotony of surprise and large-scale quest for unusual words. *Proceedings of the sixth annual international conference on Computational biology.*
- AQUASTAT. (2021) Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO). Retrieved 13 de julio from <https://www.fao.org/aquastat/en/overview/methodology/water-use>

- Aramane, P., Bhattad, A., Aithal, N., & SB, P. (2019). lot and neural network based multi region and simultaneous leakage detection in pipelines. *International Journal of Advanced Research in Engineering and Technology*, 10(6).
- Arias, J. A. C. (2020). Re-Conocimiento de la utilidad de la epidemiología descriptiva. *Medicina & Laboratorio*, 19(03-04), 109-110.
- Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., . . . Benjamins, R. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information fusion*, 58, 82-115.
- Ash, J., & Newth, D. (2007). Optimizing complex networks for resilience against cascading failure. *Physica a: statistical mechanics and its applications*, 380, 673-683.
- Asnaashari, E. A. McBean, I. Shahrour, B. Gharabaghi; Prediction of watermain failure frequencies using multiple and Poisson regression. *Water Supply* 1 March 2009; 9 (1): 9–19. doi: <https://doi.org/10.2166/ws.2009.020>
- Bai, X., Xiong, Y., Zhu, Y., & Zhu, H. (2013). Time series representation: a random shifting perspective. *International Conference on Web-Age Information Management*,
- Bakker, M., Van Duist, H., Van Schagen, K., Vreeburg, J., & Rietveld, L. (2014). Improving the performance of water demand forecasting models by using weather input. *Procedia Engineering*, 70, 93-102.
- Bakker, M., Vreeburg, J., Van Schagen, K., & Rietveld, L. (2013). A fully adaptive forecasting model for short-term drinking water demand. *Environmental Modelling & Software*, 48, 141-151.
- Bakker M, Van Schagen KM and Timmer J. 2003. Flow control by prediction of water demand. *Journal of Water Supply: Research and Technology-AQUA* Vol. 52, No. 6, pp. 417-424
- Bansal, S., Pourbohloul, B., & Meyers, L. A. (2006). A comparative analysis of influenza vaccination programs. *PLoS Med*, 3(10), e387.
- Barbosa Cubillos, C. A. (2008). Implementación del método de árbol de fallas para la confiabilidad de subestaciones de alta tensión modelado en Matlab.
- Bardet, J. P., & Little, R. (2014). Epidemiology of urban water distribution systems. *Water Resources Research*, 50(8), 6447-6465.
- Bardinet, Thierry. (1995). *Les papyrus médicaux de l'Égypte pharaonique*. Traduction intégrale et commentaire, Paris, Fayard
- Barlow, R. E., & Chatterjee, P. (1973). *Introduction to fault tree analysis*. California Univ Berkeley Operations Research Center.
- Barton, Marc. (2018). John Snow y el bróte de cólera de 1854. Retrieved from <https://www.pastmedicalhistory.co.uk/john-snow-and-the-1854-cholera-outbreak/>
- Barreno Recio, P. L. (2021). Estudio de técnicas de clustering y detección de anomalías aplicado a fresadoras CNC.
- Bailey, N. T. J. (1975). *The mathematical theory of infectious diseases and its applications*. Griffin, London, 2nd edición.
- Batish, R. (2003). A new approach to the design of intermittent water supply networks. In *World Water & Environmental Resources Congress 2003* (pp. 1-11).
- Bello, O., Abu-Mahfouz, A. M., Hamam, Y., Page, P. R., Adedeji, K. B., & Piller, O. (2019). Solving management problems in water distribution networks: A survey of approaches and mathematical models. *Water*, 11(3), 562.
- Berardi, L., Ugarelli, R., Røstum, J., & Giustolisi, O. (2014). Assessing mechanical vulnerability in water distribution networks under multiple failures. *Water Resources Research*, 50(3), 2586-2599.

- Bhaskaran, K., Gasparrini, A., Hajat, S., Smeeth, L., & Armstrong, B. (2013). Time series regression studies in environmental epidemiology. *International journal of epidemiology*, 42(4), 1187-1195.
- Bialek, J., Ciapessoni, E., Cirio, D., Cotilla-Sanchez, E., Dent, C., Dobson, I., ... & Wu, D. (2016). Benchmarking and validation of cascading failure analysis tools. *IEEE Transactions on Power Systems*, 31(6), 4887-4900
- Blackwood, J. C., & Childs, L. M. (2018). An introduction to compartmental modeling for the budding infectious disease modeler.
- Blanco, R., & Maya, M. (2001). Epidemiología Básica y principios de investigación. *Corporación para Investigaciones Biológicas, Colombia*, 4, 38-63.
- Bohlin, T. P. (2006). Practical grey-box process identification: theory and applications. Springer Science & Business Media.
- Bondu, A., Boullé, M., & Grossin, B. (2013). Saxo: An optimized data-driven symbolic representation of time series. The 2013 international joint conference on neural networks (IJCNN).
- Bonet, X. (2008). Angkor Wat. Retrieved 7 de agosto from <https://megaconstrucciones.net/?construccion=angkor-wat>
- Box, G. E., & Jenkins, G. M. (1976). Time series analysis: Forecasting and control San Francisco. Calif: Holden-Day.
- Brauer, F. (2008). Compartmental models in epidemiology. *Mathematical epidemiology*, 19-79.
- Brentan, B. M., Luvizotto Jr, E., Herrera, M., Izquierdo, J., & Pérez-García, R. (2017). Hybrid regression model for near real-time urban water demand forecasting. *Journal of Computational and Applied Mathematics*, 309, 532-541.
- Brentan, B. M., Meirelles, G., Herrera, M., Luvizotto, E., & Izquierdo, J. (2017). Correlation analysis of water demand and predictive variables for short-term forecasting models. *Mathematical Problems in Engineering*, 2017.
- Britannica, T. Editores de la Enciclopedia (2015, 4 de junio). Franca. Enciclopedia Británica. <https://www.britannica.com/place/Franca>
- Britannica, T. Editores de la Enciclopedia (2014, 23 de julio). Coro. Enciclopedia Británica. <https://www.britannica.com/place/Coro>
- Brunner, N., Starkl, M., Sakthivel, P., Elango, L., Amirthalingam, S., Pratap, C. E., . . . Parimalarenganayaki, S. (2014). Policy preferences about managed aquifer recharge for securing sustainable water supply to Chennai city, India. *Water*, 6(12), 3739-3757.
- Buchberger, S. G., & Nadimpalli, G. (2004). Leak estimation in water distribution systems by statistical analysis of flow readings. *Journal of water resources planning and management*, 130(4), 321-329.
- Caballero Nuñez, J. L. (2017). Sistema de información geográfica para mejorar la gestión técnica de agua potable en la empresa municipal de agua potable y alcantarillado Emapa-Huancavelica
- Cabrera-Bejar, J. A., & Gueorguiev Tzatchkov, V. (2012). Modeling water distribution networks with intermittent water supply. *Tecnología y ciencias del agua*, 3(2), 05-25.
- Cadenas, E., Rivera, W., Campos-Amezcuca, R., & Heard, C. (2016). Wind speed prediction using a univariate ARIMA model and a multivariate NARX model. *Energies*, 9(2), 109.
- Cai, L., Li, X., Ghosh, M., & Guo, B. (2009). Stability analysis of an HIV/AIDS epidemic model with treatment. *Journal of computational and applied mathematics*, 229(1), 313-323.
- Cai B, Liu H, Xie M (2016) A real-time fault diagnosis methodology of complex systems using object-oriented Bayesian networks. *MechSyst Signal Process* 80:31–44
- Caiado, J. (2010). Performance of combined double seasonal univariate time series models for forecasting water demand. *Journal of Hydrologic Engineering*, 15(3), 215-222.

