



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

TESIS DOCTORAL

Monitorización, detección y estimación de estados de fallo en la
calidad del agua de redes de distribución urbanas

Presentada por:

Joanna Alicia Gutiérrez Pérez

Dirigida por:

Dr. Manuel Herrera Fernández

Prof. Joaquín Izquierdo Sebastián

Dr. Idel Montalvo Arango

Valencia, España
Abril 2021



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Departamento de Ingeniería Hidráulica y Medio Ambiente

Programa de Doctorado en Ingeniería del Agua y Medioambiental

**Monitorización, detección y estimación de estados de fallo en
la calidad del agua de redes de distribución urbanas**

Por: Joanna Alicia Gutiérrez Pérez

Una Tesis Doctoral presentada en cumplimiento parcial de los
requisitos para obtener el grado de

Doctor en Ingeniería del Agua y Medioambiental

Valencia, España
Abril 2021

Dedicada al Prof. Rafael Pérez García. 
y a Patricia Gutiérrez. 

“The more security you seek, the more disturbed you will be about every
change that happens in your life”.

Sadhguru

Agradecimientos

Esta tesis ha sido posible gracias a mis directores Manuel Herrera, Joaquín Izquierdo e Idel Montalvo. Su aportación e implicación para la realización y finalización de esta tesis son incalculables. Gracias por su orientación y soporte constantes y, sobre todo, por su confianza, paciencia y amistad hacia mí.

Al Consejo Nacional de Ciencia y Tecnología de México por el apoyo recibido para el desarrollo del doctorado con su Programa de Posgrados en el Extranjero.

A la Generalitat Valenciana por su apoyo con el Programa de Formación de Investigadores.

Al grupo de investigación Fluing y al Instituto de Matemática Multidisciplinar por permitirme ser parte de ellos y por el apoyo recibido para realizar mi actividad investigadora.

A Ingeniousware y su equipo de trabajo.

“When eating fruit, remember the one who planted the tree”.
(Proverbio vietnamita)

Gracias a todas las personas que han estado involucradas de alguna manera en mi proceso de doctorado, especialmente a:

Ximo, Conchi, Rafa, Marta, Ico y Ángeles por su apoyo, por hacerme sentir como en casa, por su generosidad y por compartir grandes momentos en familia.

David, Naty, Eva, Akis, Silvia y Pauline, por tantos momentos divertidos, por su compañía y amistad y por todo lo que compartimos dentro y fuera del trabajo.

Idel y Stephi por abrirme las puertas de su casa, por su amabilidad y compartir conmigo el tesoro de las galletas.

Manu y Xia por su gran apoyo y amistad, por sus consejos y enseñanzas, por mi estancia en Cambridge y todas las atenciones recibidas.

Miguel Carrión por su amistad, por recibirme en su casa, por su gran amabilidad y hospitalidad.

Xitlali y Chucho por su apoyo los primeros días en Valencia y hacer más fácil el comienzo.

Mis compañeros del Máster por todos los días que pasamos juntos.

Ismael y Pepa por quererme y estar conmigo en mis malos y peores momentos.

¡Gracias, gracias, gracias!

ÍNDICE DE CONTENIDO

1	Introducción	1
1.1	Problemática	2
1.2	El concepto de vulnerabilidad referido a los sistemas de infraestructura crítica	6
1.3	Motivación y alcances de la investigación	9
1.4	Objetivos	10
1.5	Organización de la Tesis	11
2	El sistema de distribución de agua como infraestructura crítica	13
2.1	La vulnerabilidad de los sistemas de distribución de agua	17
2.1.1	Peligros naturales	20
2.1.2	Amenazas intencionales	23
2.2	Evaluación de la vulnerabilidad de los sistemas de distribución de agua	34
2.2.1	Herramientas y métodos	37
2.3	Prevención y alerta temprana frente a eventos de contaminación	43
2.3.1	Desarrollos orientados a la mejora de la seguridad de las RDAs	48
3	Enfoque multiobjetivo para la ubicación óptima de sensores de calidad en redes de distribución de agua	57
3.1	Fundamentos de la optimización multiobjetivo	60
3.2	Planteamiento general de los objetivos para la ubicación de sensores de calidad en RDAs	63
3.3	Definición de las funciones objetivo utilizadas	66
3.3.1	Tiempo de detección	68
3.3.2	Volumen de agua contaminada consumida	70
3.3.3	Fallo de detección	71
3.3.4	Costes de implementación	73

3.3.5 El cálculo de las matrices de contaminación y de agua contaminada consumida	73
3.4 Conclusión	76
4 Metodología para localizar sensores de calidad en redes de distribución de agua	79
4.1 La Teoría de grafos como herramienta para el análisis de RDAs	80
4.1.1 Marco conceptual y principales medidas de análisis de redes	83
4.1.2 Medidas de centralidad en redes	92
4.1.3 Detección de comunidades en redes	97
4.2 Implementación de los fundamentos de la teoría de redes para seleccionar nodos de monitoreo en una RDA	98
4.3 Optimización multiobjetivo: aplicación del algoritmo evolutivo ASO para obtener una red de sensores en una RDA	102
4.3.1 Algoritmo evolutivo ASO.....	105
4.3.2 Consideraciones para la aplicación del algoritmo evolutivo ASO en una RDA	108
4.4 La hibridación entre la teoría de redes y el algoritmo evolutivo ASO	111
5 Caso de estudio y resultados: localización de una red de sensores en una RDA	115
5.1 Caso de estudio.....	115
5.2 Obtención de los nodos-sensores de la RDA de Coro con el proceso de teoría de redes.....	118
5.3 Red de sensores aplicando la metodología: Teoría de redes – Optimización multiobjetivo	128
6 Conclusiones y trabajos futuros	147
6.1 El problema.....	147
6.2 Hacia la solución	148
6.3 Sobre la implementación de la solución.....	149
6.4 La propuesta híbrida planteada en la tesis.....	151
6.5 Direcciones futuras.....	155
Referencias bibliográficas	159

ÍNDICE DE CONTENIDO

Publicaciones propias	177
Anexos.....	184

Lista de Figuras

Figura 1.1: Esquema de la estructura de la tesis.....	12
Figura 2.1: Esquema general de un sistema de abastecimiento de agua (Iglesias Rey et al., 2003)	16
Figura 2.2: Conceptualización de la evaluación de la vulnerabilidad de las ICs.	35
Figura 2.3: Número de publicaciones por año referentes al tema de localización de sensores en RDAs. (Fuente: SCOPUS).	49
Figura 2.4: Porcentaje de publicaciones sobre localización de sensores en RDAs por área temática. (Fuente: SCOPUS, búsqueda del año 2003 al 2020).	50
Figura 3.1: Ilustración de un problema general de optimización multiobjetivo (Zitzler et al., 2004).	62
Figura 3.2. Configuración de la matriz de tiempos de contaminación para la obtención del tiempo de detección mínimo de los sensores	74
Figura 4.1: Enfoque general de la propuesta metodológica.	80
Figura 4.2: Representación de RDAs: sin dirección de los flujos (A), con dirección de los flujos (B) y no dirigida con pesos de las tuberías (C).	85
Figura 4.3: Proceso para la selección de los nodos sensores de una RDA.....	99
Figura 4.4: Frente de Pareto (tomada de Montalvo 2010).	107
Figura 4.5: Dos ejemplos de punto singular en un frente de Pareto aproximado (tomada de Montalvo et al., 2014).	107
Figura 5.1. Representación de la RDA de Coro.	116
Figura 5.2: Esquema general del proceso para la obtención de los grupos de nodos sensores de la RDA de Coro.	119
Figura 5.3: Grafo de la RDA de Coro en la interfaz del lenguaje R.	120
Figura 5.4: División de la RDA de Coro en 41 comunidades.	122
Figura 5.5. Ubicación de 41 sensores en la RDA de Coro según el valor de <i>pagerank</i> de los nodos.	124

Figura 5.6. Ubicación de 41 sensores en la RDA de Coro según el valor de intermediación de los nodos.....	125
Figura 5.7. Ubicación de 41 sensores en la RDA de Coro según el valor de centralidad del autovector.....	126
Figura 5.8. Solución del Frente de Pareto para la ubicación de 3 sensores en la RDA de Coro.....	130
Figura 5.9. Solución del Frente de Pareto para la ubicación de 10 sensores en la RDA de Coro.....	131
Figura 5.10. Solución del Frente de Pareto para la ubicación de 20 sensores en la RDA de Coro, con las soluciones de centralidad.....	132
Figura 5.11. Solución del Frente de Pareto para la ubicación de 41 sensores en la RDA de Coro con las soluciones de centralidad.....	133
Figura 5.12.. Tiempo de detección promedio vs número de fallos de detección, para el caso de 3 sensores.....	134
Figura 5.13. Tiempo de detección promedio vs número de fallos de detección, para el caso de 10 sensores.....	135
Figura 5.14. Tiempo de detección promedio vs número de fallos de detección, para el caso de 20 sensores con las soluciones de centralidad.....	135
Figura 5.15. Tiempo de detección promedio vs número de fallos de detección, para el caso de 41 sensores con las soluciones de centralidad.....	136
Figura 5.16. Fallo de detección promedio frente al consumo promedio de agua contaminada para la red de 3 sensores.....	137
Figura 5.17. Fallo de detección promedio frente al consumo promedio de agua contaminada para la red de 10 sensores.....	138
Figura 5.18. Fallo de detección promedio frente al consumo promedio de agua contaminada para la red de 20 sensores con las soluciones de centralidad.....	138
Figura 5.19. Fallo de detección promedio frente al consumo promedio de agua contaminada para la red con 41 sensores con las soluciones de centralidad.....	139
Figura 5.20. Frente de Pareto para 41 sensores, con solución elegida (círculo naranja) y soluciones de centralidad (otras figuras de colores).....	141

Figura 5.21. Red de Coro con los sensores correspondientes a la solución elegida	141
Figura 5.22. Consumo de agua contaminada vs Fallo de detección. Comparación de soluciones para 3, 10, 20 y 41 sensores.....	142
Figura 5.23. Tiempo promedio de detección vs Fallo de detección. Comparación de soluciones para 3, 10, 20 y 41 sensores.....	143
Figura 5.24. Tiempo promedio de detección vs Fallo de detección. Comparación de soluciones para 3, 10, 20 y 41 sensores.....	143
Figura 5.25. Frente de Pareto 3D para el caso de redes con 10 sensores y sus proyecciones sobre los planos coordenados (en azul los puntos tridimensionales; en verde, rojo y negro, sus proyecciones).....	144
Figura 5.26. Frentes de Pareto 3D para los números de sensores considerados.	145
Figura 5.27. Frente de Pareto 3D para el caso de 20 sensores, mostrando las soluciones (dominadas) obtenidas mediante la teoría de grafos...	146
Figura 6.1 Las soluciones de centralidad para 41 nodos mejoran a las soluciones multiobjetivo para 20 nodos.....	154

Lista de Tablas

Tabla 1.1: Definiciones de IC en los planes de protección de seis países.	2
Tabla 1.2: Cobertura de los planes sectoriales para ICs en algunos países.	4
Tabla 1.3: Definiciones de vulnerabilidad según su uso en la literatura.	6
Tabla 2.1: Peligros y amenazas asociados a los SDAPs y sus posibles efectos.	19
Tabla 2.2: Daño potencial de algunos productos químicos en los SDAPs.	29
Tabla 2.3: Clasificación de agentes biológicos, según el CDC	30
Tabla 2.4: Agentes considerados como armas biológicas y su amenaza potencial en agua.	32
Tabla 2.5: Resumen de los principales trabajos acerca de la ubicación óptima de sensores en una RDA	52
Tabla 4.1. Medidas básicas para el análisis de una red.	91
Tabla 5.1. Indicadores de la RDA de Coro, basados en la teoría de redes.	116
Tabla 5.2. Nodos – sensores con el valor más alto de <i>pagerank</i> y comunidad de pertenencia.	124
Tabla 5.3. Nodos – sensores con el valor más alto de intermediación y comunidad de pertenencia	126
Tabla 5.4. Nodos – sensores con el valor más alto de centralidad del autovector y comunidad de pertenencia.	127

I RESUMEN

La vulnerabilidad de los sistemas denominados de infraestructura crítica (IC), entre ellos los sistemas de distribución de agua potable (SDAP), ha sido uno de los temas de investigación de gran relevancia durante las últimas décadas, particularmente enmarcado en un contexto de peligro y/o amenaza. En general, los sistemas de IC proveen servicios esenciales para la sociedad actual. La seguridad, el crecimiento económico y el bienestar social de un país pueden verse comprometidos cuando alguno de esos sistemas es dañado, deshabilitado o interrumpido en su funcionamiento. Además de los accidentes y fallos característicos de los SDAPs, ciertos peligros y amenazas extraordinarios, tales como desastres naturales o actos relacionados con violencia a diferentes niveles, han generado un interés creciente entre gestores, tomadores de decisiones, ingenieros y la sociedad en general.

El abastecimiento de agua potable es uno de los servicios fundamentales para el desarrollo de una ciudad y la calidad de vida de los ciudadanos. Sin embargo, un SDAP tiene propiedades que lo hacen especialmente vulnerable en comparación con otros sistemas. La interrupción total o parcial del suministro de agua por un período de tiempo, causa graves efectos en los usuarios y en el resto de sistemas que dependen de él.

Debido a su importancia como infraestructuras, los SDAPs han sido un objetivo especialmente atrayente para ataques maliciosos, específicamente, posibles eventos de contaminación intencional. Este tipo de sucesos ha puesto de manifiesto las debilidades de estos sistemas y la necesidad de establecer medidas de prevención y mitigación que mejoren su seguridad, evitando que la salud de los usuarios sea afectada.

En la literatura se pueden encontrar diversos métodos para evaluar, prevenir y mitigar aspectos asociados a la vulnerabilidad de las ICs. Entre ellos se encuentra el enfoque de la teoría de redes complejas, basada en las técnicas de la teoría de grafos, que ha sido ampliamente utilizado, entre otros fines, para obtener parámetros relacionados con la estructura y conectividad de los elementos de una IC, para identificar los elementos más importantes, y para analizar el impacto

en el funcionamiento del sistema debido a la eliminación de uno o más elementos. Investigaciones recientes han demostrado la validez de estas técnicas para la evaluación de la vulnerabilidad de redes de diversa clase y, particularmente, de las redes de distribución de agua (RDAs). Mediante la abstracción de la RDA a un grafo con características especiales (donde los arcos son las líneas hidráulicas y los vértices son los puntos de consumo) es posible cuantificar sus atributos estructurales, facilitando su análisis a un nivel de detalle más preciso, incluida la división en sectores.

Esta tesis doctoral surge de la motivación de abordar este tema y de enfocarlo en la identificación y prevención de eventos de fallo en la calidad del agua. El estudio se centra en el análisis de la RDA que, por la ordenación y exposición de los diferentes elementos que la componen, es la parte más vulnerable de todo el sistema. Bajo el marco conceptual de la teoría de redes, se cuantifican atributos como la robustez y redundancia de la estructura de la RDA. Además, se utiliza el algoritmo evolutivo multiagente *Agent Swarm Optimization* (ASO) para identificar los lugares óptimos para ubicar sensores de monitorización de la calidad del agua, como una medida de vigilancia preventiva de cara a un evento de contaminación intencional o accidental. Al medir continuamente parámetros de la calidad del agua, entre ellos la conductividad, el pH, la concentración de diferentes sustancias y otros parámetros relacionados, directamente desde la red, una red adecuada de estos sensores ayudará a las empresas de distribución de agua a detectar la intrusión de contaminantes y a evaluar los problemas de calidad en sus redes. Decidir dónde colocar los sensores en la red y cuántos de ellos deben colocarse es un problema difícil de enfrentar, debido a la naturaleza de los objetivos que implica. Los objetivos contemplados en este trabajo son: el tiempo de detección, el volumen de agua contaminada consumida, el fallo de detección y los costes de implementación. Esta tesis presenta una forma de resolver el problema mencionado mediante la hibridación de la teoría de redes y el enfoque multiobjetivo del algoritmo ASO. Con la teoría de redes, se busca una aproximación a la búsqueda de soluciones posibles, reduciendo así el espacio de búsqueda. Con el algoritmo evolutivo multiobjetivo ASO más que una única solución precisa, la idea es encontrar un amplio espacio de soluciones que representen la mejor compensación entre todos los objetivos considerados en el problema.

Para esto, se toma como antecedente las propuestas de Gutiérrez-Pérez et al. (2017, 2012 y 2013) y Herrera et al. (2011a, 2012), en las que se demuestra la efectividad del uso de algoritmos – poco comunes en el campo de los SDAPs – como son el *PageRank* (Brin y Page, 1998a) y otras medidas interesantes en teoría de grafos. Asimismo, la tesis se fundamenta en los trabajos de Izquierdo et al. (2008), Montalvo (2010) y Montalvo et al. (2014) sobre el diseño de redes de abastecimiento con algoritmos evolutivos, específicamente con ASO.