- Caicedo Londoño, M. A. (2012). Modelo estocástico para la determinación de los caudales punta de aguas residuales en edificios de viviendas.
- Campisi-Pinto, S., Adamowski, J., & Oron, G. (2012). Forecasting urban water demand via wavelet-denoising and neural network models. Case study: city of Syracuse, Italy. *Water resources management*, 26(12), 3539-3558.
- Candelieri, A. (2017). Clustering and support vector regression for water demand forecasting and anomaly detection. *Water*, 9(3), 224.
- Candelieri, A., Giordani, I., Archetti, F., Barkalov, K., Meyerov, I., Polovinkin, A., . . . Zolotykh, N. (2019). Tuning hyperparameters of a SVM-based water demand forecasting system through parallel global optimization. *Computers & Operations Research*, 106, 202-209.
- Caparrini, F. S. (2015). *Sistemas Complejos, Sistemas Dinámicos y Redes Complejas*. Retrieved 27 de Julio from <http://www.cs.us.es/~fsancho/?e=64>
- Carvalho Cantergiani, C., & Colegios, C. (2011). Modelos basados en agentes aplicados a estudios urbanos: una aproximación teórica. *Serie Geográfica*, 17, 29-43.
- Chappuis, M. (2014). Las cañerías contaminaron el agua de la antigua Roma. Retrieved 10 de agosto from <https://www.mariachappuis.com/blog/archivos/1005>
- Chau, K.-w. (2006). A review on integration of artificial intelligence into water quality modelling. *Marine pollution bulletin*, 52(7), 726-733.
- Choi, J., Yoo, D. G., & Kang, D. (2018). Post-earthquake restoration simulation model for water supply networks. *Sustainability*, 10(10), 3618.
- Christodoulou, S.E. Water Network Assessment and Reliability Analysis by Use of Survival Analysis. *Water Resour Manage* 25, 1229–1238 (2011). <https://doi.org/10.1007/s11269-010-9679-8>
- Clayton, D., & Hills, M. (2013). *Statistical models in epidemiology*. OUP Oxford.
- Cobacho, R., Ramírez, R., Torres, D., & López-Jiménez, P. (2019). Implementación de un modelo de predicción de fallos orientado a la gestión y estrategias de mantenimiento en redes de distribución de agua potable. *Ingeniería del agua*, 23(4), 247-258.
- Cockcroft J. D. & Walton E. T. S., (1932) Experiments with High Velocity Positive Ions. (I) Further Developments in the Method of Obtaining High Velocity Positive Ions, *Proceedings of the Royal Society A*, vol. 136, pp. 619–630
- Creaco, E., & Pezzinga, G. (2018). Comparison of algorithms for the optimal location of control valves for leakage reduction in WDNs. *Water*, 10(4), 466.
- Crucitti, P., Latora, V., & Marchiori, M. (2004). Model for cascading failures in complex networks. *Physical Review E*, 69(4), 045104.
- Daoudi, A., Benbrahim, M., & Benjelloun, K. (2005). An intelligent system to classify leaks in water distribution pipes. *Small*, 7(13), 20.
- De Carvalho Cantergiani, C., & Colegios, C. (2011). Modelos basados en agentes aplicados a estudios urbanos: una aproximación teórica. *Serie Geográfica*, 17, 29-43.
- Defferrard, M., Bresson, X., & Vandergheynst, P. (2016). Convolutional neural networks on graphs with fast localized spectral filtering. *Advances in neural information processing systems*, 29.
- Delgado Fernández, T., Stuart Cárdenas, M. L., & Delgado Fernández, M. (2021). Grafos de conocimiento para gestionar información epidemiológica sobre COVID-19. *Revista Cubana de Información en Ciencias de la Salud*, 32(4).
- Di Muro, M. A., La Rocca, C. E., Stanley, H. E., Havlin, S., & Braunstein, L. A. (2016). Recovery of interdependent networks. *Scientific reports*, 6(1), 1-11.
- Di Nardo, A., Di Natale, M., Musmarra, D., Santonastaso, G. F., Tzatchkov, V., & Alcocer-Yamanaka, V. H. (2015). Dual-use value of network partitioning for water system management and protection from malicious contamination. *Journal of Hydroinformatics*, 17(3), 361-376.

- Di Nardo, A., Giudicianni, C., Greco, R., Herrera, M., & Santonastaso, G. F. (2018). Applications of graph spectral techniques to water distribution network management. *Water*, 10(1), 45.
- Dobson, I., Carreras, B. A., Lynch, V. E., & Newman, D. E. (2001, January). An initial model for complex dynamics in electric power system blackouts. In *hicss*.
- Dobson, I., Carreras, B. A., & Newman, D. E. (2003, January). A probabilistic loading-dependent model of cascading failure and possible implications for blackouts. In 36th Annual Hawaii International Conference on System Sciences, 2003. Proceedings of the (pp. 10-pp). IEEE.
- Dodds, P. S., & Watts, D. J. (2004). Universal behavior in a generalized model of contagion. *Physical review letters*, 92(21), 218701.
- Donkor, E. A., Mazzuchi, T. A., Soyer, R., & Alan Roberson, J. (2014). Urban water demand forecasting: review of methods and models. *Journal of Water Resources Planning and Management*, 140(2), 146-159.
- Donoso, M. C. (2021). La ciberseguridad en la gestión inteligente del agua. *EDITORES Y COORDINADORES*, 33.
- Dridi, L., Mailhot, A., Parizeau, M., & Villeneuve, J. P. (2005, September). A strategy for optimal replacement of water pipes integrating structural and hydraulic indicators based on a statistical water pipe break model. In Proceedings of the 8th International Conference on Computing and Control for the Water Industry (pp. 65-70).
- Du Toit W. H., Fan J., & Popescu B. C. (2015). Gestión de fallos de red de distribución de energía (España Patente Nº ES 2535247T3). Oficina Española de Patentes y Marcas. <https://patents.google.com/patent/ES2535247T3/es?q=Gesti%C3%B3n+de+fallos+de+red+de+distribuci%C3%B3n+de+energ%C3%ADa&sort=new>
- Dumss, S., Grafinger, M., Hennig, M., & Rosenberger, P. (2020). Combining multiple data sources and enriching the dataset using Industrial Edge Devices. *Procedia CIRP*, 93, 1346-1351.
- Ebbell, B. (1937). The papyrus Ebers. The greatest Egyptian Medical document. Copenhagen, Levin & Munksgaard,
- Escobar, E. P., & Sánchez, D. M. (2020). Modelos estadísticos para las predicciones de la COVID-19 en Cuba. *Revista Cubana de Higiene y Epidemiología*, 57(1), 1-13.
- Espín Leal, P. (2015). Gestión de activos: sistema de gestión integral para empresas de abastecimiento de agua.
- Espinoza R., C. A. (2010). *Predicción a corto plazo de la demanda de agua urbana en áreas densamente pobladas*
- Fecarotta, O., Carravetta, A., Morani, M. C., & Padulano, R. (2018). Optimal pump scheduling for urban drainage under variable flow conditions. *Resources*, 7(4), 73.
- Felberg, R. A., Burgin, W. S., & Grotta, J. C. (2000). Neuroprotection and the ischemic cascade. *CNS spectrums*, 5(3), 52-58.
- Foundation, M. P. (2020). What is the Matrix Profile? Retrieved 3-03-2023 from <https://matrixprofile.org/#:~:text=What%20is%20the%20Matrix%20Profile,scalable%20and%20largely%20parameter%2Dfree>.
- Font, V. G., Olivella, C. G., & Pous, H. R. (2014). Seguridad en smart cities e infraestructuras críticas. In RECSI XIII: Actas de la XIII Reunión Española sobre Criptología y Seguridad de la Información. Alicante, 2-5 de septiembre de 2014 (pp. 221-226). Servicio de Publicaciones. FUNDACIÓN AQUAE (2021). La Civilización del Indo y el Saneamiento. <https://www.fundacionaquae.org/wiki/la-civilizacion-del-indo-y-el-saneamiento-en-la-antiguedad/>

- Gagliardi, F., Alvisi, S., Franchini, M., & Guidorzi, M. (2017). A comparison between pattern-based and neural network short-term water demand forecasting models. *Water Science and Technology: Water Supply*, 17(5), 1426-1435.
- Gallardo, T. (2006). Atlas of the mean motion resonances in the Solar System. *Icarus*, 184(1), 29-38.
- García Carrasco, V. J. (2019). Modelación de la demanda urbana de agua (Doctoral dissertation, Universitat Politècnica de València).
- García, J. J. (1998). Medición del riesgo en epidemiología Primera parte. *Revista Mexicana de Pediatría*, 65(2), 76-83. <http://www.medigraphic.com/pdfs/pediat/sp-1998/sp982i.pdf>
- Garzón, J. S., & Ochoa, O. A. V. (2010). Retos y desafíos de la Epidemiología. *Revista CES Salud Pública*, 1(1), 122.
- Gato, S., Jayasuriya, N., & Roberts, P. (2007a). Forecasting residential water demand: Case study. *Journal of Water Resources Planning and Management*, 133(4), 309-319.
- Gato, S., Jayasuriya, N., & Roberts, P. (2007b). Temperature and rainfall thresholds for base use urban water demand modelling. *Journal of hydrology*, 337(3-4), 364-376.
- Ghalekhondabi, I., Ardjmand, E., Young, W. A., & Weckman, G. R. (2017). Water demand forecasting: review of soft computing methods. *Environmental monitoring and assessment*, 189, 1-13.
- Gong, M., Wang, Y., Wang, S., & Liu, W. (2017). Enhancing robustness of interdependent network under recovery based on a two-layer-protection strategy. *Scientific Reports*, 7(1), 12753.
- Gortaire, D., Ayabaca, E., Borja, F., & Valarezo, B. (2016). Modelo para el pronóstico de la demanda real de agua potable en Quito. *FIGEMPA: Investigación y Desarrollo*, 1(2), 39-50.
- Goulart, Higor (2022). "Franca cai para 5 melhor saneamento básico do país; Santos lidera". Marzo 23 de 2022. <https://sampi.net.br/franca/noticias/1708345/franca/1969/12/>
- Grossi, R., & Italiano, G. F. (1993). Suffix trees and their applications in string algorithms. *Proceedings of the 1st south American workshop on string processing*,
- Guarnaccia, C., Tepedino, C., Viccione, G., & Quartieri, J. (2020). Short-term forecasting of tank water levels serving urban water distribution networks with ARIMA models. *Frontiers in Water-Energy-Nexus—Nature-Based Solutions, Advanced Technologies and Best Practices for Environmental Sustainability: Proceedings of the 2nd WaterEnergyNEXUS Conference, November 2018, Salerno, Italy*,
- Guo, G., Liu, S., Wu, Y., Li, J., Zhou, R., & Zhu, X. (2018). Short-term water demand forecast based on deep learning method. *Journal of Water Resources Planning and Management*, 144(12), 04018076.
- Gusfield, D. (1997). Algorithms on stings, trees, and sequences: Computer science and computational biology. *Acm Sigact News*, 28(4), 41-60.
- Guzmán, C. D. (2012). Programación óptima de la renovación de tuberías en un sistema de abastecimiento urbano: Análisis de los factores de influencia.
- Hamilton, I. G., Summerfield, A. J., Lowe, R., Ruysevelt, P., Elwell, C. A., & Oreszczyn, T. (2013). Energy epidemiology: a new approach to end-use energy demand research. *Building Research & Information*, 41(4), 482-497.
- Hardiman, R. C., Kumbale, M., & Makarov, Y. V. (2003, October). Multiscenario cascading failure analysis using trellis. In *CIGRE/IEEE PES International Symposium Quality and Security of Electric Power Delivery Systems, 2003. CIGRE/PES 2003*. (pp. 176-180). IEEE.
- Hashim, H., Ryan, P., & Clifford, E. (2020). A statistically based fault detection and diagnosis approach for non-residential building water distribution systems. *Advanced Engineering Informatics*, 46, 101187.
- Head, B. W. (2014). Managing urban water crises: adaptive policy responses to drought and flood in Southeast Queensland, Australia. *Ecology and Society*, 19(2).

- Helpenstein, U. (1991). The use of transfer function models, intervention analysis and related time series methods in epidemiology. *International journal of epidemiology*, 20(3), 808-815.
- Hellerstein, J. M., Koutsoupias, E., & Papadimitriou, C. H. (1997). On the analysis of indexing schemes. *Proceedings of the sixteenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems*,
- Henneaux, P., Ciapessoni, E., Cirio, D., Cotilla-Sanchez, E., Diao, R., Dobson, I., ... & Yao, R. (2018, June). Benchmarking quasi-steady state cascading outage analysis methodologies. In *2018 IEEE International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)* (pp. 1-6). IEEE.
- Herrera, M., Abraham, E., & Stoianov, I. (2016). A graph-theoretic framework for assessing the resilience of sectorised water distribution networks. *Water Resources Management*, 30(5), 1685-1699.
- Herrera, M., García-Díaz, J. C., Izquierdo, J., & Pérez-García, R. (2011). Municipal water demand forecasting: tools for intervention time series. *Stochastic Analysis and Applications*, 29(6), 998-1007.
- Herrera, M., Ferreira, A. A., Coley, D. A., & De Aquino, R. R. (2016). SAX-quantile based multiresolution approach for finding heatwave events in summer temperature time series. *AI Communications*, 29(6), 725-732.
- Herrera, M., Izquierdo, J., Pérez-García, R., & Montalvo, I. (2012). Multi-agent adaptive boosting on semi-supervised water supply clusters. *Advances in Engineering Software*, 50, 131-136.
- Herrera, M. (2011). Improving water network management by efficient division into supply clusters. A thesis: Universitat Politecnica De Valencia.
- Herrera, M., Torgo, L., Izquierdo, J., & Pérez-García, R. (2010a). Predictive models for forecasting hourly urban water demand. *Journal of hydrology*, 387(1-2), 141-150.
- Herrera, M., Izquierdo, J., Pérez-García, R., & Ayala-Cabrera, D. (2010b). Water supply clusters by multi-agent based approach. In *Water distribution systems analysis 2010* (pp. 861-869).
- Hídricos, R. (2020). Agua y Cambio Climático. In: París, UNESCO.
- Hong, S., Zhu, J., Braunstein, L. A., Zhao, T., & You, Q. (2017). Cascading failure and recovery of spatially interdependent networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2017(10), 103208.
- Hopkins, M. (2012). Critical Node Analysis for Water Distribution System Using Flow Distribution.
- House-Peters, L. A., & Chang, H. (2011). Urban water demand modeling: Review of concepts, methods, and organizing principles. *Water Resources Research*, 47(5).
- Huang, Yuzhi & Rameezdeen, Raufdeen & Chow, Christopher & Gorjian, Nima & Li, Yangyue & Liu, Zijun & Ju, Peiqing. (2021). Monitoring the health status of water mains using a scorecard modelling approach. *Water Supply*. 22. 10.2166/ws.2021.418.
- Huang, X., Vodenska, I., Havlin, S., & Stanley, H. E. (2013). Cascading failures in bi-partite graphs: model for systemic risk propagation. *Scientific reports*, 3(1), 1-9.
- Huberman, A. M., & Miles, M. B. (2000). Métodos para el manejo y el análisis de datos. *Denman CA, Haro JA (comp.). Por los rincones. Antología de métodos cualitativos en la investigación social. Hermosillo: El Colegio de Sonora*, 253-300.
- IBGE - Instituto Brasileiro de Geografia e Estatística (2022). "IBGE atualiza a listagem dos municípios que integram os recortes territoriais brasileiros | IBGE". censos.ibge.gov.br (in Brazilian Portuguese). Retrieved 2022-09-19
- Instituto de Estudios Urbanos. (2021). Crecimiento urbano y gestión del recurso hídrico en América Latina: experiencias y desafíos. *Debates de Gobierno Urbano*, 27, 46.

- Izquierdo, J., Montalvo, I., Pérez, R., & Herrera, M. (2008). Sensitivity analysis to assess the relative importance of pipes in water distribution networks. *Mathematical and Computer Modelling*, 48(1-2), 268-278.
- Jackson, M., & Chen-Charpentier, B. M. (2017). Modeling plant virus propagation with delays. *Journal of Computational and Applied Mathematics*, 309, 611-621.
- Janke, R., F. Shang, T. Haxton, J. Burkhardt, B. Burkhart. (2021). Introduction to EPANET and Example Applications. 18th Annual EPA Drinking Water Workshop, Cincinnati, OH, August 30 - September 02, 2021.
- Jeong, H. S., Qiao, J., Abraham, D. M., Lawley, M., Richard, J. P., & Yih, Y. (2006). Minimizing the consequences of intentional attack on water infrastructure. *Computer-Aided Civil and Infrastructure Engineering*, 21(2), 79-92.
- Ji, G., Wang, J., Ge, Y., & Liu, H. (2014, May). Urban water demand forecasting by LS-SVM with tuning based on elitist teaching-learning-based optimization. In *The 26th Chinese Control and Decision Conference (2014 CCDC)* (pp. 3997-4002). IEEE.
- Jithish, J., & Sankaran, S. (2017). A neuro-fuzzy approach for domestic water usage prediction. 2017 IEEE Region 10 Symposium (TENSymp),
- Jowitt, P. W., & Xu, C. (1992). Demand forecasting for water distribution systems. *Civil Engineering Systems*, 9(2), 105-121.
- Junker, N. (2020). Community Detection with Louvain and Infomap | R-bloggers.
- Kabir G, Tesfamarian S, Loepky J, Sadiq R (2015) Integrating Bayesian linear regression with ordered weighted averaging: uncertainty analysis for predicting water main failures. *ASCE-ASME J Risk Uncertain Eng Syst* 1(3):04015007
- Kandiah, V., Binder, A. R., & Berglund, E. Z. (2017). An empirical agent-based model to simulate the adoption of water reuse using the social amplification of risk framework. *Risk Analysis*, 37(10), 2005-2022.
- Kandiah, V. K., Berglund, E. Z., & Binder, A. R. (2019). An agent-based modeling approach to project adoption of water reuse and evaluate expansion plans within a sociotechnical water infrastructure system. *Sustainable Cities and Society*, 46, 101412.
- Kao, R., Danon, L., Green, D., & Kiss, I. (2006). Demographic structure and pathogen dynamics on the network of livestock movements in Great Britain. *Proceedings of the Royal Society of London B: Biological Sciences*, 273(1597), 1999-2007.
- Karamouz, M., Saadati, S., & Ahmadi, A. (2010). Vulnerability assessment and risk reduction of water supply systems. *World Environmental and Water Resources Congress 2010: Challenges of Change*,
- Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., & Fei-Fei, L. (2014). Large-scale video classification with convolutional neural networks. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition* (pp. 1725-1732).
- Keogh, E., Chakrabarti, K., Pazzani, M., & Mehrotra, S. (2001). Locally adaptive dimensionality reduction for indexing large time series databases. *Proceedings of the 2001 ACM SIGMOD international conference on Management of data*,
- Keogh, E., Lonardi, S., & Chiu, B. Y.-c. (2002). Finding surprising patterns in a time series database in linear time and space. *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*,
- Kermack, W. O., & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772), 700-721.
- Khashei, M., & Bijari, M. (2010). An artificial neural network (p, d, q) model for timeseries forecasting. *Expert Systems with applications*, 37(1), 479-489.