Finalmente, la contribución principal de esta tesis es una metodología que ayude a las empresas de agua a establecer redes de monitoreo para detectar y estimar con rapidez posibles fallos en la calidad del agua, y que dé soporte a la toma de decisiones y a la mejora de la seguridad. Para demostrar el funcionamiento de la metodología propuesta, se presenta un caso de estudio referido a una red real, con el que se demuestran las posibilidades del acoplamiento de las técnicas propuestas. El resultado es un conjunto de posibles redes de sensores de calidad del agua, obtenido mediante un estudio *a priori* basado en el preproceso desarrollado a partir de la teoría de redes. Este conjunto es, posteriormente, ampliado y optimizado bajo el enfoque multiobjetivo de ASO, de modo que presente una cantidad suficiente de soluciones no dominadas de la frontera de Pareto del problema multiobjetivo estudiado. El fin último es dotar a los gestores del abastecimiento de una herramienta para seleccionar, mediante una toma de decisiones adecuada, aquella solución que mejor satisfaga, no solo los requerimientos técnicos y económicos, sino otros menos objetivables que puedan existir.

II ABSTRACT

The vulnerability of so-called critical infrastructure (CI) systems, including water distribution systems (WDSs), has been one of the highly relevant research topics in recent decades, particularly framed in a context of danger and/or threat. In general, CI systems provide essential services for today's society. The security, economic growth and social welfare of a country can be compromised when any of these systems is damaged, disabled or interrupted in its operation. In addition to the accidents and failures characteristic of WDSs, certain extraordinary dangers and threats, such as natural disasters or acts related to violence at different levels, have generated a growing interest among managers, decision makers, engineers and the society in general.

The supply of drinking water is one of the fundamental services for the development of a city and the quality of life of citizens. However, a WDS has properties that make it especially vulnerable compared to other systems. The total or partial interruption of the water supply for a period of time, causes serious effects on the users and on the rest of the systems that depend on it.

Due to their importance as infrastructures, WDSs have been an especially attractive target for malicious attacks, specifically potential intentional contamination events. This type of event has revealed the weaknesses of these systems and the need to establish prevention and mitigation measures that improve their safety, preventing the health of users from being affected.

Various methods can be found in the literature to assess, prevent and mitigate aspects associated with the vulnerability of CIs. The complex network theory approach, based on the techniques of graph theory, is one of them. This approach has been widely used, among other purposes, to obtain parameters related to the structure and connectivity of the elements of an IC, to identify its most critical elements, and to analyze the impact on the operation of the system due to the elimination of one or more elements. Recent research has demonstrated the validity of these techniques for assessing the vulnerability of networks of various kinds and, particularly, of water distribution networks (WDNs). By abstracting the WDN to a graph with special characteristics (where the arcs are the hydraulic

lines and the vertices are the consumption points) it is possible to quantify its structural attributes, facilitating analyses at a more precise level of detail, including the division into district metered areas.

This doctoral thesis arises from the motivation to address this issue and to focus on the identification and prevention of failure events in water quality. The study focuses on the analysis of the WDN, which, due to the arrangement and exposure of the different elements integrating it, is the most vulnerable part of the entire system. Under the conceptual framework of the network theory, attributes such as robustness and redundancy of the WDN structure are quantified. In addition, the evolutionary multi-agent Agent Swarm Optimization (ASO) algorithm is used to identify optimal places to locate water quality monitoring sensors, as a preventive surveillance measure to face any accidental or intentional contamination event. By continuously measuring water quality parameters, including conductivity, pH, concentration of different substances and other related parameters, directly from the network, a suitable network of these quality sensors will help water distribution companies detect the intrusion of pollutants and assess quality problems in their networks. Deciding where to place the sensors on the network and how many of them should be placed is a difficult problem to face, due to the nature of the objectives involved. The objectives considered in this work are: detection time, volume of consumed contaminated water, detection failure and implementation costs. This thesis presents a way to solve the aforementioned problem through the hybridization of the network theory and the multi-objective approach of the ASO algorithm. With the graph theory, an approach to the search for possible solutions is sought, thus reducing the search space. With the ASO multi-objective evolutionary algorithm, instead of a unique specific solution, the idea is to find a wide range of solutions that represent the best tradeoff among all the objectives considered in the problem.

To this end, the proposals by Gutiérrez-Pérez et al. (2017, 2012 and 2013) and Herrera et al. (2011a, 2012), which demonstrate the effectiveness of the use of algorithms - uncommon in the field of WDSs - such as PageRank (Brin and Page, 1998a) and other interesting measures in graph theory, are considered. Likewise, the thesis is based on the works by Izquierdo et al. (2008), Montalvo

(2010), and Montalvo et al. (2014) on the design of supply networks with evolutionary algorithms, specifically with ASO.

Finally, the main contribution of this thesis is a methodology intended to help water companies to set monitoring networks to quickly detect and estimate possible failures in water quality, and to support decision-making and improvement of water safety. To demonstrate the operation of the proposed methodology, a case study related to a real-world network is presented, with which the possibilities of coupling the proposed techniques are demonstrated. The result is a set of water quality sensor networks, obtained through an *a priori* study based on the pre-process developed from the network theory. This set is later expanded and optimized under the ASO multi-objective approach, so that it presents a sufficient number of non-dominated solutions of the Pareto front of the studied multi-objective problem. The ultimate goal is to provide supply managers with a tool to select, through appropriate decision-making, the solution that best satisfies, not only technical and economic requirements, but also other less objective objectives that may exist.

III RESUM

La vulnerabilitat dels anomenats sistemes d'infraestructura crítica (IC), inclosos els sistemes de distribució d'aigua potable (SDAP), ha estat un dels temes de recerca altament rellevants de les darreres dècades, especialment emmarcat en un context de perill i/o amenaça. En general, els sistemes IC ofereixen serveis essencials per a la societat actual. La seguretat, el creixement econòmic i el benestar social d'un país es poden comprometre quan algun d'aquests sistemes es danya, es desactiva o s'interromp en el seu funcionament. A més dels accidents i falles característics dels SDAPs, certs perills i amenaces extraordinaris, com ara desastres naturals o actes relacionats amb la violència a diferents nivells, han generat un interès creixent entre els administradors, els responsables de la presa de decisions, els enginyers i la societat en general.

El abastiment d'aigua potable és un dels serveis fonamentals per al desenvolupament d'una ciutat i la qualitat de vida dels ciutadans. No obstant això, un SDAP té propietats que el fan especialment vulnerable en comparació amb altres sistemes. La interrupció total o parcial de l'abastiment d'aigua durant un període de temps provoca greus efectes en els usuaris i en la resta de sistemes que en depenen.

A causa de la seva importància com a infraestructures, els SDAPs han estat un objectiu especialment atractiu per a atacs maliciosos, específicament possibles esdeveniments de contaminació intencionals. Aquest tipus d'esdeveniments han posat de manifest les debilitats d'aquests sistemes i la necessitat d'establir mesures de prevenció i mitigació que millorin la seva seguretat, evitant que es vegi afectada la salut dels usuaris.

Es poden trobar diversos mètodes a la literatura per avaluar, prevenir i mitigar aspectes associats a la vulnerabilitat de les ICs. Entre ells es troba l'enfocament complex de la teoria de xarxes, basat en les tècniques de la teoria de grafs, que s'ha utilitzat àmpliament, entre altres finalitats, per obtenir paràmetres relacionats amb l'estructura i la connectivitat dels elements d'una IC, per identificar-ne els elements més crítics, i analitzar l'impacte en el funcionament del sistema a causa de l'eliminació d'un o més elements. Investigacions recents

han demostrat la validesa d'aquestes tècniques per avaluar la vulnerabilitat de xarxes de diversos tipus i, en particular, de xarxes de distribució d'aigua (XDAs). Abstraient la XDA a un gràfic amb característiques especials (on els arcs són les línies hidràuliques i els vèrtexs són els punts de consum) és possible quantificar-ne els atributs estructurals, facilitant les anàlisis a un nivell de detall més precís, inclosa la divisió en sectors.

Aquesta tesi doctoral sorgeix de la motivació per abordar aquesta qüestió i centrar-se en la identificació i prevenció d'esdeveniments de fallida en la qualitat de l'aigua. L'estudi se centra en l'anàlisi de la XDA que, a causa de la disposició i exposició dels diferents elements que el componen, és la part més vulnerable de tot el sistema. Sota el marc conceptual de la teoria de xarxes, es quantifiquen atributs com la robustesa i la redundància de l'estructura de la XDA. A més, s'utilitza l'algorisme evolutiu multi-agent *Agent Swarm Optimization* (ASO) per identificar els llocs òptims per localitzar els sensors de control de la qualitat de l'aigua, com a mesura de vigilància preventiva davant d'un esdeveniment de contaminació intencional o accidental. Mesurant contínuament els paràmetres de qualitat de l'aigua, inclosa la conductivitat, el pH, la concentració de diferents substàncies i altres paràmetres relacionats, directament des de la xarxa, una xarxa adequada d'aquests sensors ajudarà les empreses de distribució d'aigua a detectar la intrusió de contaminants i avaluar els problemes de qualitat a les seves xarxes. Decidir on col·locar els sensors a la xarxa i quants d'ells s'hauria de col·locar és un problema difícil d'afrontar, a causa de la naturalesa dels objectius implicats. Els objectius considerats en aquest treball són: temps de detecció, volum d'aigua contaminada consumida, fallada de detecció i costos d'implementació. Aquesta tesi presenta una manera de resoldre el problema esmentat mitjançant la hibridació de la teoria de xarxes i l'enfocament multi-objectiu de l'algorisme ASO. Amb la teoria de grafs es busca un enfocament a la cerca de possibles solucions, reduint així l'espai de cerca. Amb l'algorisme evolutiu multi-objectiu ASO, en lloc d'una solució específica única, la idea és trobar una àmplia gamma de solucions que representin el millor compromís entre tots els objectius considerats en el problema.

Amb aquesta finalitat, es consideren les propostes de Gutiérrez-Pérez et al. (2017, 2012 i 2013) i Herrera et al. (2011a, 2012), que demostren l'eficàcia de

l'ús d'algoritmes - poc freqüents en el camp dels SDAP - com PageRank (Brin i Page, 1998a) i altres mesures interessants en teoria de grafs. Així mateix, la tesi es basa en els treballs d'Izquierdo et al. (2008), Montalvo (2010) i Montalvo et al. (2014) sobre el disseny de xarxes d'abastiment d'aigua amb algorismes evolutius, específicament amb ASO.

Finalment, la principal contribució d'aquesta tesi és una metodologia destinada a ajudar les empreses d'aigua a establir xarxes de control per detectar i estimar ràpidament possibles fallades en la qualitat de l'aigua, i donar suport a la presa de decisions i la millora de la seguretat de l'aigua. Per demostrar el funcionament de la metodologia proposada, es presenta un cas d'estudi relacionat amb una xarxa real, amb la qual es demostren les possibilitats d'acoblament de les tècniques proposades. El resultat és un conjunt de xarxes de sensors de qualitat de l'aigua, obtingudes mitjançant un estudi a priori basat en el pre-procés desenvolupat a partir de la teoria de xarxes. Aquest conjunt s'amplia i s'optimitza després amb l'enfocament multi-objectiu ASO, de manera que presenta un nombre suficient de solucions no dominades del front de Pareto del problema multi-objectiu estudiat. L'objectiu final és proporcionar als gestors de subministraments una eina per seleccionar, mitjançant una presa de decisions adequada, la solució que millor satisfaci, no només els requisits tècnics i econòmics, sinó també d'altres menys objectivables que puguin existir.

1 Introducción

El estudio de la vulnerabilidad de los sistemas considerados como infraestructuras críticas (ICs) ha sido uno de los temas de investigación de gran relevancia durante las últimas décadas, enmarcado principalmente en un ambiente de conflicto. Además de los accidentes y fallos propios de los sistemas, ciertos eventos, como son los desastres naturales y los actos relacionados con la violencia a diferentes niveles, han generado un interés creciente entre políticos, gestores, ingenieros y la sociedad en general. Los sistemas de ICs proveen servicios esenciales para la sociedad actual, por lo que la seguridad, el crecimiento económico y el bienestar social de un país pueden verse comprometidos cuando alguno de esos sistemas es dañado, deshabilitado o interrumpido en su funcionamiento.

Considerado como una IC, el Sistema de Distribución de Agua Potable (SDAP) es uno de los servicios fundamentales para el desarrollo de una ciudad y la calidad de vida de sus habitantes. Sin embargo, un SDAP tiene propiedades que lo hacen especialmente vulnerable en comparación con otros sistemas. El suministro de agua puede verse gravemente comprometido por varios peligros o amenazas y, en consecuencia, ser interrumpido parcial o totalmente por un período de tiempo, lo que causaría graves efectos en los usuarios y el resto de sistemas que dependen de él. En este contexto, un SDAP resulta especialmente atractivo para ataques de contaminación intencionales. Este tipo de sucesos pone de manifiesto las debilidades del sistema y la necesidad de evaluar y prevenir contingencias, mitigar daños y establecer mejoras en la seguridad.

Existen diversos métodos para enfrentar las cuestiones antes expuestas y dar una solución que se adapte, según el tipo de sistema, al objetivo de análisis y a los datos disponibles¹. Uno de los principales enfoques es la teoría de redes complejas, la cual se fundamenta en las técnicas de teoría de grafos. La teoría de redes ha sido ampliamente utilizada, entre otros fines, para obtener parámetros relacionados con la estructura y conectividad de los elementos de una IC, para

¹ Estos métodos, de forma general, están enmarcados en categorías de homogeneidad: Análisis estadístico, Modelación probabilística, Análisis de riesgo, Teoría de redes complejas, Modelación y simulación multiagente y Teoría dinámica de control del sistema (Kröger and Zio 2011a).

identificar los elementos más importantes y para analizar el impacto en el funcionamiento del sistema por causa de la eliminación de uno o más elementos. Investigaciones recientes han demostrado la validez de estas técnicas para la evaluación de la vulnerabilidad de redes de diversa clase y, particularmente, de las redes de distribución de agua (RDAs). Mediante la abstracción de la RDA a un grafo² con características especiales — donde los arcos son las líneas hidráulicas y los vértices o nodos son los puntos de consumo — es posible cuantificar sus atributos estructurales y facilitar su análisis a un nivel de detalle más preciso como, por ejemplo, su división en clústeres.

1.1 Problemática

En su definición más amplia, una infraestructura crítica es un bien físico o intangible cuya destrucción o paralización quebrantaría seriamente la salud pública, el orden social y el cumplimiento de las principales responsabilidades del gobierno de una nación. Bajo esta premisa, las ICs han recibido una gran atención dentro de las políticas de inversión y los planes de desarrollo de diversos países. Lo anterior, con el fin de crear estrategias de seguridad y programas de protección que consideren, principalmente, la función que desempeñan o el servicio que proporcionan. Como muestra de ello, en la Tabla 1.1, se incluyen las definiciones de IC establecidas por algunos de los países más representativos, según la OECD (*Organisation for Economic Co-operation and Development*) (Gordon y Dion, 2008).

Tabla 1.1: Definiciones de IC en los planes de protección de seis países.

EE. UU.	La definición dentro del plan general de infraestructuras críticas de los Estados Unidos es: “sistemas y bienes, ya sean físicos o virtuales, tan vitales para los Estados Unidos que la incapacidad o destrucción de esos sistemas y bienes, tendría un impacto debilitador en la seguridad, la estabilidad económica nacional, la salud pública nacional o los servicios
---------	--

² Los grafos son estructuras formadas por un conjunto de vértices y un conjunto de aristas. Pero, desde el punto de vista del análisis de la RDA, también serán utilizados los términos nodos para los vértices y arcos o enlaces para las aristas, que son las conexiones entre dos o más nodos.