- Khakzad, N., & Reniers, G. (2015). Using graph theory to analyze the vulnerability of process plants in the context of cascading effects. *Reliability Engineering & System Safety*, 143, 63-73.
- Khashei, M., & Bijari, M. (2011). A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. *Applied soft computing*, 11(2), 2664-2675.
- Khorasgani, H. (2018). Model-and data-driven approaches to fault detection and isolation in complex systems. Vanderbilt University.
- Khorasgani, H., Hasanzadeh, A., Farahat, A., & Gupta, C. (2019). Fault detection and isolation in industrial networks using graph convolutional neural networks. 2019 IEEE International Conference on Prognostics and Health Management (ICPHM),
- 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. *Journal of Pipeline Systems Engineering and Practice*, 6(4), 04015005.
- Korkali, M., Veneman, J. G., Tivnan, B. F., Bagrow, J. P., & Hines, P. D. (2017). Reducing cascading failure risk by increasing infrastructure network interdependence. *Scientific reports*, 7(1), 44499.
- Kuhn, T. S. (2011). *La estructura de las revoluciones científicas*. Fondo de cultura económica.
- Kundur, P., Taylor, C., & Pourbeik, P. (2007). Blackout experiences and lessons, best practices for system dynamic performance, and the role of new technologies. IEEE Task Force Report.
- Larsen, R. J., & Marx, M. L. (2005). An introduction to mathematical statistics. Prentice Hall.
- Leick, G. (2010). The A to Z of Mesopotamia. Scarecrow Press.
- Liao, T. W. (2005). Clustering of time series data—a survey. *Pattern recognition*, 38(11), 1857-1874.
- Lin, J., Keogh, E., Wei, L., & Lonardi, S. (2007). Experiencing SAX: a novel symbolic representation of time series. *Data Mining and knowledge discovery*, 15(2), 107-144.
- Liu, C., Li, D., Zio, E., & Kang, R. (2014). A modeling framework for system restoration from cascading failures. *PloS one*, 9(12), e112363.
- Liu, J., & Zhang, T. (2011). Epidemic spreading of an SEIRS model in scale-free networks. *Communications in Nonlinear Science and Numerical Simulation*, 16(8), 3375-3384.
- Liu, T., Li, X., & Liu, X. (2010). Integration of small world networks with multi-agent systems for simulating epidemic spatiotemporal transmission. *Chinese Science Bulletin*, 55, 1285-1293.
- Lkhagva, B., Suzuki, Y., & Kawagoe, K. (2006). Extended SAX: Extension of symbolic aggregate approximation for financial time series data representation. *DEWS2006 4A-i8*, 7.
- Lloyd, A. L., & Valeika, S. (2007). Network models in epidemiology: an overview. *Complex population dynamics: nonlinear modeling in ecology, epidemiology and genetics*, 189-214.
- Lonardi, S. (2001). Global detectors of unusual words: design, implementation, and applications to pattern discovery in biosequences [Purdue University].
- López, F. (2002). La gestión de la demanda de agua como instrumento para el desarrollo sostenible de regiones áridas y semiáridas. El agua en Iberoamérica, de la escasez a la desertificación. https://www.produccionanimal.com.ar/agua_cono_sur_de_america/04-demanda_agua_como_instrumento_desarrollo.htm.
- López-Moreno, S., Garrido-Latorre, F., & Hernández-Ávila, M. (2000). Desarrollo histórico de la epidemiología: su formación como disciplina científica. *salud pública de méxico*, 42(2), 133-143.
- López-Vera, F. LA GESTIÓN DE LA DEMANDA DE AGUA COMO INSTRUMENTO PARA UN DESARROLLO SOSTENIBLE DE REGIONES ARIDAS Y SEMIARIDAS. EL AGUA EN IBEROAMÉRICA, 105.
- Luque, R. L., Calvo, I. P., Estrada, J. C. G., & Cañas, J. R. (2002). Técnicas de predicción a corto plazo de la demanda de agua. Aplicación agrícola. *Ingeniería del agua*, 9(3), 319-332.
- Macal, C. M. (2016). Everything you need to know about agent-based modelling and simulation. *Journal of Simulation*, 10(2):144–156.

- Machiwal, D., & Jha, M. K. (2012). *Hydrologic time series analysis: theory and practice*. Springer Science & Business Media.
- Madar, N., Kalisky, T., Cohen, R., Ben-avraham, D., & Havlin, S. (2004). Immunization and epidemic dynamics in complex networks. *The European Physical Journal B*, 38, 269-276.
- Maidment, D. R., Miaou, S. P., & Crawford, M. M. (1985). Transfer function models of daily urban water use. *Water Resources Research*, 21(4), 425-432.
- Mailhot, A., Pelletier, G., Noël, J. F., & Villeneuve, J. P. (2000). Modeling the evolution of the structural state of water pipe networks with brief recorded pipe break histories: Methodology and application. *Water Resources Research*, 36(10), 3053-3062.
- Makaya, E., & Hensel, O. (2015). Modelling flow dynamics in water distribution networks using artificial neural networks-A leakage detection technique. *International Journal of Engineering, Science and Technology*, 7(1), 33-43.
- Manzano Castro, M., Calle Ortega, E., Ripoll i Misse, J., Manolova, A., Fagertun, V., Torres Padrosa, V., ... & Scoglio, C. (2014). Epidemic and Cascading Survivability of Complex Networks. In © *Reliable Networks Design and Modeling (RNDM)*, 2014 6th International Workshop on, 2014, p. 187-193. Institute of Electrical and Electronics Engineers (IEEE).
- Márquez Grajales, A. (2019). *Discretización de series temporales usando un algoritmo multi-objetivo y árboles de decisión*.
- Martínez-Álvarez, F. (2011). Clustering preprocessing to improve time series forecasting. *Ai Communications*, 24(1), 97-98.
- Martínez Codina, Á. (2015). *Metodología de análisis estadístico de roturas en redes de distribución de agua* (Doctoral dissertation, Caminos).
- McCraith, E. M. (1976). A space-economical suffix tree construction algorithm. *Journal of the ACM (JACM)*, 23(2), 262-272.
- McKibben, J., & Davis, D. (2002). Integration of GIS with Computerized Maintenance Management Systems (CMMS) and Asset Management Systems. URL: <http://proceedings.esri.com/library/userconf/proc02/pap0554/p0554.htm> Search in.
- McLaurin, E., McDonald, A. D., Lee, J. D., Aksan, N., Dawson, J., Tippin, J., & Rizzo, M. (2014). Variations on a theme: Topic modeling of naturalistic driving data. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*,
- Menichella, M., Paolicchi, P., & Farinella, P. (1996). The main belt as a source of near-Earth asteroids. In *Worlds in Interaction: Small Bodies and Planets of the Solar System* (pp. 133-149). Springer.
- Maslia, M. L., Sautner, J. B., Aral, M. M., Reyes, J. J., Abraham, J. E., & Williams, R. C. (2000). Using water-distribution system modeling to assist epidemiologic investigations. *Journal of Water Resources Planning and Management*, 126(4), 180-198.
- Mazhar, M. A., Khan, N. A., Ahmed, S., Khan, A. H., Hussain, A., Changani, F., ... & Vambol, V. (2020). Chlorination disinfection by-products in municipal drinking water—a review. *Journal of Cleaner Production*, 273, 123159.
- Mei, S., van de Vijver, D., Xuan, L., Zhu, Y., & Sloot, P. M. (2010). Quantitatively evaluating interventions in the influenza A (H1N1) epidemic on China campus grounded on individual-based simulations. *Procedia Computer Science*, 1(1), 1675-1682.
- Molino, B., Rasulo, G., & Tagliatela, L. (1996). Forecast model of water consumption for Naples. *Water resources management*, 10(4), 321-332.
- Molle, F. (2009). River-basin planning and management: The social life of a concept. *Geoforum*, 40(3), 484-494.
- Mora-Rodríguez, J., Ramos, H. M., & López-Jiménez, P. A. (2013). Eventos de intrusión patógena en sistemas de distribución de agua potable. *Tecnología y ciencias del agua*, 4(3), 5-25.

- Mouatadid, S., & Adamowski, J. (2017). Using extreme learning machines for short-term urban water demand forecasting. *Urban water journal*, 14(6), 630-638.
- Mu, L., Zheng, F., Tao, R., Zhang, Q., & Kapelan, Z. (2020). Hourly and daily urban water demand predictions using a long short-term memory based model. *Journal of Water Resources Planning and Management*, 146(9), 05020017.
- Muloiwa, M., Dinka, M., & Nyende-Byakika, S. (2022). Analysis of domestic water consumption in peri-urban South Africa: The case study of Thohoyandou in Limpopo province, South Africa. *African Journal of Science, Technology, Innovation and Development*, 14(6), 1546-1559.
- Muhammad, A. U., Li, X., & Feng, J. (2019). Artificial intelligence approaches for urban water demand forecasting: A review. *International Conference on Machine Learning and Intelligent Communications*.
- Msiza, I. S., Nelwamondo, F. V., & Marwala, T. (2008). Water demand prediction using artificial neural networks and support vector regression.
- Nasser, A. A., Rashad, M. Z., & Hussein, S. E. (2020). A two-layer water demand prediction system in urban areas based on micro-services and LSTM neural networks. *IEEE Access*, 8, 147647-147661.
- Navarrete-López, C., Herrera, M., Brentan, B. M., Luvizotto Jr, E., & Izquierdo, J. (2019). Enhanced water demand analysis via symbolic approximation within an epidemiology-based forecasting framework. *Water*, 11(2), 246.
- Navarrete López, C. F. (2017) Análisis inteligente de datos aplicados al estudio y propagación de contaminantes en redes de abastecimiento de agua urbana.
- Navarrete-López, C., Calderón Rivera, D., Díaz Arévalo, J. L., Herrera, M., & Izquierdo, J. (2017). Modelos geoestadísticos para la predicción de fallos de una zona de la red de abastecimiento de agua de Bogotá, integrando algoritmos de Machine Learning (Geostatistical Models for the Prediction of Water Supply Network Failures in Bogotá, Integrating Machine Learning Algorithms). *Ibero-American Seminar on Water and Drainage Networks (SEREA 2017)*
- Nel, D., & Haarhoff, J. (1996). Sizing municipal water storage tanks with Monte Carlo simulation. *Aqua- Journal of Water Supply: Research and Technology*, 45(4), 203-212.
- Newman, M. (2018). *Networks*, volumen 1. Oxford University Press. doi:10.1093/oso/9780198805090.001.0001.
- Nhapi, I., Holch, W., Mazvimavi, D., Mashauri, D., Jewitt, G., Mudege, N., . . . Beukman, R. (2005). Integrated water resources management (IWRM) and the millennium development goals: Managing water for peace and prosperity. *Physics and Chemistry of the Earth*, 30(11-16), 623-624.
- Nguyen, K. A., Stewart, R. A., Zhang, H., Sahin, O., & Siriwardene, N. (2018). Re-engineering traditional urban water management practices with smart metering and informatics. *Environmental modelling & software*, 101, 256-267.
- Odan, F. K., & Reis, L. F. R. (2012). Hybrid water demand forecasting model associating artificial neural network with Fourier series. *Journal of Water Resources Planning and Management*, 138(3), 245-256.
- Ojeda de la Cruz, A., Narváez Tijerina, A. B., & Quintana Pacheco, J. (2014). Management of Urban Residential Water in Hermosillo (Sonora, Mexico). *Cuadernos de Geografía: Revista Colombiana de Geografía*, 23(1), 147-164.
- Organization, U. N. C. s. F. a. W. H. (2019). Progress on household drinking water, sanitation and hygiene 2000–2017: special focus on inequalities. In: *United Nations Children's Fund (UNICEF) and World Health Organization New York*.

- Ortega, M. d. I. C. (2014). *Reparación, rehabilitación y renovación de redes*. . Retrieved 27 de julio from <https://www.interempresas.net/Agua/Articulos/121692-Reparacion-rehabilitacion-y-renovacion-de-redes.html>
- Pacchin, E., Alvisi, S., & Franchini, M. (2017). A short-term water demand forecasting model using a moving window on previously observed data. *Water*, 9(3), 172.
- Pacchin, E., Gagliardi, F., Alvisi, S., & Franchini, M. (2019). A comparison of short-term water demand forecasting models. *Water resources management*, 33, 1481-1497.
- Pahwa, S., Scoglio, C., & Scala, A. (2014). Abruptness of cascade failures in power grids. *Scientific reports*, 4(1), 1-9.
- Palacios Ochoa, L. J., & Trelles Cabrera, F. D. (2018). *Diseño de un prototipo de sistema SCADA para el monitoreo y control de consumo de agua en viviendas residenciales* (Bachelor's thesis).
- Panagopoulos, G. P., Bathrellos, G. D., Skilodimou, H. D., & Martsouka, F. A. (2012). Mapping urban water demands using multi-criteria analysis and GIS. *Water Resources Management*, 26, 1347-1363.
- Pandolfi, D., Alba Torres, E., Villagra, A., & Leguizamón, G. (2018). Redes inteligentes de agua: factores y métodos para la predicción del consumo residencial de agua potable. XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste).
- Parida B P, Moalafhi D B and Kenabatho P. K. (2006). Forecasting runoff coefficients using ANN for water resources management: The case of Notwane catchment in Eastern Botswana, *Physics and Chemistry of the Earth Vol. 31*, pp. 928-934
- Pastor-Satorras, R., Castellano, C., Van Mieghem, P., & Vespignani, A. (2015). Epidemic processes in complex networks. *Reviews of modern physics*, 87(3), 925.
- Pedersén, O., Sinclair, P., Hein, I., & Andersson, J. (2010). Cities and urban landscapes in the Ancient Near East and Egypt with special focus on the city of Babylon.
- Pesantez, J. E., Berglund, E. Z., & Kaza, N. (2020). Smart meters data for modeling and forecasting water demand at the user-level. *Environmental Modelling & Software*, 125, 104633.
- Pedregal Mateos, B. (2004). Estimación de la demanda de agua urbana en los planes marco: el caso español a la luz de la experiencia del Plan Hidrológico de California. *Boletín de la Asociación de Geógrafos Españoles*, 37, 183-204.
- Pereira Martín, M., Blesa, J., Puig, V., & Muñoz De la Peña, D. (2019). Detección de fallos en redes de presión pertenecientes a redes de suministro de agua mediante filtros estáticos lineales basados en PCA. In *XL Jornadas de Automática* (pp. 568-573). Universidade da Coruña, Servizo de Publicacións.
- Perkins, S. E., Cagnacci, F., Stradiotto, A., Arnoldi, D., & Hudson, P. J. (2009). Comparison of social networks derived from ecological data: implications for inferring infectious disease dynamics. *Journal of Animal Ecology*, 78(5), 1015-1022.
- Pescaroli, G., & Alexander, D. (2016). Critical infrastructure, panarchies and the vulnerability paths of cascading disasters. *Natural Hazards*, 82, 175-192.
- Pietrucha-Urbanik, K. (2015). Failure Prediction in Water Supply System – Current Issues. In: Zamojski, W., Mazurkiewicz, J., Sugier, J., Walkowiak, T., Kacprzyk, J. (eds) *Theory and Engineering of Complex Systems and Dependability*. DepCoS-RELCOMEX 2015. *Advances in Intelligent Systems and Computing*, vol 365. Springer, Cham. https://doi.org/10.1007/978-3-319-19216-1_33
- Pintelas, E., Livieris, I. E., & Pintelas, P. (2020). A grey-box ensemble model exploiting black-box accuracy and white-box intrinsic interpretability. *Algorithms*, 13(1), 17.
- Piqué, T., & Cejalvo, A. (1992). NTP 333: Análisis probabilístico de riesgos: Metodología del Árbol de fallos y errores. Instituto Nacional de Seguridad e Higiene En El Trabajo, 9

- Plietzsch, A., Schultz, P., Heitzig, J., & Kurths, J. (2016). Local vs. global redundancy–trade-offs between resilience against cascading failures and frequency stability. *The European Physical Journal Special Topics*, 225, 551-568.
- Polo, J. P., Candezano, M. A. C., & Núñez, L. N. (2020). Dos enfoques matemáticos epidemiológicos para modelar el comportamiento de los decesos causados por el COVID-19. *Investigación e Innovación en Ingenierías*, 8(2), 75-86.
- Porta, M. E. (2008). *A Dictionary of Epidemiology* (5th ed.). Oxford University Press.
- Postel, S. (1993). El último oasis: cómo afrontar la escasez de agua. Apóstrofe.
- Puust, R., Kapelan, Z., Savic, D., & Koppel, T. (2010). A review of methods for leakage management in pipe networks. *Urban Water Journal*, 7(1), 25-45.
- Quattrociocchi, W., Caldarelli, G., & Scala, A. (2014). Self-healing networks: redundancy and structure. *PloS one*, 9(2), e87986.
- Quevedo, J., Puig, V., Cembrano, G., Blanch, J., Aguilar, J., Saporta, D., . . . Molina, A. (2010). Validation and reconstruction of flow meter data in the Barcelona water distribution network. *Control Engineering Practice*, 18(6), 640-651.
- Rajani, B., & Kleiner, Y. (2001). Comprehensive review of structural deterioration of water mains: physically based models. *Urban water*, 3(3), 151-164.
- Rajani, B., Kleiner, Y., & Sink, J. E. (2012). Exploration of the relationship between water main breaks and temperature covariates. *Urban Water Journal*, 9(2), 67-84.
- Rajaraman, A., & Ullman, J. D. (2011). *Mining of massive datasets*. Cambridge University Press.
- Rasheed, F., Alshalalfa, M., & Alhaji, R. (2010). Efficient periodicity mining in time series databases using suffix trees. *IEEE Transactions on Knowledge and Data Engineering*, 23(1), 79-94.
- Ramsay MA. (2006) John Snow, MD: anaesthetist to the Queen of England and pioneer epidemiologist. *Proc (Bayl Univ Med Cent)*. 2006 Jan;19(1):24-8.
- Rechy-Ramírez, F., Mesa, H.-G. A., Mezura-Montes, E., & Cruz-Ramírez, N. (2011). Times series discretization using evolutionary programming. *Mexican International Conference on Artificial Intelligence*,
- Reynoso, C. (2008). Hacia la complejidad por la vía de las redes: nuevas lecciones epistemológicas. *Desacatos* (28), 17-40.
- Riera, G. E. (2019). El derecho al agua y su relación con el Apartheid. *Espacios transnacionales: revista latinoamericana-europea de pensamiento y acción social*, 7(13), 12-21.
- Rinaldi, S. M., Peerenboom, J. P., & Kelly, T. K. (2001). Identifying, understanding, and analyzing critical infrastructure interdependencies. *IEEE control systems magazine*, 21(6), 11-25.
- Rivera Muñoz, D. S. (2021). Inversión en redes de transmisión considerando fallas en cascada y seguridad probabilística.
- Riza, L. S., Fazanadi, M. N., Utama, J. A., Samah, K. A. F. A., Hidayat, T., & Nazir, S. (2022). SAX and Random Projection Algorithms for the Motif Discovery of Orbital Asteroid Resonance Using Big Data Platforms. *Sensors*, 22(14), 5071.
- Rodà de Llanza, I. (2014). Urbanisme i monumentalització: Roma i les ciutats a Catalunya. *Urbanisme i monumentalització*, 1-168.
- Rodolfo, R. G. (2015). La reinención de la epidemiología a la luz de las nuevas tecnologías / The Reinvention of Epidemiology in the Light of New Technologies / A reinvenção da epidemiologia à luz das novas tecnologias [research article]. *Revista Ciências de la Salud* (2), 283. <https://doi.org/10.12804/revsalud13.02.2015.13>
- Rodríguez, J. B. M. (2003). El Concepto de garantía en las redes de suministro de agua potable. *Ingeniería Hidráulica y Ambiental*, 24(3), 49-55
- Rogers, P. D. (2011). Prioritizing water main renewals: case study of the Denver water system. *Journal of Pipeline Systems Engineering and Practice*, 2(3), 73-81.