	<p>sanitarios, o cualquier combinación de estos temas.” Para fines de política de inversión, esta definición es más precisa: “sistemas y bienes, ya sean físicos o virtuales, tan vitales para los Estados Unidos que la incapacidad o destrucción de esos sistemas y bienes tendría un impacto debilitador en la seguridad nacional.”</p>
Canadá	<p>“Las infraestructuras críticas de Canadá consisten en aquellas instalaciones físicas y tecnologías de información, redes, servicios y bienes, los cuales, si se interrumpen o destruyen, tendrían un severo impacto en los servicios sanitarios, en la estabilidad o el bienestar económico de los canadienses o en el funcionamiento eficaz de los gobiernos en Canadá.”</p>
Países de la Unión Europea	<p>“Las instalaciones críticas son organizaciones o instalaciones de mayor importancia para la comunidad cuyo fallo o deterioro causaría una prolongada escasez de los suministros, importantes perturbaciones del orden público u otras consecuencias dramáticas.”</p> <p>“La Infraestructura Crítica Nacional comprende aquellos bienes, servicios y sistemas que soportan la vida económica, política y social del país cuya importancia es tal que su pérdida podría: 1) causar grandes pérdidas de vida; 2) tener un serio impacto en la economía nacional; 3) tener otras consecuencias sociales graves para la comunidad; o 4) ser de interés inmediato para el gobierno nacional.”</p>
Australia	<p>“Infraestructura crítica es definida como aquella instalación física, cadena de suministro, tecnología de información y red de comunicación cuya destrucción, degradación o inutilización por un largo período de tiempo, impactaría significativamente en el bienestar económico y social de la nación, o afectaría la capacidad de Australia para conducir la defensa nacional y garantizar la seguridad nacional.”</p>

Durante los últimos años la definición general y el criterio para determinar qué es una IC han cambiado. En un principio, se consideraron las infraestructuras fundamentales para la defensa, la seguridad económica de la nación y la

continuidad del gobierno. Posteriormente, se pensó en aquellas vitales para la seguridad y la salud pública, y para la moral de una nación (Moteff et al., 2003). Incluso, algunos países han ampliado su visión a sectores como alimentación, sanidad, gobierno y finanzas, tal como aparece en la Tabla 1.2 de la OCDE (Gordon y Dion, 2008). Esto demuestra la importancia y la trascendencia de las ICs para el funcionamiento óptimo de los servicios que se prestan al ciudadano y también para los sistemas de producción, en general. Toda actividad humana en la sociedad actual depende fuertemente, de una u otra forma, de alguno de estos sectores estratégicos.

Tabla 1.2: Cobertura de los planes sectoriales para ICs en algunos países.

Sector	Unión Europea	EE. UU	Canadá	Australia
Energía (incluyendo la nuclear)	X	X	X	X
IT y Comunicación	X	X	X	X
Finanzas	X	X	X	X
Salud	X	X	X	X
Alimentación	X	X	X	X
Agua	X	X	X	X
Transporte	X	X	X	X
Seguridad	Serv. de emergencia	Serv. de emergencia	X	Serv. de emergencia
Gobierno	X	X	X	
Químicos		X	X	
Defensa industrial		X		
Otros sectores y actividades		Reuniones públicas, íconos nacionales		

En ese contexto, y como respuesta al supuesto de una creciente amenaza de terrorismo, a finales de los años 90 el Gobierno Federal de los EE. UU., y bajo la presidencia de Bill Clinton, estableció la *President's Commission on Critical Infrastructure Protection* (PCCIP). En la orden ejecutiva publicada (E.O. 13010)

se definió al término infraestructura como: “El marco de redes interdependientes y de sistemas que abarcan determinadas industrias, instituciones (incluyendo personas y procedimientos) y capacidades de distribución que proporcionan un flujo confiable de productos y de servicios esenciales para la defensa y la seguridad económica de los EE. UU., el buen funcionamiento del gobierno en todos los niveles, y de la sociedad en conjunto”. Posteriormente, debido a que la seguridad nacional de los EE. UU. se calificó como de alta prioridad nacional, el término IC se fue desarrollando dentro de una mayor preocupación política. Por ello, en la misma orden ejecutiva, se estableció: “Ciertas infraestructuras nacionales son tan vitales que su incapacidad o destrucción tendría un impacto debilitante en la defensa o seguridad económica de los Estados Unidos”. Asimismo, se definieron ocho sectores de prestación de servicios esenciales para la sociedad, que son:

1. Telecomunicaciones
2. Sistemas de energía eléctrica
3. Almacenamiento y transporte de gas y petróleo.
4. Banca y finanzas
5. Transporte
6. Sistemas de abastecimiento de agua
7. Servicios de emergencia (incluyendo servicios médicos, de rescate, policía y bomberos)
8. Continuidad del Gobierno

Del mismo modo, se destacó otro grupo de bienes no clasificados dentro de los sectores anteriores que son: los monumentos nacionales, las plantas de energía nuclear, las instalaciones de gobierno y los bienes comerciales.

En este sentido, tanto en los EE. UU. como en otros países, una IC abarca una amplia gama de bienes y sistemas. Ningún sistema es más importante que otro, sino las características de interdependencia que existen entre ellos para su funcionamiento. Pues si un sistema es interrumpido, existe el potencial para fallos secundarios o en cadena. En consecuencia, existe una necesidad creciente de conocer, medir y supervisar sus fortalezas y debilidades, tanto potenciales como sistemáticas, para la identificación de criticidades y su refuerzo. Dada la

fuerte presencia de diversos sistemas de IC en el mundo industrializado, es fundamental tener una visión de las técnicas capaces de identificar las debilidades más generalizadas de los elementos que las componen para ser protegidas o reforzadas.

Los conceptos de fiabilidad y vulnerabilidad son especialmente importantes cuando se analiza la capacidad de una IC para mantener la continuidad de su funcionamiento (Murray y Grubestic, 2007). La fiabilidad es la probabilidad de no fallo de los elementos de la infraestructura, con base en una serie de patrones establecidos o normas de funcionamiento. Contrario a la fiabilidad, la vulnerabilidad es un concepto más amplio y con implicaciones más extensas. Mientras que la fiabilidad se enfoca en la posibilidad de mantener el funcionamiento de los elementos de la infraestructura, la vulnerabilidad se centra en el potencial para alterar esos elementos o degradarlos hasta un punto donde el funcionamiento es interrumpido. Sin embargo, vulnerable no siempre significa poco fiable, y a su vez, poco fiable no necesariamente es ser vulnerable.

1.2 El concepto de vulnerabilidad referido a los sistemas de infraestructura crítica

En su significado general, vulnerabilidad es la cualidad o estado de ser vulnerable en diferentes maneras. Esta definición se extiende a diversas aplicaciones y campos del conocimiento. Como ejemplo, en la Tabla 1.3 se muestran algunas definiciones en contextos específicos.

Tabla 1.3: Definiciones de vulnerabilidad según su uso en la literatura.

Definición	Autor/Fuente
<p>Vulnerabilidad son las características y circunstancias de una comunidad, sistema o un bien que los hacen susceptibles a los efectos dañinos de una amenaza.</p> <p>Son las condiciones determinadas por factores o procesos físicos, sociales, económicos y ambientales, que aumentan la susceptibilidad de una comunidad al impacto de amenazas.</p>	<p>(EIDR, 2009)</p>

<p>Vulnerabilidad es el potencial para la pérdida o daño a los bienes expuestos, en gran medida, debido a las complejas interacciones entre los procesos naturales, las decisiones del uso de la tierra, y la resiliencia de las comunidades.</p>	<p>(NOAA, 2010)</p>
<p>Vulnerabilidad es el grado al cual un sistema es susceptible y es incapaz de hacer frente a los efectos adversos del cambio climático. La vulnerabilidad está en función de la condición, magnitud y rapidez del cambio climático y la variación a la que un sistema está expuesto, la sensibilidad y la capacidad de adaptación de dicho sistema.</p>	<p>(Adger et al., 2007)</p>
<p>La vulnerabilidad es definida como el grado al cual es probable que sea dañada o interrumpida una comunidad, estructura o área geográfica por el impacto del daño de un fenómeno particular con un determinado nivel de severidad.</p>	<p>(Coburn et al., 1994)</p>
<p>La vulnerabilidad de un sistema industrial se refiere a los inmuebles, instalaciones y equipo de producción, incluyendo sus recursos humanos, organización y todo el <i>software</i>, <i>hardware</i> y mercancía neta, que pueden debilitarse o limitar su capacidad para soportar amenazas y sobrevivir a sucesos accidentales, los cuales se originan tanto dentro como fuera de las fronteras del sistema.</p>	<p>(Einarsson y Rausand, 1998)</p>
<p>El concepto de vulnerabilidad requiere de otros dos que son el peligro y el azar. Con respecto a un riesgo natural, el peligro es un evento potencialmente peligroso, un fenómeno natural o físico que es representado, por lo general, por su período de retorno y sus características físicas como, por ejemplo, su magnitud e intensidad.</p>	<p>(Boudou, 2006)</p>

En el ámbito de los sistemas de ICs, se pueden encontrar varias acepciones del término vulnerabilidad. Algunas de ellas están relacionadas con la “susceptibilidad a” y, en otros casos, el término es utilizado indistintamente

como “riesgo” En relación con esto último, Ezell (2007) apunta que mientras la vulnerabilidad resalta la noción de susceptibilidad a un escenario, el riesgo se enfoca en la severidad de las consecuencias dentro del contexto de un escenario. Por lo tanto, la vulnerabilidad es una condición del sistema que debe ser evaluada dentro del contexto de un escenario, o bien, una medida de la susceptibilidad frente a uno o varios escenarios. Por otro lado, Apostolakis y Lemon (2005) definen la vulnerabilidad como la manifestación de los estados inherentes del sistema (físico, técnico, organizacional y cultural) que pueden ser utilizados por un adversario para averiar o dañar el sistema. Por su parte, Bouchon (2006) señala que la vulnerabilidad, vista como una propiedad del sistema, se enfoca en tres elementos principales: 1) el grado de pérdidas y daños³, 2) el grado de exposición a los peligros⁴ y 3) el grado de resiliencia⁵.

En conclusión, como señalan Kröger y Zio (2011a), el concepto de vulnerabilidad referido a los sistemas de ICs se define como un defecto o debilidad (característica inherente, incluyendo la capacidad de resiliencia) en el diseño, implementación, operación y gestión de la infraestructura de un sistema o de sus elementos. Como consecuencia, el sistema se hace susceptible a la destrucción o incapacitación cuando está expuesto a un peligro o amenaza, o bien, se reduce su capacidad para restablecerse a nuevas condiciones estables. Así, el concepto de vulnerabilidad es utilizado para caracterizar la falta de robustez y resiliencia de un sistema. En este sentido, robustez significa que el sistema conservará su estructura intacta (se mantiene sin cambios, o casi sin cambios) cuando está expuesto a perturbaciones. Resiliencia representa la adaptación del sistema para volver a un estado de normalidad (recuperarse o retornar a su estado original) después de haber sido alterado. La resiliencia es la habilidad de un sistema para operar cerca de su funcionamiento óptimo (técnica y/o institucionalmente) en un lapso corto, después de un evento adverso. Unidos, los conceptos de robustez y resiliencia, son complementos de la vulnerabilidad y, asimismo, los conceptos

³ debido al impacto de un peligro (dimensión técnica).

⁴ es decir, la probabilidad de estar expuesto a peligros de un cierto grado y la susceptibilidad de un elemento en riesgo de sufrir pérdidas y daños (el elemento en riesgo podría ser un sistema técnico).

⁵ es decir, la capacidad de un sistema para anticipar, afrontar o absorber, resistir y recuperarse del impacto de un peligro (técnico) o el desastre (social). Implica que el sistema puede adaptarse para recuperar una nueva posición estable (recuperar o volver a, o cerca de, su estado original) después de las perturbaciones.

de redundancia y fiabilidad juntos. Redundancia se entiende como la capacidad del sistema para continuar operando a pesar de fallos en alguno de sus componentes, lo que implica que el sistema tenga los componentes suficientes o equipos en paralelo para poder sustituir al que falló. Fiabilidad se puede definir como la capacidad del sistema para cumplir con la función para la cual fue diseñado⁶. Por lo tanto, la relación que existe es que a mayor redundancia se obtiene mayor fiabilidad.

Con referencia a los SDAPs, Li (2007) define vulnerabilidad como una propiedad asociada a un componente, a un subsistema o al sistema completo que representa la posibilidad de daño, ya sea por peligros o amenazas, con probabilidades y severidades dadas. Esta propiedad es determinada por los atributos y las condiciones propios del SDAP y varía con el tiempo y los cambios de los diferentes peligros y amenazas.

Conocer el alcance de la vulnerabilidad será clave para determinar las medidas de mitigación y los planes de emergencia que sea necesario implementar. Con ello, es posible dar una respuesta eficaz al impacto de una determinada amenaza.

1.3 Motivación y alcances de la investigación

Este trabajo de tesis surge de la motivación de abordar el tema de la vulnerabilidad de los SDAPs y de enfocarlo en la identificación y prevención de eventos de fallo en la calidad del agua. El estudio se centra en el análisis de la RDA que, por la ordenación y exposición de los diferentes elementos que la componen, es la parte más vulnerable de todo el sistema. Bajo el marco conceptual de la teoría de grafos se cuantifican atributos como la conectividad y la centralidad, que se reflejan en la robustez y redundancia de la estructura de la RDA. Además, se propone la aplicación del algoritmo ASO (*Agent Swarm Optimization*) (Montalvo, 2010) para identificar los puntos propicios para ubicar sensores de monitorización, como una medida de vigilancia preventiva de cara a un evento de contaminación intencional. Para esto, se toma como antecedente

⁶ En el caso específico de los SDAP, se puede decir que es suministrar agua en las cantidades requeridas, con la calidad óptima y la presión adecuada.

los trabajos de Gutiérrez-Pérez et al. (2012 y 2013); Herrera et al. (2011a) y Yazdani y Jeffrey (2011) en los que se demuestra la efectividad del uso de algoritmos poco comunes en el campo de los SDAPs como son el *PageRank* (Brin y Page, 1998). Asimismo, los trabajos de Montalvo (2008) y Montalvo et al. (2014) en diseño de redes de abastecimiento de agua con algoritmos evolutivos, específicamente con PSO (*Particle Swarm Optimization*) dentro de ASO.

En este trabajo se busca vincular las técnicas y algoritmos propuestos para la formulación de una metodología híbrida. Se plantea que la adaptación de la teoría de grafos en el análisis de una RDA, proporcionará información suficiente para la cuantificación de los principales conceptos asociados a su vulnerabilidad y dará una primera solución simplificada al problema de la localización de sensores. Asimismo, con la aplicación del algoritmo de optimización, se buscará un conjunto de soluciones óptimas para seleccionar el conjunto de puntos de monitorización en la RDA de estudio. Finalmente, la contribución principal de esta tesis es una metodología que ayude a los responsables del abastecimiento de agua, por ejemplo, las empresas de agua, a monitorear sus redes para detectar y estimar con rapidez posibles fallos en la calidad del agua, y que dé soporte a la toma de decisiones y a la mejora de la seguridad. Esto, permitirá proponer medidas de vigilancia preventiva que ayuden a mitigar los efectos de un evento de contaminación intencional.

1.4 Objetivos

El objetivo principal de esta tesis es:

Crear una metodología para ubicar sensores que permitan monitorizar, detectar y estimar estados de fallo en las redes de distribución urbanas, que permita tomar decisiones y acciones oportunas para mejorar su seguridad y reducir posibles consecuencias.

Para lograr dicho objetivo, se han planteado los objetivos específicos siguientes:

- Definir con claridad el problema de la ubicación óptima de sensores, seleccionando los objetivos adecuados.

- Plantear el problema desde un punto de vista topológico y obtener soluciones iniciales basadas en la teoría de grafos.
- Optimizar las soluciones obtenidas y obtener otras nuevas mediante un enfoque multiobjetivo utilizando una técnica de optimización adecuada.
- Implementar un algoritmo conjunto para optimizar el número y localización de sensores de calidad del agua.
- Elaborar conclusiones que puedan servir de referencia a los gestores de un SDAP en términos de detección temprana de intrusiones patógenas en el sistema.

1.5 Organización de la Tesis

El documento de tesis está estructurado en seis capítulos (Figura 1.1). Los Capítulos 1 y 2 se refieren a la introducción, a los objetivos de investigación y al contexto en el que se enmarca la vulnerabilidad y seguridad de los SDAPs. En el Capítulo 3, se exponen las bases teóricas y principales consideraciones para abordar el problema de la localización de sensores en las RDAs desde un punto de vista ingenieril; asimismo, se establecen las expresiones matemáticas del problema tal como se aborda en esta tesis. El Capítulo 4, se refiere a la metodología propuesta para la solución del problema de la localización de sensores en una RDA. Este capítulo plantea la utilización conjunta de la teoría de grafos y la optimización multiobjetivo, para encontrar un conjunto de soluciones óptimo al problema de la localización de sensores. En el Capítulo 5 se desarrolla un caso de estudio de una red real, y se presentan los resultados obtenidos. El documento se cierra con el Capítulo 6, que contiene las conclusiones principales de los resultados obtenidos y las aportaciones logradas, de las que se pueden elaborar algunas directrices básicas para el diseño de sistemas de detección temprana de intrusión de contaminantes. También se esbozan las múltiples líneas de trabajo existentes en el campo de la vulnerabilidad, la monitorización y la protección de los SDAPs, que podrían abordarse en trabajos futuros posibles.

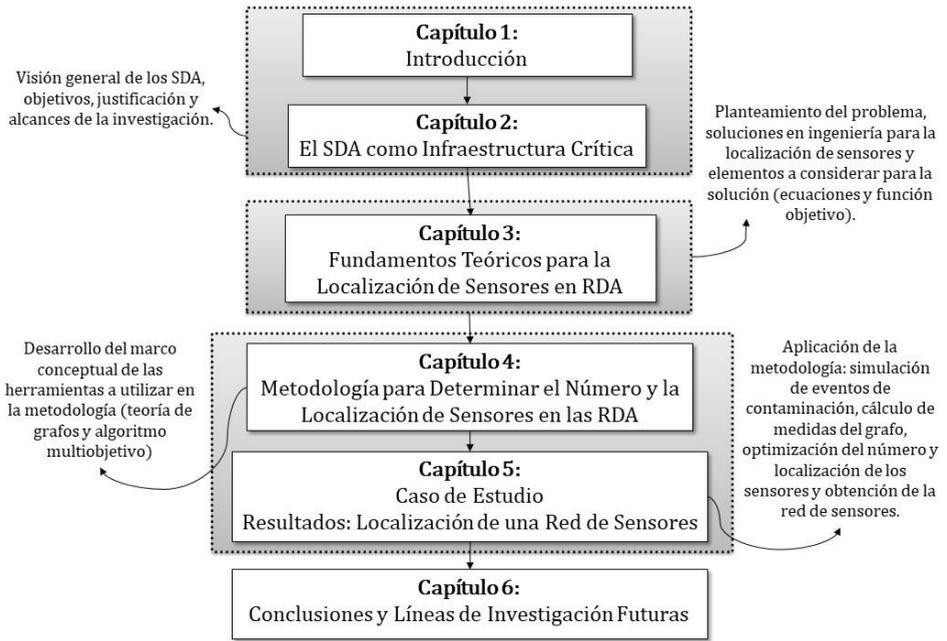


Figura 1.1: Esquema de la estructura de la tesis

2 El sistema de distribución de agua como infraestructura crítica

Este capítulo aborda la problemática general de los SDAPs, su tratamiento como IC y sus principales características, y se describen de forma general los principales conceptos asociados a su vulnerabilidad.

En términos generales, un SDAP es una agrupación de subsistemas interdependientes, cada uno sirviendo a poblaciones en un área geográfica limitada (Haimés et al., 1998). El objetivo principal de un SDAP es suministrar agua en cantidad requerida y con la calidad óptima hasta los puntos de consumo, así como realizar la gestión de la evacuación de las aguas residuales. Para lograr tal objetivo, el SDAP está compuesto de elementos que permiten la captación desde las fuentes (estaciones de bombeo), el tratamiento (estación de tratamiento de agua potable), el transporte (tuberías), la regulación y el control (válvulas), el almacenamiento (depósitos) y la medición (caudalímetros); todo funcionando bajo reglas de operación y mediante la intervención de personas.

Las tuberías o líneas de conducción son los elementos principales de la RDA, así como los más numerosos. La gran mayoría de los problemas de diseño, operación, mantenimiento y rehabilitación se encuentran relacionados directa o indirectamente con estos elementos. Las tuberías varían dependiendo de la función dentro del SDAP y su ubicación, son de rigidez variable, generalmente están enterradas, y para su instalación se requiere de otros accesorios (juntas y válvulas). Las aducciones son las líneas de conducción que llevan el agua desde las fuentes de abastecimiento a la estación de tratamiento o potabilizadora, al depósito de regulación o estación de bombeo; estas líneas también llevan el agua hasta la red y sus depósitos. Conectadas a las aducciones en la RDA, están las líneas principales o arterias, las cuales son de mayor diámetro y alimentan a las líneas secundarias. Por lo general no hay acometidas en estas líneas. Las líneas secundarias son las que llevan el agua de las arterias a las tuberías de distribución y en ellas se hacen tomas a los principales consumidores. En las líneas de distribución están la mayoría de las acometidas. Estas son un conjunto de tuberías, válvulas y accesorios que conectan la red pública con la instalación interior de las edificaciones. Dependiendo del tamaño de la red, la clasificación

puede llegar a reducirse a tuberías de distribución y acometidas (Mora-Rodríguez, 2011).

Los materiales más comunes en la fabricación de tuberías son el polietileno (PE), el policloruro de vinilo (PVC), el poliéster reforzado con fibra de vidrio (PRFV), el hierro fundido dúctil, el acero, el fibrocemento y el hormigón. Los materiales plásticos como el PE, el PVC y el PRFV son resistentes a la corrosión, presentan una rugosidad menor y, por lo tanto, no generan incrustaciones. El PE es un material plástico inerte. Este material es inodoro, insípido, inoxidable, inalterable a la acción de terrenos corrosivos, es resistente, con bajo factor de fricción, es aislante térmico (a 0°C se dilata y recupera su dimensión al descongelarse el líquido), tiene bajo módulo de elasticidad (aminora transitorios hidráulicos) y es insoluble en agua. El PVC presenta buenas propiedades mecánicas, es un material termoplástico por lo que puede ser fácilmente manipulado. El PVC es inodoro, insípido, no tóxico y químicamente inerte. El PRFV es un material poco utilizado desde hace 30 años; se caracteriza por su elevada resistencia, poco peso y fácil instalación. Es recomendable para la distribución a temperaturas inferiores a 35°C. El hierro fundido dúctil es uno de los materiales más utilizados, que sustituyó a la fundición gris. Presenta buena resistencia mecánica en instalaciones enterradas, así como a movimientos del terreno. Su revestimiento le proporciona resistencia a la corrosión, pero su revestimiento interior debe ser adecuado para el uso en agua de consumo. El acero tiene características similares al hierro fundido. Se usa en grandes diámetros o en conducciones sobre la superficie, pero requiere de un tratamiento especial (galvanizado) para hacerlo apto como conductor de agua potable. En cuanto al hormigón, éste se utiliza sobre todo en diámetros grandes o cuando requieren soportar altas presiones. Finalmente, las tuberías de fibrocemento están casi en desuso y han sido sustituidas por otros materiales que ofrecen más ventajas.

Todos los materiales tienen ventajas e inconvenientes y la selección del tipo de tubería dependerá del coste, presión de trabajo, condiciones climáticas, características del agua a distribuir, mantenimiento, tiempo de vida útil, las propiedades del suelo y las propiedades técnicas de instalación, e incluso, de

aspectos como el transporte, el montaje, la protección requerida y su resistencia a cargas internas y externas (Mora-Rodríguez, 2011).

Otro de los elementos más numerosos en las RDAs son las válvulas, que se emplean, en su mayoría, para aislar determinados tramos de la red en caso de que sea necesario realizar algún trabajo de rehabilitación, reparación o mantenimiento. En cuanto a los elementos de medida y control, su instalación a lo largo de la RDA es necesaria para tener vigilado y controlado gran parte del abastecimiento; la finalidad es conocer los valores de las diferentes variables hidráulicas que indican el estado de la red. Dentro de los elementos de medida y control, los más importantes son los contadores domésticos. Finalmente, los elementos encargados de generar el consumo final del sistema son toda una serie de componentes que constituyen el tramo final donde termina la red de distribución, los cuales son: desagües, bocas o hidrantes de riego, hidrantes de extinción de incendios y acometidas.

El proceso completo del abastecimiento de agua, desde la fuente hasta los puntos de consumo final, es un proceso amplio que se puede clasificar en dos etapas principales: la producción y la distribución (Figura 2.1). La etapa de producción incluye las instalaciones de captación, la Estación de Tratamiento de Agua Potable (ETAP), los depósitos de regulación y los sistemas de bombeo, además de las conducciones que conectan a estos elementos. Mientras que en la etapa de distribución se consideran las instalaciones de conducción desde los depósitos o estaciones de bombeo de agua tratada, la red arterial o primaria de abastecimiento, la red de distribución o red secundaria, los depósitos dentro de la red y las estaciones de bombeo.

Dentro de la etapa de producción, la calidad del agua varía según el tipo de captación y el proceso establecido en la ETAP. Dependiendo de las características fisicoquímicas y biológicas que presente el agua en estado natural, se determina el tipo de tratamiento que se llevará a cabo. Por eso, la ETAP se considera como la instalación fundamental en relación con la calidad del agua.

El suministro del agua se lleva a cabo mediante la RDA, que es la parte del sistema de abastecimiento que transporta el agua directamente hacia los puntos de

consumo (edificios, industrias, bocas de riego e incendio, etc.). Está constituida por todo un conjunto de tuberías, piezas especiales y elementos dispuestos y ordenados de forma conveniente para garantizar el abastecimiento en las condiciones de caudal, presión y calidad suficientes para cada uno de los usos. Los elementos más característicos que conforman la red de distribución, en su etapa final de abastecimiento, son las tuberías, las válvulas y los elementos de medición y control del sistema.

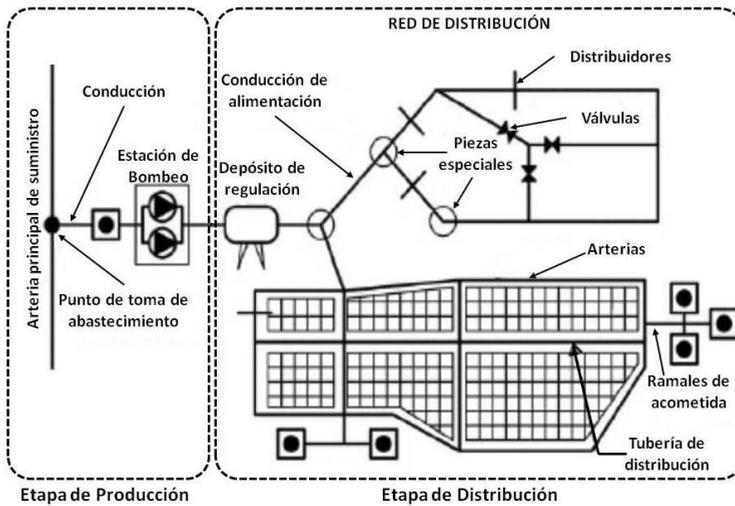


Figura 2.1: Esquema general de un sistema de abastecimiento de agua (Iglesias Rey et al., 2003)

Como señala Mora-Rodríguez (2011), la calidad del agua está relacionada con todos los elementos que conforman el abastecimiento y de la integridad de la red de distribución. Después del tratamiento en la ETAP, la calidad del agua estará influenciada por las condiciones en las que se encuentren los elementos de transporte, almacenamiento y distribución; sin embargo, se deben mantener las condiciones idóneas para su consumo. Por ello, en la etapa de distribución se instalan estaciones de cloración que se encargan de preservar la desinfección en zonas donde el contenido de cloro no sea suficiente para mantener el nivel óptimo. Así, los niveles de calidad del agua potable dependen de dos aspectos principales: primero, de la calidad que presente a la salida de la ETAP y segundo, de las condiciones existentes durante la distribución del agua que hacen variar su calidad a lo largo del trayecto hasta el punto de consumo.

Son también considerados como parte del sistema de abastecimiento de agua, los sistemas de medición, de control y de comunicación, es decir, todos los dispositivos de medición y monitorización de flujos, presión, calidad del agua y otros parámetros operacionales, incluyendo también elementos de control y *software* computacional como el SCADA (*Supervisory Control and Data Acquisition*).

2.1 La vulnerabilidad de los sistemas de distribución de agua

Cada uno de los componentes y subsistemas que constituyen un SDAP es un punto potencialmente frágil para su seguridad global. Su distribución espacial y accesibilidad influyen en el grado de vulnerabilidad frente a los diferentes peligros y amenazas a los que está expuesto. De esos subsistemas, la RDA es una de las partes más vulnerables, ya que un desajuste en su funcionamiento o una ruptura o avería de sus componentes podría deshabilitar o interrumpir el suministro de agua. Además, por la RDA es posible que se lleve a cabo la propagación de un contaminante, ya sea de forma accidental o intencional, y afectar la salud de los usuarios. En este sentido, cualquier persona podría forzar la entrada de un contaminante dentro de la RDA mediante el uso de un equipo relativamente pequeño y asequible (por ejemplo, una bomba) y desde lugares ventajosamente accesibles y privados como una casa, un hidrante de extinción de incendios o una estación de bombeo. Los efectos de un suceso como ese, dependerán del diseño hidráulico de la RDA y de sus condiciones de demanda.

De forma general, los daños en un SDAP pueden incluirse dentro de tres grandes categorías (Pérez-García et al., 2009):

- 1) Daños físicos en el sistema, que pueden dejar fuera de funcionamiento la totalidad o parte de este; están incluidos los daños en las fuentes de abastecimiento, las plantas de tratamiento, los depósitos de almacenamiento, las estaciones de bombeo, y los sistemas de conducción, transporte y distribución.
- 2) Daños en los sistemas de información y comunicaciones, que pueden ser elementos de medida y seguridad, líneas de transmisión y todo el *software* y *hardware* empleado para la operación y control del sistema de abastecimiento.

3) Daños a los usuarios del sistema, incluidas todas las anomalías en la calidad del agua servida, bien sea por fuentes biológicas, químicas o radiactivas.

El amplio conjunto de peligros y amenazas asociado a los SDAPs hacen evidentes las debilidades del sistema, siendo casi imposible su protección total contra todo tipo de contingencias. Sin embargo, sí es posible analizar y estudiar las posibilidades de fallo y sus consecuencias para prevenir los efectos adversos de cualquier suceso que implique un riesgo a los usuarios, tanto si proviene de una causa accidental como provocada.

En general, los peligros y amenazas son eventos extremos, naturales o intencionales, que influyen negativamente en el funcionamiento normal de los SDAPs. Las amenazas y los peligros son la causa de posibles daños o situaciones con un potencial de daño. Los peligros están relacionados con los efectos de accidentes, mientras que las amenazas se refieren a eventos deliberados. Tanto los peligros naturales como las amenazas humanas son representados en términos de su probabilidad (ocurrencia) y su severidad (consecuencias).

Coburn et al. (1994) señalan que una amenaza es un evento anormal o extremo, ya sea en el entorno natural o en uno creado por el hombre, que tiene efectos adversos en los sistemas de ingeniería e incluso en la vida humana. Las amenazas suelen encadenarse unas con otras e incrementar la posibilidad de causar daños. Un peligro latente está asociado a un fenómeno de origen natural o antropogénico, que puede manifestarse en un sitio específico y en un tiempo determinado, y producir efectos adversos en las personas, los bienes, los servicios y el medio ambiente. Por lo tanto, un peligro puede ser expresado como la probabilidad de ocurrencia de una amenaza, con una cierta intensidad, en un sitio determinado y en un período de tiempo específico. Sin embargo, en los SDAPs los peligros o amenazas no necesariamente están asociados a eventos de tipo anormal o extremo, ya que el riesgo puede ser alto incluso si el peligro es moderado debido a las grandes vulnerabilidades de los componentes del sistema (Li, 2007).

Las causas naturales pueden afectar a una porción importante del sistema, pero la extensión e intensidad del daño está íntimamente ligada a la probabilidad de ocurrencia del fenómeno que conduce al daño (Matalas, 2005). Esto hace que el fenómeno sea dependiente de situaciones pasadas y, así, permite valorar la probabilidad del suceso y su período de retorno. Es por esto que los peligros naturales han sido estudiados extensivamente y existe información suficiente y disponible, para poder calcular su probabilidad de ocurrencia con cierto grado de fiabilidad, así como también calcular y valorar el riesgo producido. En cuanto a las amenazas por actos humanos, envuelven incertidumbres inherentes y, en el caso de los daños producidos de forma deliberada, no existe la posibilidad de establecer un patrón probabilístico de ocurrencia debido a que, pese a la repercusión de los actos deliberados (entre ellos los actos terroristas), el impacto real es muy inferior a su impacto psicológico. Además, los ataques intencionales constituyen un conjunto de eventos independientes entre sí, más condicionados por la coyuntura socio-política y por la oportunidad que por la cronología previa de otros ataques (Pérez-García et al., 2009).

De acuerdo con Mays (2004) y Grigg (2003), las posibles consecuencias de los peligros y amenazas antes mencionados se presentan resumidas en la Tabla 2.1.

Tabla 2.1: Peligros y amenazas asociados a los SDAPs y sus posibles efectos.

Naturales	Consecuencias
Terremotos	<ul style="list-style-type: none"> ▪ Rotura de tuberías. ▪ Pérdida de presión. ▪ Colapso de la estructura.
Inundaciones	<ul style="list-style-type: none"> ▪ Pérdidas de la planta de tratamiento ▪ Contaminación del sistema de distribución
Sequías	<ul style="list-style-type: none"> ▪ Escasez de agua. ▪ Exceso de polvo. ▪ Problemas de la calidad del agua.
Viento (huracanes, tornados)	<ul style="list-style-type: none"> ▪ Problemas inducidos por las inundaciones. ▪ Daños en la estructura. ▪ Pérdida de energía.
Clima severo (frío, calor, nieve, hielo)	<ul style="list-style-type: none"> ▪ Tuberías heladas. ▪ Interrupciones y pérdidas (fugas). ▪ Aumento de la demanda del agua.