- Rojas, R. (2013). *Neural networks: a systematic introduction*. Springer Science & Business Media.
- Romano, M., & Kapelan, Z. (2014). Adaptive water demand forecasting for near real-time management of smart water distribution systems. *Environmental Modelling & Software*, 60, 265-276.
- Rossman, L. (2000). *EPANET User's manual*. Cincinnati. Environmental Protection Agency
- Rossman, L. A. (1994). *EPANET User's manual*. Cincinnati. Environmental Protection Agency
- Ruefenacht, H., & Guibentif, H. (1997). A model for forecasting water consumption in Geneva canton, Switzerland. *Aqua- Journal of Water Supply: Research and Technology*, 46(4), 196-201.
- Ruiz Ruano García, A. M., & López Puga, J. (2020). Modelos gráficos y redes en Psicología.
- SABESP - Companhia de Saneamento Básico do Estado de São Paulo (2023). SABESP en el municipio de Franca. (sabes.com.br) Recuperado el 25 03 2023 de <https://site.sabesp.com.br/site/interna/Municipio.aspx?secaold=18&id=230>
- Sala, Laura. (18 de junio de 2018). La Historia del agua. Fundación Aquea. <https://www.fundacionaquea.org/wiki/la-historia-del-agua/>
- Salam, A. E. U., Tola, M., Selintung, M., & Maricar, F. (2014). On-line monitoring system of water leakage detection in pipe networks with artificial intelligence. *ARNP Journal of Engineering and Applied Sciences*, 9(10), 1817-1822.
- Sánchez Llorens, M., & Fernández Alonso, F. (2017). Las maneras de fluir del agua visible e invisible en las ciudades Implicaciones de la presencia del agua en el futuro urbano. In: España: Universidad Politécnica de Madrid.
- Santos, C. A. (2019). Paradigmas de la gestión integrada del agua (I): una crítica evolutiva a la gestión integrada de los recursos hídricos (GIRH). *Revista Científica ECOCIENCIA*, 6(2), 1-21.
- Santos, J.R., Haimes, Y.Y., 2004 (2004). Modeling the demand reduction input-output (I-O) inoperability due to terrorism of interconnected infrastructures. *Risk Anal.: Int. J.* 24, 1437–1451
- Saporta, D., & Muñoz, M. (1994). El consumo en redes de distribución. Predicción diaria de la demanda. *Mejora del rendimiento y de la fiabilidad en sistemas de distribución de agua. Aguas de Valencia y UD Mecánica de Fluidos (UPV)*, 2, 43-75.
- Scarpa, F., Lobba, A., & Becciu, G. (2016). Elementary DMA design of looped water distribution networks with multiple sources. *Journal of Water Resources Planning and Management*, 142(6), 04016011.
- Schäfer, B., Witthaut, D., Timme, M., & Latora, V. (2018). Dynamically induced cascading failures in power grids. *Nature communications*, 9(1), 1975.
- Schrader-King, K. (2020). Desarrollo urbano. Retrieved 20 Octubre from <https://www.bancomundial.org/es/topic/urbandevlopment/overview>
- Sebri, M. (2016). Forecasting urban water demand: A meta-regression analysis. *Journal of environmental management*, 183, 777-785.
- Shafiee, M. E., & Berglund, E. Z. (2017). Complex adaptive systems framework to simulate the performance of hydrant flushing rules and broadcasts during a water distribution system contamination event. *Journal of Water Resources Planning and Management*, 143(4), 04017001.
- Shafiee, M. E., Berglund, E. Z., & Lindell, M. K. (2018). An agent-based modeling framework for assessing the public health protection of water advisories. *Water resources management*, 32(6), 2033-2059.
- Shi, Y. (2003). Stochastic dynamic model of SARS spreading. *Chinese Science Bulletin*, 48, 1287-1292.
- Shirkooji, M. G., Doghri, M., & Duchesne, S. (2021). Short-term water demand predictions coupling an artificial neural network model and a genetic algorithm. *Water Supply*, 21(5), 2374-2386.

- Shuang, Q., Zhang, M., & Yuan, Y. (2014). Node vulnerability of water distribution networks under cascading failures. *Reliability Engineering & System Safety*, 124, 132-141.
- Shuang, Q., Liu, Y., Tang, Y., Liu, J., & Shuang, K. (2017). System reliability evaluation in water distribution networks with the impact of valves experiencing cascading failures. *Water*, 9(6), 413.
- Silva, S., Fraga, S., & Lunet, N. (2012). *Qualitative research in epidemiology*. INTECH Open Access Publisher.
- Sillero, J. P. (2021). *La gestión participativa del agua - un futuro sostenible y seguro*. Retrieved 13 de agosto from <https://www.iucncongress2020.org/es/noticias/todos-noticias/la-gestion-participativa-del-agua-un-futuro-sostenible-y-seguro>
- Simonsen, I., Buzna, L., Peters, K., Bornholdt, S., & Helbing, D. (2008). Transient dynamics increasing network vulnerability to cascading failures. *Physical review letters*, 100(21), 218701.
- Snider, B., & McBean, E. A. (2020). Improving urban water security through pipe-break prediction models: Machine learning or survival analysis. *Journal of Environmental Engineering*, 146(3), 04019129.
- Solano H., H. (2008). *Análisis de supervivencia en fiabilidad. Predicción en condiciones de alta censura y truncamiento: el caso de las redes de suministro de agua potable*
- Soldi, D., Candelieri, A., & Archetti, F. (2015). Resilience and vulnerability in urban water distribution networks through network theory and hydraulic simulation. *Procedia Engineering*, 119, 1259-1268.
- Souza Groppo, G., Costa, M. A., & Libânio, M. (2019). Predicting water demand: A review of the methods employed and future possibilities. *Water Supply*, 19(8), 2179-2198.
- Suhartono, S., Isnawati, S., Salehah, N. A., Prastyo, D. D., Kuswanto, H., & Lee, M. H. (2018). Hybrid SSA-TSR-ARIMA for water demand forecasting. *International Journal of Advances in Intelligent Informatics*, 4(3), 238-250.
- Sun, P. G., & Ma, X. (2017). Controllability and observability of cascading failure networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2017(4), 043404.
- Tanaka, H., Fan, L. T., Lai, F. S., & Toguchi, K. (1983). Fault-tree analysis by fuzzy probability. *IEEE Transactions on reliability*, 32(5), 453-457.
- Taormina, R., Galelli, S., Tippenhauer, N. O., Salomons, E., Ostfeld, A., Eliades, D. G., . . . Banks, M. K. (2018). Battle of the attack detection algorithms: Disclosing cyber attacks on water distribution networks. *Journal of Water Resources Planning and Management*, 144(8), 04018048.
- Tian, D., Martinez, C. J., & Asefa, T. (2016). Improving short-term urban water demand forecasts with reforecast analog ensembles. *Journal of Water Resources Planning and Management*, 142(6), 04016008.
- Tisue, S., & Wilensky, U. (2004, May). Netlogo: A simple environment for modeling complexity. In *International conference on complex systems* (Vol. 21, pp. 16-21).
- Tiwari, M., Adamowski, J., & Adamowski, K. (2016). Water demand forecasting using extreme learning machines. *Journal of Water and Land Development*, 28(1), 37.
- Toshniwal, D. (2009). Feature extraction from time series data. *Journal of Computational Methods in Sciences and Engineering*, 9(s1), S99-S110.
- Tucci, Carlos E. M. Águas urbanas. *Estudos Avançados* [online]. 2008, v. 22, n. 63, pp. 97-112. Disponível em: <https://doi.org/10.1590/S0103-40142008000200007>
- Tzatchkov, V., & Cabrera-Bejar, J. (2008). Una forma de modelar el suministro intermitente de una red de distribución de agua potable. *Proc VIII Seminario Ibero-americano SEREA*.
- Tzatchkov, V., & Izurieta Dávila, J. (1996). Sistema de cómputo para el análisis y diseño óptimo de redes de distribución de agua potable.