Por causa humana		Consecuencias
Ataques	Físicos	<ul style="list-style-type: none"> ▪ Destrucción física de los bienes del sistema o interrupción del abastecimiento de agua. ▪ La pérdida de presión del agua compromete las capacidades de los servicios contra incendios y podría llevar a una posible acumulación de bacterias en el sistema. ▪ Potencial para crear un fenómeno transitorio mediante la apertura y cierre de las válvulas de control más importantes y encendido y apagado muy rápido de las bombas que pudiera provocar roturas simultáneas de la tubería.
	Cibernéticos	<ul style="list-style-type: none"> ▪ Interrupción física de la red SCADA. ▪ Ataques en el sistema de control central para provocar fallos simultáneos. ▪ Ataques usando virus y gusanos informáticos. ▪ Cambiar u ocultar datos para neutralizar la cloración o no añadir desinfectantes permitiendo la aparición de microbios.
	Químicos y Biológicos	<ul style="list-style-type: none"> ▪ Afectación grave a la calidad del agua. ▪ Problemas de salud y muertes entre los usuarios del sistema. ▪ Pánico. ▪ Pérdida de confianza de los consumidores.
Accidentes		<ul style="list-style-type: none"> ▪ Contaminación de fuentes de agua. ▪ Pérdida de instalaciones y estructuras vitales. ▪ Roturas e interrupciones del suministro.
Fallos		<ul style="list-style-type: none"> ▪ Pérdida de la capacidad y funcionalidad del sistema. ▪ Pérdida del control del sistema.
Amenazas psicológicas		<ul style="list-style-type: none"> ▪ Pérdida de la confianza pública. ▪ Incitación al pánico.

2.1.1 Peligros naturales

Dentro de los peligros originados por fenómenos de la naturaleza y sus elementos, los terremotos, las inundaciones y las sequías son los eventos más

importantes que pueden causar grandes daños y pérdidas en el servicio de abastecimiento.

- Terremotos

Los principales peligros de un terremoto incluyen fallos en el terreno debido a la licuación o deslizamiento de tierras. Además, producen asentamientos que traen efectos negativos sobre las instalaciones y el equipo por el temblor de la tierra y la ruptura superficial de la falla. La mayoría de los daños son producidos a las tuberías de transmisión y distribución en áreas que experimentaron deformación en el terreno, como resultado de deslizamiento o ruptura de la falla. Las tuberías fabricadas de materiales quebradizos como por ejemplo, asbestos, cemento o concreto son las que sufren más averías que las tuberías soldadas o de acero dúctil (National Research Council, 2004).

En un terremoto, las instalaciones de la planta potabilizadora también experimentan daños, pero son mucho menores que el daño producido a las tuberías. Sin embargo, el potencial para liberar gas clorhídrico puede ser una importante preocupación de seguridad en las plantas de tratamiento. Los SDAPs son esenciales cuando ocurren incendios a causa de terremotos, porque pueden ser necesarias grandes cantidades de agua para extinguirlos. En relación con esto, uno de los aspectos más importantes a considerar, es la respuesta de las paredes de las tuberías de gran diámetro a la propagación de las ondas sísmicas. Por ello, son necesarios los métodos para mejorar la caracterización de la interacción suelo-tubería junto con la validación de los ensayos a gran escala.

- Inundaciones

Las inundaciones son fenómenos naturales que tienen como factores a la lluvia o el crecimiento anormal de la marea, provocando el aumento de los niveles de los ríos y los cuerpos de agua. En general, los daños ocasionados por inundaciones a los SDAPs pueden asociarse, entre otros, a la destrucción total o parcial de captaciones localizadas en ríos o barrancos, al azolve y colmatación de componentes por arrastre de sedimentos, a la pérdida de captación por cambio del cauce del río, a la rotura de tuberías expuestas en la zona de captación, a la rotura de tuberías de distribución y conexiones en las áreas costeras debido a la agresión de marejadas y en áreas próximas a cauces, a la contaminación del agua

por intrusión en las tuberías debido a roturas, daños a los equipos de bombeo por contacto con el agua., etc. Colateralmente, hay impactos indirectos como la suspensión de energía eléctrica y comunicaciones.

- Sequías

Las sequías son períodos prolongados de escasez de agua en ciclos climáticos naturales, que se originan en un conjunto complejo de elementos meteorológicos que actúan en el suelo y en la atmósfera. Esto determina la alteración en el balance hídrico de una zona o localidad, haciendo que el recurso agua sea insuficiente para satisfacer los requerimientos de consumo para riego, para generación de energía eléctrica y para consumo doméstico. La escasez de agua tiene diferentes consecuencias dependiendo de factores como el clima y la capacidad de regulación natural, dentro del ciclo hidrológico particular en estudio. También influyen de manera notable la infraestructura hidráulica y la gestión realizada. Cuantificar los daños que provocan las sequías es complejo, ya que tales daños no son, generalmente, estructurales y se dejan notar en diferentes sectores. Debido a la diversidad de los impactos y a la imprecisión en el tiempo de inicio y fin del fenómeno, la concepción de las sequías se plantea con términos poco precisos, los cuales son adaptados a las necesidades de los diferentes gestores y planificadores. Los efectos esperados en los SDAPs son: la pérdida o disminución del caudal de agua superficial y/o subterránea, el racionamiento y/o suspensión del servicio de abastecimiento, el incremento en los costos del agua debido a que el usuario tiene que comprarla ya sea embotellada o en camiones tanque, con la consecuente pérdida de su calidad, etc.

La experiencia de los servicios de abastecimiento de agua en cuanto a inundaciones y terremotos es extensa, y la mayoría de las consecuencias han sido identificadas y estudiadas. El viento es un peligro común, pero el daño más importante al servicio debido al viento, son las inundaciones provocadas por este. Otros desastres naturales, como el clima extremo y los rayos o relámpagos de las tormentas, también amenazan al sistema, pero no tan significativamente como un terremoto.

2.1.2 Amenazas intencionales

A través de la historia, el recurso hídrico ha sido un objetivo estratégico en diversos conflictos. Estos conflictos han sido, según la clasificación de Gleick y Heberger (2011), por las siguientes razones:

- a) por el control de los recursos hídricos cuando el abastecimiento de agua o el acceso a ella es el origen del problema;
- b) por el uso del agua como herramienta militar, debido a que los recursos hídricos o los sistemas de agua son usados por una Nación o Estado como un arma durante una acción militar;
- c) por el uso del agua como una herramienta política, donde los recursos hídricos o los propios sistemas de agua son utilizados por una Nación, Estado o ente no estatal como un objetivo político;
- d) por actos relacionados con el terrorismo, cuando los recursos hídricos o los sistemas de abastecimiento de agua, o bien, son objetivos de ataque, o son instrumentos de violencia o de coerción por parte de entes no estatales;
- e) porque el agua es considerada un blanco militar, pues los sistemas de recursos hídricos o los sistemas de abastecimiento de agua son blancos de acciones militares de Naciones o Estados; y
- f) por controversias Estatales y no Estatales; en las que los recursos hídricos o los sistemas de abastecimiento son un motivo importante de conflicto en el contexto del desarrollo económico y social de una nación.

El concepto de servicio estratégico referido al abastecimiento de agua prevalece en prácticamente todos los países (Pérez-García et al., 2009). En relación con esto, en 1941 el entonces director del *Federal Bureau of Investigation* (FBI) de los EE. UU., J. Edgar Hoover señaló: “Se ha reconocido desde hace mucho que, entre los servicios públicos, los sistemas de abastecimiento de agua son un punto particularmente vulnerable para el ataque de agentes extranjeros, debido a la posición estratégica que ocupan, manteniendo en marcha las ruedas de la industria y preservando la salud y la moral del pueblo americano”. Asimismo, en diciembre de 2004 el Consejo Europeo aprobó, a propuesta de la Comisión Europea, la iniciativa del *European Program for Critical Infrastructure*

Protection (EPCIP), siendo el SDAP una de las infraestructuras incluidas en la iniciativa. Posteriormente, mayo de 2007, el Gobierno español aprobó iniciativas como el Plan Nacional para la Protección de las Infraestructuras Críticas (PNPIC), la elaboración de un primer Catálogo Nacional de Infraestructuras Estratégicas, la aprobación de un Acuerdo sobre Protección de Infraestructuras Críticas y finalmente, en noviembre de 2007, la creación del Centro Nacional de Protección de Infraestructuras Críticas (CNPIC), que se ocupa de un amplio catálogo de instalaciones o sistemas considerados ICs. Aunque la lista es secreta, debido a la lógica discreción que rige las actuaciones en materia de seguridad, es obvio suponer que la lista considera los SDAPs, incluyendo los procesos de captación y potabilización y al sistema de distribución.

La infraestructura de abastecimiento de agua está expuesta a daños por causas naturales y por actos humanos malintencionados. Si bien el agua es potabilizada en la ETAP antes de su suministro, su calidad puede deteriorarse dentro del sistema de distribución y originar daños en la salud de la población que la consume. Dentro de las principales causas que resultan en brotes de enfermedades transmitidas por el agua, además de los peligros naturales, están los fallos en el sistema, el desarrollo de biopelícula (Ramos-Martínez et al., 2014), la entrada de contaminantes en el sistema de distribución, errores de operación y ataques malintencionados.

De forma general las amenazas por actos humanos son clasificadas en cuatro grandes grupos que son: accidentes, fallos, amenazas psicológicas y ataques (Grigg, 2003).

- Accidentes

Los accidentes envuelven acciones no intencionales contra el servicio de abastecimiento, tales como transporte, construcción, liberación de materiales tóxicos, incendios, etc., pudiendo causar importantes daños o desperfectos. Son ejemplos, la contaminación producida por un accidente en una planta nuclear; un accidente de operación o construcción que deshabilite un embalse, una tubería principal de conducción o una planta de potabilización; la contaminación de una fuente de abastecimiento debido a un vertido accidental durante el transporte de sustancias tóxicas, entre otros.

- Fallos

En cuanto a los fallos en un SDAP están incluidos entre otros, las roturas de algunos elementos como tuberías y válvulas; y los fallos en el suministro de energía o en el sistema de cómputo. Los fallos son amenazas de origen humano en el sentido de que, pueden deberse a una mala operación por desconocimiento del funcionamiento del propio sistema y de sus elementos, lo que causa una disminución de la eficiencia y el rendimiento deseado en el suministro.

- Amenazas psicológicas

Las amenazas psicológicas si bien no son la principal preocupación, deben ser controladas antes de provocar pánico en los consumidores. Son aquellos actos que incitan al caos debido a la desinformación. Las falsas alarmas y llamadas de emergencia relacionadas, principalmente, con la calidad del agua o bien, con la disponibilidad del recurso.

- Ataques

Los ataques a los SDAPs son las amenazas de mayor importancia y preocupación. Estos pueden ser causados por actos de terrorismo, vandalismo o sabotaje. Los ataques incluyen el uso de varios grados de violencia, fuerza, terror, intimidación o medios tecnológicos para amenazar o dañar ilícitamente el suministro de agua. La protección básica contra dichos ataques (o amenazas de ataques) es el sistema de seguridad. Esto incluye, la evaluación del peligro, la evaluación de la vulnerabilidad, la planificación de medidas de mitigación y las acciones de respuesta específicas en caso de crisis. Por lo tanto, las claves de la seguridad son la detección, el retraso y la respuesta.

Los ataques a los SDAPs se clasifican en tres tipos: ataques físicos, ataques cibernéticos y ataques químicos o biológicos (Grigg, 2003; Haines, 2006; Haines, et al. 1998; Mays, 2004). Cada escenario de ataque tiene características únicas y requiere de una evaluación de la vulnerabilidad y de esfuerzos de mitigación diferentes. Las consecuencias potenciales de esos tipos de ataques no son necesariamente mutuamente excluyentes, ya que la destrucción física de uno o más componentes podría resultar en una catástrofe.

a) Ataques Físicos:

Son blancos de los ataques físicos, principalmente, las tuberías, las estaciones de bombeo, tanques de almacenamiento y otras instalaciones, lo que puede alterar el diseño del sistema, como la RDA, por la destrucción o degradación de sus componentes. Los eventos relacionados con ataques físicos que destruyen o alteran los componentes de un sistema de agua son más probables en comparación con otros, porque para su realización se requiere un mínimo o ningún conocimiento, ya sea para el uso de materiales explosivos o herramientas destructivas.

b) Ataques Cibernéticos:

Los avances en las tecnologías de cómputo, especialmente durante la década de 1990 han ayudado a los SDAPs a automatizar sus diferentes procesos, como los de potabilización y depuración, el suministro y bombeo, y los sistemas del manejo de la información (Panguluri et al. 2004). Sin embargo, esto ha llevado a que los sistemas estén cada vez más expuestos a las amenazas de tipo informático.

Un ataque cibernético ha sido definido como aquel que debilita la confidencialidad, integridad o disponibilidad del sistema de información. La protección del sistema debe ser contra diferentes modos de ataque, desde los más simples y comunes, como el uso inadecuado del equipo, hasta la inhabilitación total o parcial del sistema debido a un fallo ocasionado por un ataque directo al sistema principal, o un ataque hecho directamente por alguien no autorizado (*hacker*) que accede ilícitamente al sistema informático.

La importancia del sistema informático en un SDAP es que está vinculado a otros sistemas o bases de datos que contienen información relevante como, por ejemplo, al sistema financiero y de recursos humanos, al SCADA, al Sistema Computarizado de Mantenimiento y Gestión, entre otros. Dichos sistemas son considerados una parte de la infraestructura de IT (*Information Technology*) de los SDAPs. Los ciber-ataques a la infraestructura de IT pueden causar importantes daños financieros e interrupciones de las operaciones internas del

servicio; pero esos daños no se presentan inmediatamente en el abastecimiento de agua. Sin embargo, los ataques al sistema SCADA podrían tener un grave impacto inmediato en el abastecimiento de agua. Asimismo, una prolongada interrupción de la infraestructura de la IT podría también provocar complicaciones en el abastecimiento de agua.

Los ataques cibernéticos son principalmente al SCADA, ya que este permite a los operadores del servicio supervisar y controlar los procesos entre diversos sitios remotos (Qiao, 2005). Asimismo, los ataques son a la información y al sistema de gestión de la infraestructura del agua, para provocar pérdida de datos, dañar la información y los equipos.

Diversas metodologías y herramientas han sido desarrolladas para determinar las debilidades potenciales en la infraestructura del sistema de cómputo del servicio de abastecimiento de agua. Esas metodologías ayudan a evaluar la susceptibilidad ante ciertos ataques e identificar acciones correctivas que puedan mitigar el riesgo y la gravedad de las consecuencias (Panguluri et al., 2004). Al respecto, la *Risk Assessment Methodology for Water* (RAM-W) fue desarrollada específicamente para los SDAPs por la *American Water Works Association Research Foundation* (AWWA) y *Sandia National Laboratories*. En esta metodología el “árbol de fallos” del SCADA es uno de los componentes del proceso, el cual es una representación gráfica que muestra cada punto de vulnerabilidad que puede ser usado por un ciber-atacante para destruir y/o deshabilitar los componentes críticos del SCADA o interferir con la operación normal del servicio de abastecimiento.

Es esencial asegurar los sistemas de IT junto con el SCADA, especialmente si están conectados entre sí, ya que las debilidades en los sistemas de IT pueden ser utilizados para deshabilitar o interrumpir las operaciones del SCADA. El objetivo de tales metodologías es conducir al desarrollo, seguimiento y refuerzo de una seguridad efectiva, la cual es única para cada sistema, ya que una misma medida para todos no sería eficaz. Además, aunque se invierta una gran cantidad de tiempo y dinero en evaluar la vulnerabilidad de un sistema, a menudo es casi imposible garantizar todas las vulnerabilidades en un sistema. Por lo tanto, es

importante determinar las consecuencias de tipos específicos de ciber-amenazas y asegurar el sistema para esas amenazas específicas.

c) Ataques químicos y biológicos:

Los ataques químicos y biológicos generalmente no afectan a la estructura del sistema de abastecimiento, pero sí amenazan la salud de los usuarios debido a que los agentes se propagan con facilidad a través del agua. Sin embargo, las estrategias que reducen la vulnerabilidad para un tipo de ataque pueden incrementar la vulnerabilidad de otro. Por ejemplo, el aumento de la redundancia para disminuir las consecuencias de un fallo o ataque físico, podría aumentar la relación entre los elementos de la red de distribución y el riesgo de contaminación intencional.

La contaminación química o biológica en los SDAPs es considerada como la amenaza potencial más grave para los consumidores. Los agentes biológicos pueden ser dispersados a través de la RDA y llegar hasta los puntos de consumo ocasionando daños a la salud e incluso la muerte, ya que algunos contaminantes no son detectados hasta que se presentan una serie de síntomas. Aún sin haber impactos severos en la salud, el hecho de saber que ha sido violada la seguridad del sistema, trae como consecuencia la pérdida de confianza de los usuarios al servicio.

Existen diversas sustancias que pueden contaminar el agua, siendo de especial interés aquellas que no son detectadas fácilmente por el consumidor, capaces de originar una contaminación de un gran volumen de agua en comparación con el volumen suministrado en un área pequeña, y que ocasionen un grado significativo de morbilidad o mortalidad en los consumidores. Cabe diferenciar entre contaminantes que ocasionan un efecto agudo y aquellos de efecto crónico. Los primeros producen daños a corto plazo y los segundos a largo plazo. Los contaminantes que producen una respuesta aguda son de mayor preocupación debido a que presentan las características necesarias para ser utilizados como armas en un ataque malintencionado, es decir, alta solubilidad, ser incoloros, ser inodoros, insípidez, alta toxicidad, resistencia a desinfectantes en

concentraciones normales, resistencia a la ebullición durante uno a tres minutos, fácil adquisición y transporte y no requerir de equipos sofisticados para su uso.

El desarrollo de la industria química ha generado un gran número de sustancias que presentan características altamente peligrosas, que las convierte en posibles “armas químicas de guerra”, utilizadas para amenazar la seguridad y la salud de una población. Estos químicos son clasificados, de acuerdo con su efecto fisiológico, en agentes mortales y agentes de hostigamiento. Los agentes mortales, creados para causar un gran número de muertes, son clasificados en cuatro tipos: neurotóxicos, vesicantes, asfixiantes y sanguíneos. Los denominados de hostigamiento, no necesariamente causan muertes, pero sí provocan incapacitación de distintas maneras. Entre estos agentes se encuentran los siguientes: lacrimógenos, vomitivos, depresores, alucinógenos y estimulantes (Chastain, 2004).

En la Tabla 2.2, según (Mays, 2004), se presentan algunos de los agentes y venenos químicos industriales más importantes que causan daños en la salud. Las concentraciones agudas hacen referencia a medio litro agua consumida; asimismo, establece los umbrales máximos recomendados en caso de un consumo continuado por día.

Tabla 2.2: Daño potencial de algunos productos químicos en los SDAPs.

Agentes químicos (mg/L a menos que se indique lo contrario)	Concentración aguda 0,5 L	Máximo recomendado	
		5 L/día	15 L/día
▪ Agentes químicos de guerra:			
Cianuro de Hidrógeno	25	6,0	2,0
Tabun (GA, µg/L)	50	70,0	22,5
Sarin (GB, µg/L)	50	13,8	4,6
Soman (GD, µg/L)	50	6,0	2,0
VX (µg/L)	50	7,5	2,5
Lewisite (fracción de arsénico)	100-130	80,0	27,0
Sulfuro de Mostaza (µg/L)		140,0	47,0
Bencilato de 3-quinuclidinilo (BZ, µg/L)		7,0	2,3

Dietielamida de Ácido Lisérgico	0,050		
▪ Venenos químicos Industriales			
Cianuros	25	6,0	2,0
Arsénico	100-130	80,0	27,0
Fluoruro	3000		
Cadmio	15		
Mercurio	75-300		
Dieldrina	5000		
Fluoroacetato de Sodim		No proporcionado	
Parathion		No proporcionado	

En cuanto a las amenazas biológicas, existe una gran variedad de agentes patógenos clasificados principalmente en: bacterias (rickettsias), protozoos, virus y toxinas (tales como biotoxina botulínica, aflatoxina, ricina, entre otras) que son potencialmente resistentes a la desinfección por cloración, además de ser relativamente estables en agua durante largos períodos de tiempo. Aunque el agua proporciona dilución, una partícula flotante de cualquier tamaño podría usarse para dispersar patógenos en los sistemas de agua potable. Otros sistemas más sofisticados como las microcápsulas, también podrían ser usados para dispersar patógenos en los SDAPs.

El *Center for Disease Control and Prevention* (CDC) de EE. UU. (CDC, 2014) divide a los agentes biológicos en tres clases: A, B y C. Esta clasificación está relacionada con su peligrosidad y su probabilidad de uso, siendo los más dañinos los de clase A. La Tabla 2.3 detalla cada una de las clases y los agentes que están incluidos.

Tabla 2.3: Clasificación de agentes biológicos, según el CDC

Categoría A
<ul style="list-style-type: none"> • Agentes de alta prioridad, incluyen organismos que suponen un riesgo para la seguridad nacional debido a que: <ul style="list-style-type: none"> • pueden ser fácilmente diseminados o transmitidos de persona a persona, • causan altas tasas de mortalidad, y tienen el potencial de provocar un gran impacto en la salud pública, • pueden causar pánico y disturbios sociales,

-
-
- requieren de acciones especiales para lograr que el sistema de salud pública esté preparado para enfrentarlos.
 - Agentes (Enfermedad/agente):
 - Carbunco (Antrax) (*Bacillus anthracis*)
 - Botulismo (Toxina de *Clostridium botulism*)
 - Peste (*Yersinia pestis*)
 - Viruela (variola mayor)
 - Tularemia (*Francisella tularensis*)
 - Fiebre hemorrágica viral (Filovirus, p.ej. Ebola; Arenavirus, p.ej. virus Lassa y Machupo; Bunyavirus, Flavivirus)
-

Categoría B

- Agentes de segunda prioridad, se incluyen aquellos que:
 - son moderadamente fáciles de dispersar,
 - causan moderadas tasas de morbilidad y bajas tasas de mortalidad,
 - requieren mejoras específicas en la capacidad de diagnóstico del CDC y sistemas mejorados para la vigilancia de enfermedades.
 - Agentes (Enfermedad/Agente):
 - Brucelosis (especies de *Brucella*)
 - Toxina Epsilon de *Clostridium perfringens*
 - Amenazas para la seguridad de los alimentos (especies de *Salmonella*, *E. Coli* O157:H7, *Shigella*)
 - Muermo (*Burkholderia pseudomallei*)
 - Psitacosis (*Chlamydia psittaci*)
 - Fiebre Q (*Coxiella burnetii*)
 - Toxina de ricino (*Ricinus communis*-semillas de ricino)
 - Enterotoxina estafilocócica B
 - Tifus (*Rickettsia prowazekii*)
 - Encefalitis viral (alfavirus, p.ej., la encefalitis equina venezolana, la encefalitis equina oriental, la encefalitis equina occidental)
 - Amenazas contra la seguridad del agua (*Vibrio cholerae*, *Cryptosporidium parvum*)
-

Categoría C

- Agentes que tienen la tercera más alta prioridad, incluyen los patógenos emergentes que pueden ser manipulados para su diseminación masiva en el futuro, debido a su:
-

- disponibilidad,
- facilidad de producción y dispersión,
- potencial de causar altas tasas de morbilidad y mortalidad y un gran impacto en la salud pública.
- Agentes (Enfermedad/Agente):
 - Enfermedades infecciosas emergentes como el virus Nipah y el Hantavirus
 - Tuberculosis multirresistente
 - Virus de la encefalitis transmitida por garrapatas
 - Virus de la fiebre hemorrágica transmitida por las garrapatas
 - Fiebre amarilla

De acuerdo con la clasificación anterior, los agentes de mayor interés corresponden a los considerados dentro de la Categoría B, ya que están incluidos los patógenos transmitidos por agua y alimentos.

A modo de resumen, la Tabla 2.4 presenta información relacionada con el daño potencial de algunos de los agentes biológicos contaminantes más relevantes. Aunque existe información sobre éstos y otros agentes, aún se requiere investigar más sobre sus propiedades para caracterizar completamente su impacto, su estabilidad en el agua y su tolerancia a la desinfección, principalmente al cloro (Clark y Deininger, 2000).

Tabla 2.4: Agentes considerados como armas biológicas y su amenaza potencial en agua.

Agente	Tipo	Amenaza en agua	Estabilidad en agua	Tolerancia al Cloro
Antrax	Bacteria	Si	2 años (esporas)	Resistente (esporas)
Brucelosis	Bacteria	Probable	20-72 días	Desconocida
C. Perfringens	Bacteria	Probable	Común en alcantarillado	Resistente
Tularemia	Bacteria	Si	Hasta 90 días	Desactivada, 1 ppm, 5 min.
Glanders	Bacteria	Poco probable	Hasta 30 días	Desconocida
Melioidosis	Bacteria	Poco probable	Desconocido	Desconocida

Agente	Tipo	Amenaza en agua	Estabilidad en agua	Tolerancia al Cloro
Shigellosis	Bacteria	Si	2-3 días	Desactivada, 0,05 ppm, 10 min.
Cólera	Bacteria	Si	Buena supervivencia	Fácil eliminar
Salmonella	Bacteria	Si	8 días, agua dulce	Desactivada
Peste	Bacteria	Si	16 días	Desconocida
Fiebre Q	Rickettsia	Posible	Desconocida	Desconocida
Tifus	Rickettsia	Poco probable	Desconocida	Desconocida
Psitacosis	Rickettsia	Posible	12-24 horas, agua marina	Desconocida
Encefalomielitis	Virus	Poco probable	Desconocida	Desconocida
Fiebre hemorrágica	Virus	Poco probable	Desconocida	Desconocida
Viruela	Virus	Posible	Desconocida	Desconocida
Hepatitis A	Virus	Si	Desconocida	Desactivada, 0.4 ppm, 30 min.
Cryptosporidiosis	Protozoo	Si	Estable durante días o más	Oocistos resistentes
Toxinas botulimicas	Biotoxina	Si	Estable	Desactivada, 6 ppm, 20 min.
T-2 micotoxina	Biotoxina	Si	Estable	Resistente
Aflatoxina	Biotoxina	Si	Probablemente estable	Probablemente tolerante
Ricina	Biotoxina	Si	Desconocida	Resistente a 10 ppm
Enterotoxinas Estafilocócicas	Biotoxina	Si	Probablemente estable	Desconocida
Microcistinas	Biotoxina	Si	Probablemente estable	Resistente 100 ppm
Anatoxina A	Biotoxina	Probable	Desactivada en días	Desconocida
Tetrodotoxina	Biotoxina	Si	Desconocida	Desactivada, 50 ppm
Saxitoxina	Biotoxina	Si	Estable	Resistente a 10 ppm

En la mayoría de los SDAPs el riesgo potencial de los fallos en la calidad del agua nunca es cero, ya que los fallos pueden suceder de diferentes maneras. Por lo tanto, el diseño y mantenimiento de sistemas robustos son importantes para garantizar agua segura incluso cuando múltiples fallos ocurran simultáneamente. En relación con esto último, el concepto de “agua potable segura” significa que la gente no debe morir o enfermar gravemente por consumirla, pero esto no implica que no exista un riesgo asociado a la misma. La seguridad no es una condición absoluta, sino más bien relativa y en la práctica, la premisa de que el agua es segura, significa que el agua está a un nivel de riesgo tan bajo que los consumidores no tienen que preocuparse por contraer una enfermedad a través de ella (Hrudey y Krewski, 1995).

2.2 Evaluación de la vulnerabilidad de los sistemas de distribución de agua

Como ya se ha comentado en apartados anteriores, las infraestructuras críticas presentan características de sistemas complejos que requieren de estudio específico. Un gran número de métodos pueden ser considerados para evaluar la vulnerabilidad de las diferentes ICs que existen y, en particular, de los SDAPs. La elección del método depende del tipo de sistema, del objetivo de análisis y de la información disponible.

Según Akgun et al. (2010), los diferentes enfoques para abordar el análisis de la vulnerabilidad pueden clasificarse de forma general en dos grandes grupos: métodos cualitativos y métodos cuantitativos. Los métodos cualitativos, por lo general, se aplican en el subproceso de los estudios de la evaluación del riesgo. Estos métodos permiten la clasificación de los elementos en orden de importancia o la separación en categorías descriptivas de la vulnerabilidad y pueden ser utilizados para realizar una preevaluación de la misma. Sin embargo, la creciente importancia de la medición de la vulnerabilidad de una IC requiere de una medición más exacta para poder hacer frente a peligros y amenazas.

Según Kröger y Zio (2011b) la conceptualización de la evaluación de la vulnerabilidad enfocada a las ICs, puede generalizarse en tres procesos principales, que son:

1. Análisis del sistema y de sus propiedades
2. Cuantificación de los indicadores de la vulnerabilidad del sistema y la identificación de los elementos importantes
3. Aplicación de las mejoras técnicas y organizacionales

En relación con lo anterior, de forma general, los dos principales resultados que se obtienen de un análisis de vulnerabilidad son los indicadores de la cuantificación de la vulnerabilidad y la identificación de los elementos críticos o esenciales (Figura 2.2). Los indicadores proporcionan información complementaria ya que son parámetros que engloban las características estáticas y/o dinámicas del sistema completo. Por otro lado, la identificación de los elementos críticos está influenciada por el *ranking* con respecto a su eficiencia de conectividad y/o la contribución de los elementos en la propagación de los efectos por fallos en el sistema.

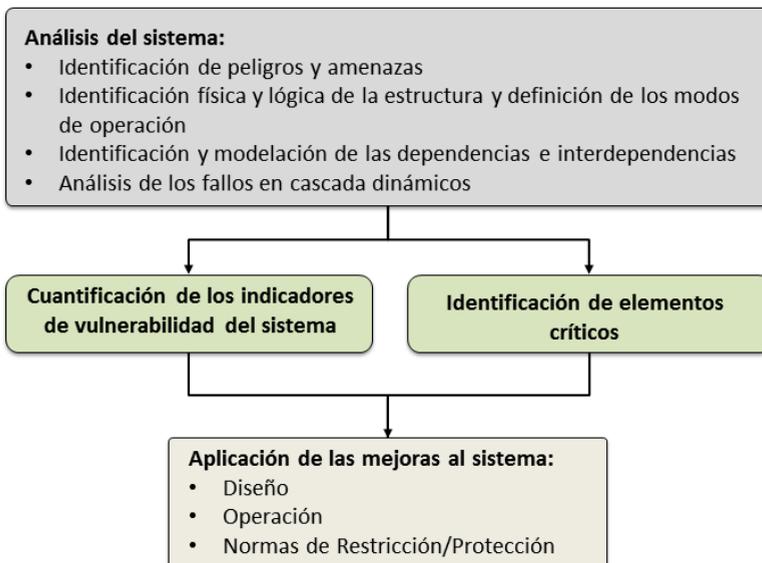


Figura 2.2: Conceptualización de la evaluación de la vulnerabilidad de las ICs

Como consecuencia del gran desafío que representa la evaluación de la vulnerabilidad de sistemas de IC, en los últimos años se han desarrollado diversos modelos y metodologías para el análisis y valoración de la vulnerabilidad de los SDAPs. Estos métodos han sido orientados principalmente hacia la obtención de parámetros y a la caracterización y rehabilitación de las tuberías. Sin embargo, el proceso de la medición de la vulnerabilidad resulta complejo debido a que varía en función del diseño y operación de cada sistema.

La *National Water Resources Association* (NWRA) define al proceso de evaluación de la vulnerabilidad como la identificación de las debilidades en la seguridad de un sistema, enfocándose en identificar las amenazas que puedan comprometer la capacidad para proporcionar el servicio de abastecimiento. Por su lado, Baker (2005) señala que la evaluación de la vulnerabilidad es un importante subconjunto de los procesos de evaluación del riesgo, que incluye la búsqueda en el sistema de elementos frágiles y sus modos de fallo, basados en un conjunto dado de amenazas.

Mays (2004c) señala los siguientes elementos comunes en todo proceso de la evaluación de la vulnerabilidad de los SDAPs:

- Caracterización del sistema de abastecimiento, incluidos su misión y objetivos.
- Identificación y priorización de las consecuencias adversas a evitar.
- Determinación de los bienes críticos que podrían ser objeto de actos malintencionados y que originen consecuencias no deseadas,
- Estimación de la probabilidad (probabilidad cualitativa) de tales actos malintencionados.
- Valoración de las medidas existentes para contrarrestar los eventos.
- Análisis de los riesgos actuales y el desarrollo de un plan de prioridades para la reducción de riesgos.