- UN-Water, FAO. (2007) Coping with water scarcity. Challenge of the twenty-first century, 2007
- Umoh, E. A., Umar, M., & Umar, M. N. (2022). Chaos Theory Applied to Cascading Disaster Dynamics, Modelling and Control. 2022 IEEE Nigeria 4th International Conference on Disruptive Technologies for Sustainable Development (NIGERCON),
- Vammen, K. (2015). Desafios del agua urbana en las Americas: Perspectivas de las academias de ciencias. *Agricultura, sociedad y desarrollo*, 12(3), 475-478.
- Välimäki, N., Gerlach, W., Dixit, K., & Mäkinen, V. (2007). Compressed suffix tree—a basis for genome-scale sequence analysis. *Bioinformatics*, 23(5), 629-630.
- Vargas Ccahuana, J., & Huyhua Montes, H. (2020). Ampliación y mejoramiento integral de los sistemas de agua potable y desague en las comunidades de Ccollotaro y Ccoyaraqui del distrito de Caicay, provincia de Paucartambo, región Cusco.
- Vázquez Alcalá, E. F. (2022). Simulación de epidemias bajo medidas de contingencia (Doctoral dissertation, Universidad Autónoma de Nuevo León)
- Viccione, G., Guarnaccia, C., Mancini, S., & Quartieri, J. (2020). On the use of ARIMA models for short-term water tank levels forecasting. *Water Supply*, 20(3), 787-799.
- Vijayalaksmi, D., & Babu, K. J. (2015). Water supply system demand forecasting using adaptive neuro-fuzzy inference system. *Aquatic Procedia*, 4, 950-956
- Wang, H., Rong, F., Ke, F., & Bai, Y. (2003). Control dynamics of severe acute respiratory syndrome transmission. *Chinese Science Bulletin*, 48, 2684-2687.
- Winkler, D., Haltmeier, M., Kleidorfer, M., Rauch, W., & Tscheikner-Gratl, F. (2018). Pipe failure modelling for water distribution networks using boosted decision trees. *Structure and Infrastructure Engineering*, 14(10), 1402-1411.
- Yi, B.-K., & Faloutsos, C. (2000). Fast time sequence indexing for arbitrary Lp norms.
- Yuan, J., Wang, Z., Han, M., & Sun, Y. (2015). A lazy associative classifier for time series. *Intelligent Data Analysis*, 19(5), 983-1002.
- Zabaleta, A. P. (2015). *Gestión del agua en una Smart City*. Retrieved 23 de julio de 2022 from
- Zarza, i. L. F. (2018). *WatEner, una mayor eficiencia en la gestión de redes de abastecimiento* Retrieved 21 de Julio de 2022 from <https://www.iagua.es/noticias/grupo-inclam/watener-mayor-eficiencia-gestion-redes-abastecimiento>
- Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159-175.
- Zhang, Y., Li, Z., Aviso, K. B., Jia, X. X., Zhang, P., Tan, R. R., ... & Jia, X. (2022). Industrial water network vulnerability analysis using dynamic inoperability input-output model. *Journal of Environmental Management*, 314, 115015.
- Zubaidi, S. L., Al-Bugharbee, H., Ortega-Martorell, S., Gharghan, S. K., Olier, I., Hashim, K. S., . . . Kot, P. (2020). A novel methodology for prediction urban water demand by wavelet denoising and adaptive neuro-fuzzy inference system approach. *Water*, 12(6), 1628.
- Zhou, L., Xia, J., Yu, L., Wang, Y., Shi, Y., Cai, S., & Nie, S. (2016). Using a hybrid model to forecast the prevalence of schistosomiasis in humans. *International journal of environmental research and public health*, 13(4), 355.
- Zhou, S., McMahon, T., Walton, A., & Lewis, J. (2002). Forecasting operational demand for an urban water supply zone. *Journal of hydrology*, 259(1-4), 189-202.
- Zio, E., & Sansavini, G. (2011). Component criticality in failure cascade processes of network systems. *Risk Analysis: An International Journal*, 31(8), 1196-1210.
- Zubaidi, S. L., Ortega-Martorell, S., Al-Bugharbee, H., Olier, I., Hashim, K. S., Gharghan, S. K., . . . Al-Khaddar, R. (2020). Urban water demand prediction for a city that suffers from climate change and population growth: Gauteng province case study. *Water*, 12(7), 1885.

Zubaidi, S. L., Dooley, J., Alkhaddar, R. M., Abdellatif, M., Al-Bugharbee, H., & Ortega-Martorell, S. (2018). A Novel approach for predicting monthly water demand by combining singular spectrum analysis with neural networks. *Journal of hydrology*, 561, 136-145.

ANEXOS

La información que se anexa, recursos de la tesis y algunos enlaces de interés, se encuentran depositados en el siguiente website: clafenalo.byethost9.com.

Anexo 1

Código fuente del modelo SIR para el caso de estudio de la red de Coro.

El código desarrollado y ejecutado en NetLogo es el siguiente:

```
extensions [vid]

globals [
  num-valves      ;;
  colors          ;;
  source-colors  ;;
  init-colors    ;;
  tmp-colors     ;;
  mains-min-size ;;
  nb-clusters    ;;
  csv fileList  ;;
]

;; Los nodos representan empalmes, tanques y reservorios.

breed [junctions junction] ;; consumidores (nodos)
junctions-own [
  node_id
  elevation
  junction_demand
  junction_pattern
  junction_xcor
  junction_ycor
  Presion1
  presion2
  Status
  infected?      ;; if true, the turtle is infectious
  resistant?    ;; if true, the turtle can't be infected
  susceptible?  ;; if true, the turtle is susceptible
  virus-check-timer ;; number of ticks since this turtle's last
  virus-check
]
```

```

breed [reservoirs reservoir]
reservoirs-own [
  node_id
  elevation
  reservoir_head
  reservoir_xcor
  reservoir_ycor
]

breed [tanks tank]
tanks-own [
  node_id
  elevation
  tank_initLevel
  tank_minLevel
  tank_maxLevel
  tank_diameter
  tank_minVol
  tank_xcor
  tank_ycor
]

;;Los enlaces representan tuberías, bombas y válvulas de control

undirected-link-breed [pipes pipe]
pipes-own [
  pipe_id
  pipe_node1
  pipe_node2
  pipe_length
  pipe_diameter
  pipe_roughness
  pipe_minorLoss
  pipe_status weight
]

directed-link-breed [pumps pump]
pumps-own [
  pump_id
  pump_parameters
  weight
]

undirected-link-breed [valves valve]
valves-own []

;;

to setup
  clear-all
  set mains-min-size 18

```

```

ask patches [set pcolor white]
set-default-shape junctions "circle"
set-default-shape reservoirs "reservoir"
set-default-shape tanks "tank"
set-default-shape pumps "pump"
ask tanks [set color blue set size 1]
ask reservoirs [set color blue set size 1]
ask junctions [set color black set size 1]
set init-colors [brown green blue yellow pink orange lime cyan
sky violet magenta turquoise gray]; etc.
set colors init-colors
set source-colors []

reset-ticks
end

to infect
ask junctions
[ if Presion1 > Factor_Presion ;; parcelas en el lado derecho
  [ set color orange
    set status 0
  ]
  ask n-of Fallo_inicial junctions [if Presion1 > Factor_Presion
    [ become-infected ]]
  ;reset-ticks
]

end

to go
;;
;;;;;ver en rojo los nodos con presiones mayor a 45 mca;;;;;

if
all? junctions [not infected?]
[ stop ]
ask junctions
[
  set virus-check-timer virus-check-timer + 1
  if virus-check-timer >= Frecuencia_control_fallo
  [ set virus-check-timer 0 ]
]
spread-virus
do-virus-checks

export-world
"C:/Users/cfnav/Desktop/TESIS_2023/Tesis_Netlogo/EpanetExport-
master/resultados.xls"

if vid:recorder-status = "inactive" [
  vid:start-recorder