En resumen, el análisis de vulnerabilidad de un SDAP evalúa la resistencia de cada uno de sus componentes ante el impacto de ciertas amenazas (o fallos) en una determinada área. Además, la determinación de la criticidad detecta los componentes principales de la red, cuyo deterioro o remoción podría resultar en

importantes consecuencias negativas. El resultado final, definirá las medidas de mitigación necesarias y los procedimientos de emergencia o respuesta al impacto. Por lo tanto, analizar adecuadamente la vulnerabilidad es fundamental para establecer planes de emergencia y de mitigación adecuados que permitan organizar la preparación, las acciones a llevar a cabo durante la emergencia y los protocolos de rehabilitación de los sistemas. Finalmente, la mejora de la seguridad reduce la vulnerabilidad mediante la modificación de la topología del sistema y/o la asignación de recursos de seguridad disponibles.

2.2.1 Herramientas y métodos

Este apartado presenta una breve revisión de las herramientas y métodos más relevantes para la evaluación de los SDAPs, algunos apoyados por la *United States Environmental Protection Agency* (US-EPA). La US-EPA ha enfatizado en la prevención de actos de tipo intencional o terrorista y ha impulsado el desarrollo de instrumentos para reforzar los programas de protección que proporcionan información y directrices para mejorar la seguridad y reducir las diferentes debilidades de los SDAPs. Estos desarrollos están fundamentados, principalmente, en: los objetivos de la gestión del abastecimiento de agua; y en términos de prevención, detección, respuesta y recuperación frente a eventos de tipo natural o antrópico.

El programa de seguridad de la US-EPA contempla una metodología básica para la evaluación de la vulnerabilidad. La metodología se basa en un proceso paso a paso que valora las funciones de un SDAP y su capacidad para reducir el riesgo de diferentes peligros y amenazas. Este proceso permite identificar las debilidades de seguridad del SDAP y se enfoca en aquellas amenazas que le pueden impedir continuar proporcionando un suministro seguro y confiable. Los pasos que componen el proceso son los que a continuación se enlistan (Mays 2004c):

1. Evaluación del SDAP y sus componentes, mediante la identificación de las relaciones críticas entre ellos, así como los grupos de usuarios y sus necesidades de consumo específicas.

2. Identificación de las posibles amenazas a todo el sistema y a elementos específicos.
3. Consideración de las consecuencias, en función de las amenazas antes identificadas.
4. Cuantificación de la probabilidad de las consecuencias negativas.
5. Evaluación de las medidas de protección ya existentes, es decir, qué métodos son utilizados para detectar determinadas amenazas.
6. Planes de acción para reducir los riesgos. Implica el desarrollo de una lista de acciones que reduzcan la vulnerabilidad del sistema a las amenazas antes identificadas.

En relación con incidentes de contaminación, el *National Homeland Security Research Center* (NHSRC) de la US-EPA, en colaboración con un grupo interdisciplinario de científicos del *Sandia National Laboratories*, el *Argonne National Laboratory* y la Universidad de Cincinnati, plantearon un marco probabilístico para el análisis de la vulnerabilidad de SDAPs denominado TEVA (*Threat Ensemble Vulnerability Assessment*). El objetivo principal de este grupo consistió en investigar el potencial de los impactos de eventos de contaminación en los SDAPs, así como implementar nuevas técnicas y métodos para reducirlos. Los módulos centrales de TEVA son: la simulación de incidentes de contaminación, la evaluación de las consecuencias y, el diseño y evaluación de las estrategias de mitigación. Estos módulos están fundamentados en los seis pasos de la US-EPA y, en conjunto, permiten desarrollar una visión integrada de la vulnerabilidad de un SDAP frente a una variedad de amenazas de contaminación, así como el potencial para disminuir su vulnerabilidad.

Dentro del marco de TEVA, han sido desarrolladas e integradas otras herramientas como SPOT (*Sensor Placement Optimization Tool*) y CANARY⁷. Estas herramientas se desarrollaron para el diseño y optimización de un sistema de alerta temprana de contaminación. El *software* de optimización SPOT incorpora una variedad de métodos de solución para determinar y evaluar la

⁷ El nombre CANARY no es un acrónimo, pero sugiere un paralelismo con la historia del "canario en la mina de carbón", enfoque de detección de eventos en el que los mineros utilizan canarios para detectar indicios de gas venenoso. De manera análoga, el *software* CANARY evalúa los datos de calidad del agua en tiempo real obtenidos de diversos instrumentos y usa técnicas matemáticas y estadísticas para identificar la aparición de eventos de calidad de agua anómalos (Hart et al. 2007).

localización de sensores. La aplicación SPOT se compone de tres módulos principales. El primero, simula un conjunto de eventos de contaminación dentro del gran conjunto de eventos probables. El segundo módulo calcula el potencial de los efectos. Y el tercer módulo se encarga de la optimización para la ubicación de los sensores. TEVA-SPOT se ha utilizado para determinar el número y sitios óptimos de monitorización en un SDAP. Este *software* permite especificar un amplio rango de objetivos, siendo los más importantes: las medidas de salud basadas en la población, el tiempo de detección, la propagación de la contaminación, el volumen de agua consumida y el número detectado de eventos de contaminación. En relación con CANARY, es un algoritmo que evalúa la respuesta de los sensores de calidad y analiza en tiempo real los datos de salida. Asimismo, identifica cambios anormales en la calidad del agua que pueden indicar un evento de contaminación. Las posibilidades que ofrece CANARY son:

- 1) el uso de un formato estándar para operaciones de entrada y salida de datos de calidad de agua,
- 2) la capacidad de seleccionar diferentes algoritmos de detección (el programa contiene tres enfoques matemáticos diferentes para analizar los datos),
- 3) la posibilidad de seleccionar distintos sistemas y opciones de configuración de localización específica,
- 4) un modo de operaciones vía *online*, un modo de evaluación/entrenamiento vía *off-line*; y
- 5) la capacidad de generar los datos necesarios para establecer parámetros de rendimiento (por ejemplo, índices de falsas alarmas).

Este enfoque algorítmico mejora la sensibilidad de detección de los equipos de campo y al mismo tiempo, reduce alarmas de falsos positivos. Para más detalle sobre TEVA, SPOT y CANARY, se puede consultar a Berry et al. (2012), Hart et al. (2007 y 2008) y Murray et al. (2010).

Dentro del mismo contexto del análisis de la seguridad de los SDAPs, a continuación, se presentan algunos de los modelos más importantes desarrollados en los últimos años, compendiados por Li (2007), Moglia et al. (2006) y Pérez-García et al. (2009). Esto no pretende ser una lista de todas las

herramientas y programas disponibles, sino una visión general de los desarrollos más representativos.

- UTILNETS: Es un Sistema de Apoyo a la Decisión (DSS) para la planificación de la rehabilitación y para la optimización del mantenimiento de las tuberías de un sistema de abastecimiento. El modelo valora la fiabilidad de las tuberías a partir de predicciones estadísticas sobre su vida útil, y sirve para determinar las consecuencias del mantenimiento para optimizar la política de rehabilitación. El modelo fue producto de un proyecto europeo de investigación dentro del programa BRITE-EURAM, que comenzó aproximadamente en 1993.
- KANEW: El modelo está dirigido a los mismos objetivos de planificación de la rehabilitación del sistema de conducciones, basado en la edad y material de las tuberías. El modelo es producto de un proyecto de investigación de la *American Water Works Association Research Foundation (AWWARF)* de EE.UU. (1998).
- WARP: Se trata de una aplicación prototipo desarrollada inicialmente en el entorno *MS-Excel* y posteriormente convertida en un programa independiente. El modelo fue desarrollado en el año 2000 por el *National Research Council (NRC)* de Canadá. La aplicación D-WARP (*Distribution Water mAins Renewal Planner*) permite analizar estadísticas de rotura de tuberías y proyectar futuras tasa de roturas, así como estimar costes y planificar escenarios. El modelo tiene en cuenta factores temporalmente dependientes, tales como la temperatura, humedad del suelo y precipitación.
- PARMS: Acrónimo de *Pipeline Asset and Risk Management System*, es un conjunto de modelos para el apoyo de la gestión de sistemas de abastecimiento, desarrollado en el 2002 por la *Commonwealth Scientific and Industrial Research Organization (CSIRO)* de Australia. Utiliza modelos estadísticos para predecir fallos en las tuberías a partir de factores como el tipo de suelo, presión del agua, edad, material y diámetro de la tubería y otros.

- CARE-W: Este modelo nació en el año 2002 de una iniciativa de investigación sobre la rehabilitación de sistemas de tuberías dentro del V Programa Marco Europeo de I+D, con la participación de investigadores de varios países europeos. El modelo plantea un enfoque proactivo para el reemplazo de tuberías (mantenimiento preventivo) mediante el análisis de fiabilidad de las tuberías y el planteamiento de una metodología de priorización. Cabe destacar que, dentro del marco de este mismo proyecto, en el año 2003 se desarrolló el modelo CARE-S para sistemas de saneamiento.
- RAM-WSM (*Risk Assessment Methodology for Water Utilities*): Metodología para la medición del riesgo, desarrollada por el Departamento de Energía de los *Sandia National Laboratories* y financiado por la US-EPA. Esta herramienta compara entre sí los componentes del sistema para determinar cuáles son los más críticos. El método define la vulnerabilidad como una debilidad o deficiencia explotable de la seguridad de una instalación. La medición de la vulnerabilidad la realiza con base en el análisis de las características de los eventos indeseables, la accesibilidad de los eventos indeseables, las características de la seguridad y de las políticas, las medidas de protección, además de otras consideraciones. Los resultados de la evaluación se utilizan para obtener la valoración del riesgo de los componentes del sistema.
- VSAT (*Vulnerability Self-Assessment Tool*): Esta herramienta fue desarrollada por la *Association of the Metropolitan Sewerage Agencies* (AMSA) en colaboración con dos empresas de consultoría y con financiamiento de la US-EPA. El método se basa en la evaluación cualitativa del riesgo y evalúa la vulnerabilidad en una escala cualitativa (por ejemplo, muy alto, alto, moderado y bajo) mediante la consideración de unas medidas ya establecidas. Posteriormente, los resultados se utilizan para generar clasificaciones de criticidad a partir de las cuales se evalúan los niveles de riesgo de las instalaciones. Este método ha sido utilizado para la planeación de acciones de respuesta frente a eventos extremos para reestablecer las condiciones normales de funcionamiento.

- La *National Rural Water Association* (NRWA) y la *Association of State Drinking Water Administrators* (ASDWA), con asistencia de la US-EPA, desarrollaron la Guía de autoevaluación de la seguridad para sistemas de abastecimiento pequeños, de entre 3.300 y 10.000 consumidores. Esta guía proporciona un inventario de los componentes críticos del sistema y resuelve cuestiones generales sobre la evaluación de la vulnerabilidad para, entonces, obtener la vulnerabilidad de los componentes y priorizar las acciones que se llevarán a cabo.
- ASSET (*Automated Security Survey & Evaluation Tool*): El modelo ha sido desarrollado por la *New England Water Works Association* (NEWWA) en conjunto con la US-EPA y otras consultoras privadas. Esta herramienta está dirigida a sistemas que sirven entre 3.300 y 50.000 usuarios (sistemas pequeños y medianos). Es una herramienta *software* que se implementó en todos los sistemas de abastecimiento de Nueva Inglaterra. Consiste en un programa de autoguía diseñado para ayudar a obtener una medición completa de la vulnerabilidad, así como para mejorar la seguridad frente a ciertas amenazas. La evaluación comprende ocho pasos que son: 1) la recolección de información, 2) la identificación de la misión y los objetivos, 3) determinación de los componentes críticos, 4) evaluación de la amenaza, 5) evaluación de la seguridad física y de las medidas existentes, 6) análisis del riesgo, 7) plan de prioridades para la reducción del riesgo y 8) realización de un reporte final.
- AWARE-P. El *software* de gestión y planeación de infraestructuras (*Infrastructure Asset Management - IAM*) AWARE-P, para sistemas de agua potable, aguas residuales y pluviales está integrado en el marco del proyecto AWARE-P como proyecto I+D financiado por el Espacio Económico Europeo. El proyecto ha sido desarrollado y cofinanciado por un consorcio liderado por el Laboratorio Nacional de Ingeniería Civil (LNEC, Portugal), el Instituto Superior Técnico de Lisboa (IST), el SINTEF (de Noruega), Addition (Portugal) y YDreams (Portugal), así como por la Empresa Reguladora de Servicios de Agua y Residuos (ERSAR, Portugal). Es un entorno *software* organizado en el que se evalúan y comparan alternativas de planeación o proyectos en competencia, a través de medidas de desempeño, riesgo y

costes. Comprende un conjunto de herramientas de análisis que pueden ser utilizadas individualmente con fines de diagnóstico y sensibilidad, o como parte del proceso de planificación integrada del programa AWARE-P-IAM. Permite visualizar y desplegar información gráfica (en 2D y 3D), diagnosticar y evaluar a nivel global y local, cualquier sistema de abastecimiento de agua a través de su conjunto de modelos de desempeño, riesgo y costes. También, es posible comparar un sistema con un determinado número de alternativas de planeación o proyectos, utilizando métodos estandarizados que facilitan la selección y toma de decisiones, contrastar el estado actual del sistema y bajo diferentes escenarios. El *software* almacena diferentes *plugins* que son herramientas útiles: para la producción de índices de desempeño; para la creación de modelos estadísticos que permitan predecir futuros fallos de una red en un contexto de la estimación del riesgo y costes; para cálculo de la importancia de las tuberías de la red, en función del impacto de un fallo en un nodo de consumo; para el cálculo del riesgo de la interrupción del servicio, expresada como el volumen de la demanda no entregada en un sistema por un período de un año y, por último, para obtener un índice del grado de deterioro de la infraestructura.

2.3 Prevención y alerta temprana frente a eventos de contaminación

Como se ha mencionado en apartados anteriores, los SDAPs son geográficamente dispersos y con múltiples puntos de acceso que facilitan ataques de diversos tipos. Entre ellos, las amenazas físicas ocasionan grandes pérdidas económicas, pero su efecto en la salud de los consumidores es limitado. En cambio, en el caso de las intrusiones contaminantes intencionales, ya sean químicas o biológicas, las sustancias contaminantes pueden propagarse a lo largo de la RDA y perjudicar directamente la salud de los consumidores. Por esta razón, la contaminación intencional es considerada como la amenaza más seria para los SDAPs.

Sin embargo, para que los contaminantes consigan llegar hasta los grifos de los consumidores, la intrusión deliberada debe ocurrir en las fuentes de abastecimiento o dentro de la red de tuberías de distribución. Un contaminante

puede ser inyectado en la RDA mediante una bomba o un depósito móvil presurizado, capaz de superar la presión de todo el SDAP. La capacidad de monitorear y controlar la calidad del agua dentro del sistema es una de las mayores preocupaciones. Limitar el acceso físico a todos los componentes clave del sistema, además de ser prácticamente imposible, no es suficiente para reducir efectivamente el riesgo asociado a un evento que puede ocurrir en cualquier momento y lugar en todo el sistema.

En este contexto, diversas investigaciones en el ámbito de la seguridad de los SDAPs se han enfocado en la mejora de los sistemas de monitorización y detección temprana de contaminantes, con el objetivo de minimizar los posibles impactos, tanto económicos como en la salud pública. Es posible encontrar métodos, procedimientos no formales o guías para determinar la ubicación de estaciones de monitorización de la calidad del agua y establecer un sistema de alerta temprana (*Early Warning Contamination System - EWCS*). Un EWCS es un conjunto de estaciones de monitorización para el control de la calidad del agua en una RDA, que permite detectar una variedad de sustancias tóxicas y microorganismos infecciosos. El objetivo principal de un EWCS es identificar, en tiempo y de manera fiable, eventos de contaminación de baja probabilidad y de alto impacto, como son los ataques deliberados en la RDA (fuentes, depósitos y puntos de consumo); para permitir una respuesta local eficaz que reduzca o evite completamente los impactos que puedan resultar de un evento de tales características. En este sentido, un EWCS debe ser una parte integral del funcionamiento de un SDAP capaz de detectar no sólo contaminaciones intencionales, sino también contaminación por accidente o como resultado de eventos naturales.

Se considera que un EWCS es eficiente, si es capaz de analizar e interpretar los resultados obtenidos y, a su vez, utilizarlos para la toma de decisiones y el diseño de un plan acciones de respuesta para proteger la salud de la población y la infraestructura propia del SDAP. Según Grayman et al. (2004), un EWCS debe incluir un amplio rango de mecanismos que deben trabajar en conjunto para prevenir la entrada de contaminantes al sistema de distribución de agua; esos componentes son:

- Un mecanismo para detectar la probable presencia de un contaminante en la fuente de agua.
- Un medio de confirmación de la presencia de contaminación, determinando la naturaleza del evento y capaz de predecir cuándo afectará a la fuente en los sitios de toma y la intensidad de la contaminación en la toma.
- Un marco institucional conformado generalmente por una unidad central que coordine los esfuerzos asociados con la gestión del evento de contaminación.
- Vínculos de comunicación para la transferencia de información.
- Diversos mecanismos para responder a la presencia de contaminación en la fuente de agua, con el fin de mitigar su impacto.