```

```

]
vid:record-view
tick
if ticks = 300 [
  vid:save-recording "filme.mp4"
  print vid:recorder-status
  stop
]
]

end

to-report status-value
  report junctions
  print status-value
end

;;;;;; leyendo archivos e importando los datos de la red dentro de
NetLogo ;;;;;;

to import-network
  import-nodes
  import-links
end

to import-nodes
  file-open (word Nombre_red "_junctions_with_coords2.dat")
  while [not file-at-end?]
  [
    let items read-from-string ( word "[" file-read-line "]" )
    create-junctions 1 [
      set shape "circle"
      set color black
      set heading 0
      set node_id item 0 items
      set elevation item 1 items
      set junction_demand item 2 items
      set junction_pattern item 3 items
      set junction_xcor item 4 items
      set junction_ycor item 5 items
      set presion1 item 6 items
      ;set presion2 item 7 items
      set xcor junction_xcor * Escala
      set ycor junction_ycor * Escala
      set Status 1
      set virus-check-timer random Frecuencia_control_fallo
      set infected? false
      set resistant? false
      set susceptible? true
      set label who ;
    ]
  ]
]

```

```

file-close

file-open (word Nombre_red "_reservoirs_with_coords.dat")
while [not file-at-end?]
[
  let items read-from-string (word "[" file-read-line ")")
  create-reservoirs 1 [
    set shape "reservoir"
    set heading 0
    let the-color first colors
    set color the-color
    set colors but-first colors
    set source-colors lput the-color source-colors
    set node_id item 0 items
    set elevation item 1 items
    set reservoir_head item 2 items
    set reservoir_xcor item 3 items
    set reservoir_ycor item 4 items
    set xcor reservoir_xcor * Escala
    set ycor reservoir_ycor * Escala
    set label item 0 items ;
  ]
]
file-close

```

```

file-open (word Nombre_red "_tanks_with_coords.dat")
while [not file-at-end?]
[
  let items read-from-string (word "[" file-read-line ")")
  create-tanks 1 [
    set shape "tank"
    set heading 0
    let the-color first colors
    set color the-color
    set colors but-first colors
    set source-colors lput the-color source-colors
    set node_id item 0 items
    set elevation item 1 items
    set tank_initLevel item 2 items
    set tank_minLevel item 3 items
    set tank_maxLevel item 4 items
    set tank_diameter item 5 items
    set tank_minVol item 6 items
    set tank_xcor item 7 items
    set tank_ycor item 8 items
    set xcor tank_xcor * Escala
    set ycor tank_ycor * Escala
    set label item 0 items ;
  ]
]
file-close

```

end

```

to import-links
  file-open (word Nombre_red "_links.dat" )
  while [not file-at-end?]
  [
    let items read-from-string (word "[" file-read-line ")")
    ask get-node (item 1 items)
    [
      create-pipe-with get-node (item 2 items) ;
      [ set color black
        set pipe_id item 0 items
        set pipe_length item 3 items
        set pipe_diameter item 4 items
        set pipe_roughness item 5 items
        set pipe_minorLoss item 6 items
        set pipe_status item 7 items
      ]
    ]
  ]
  file-close

  file-open (word Nombre_red "_pumps4NetLogo.dat")
  while [not file-at-end?]
  [
    let items read-from-string (word "[" file-read-line ")")
    ask get-node (item 1 items)
    [
      create-pump-to get-node (item 2 items)
      [ set color black
        set pump_id item 0 items
        set pump_parameters item 3 items
        set shape "pump"
        set label item 0 items;
      ]
    ]
  ]
  file-close

end

to-report get-node [id]
  report one-of turtles with [node_id = id]
end

to become-infected
  set infected? true
  set resistant? false
  set susceptible? false
  set color red
end

to become-susceptible

```

```

    set infected? false
    set resistant? false
    set susceptible? true
    set color green
end

to become-resistant
  set infected? false
  set resistant? true
  set susceptible? false
  set color blue
  ask my-pipes [ set color blue - 2 ]
end

to spread-virus
  ask junctions with [infected?]
    [ ask junctions with [not resistant?]
      [ if random-float 10 < Probabilidad_propagacion_fallo
        [ become-infected ] ] ]
end

to do-virus-checks
  ask junctions with [infected? and virus-check-timer = 0]
  [
    if random 100 < Oportunidad_recuperacion
    [
      ifelse random 100 < Ganancia_resistencia
        [ become-resistant ]
        [ become-susceptible ]
    ]
  ]
end

```

Anexo 2

Tablas de resultados de las simulaciones en el caso de estudio de la red de Coro.

Las tablas con los resultados de las simulaciones son las siguientes:

	MOMENTO_1					
	SUSCEPTIBLE		INFECTADO		RECUPERADO	
	True	False	True	False	True	False
SIMULACION 1	2680	206	203	2684	3	2884
SIMULACION 2	2742	144	142	2745	2	2885
SIMULACION 3	2759	127	126	2761	1	2886
SIMULACION 4	2719	167	167	2720	0	2887
SIMULACION 5	2751	135	135	2752	0	2887
PROMEDIO	2730	156	155	2732	1	2886

	MOMENTO_2					
	SUSCEPTIBLE		INFECTADO		RECUPERADO	
	True	False	True	False	True	False
SIMULACION 1	96	2791	2395	491	395	2492
SIMULACION 2	102	2785	2373	513	411	2476
SIMULACION 3	92	2795	2399	488	395	2491
SIMULACION 4	84	2803	2421	465	381	2506
SIMULACION 5	79	2808	2393	493	414	2473
PROMEDIO	91	2796	2396	490	399	2488

	MOMENTO_3					
	SUSCEPTIBLE		INFECTADO		RECUPERADO	
	True	False	True	False	True	False
SIMULACION 1	55	2832	1695	1191	1136	1751
SIMULACION 2	57	2830	1700	1186	1129	1758
SIMULACION 3	65	2822	1740	1147	1081	1805
SIMULACION 4	62	2825	1721	1165	1103	1784
SIMULACION 5	60	2827	1714	1172	1112	1775
PROMEDIO	60	2827	1714	1172	1112	1775

	MOMENTO_4					
	SUSCEPTIBLE		INFECTADO		RECUPERADO	
	True	False	True	False	True	False
SIMULACION 1	20	2867	746	2141	2120	766
SIMULACION 2	38	2849	717	2169	2131	756
SIMULACION 3	23	2864	756	2131	2107	779
SIMULACION 4	24	2863	759	2127	2103	784
SIMULACION 5	32	2855	768	2119	2086	800
PROMEDIO	27	2860	749	2137	2109	777

	MOMENTO_5					
	SUSCEPTIBLE		INFECTADO		RECUPERADO	
	True	False	True	False	True	False
SIMULACION 1	4	2883	152	2732	2730	156
SIMULACION 2	7	2880	132	2755	2747	139
SIMULACION 3	6	2881	139	2748	2741	145
SIMULACION 4	10	2877	157	2730	2719	167
SIMULACION 5	7	2880	137	2750	2743	144
PROMEDIO	7	2880	143	2743	2736	150

	MOMENTO_6					
	SUSCEPTIBLE		INFECTADO		RECUPERADO	
	True	False	True	False	True	False
SIMULACION 1	2	2885	64	2823	2820	66
SIMULACION 2	3	2884	44	2843	2839	47
SIMULACION 3	5	2882	59	2828	2822	64
SIMULACION 4	3	2884	58	2829	2825	61
SIMULACION 5	4	2883	51	2836	2831	55
PROMEDIO	3	2884	55	2832	2827	59

	MOMENTO_7					
	SUSCEPTIBLE		INFECTADO		RECUPERADO	
	True	False	True	False	True	False
SIMULACION 1	7	2880	15	2872	2864	22
SIMULACION 2	4	2883	15	2872	2867	19
SIMULACION 3	3	2884	25	2862	2858	28
SIMULACION 4	5	2882	21	2866	2860	26
SIMULACION 5	5	2882	11	2876	2870	16
PROMEDIO	5	2882	17	2870	2864	22

	MOMENTO_8					
	SUSCEPTIBLE		INFECTADO		RECUPERADO	
	True	False	True	False	True	False
SIMULACION 1	4	2883	6	2881	2876	10
SIMULACION 2	6	2881	6	2881	2874	12
SIMULACION 3	1	2886	15	2872	2870	16
SIMULACION 4	3	2884	11	2876	2872	14
SIMULACION 5	5	2882	2	2885	2879	7
PROMEDIO	4	2883	8	2879	2874	12

	MOMENTO_9					
	SUSCEPTIBLE		INFECTADO		RECUPERADO	
	True	False	True	False	True	False
SIMULACION 1	4	2883	2	2885	2880	6
SIMULACION 2	4	2883	2	2885	2880	6
SIMULACION 3	6	2881	5	2882	2875	11
SIMULACION 4	5	2882	4	2883	2877	9
SIMULACION 5	6	2881	0	2887	2880	6
PROMEDIO	5	2882	3	2884	2878	8

	MOMENTO_10					
	SUSCEPTIBLE		INFECTADO		RECUPERADO	
	True	False	True	False	True	False
SIMULACION 1	4	2883	0	2887	2882	4
SIMULACION 2	5	2882	0	2887	2881	5
SIMULACION 3	4	2883	3	2884	2879	7
SIMULACION 4	6	2881	0	2887	2880	6
SIMULACION 5	NA	NA	NA	NA	NA	NA
PROMEDIO	5	2882	1	2886	2881	6