En otro sentido, Clark et al. (2004) apunta que las características de un EWCS ideal y los requerimientos mínimos que éste debe cumplir son los siguientes:

- a) **Tiempo de respuesta rápido.** Advertir con tiempo suficiente para responder al evento de contaminación y prevenir la exposición de los usuarios al contaminante. El tiempo de respuesta de un EWCS es el instante desde que el contaminante tiene contacto con el sensor hasta el momento en que se produce un reporte y se inicia una acción de respuesta. El tiempo de respuesta dependerá de factores como: el sitio en el que el contaminante es introducido en el sistema, las características del sistema de tratamiento, los tiempos de retención y las propiedades del contaminante. En algunos casos no es necesario un tiempo de respuesta rápido, mientras que en otros el tiempo de respuesta debe ser inmediato.
- b) **Automatización.** Un EWCS debe requerir poca o nula intervención de un operador, así como también funcionar continuamente durante 24 horas. Además, el sistema debe permitir la operación a distancia, de tal manera que una respuesta en un sensor lejano pueda ser transmitida de forma inmediata a un sistema central de manejo de datos. En el caso de que un contaminante sea detectado, el EWCS debe inmediatamente de activar

una alarma y avisar al operador por vía electrónica o algún otro medio de comunicación.

- c) Cobertura de un amplio rango de contaminantes. Debido a que existe un gran número de sustancias que son una amenaza potencial, es casi imposible saber con anticipación cuál agente contaminante puede ser usado en un sistema de abastecimiento. Por ello, un EWCS debe tener la capacidad de abarcar un amplio rango de sustancias. Sin embargo, puede suceder que los métodos que son efectivos para determinados agentes, no lo son para distinguir entre sustancias peligrosas y sustancias inocuas. Por lo tanto, debe haber un balance entre la necesidad y la especificidad.
- d) Especificidad para contaminantes de interés. Un EWCS debe identificar positivamente agentes específicos que suponen una amenaza para la salud de los usuarios y, al mismo tiempo, ser capaz de diferenciar entre esas sustancias y otras relacionadas o muy similares, aunque no dañinas.
- e) Identificación del punto en el cual fue introducido el contaminante.
- f) Sensibilidad suficiente. El EWCS debe proporcionar un nivel de detección y cuantificación de un agente específico, de acuerdo con el límite de concentración más bajo considerado como una amenaza para la salud.
- g) Baja ocurrencia de resultados falsos. Se consideran resultados falsos cuando un EWCS indica un falso positivo o un falso negativo. En el caso de un falso positivo, el EWCS señala la presencia de un contaminante que en ese momento no está presente; por el contrario, un falso negativo es cuando un contaminante no es detectado pero que está presente a niveles que son significativos para la salud. Para ambos errores el porcentaje de ocurrencia debería ser cero, sin embargo, dado que siempre existe la probabilidad de que cualquier tipo de error ocurra, es

necesario establecer un porcentaje de aceptación para el error, tal que no afecte en la toma de decisiones.

- h)** Fiable y Robusto. El sistema de monitorización remoto debe ser capaz de soportar diversos factores externos. Es decir, el sistema puede verse expuesto a situaciones tales como cortes de energía, eventos climáticos extremos, actos de vandalismo o el continuo ensuciamiento de los componentes de los sensores.
- i)** Dotar de vigilancia continua durante todo el año.
- j)** Producción de resultados con una precisión aceptable.
- k)** Mínimo conocimiento y entrenamiento. El equipo que se utilice en el EWCS no debe requerir de un excesivo entrenamiento y conocimiento para la operación, el mantenimiento y la interpretación de los resultados.
- l)** Costo asequible. Que sea asequible para la mayoría de las empresas gestoras de abastecimientos de agua.

Sin embargo, un EWCS con todas esas características es poco viable. No obstante, existen algunas tecnologías que pueden ser utilizadas para construir un sistema de alerta temprana que cumpla con ciertos criterios mínimos como: dar una rápida respuesta, mantener suficiente sensibilidad y funcionar como un sistema automatizado que permita la monitorización remota.

En la actualidad se dispone de monitores de la calidad del agua, incluidos los análisis físico, químico, radiológico y microbiológico, así como los sistemas de bio-monitorización, que utilizan organismos vivos como indicadores de cambios en la calidad del agua. Algunos de los métodos de monitorización físicos y químicos más comunes que han sido propuestos para usarse en un EWCS, incluyen pruebas puntuales simples como turbidez, pH, temperatura, olor, conductividad, oxígeno disuelto; pruebas conjuntas como por ejemplo inmuno-ensayos para herbicidas; y monitorización más avanzada para químicos como

fluorescencia para aceites, cromatografía para petróleo y sus componentes, productos químicos orgánicos volátiles y fenoles (Hasan et al., 2004).

Existe un importante esfuerzo de investigación en el desarrollo de sensores más robustos y confiables para implementarse en un SDAP. Un ejemplo es la utilización de la tecnología basada en los satélites, ya que pueden ser útiles para la identificación de agentes químicos y biológicos en las fuentes de agua o para la transmisión rápida de datos desde las estaciones de monitorización. Asimismo, se desarrollan otras tecnologías como, por ejemplo, el bio-análisis por medio de sensores de *chip*, que ofrecen mejorar el acceso automatizado a datos en tiempo real para su inclusión en un EWCS.

La red de distribución representa la mayor vulnerabilidad de todo un sistema de abastecimiento de agua potable. El riesgo de contaminación (intencional o accidental) es alto y las consecuencias pueden ser graves para la población consumidora. Por ello, el mecanismo de protección más viable es un sistema de control y vigilancia de la calidad del agua. Sin embargo, su implementación y funcionalidad dependen de las características del sistema, de su distribución geográfica, del nivel de conocimiento de sus variables hidráulicas, de la variabilidad de la calidad del agua, de la forma de alimentación de la red (por una o por múltiples fuentes), de la suciedad y obstrucción de las tuberías, de la presencia de desinfectantes, de la localización de las estaciones de monitorización y de otras cuestiones como el factor económico.

2.3.1 Desarrollos orientados a la mejora de la seguridad de las RDAs

Por la complejidad y dinámica de los SDAPs, el principal reto en el diseño e implementación de un EWCS es la localización y densidad convenientes de sensores. La ubicación estratégica de los sensores es fundamental para una cobertura amplia y para la rápida detección de los contaminantes. Esta tarea se dificulta debido a que puede estar condicionada por temas económicos, tecnológicos o del funcionamiento del propio SDAP. El objetivo es que el sistema de alerta temprana logre integrar los datos obtenidos desde los sensores, junto con otros medios de detección (como, por ejemplo, con los sistemas de vigilancia

de las instituciones de salud pública y de seguridad, o con los registros de las reclamaciones de los clientes del servicio de abastecimiento). Con esto, se facilita la toma de decisiones y las acciones de mitigación para una respuesta eficaz a los incidentes de contaminación.

Como respuesta al complejo problema de la ubicación de sensores en una RDA, durante las últimas dos décadas se ha incrementado el interés y el número de trabajos relacionados con el tema. Muestra de ello, es el número de publicaciones encontrado en fuentes como SCOPUS y ScienceDirect, bajo los términos de búsqueda “*sensor placement in water distribution networks*”, siendo el resultado 200 y 131 publicaciones, respectivamente. Estos resultados recogen trabajos de ubicación de sensores para: la monitorización de la calidad del agua, la detección de la fuente contaminante, la detección y localización de anomalías en la presión, y en el SCADA. En el caso específico de sensores para calidad de agua, se encontraron 97 publicaciones en SCOPUS y 95 en ScienceDirect. Según la búsqueda realizada en SCOPUS, la trayectoria que ha seguido la investigación sobre el tema se puede observar en el gráfico de la Figura 2.3. El gráfico corresponde al total de 200 publicaciones (que incluyen a las 97 referidas específicamente a sensores de calidad de agua).

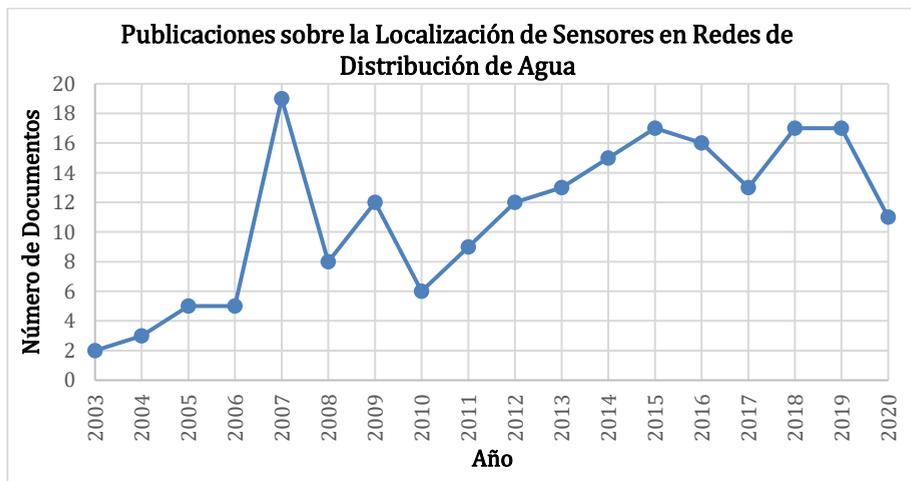


Figura 2.3: Número de publicaciones por año referentes al tema de localización de sensores en RDAs. (Fuente: SCOPUS).

La Figura 2.4 muestra la distribución porcentual de las diversas áreas del conocimiento en las que se ha abordado la ubicación de sensores en RDAs. Al observar dicho gráfico, cabe destacar la multidisciplinariedad que comprende el tema, siendo el área de Ingeniería la más significativa con un 29,2% de trabajos.

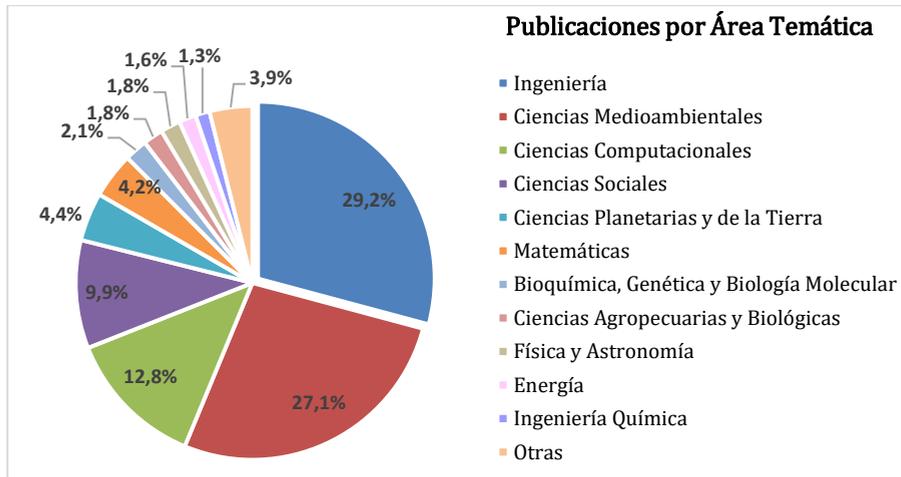


Figura 2.4: Porcentaje de publicaciones sobre localización de sensores en RDAs por área temática. (Fuente: SCOPUS, búsqueda del año 2003 al 2020).

Según Hart y Murray (2010), los diversos enfoques y soluciones referentes a la localización de sensores en las RDAs pueden ser agrupadas de forma general, en tres categorías que son: 1) opinión de expertos (Berry et al. 2005b), 2) métodos de *ranking* y análisis topológico (Bahadur et al., 2003; Chastain, 2004; 2006; Ghimire y Barkdoll, 2006a; Xu et al., 2008), y 3) métodos de optimización. En relación con estos últimos, se han generado diversos modelos y algoritmos de optimización, bajo suposiciones simplificadas, acerca de los objetivos de diseño de una red de sensores, el transporte de contaminantes, el tiempo de respuesta del sensor, la detección de eventos, la respuesta de emergencia, los costes de instalación y mantenimiento, etc. De esos trabajos, cabe resaltar las contribuciones surgidas de *The Battle of the Water Sensor Networks* (BWSN) propuesta por Ostfeld et al. (2008), en donde se analizaron y compararon aproximadamente 15 enfoques para el emplazamiento de sensores y se consideraron cerca de 90 publicaciones enfocadas, en su mayoría, en métodos de optimización. Algunas de esas contribuciones son: Alzamora y Ayala (2006), que propusieron un marco general para la localización de sensores aplicando

algoritmos topológicos. Berry et al. (2006) propusieron una formulación *p-mediana* adaptado de la teoría de localización discreta para definir el problema de localización de sensores, que fue resuelto utilizando adicionalmente un método heurístico. Dorini et al. (2006) plantearon un marco de optimización multiobjetivo restringido, basado en la metodología de entropía cruzada, titulado *noisy cross-entropy sensor locator (nCESL) algorithm*. Eliades y Polycarpou (2006) formularon una solución multiobjetivo mediante un algoritmo de profundización iterativa de soluciones de Pareto. Ghimire y Barkdoll (2006a; 2006b) propusieron un enfoque heurístico basado en la demanda, en el que los sensores fueron localizados en los nodos con demandas altas, o con el valor más alto de masa liberada. Aral et al. (2010) desarrolló una metodología de optimización mediante un algoritmo genético, basada en una sola función objetivo en la que fueron integrados cuatro objetivos de diseño. Gueli (2006) sugirió un modelo “depredador-presa” aplicado a optimización multiobjetivo, con base en un proceso de evolución. Huang et al. (2006) propuso un algoritmo genético multiobjetivo acoplado con minería de datos. Krause et al. (2006) aplicó un algoritmo para la ubicación de sensores definiendo una función sub-modular de maximización. Con esto, identificó una limitación en la formulación propuesta en BWSN y señaló que las funciones objetivo sólo se evalúan contra los escenarios que se han detectado, pero no se consideran los efectos de los escenarios no detectados que pudieran ser críticos. Ostfeld y Salomons (2006) y Preis y Ostfeld (2008) usaron el esquema del *Multiobjective Nondominated Sorted Genetic Algorithm-II* (NSGA-II). Propato y Piller (2006) y Propato (2006) utilizaron programación lineal mixta-entera, para resolver la localización óptima de los sensores. Trachtman (2006) propuso el enfoque denominado “hombre de paja” (*straw man*) para la localización de los sensores, en el que consideró la distribución de la población, patrones de la presión y flujo del sistema, la localización de consumidores específicos, entre otros factores. Wu y Walski (2006) utilizó una formulación de optimización multiobjetivo, la cual fue resuelta mediante un algoritmo genético, con eventos de contaminación generados aleatoriamente con un esquema de Monte Carlo.

Hart y Murray (2010) consideran que la formulación y la solución al problema de optimización de la localización de sensores, deben incluir los siguientes aspectos:

- Definición del riesgo de contaminación que es necesario minimizar;
- Descripción de las características de los sensores utilizados en el EWCS;
- Determinación de los objetivos de optimización;
- Formulación del modelo de optimización;
- Aplicación de una estrategia de optimización apropiada;
- Implementación del diseño.

Las publicaciones más relevantes relacionadas con cada uno de los puntos anteriores se presentan resumidas en la Tabla 2.5, la cual ha sido adaptada de Hart y Murray (2010). En ella se muestran cinco grupos, según el enfoque de los trabajos, que son:

- (1) cálculo del riesgo con simulaciones del transporte del contaminante,
- (2) modelación de fallos de los sensores o sensores imperfectos,
- (3) diseño multiobjetivo durante el proceso de optimización,
- (4) tipo de objetivo de optimización y
- (5) modelación de la incertidumbre de los datos.

Tabla 2.5: Resumen de los principales trabajos acerca de la ubicación óptima de sensores en una RDA

Autores	Grupo según el enfoque del trabajo				
	(1)	(2)	(3)	(4)	(5)
Al-Zahrani y Moied (2001, 2003); Kessler et al. (1998); Kumar et al. (1999); Lee y Deininger (1992); Lee et al. (1991); Uber et al. (2004)	no	no	no	cobertura	no
Berry et al. (2005a; 2005c); Krause y Guestrin (2009); Shastri y Diwekar (2006)	no	no	no	promedio	no
Carr et al. (2004, 2006)	no	no	no	robusto	si
Watson et al. (2004)	no	no	sí	promedio	no
Chastain (2004, 2006); Cozzolino et al. (2006); Ostfeld y Salomons (2003, 2004, 2005a; 2005b)	sí	no	no	cobertura	si

Autores	Grupo según el enfoque del trabajo				
	(1)	(2)	(3)	(4)	(5)
Berry et al. (2004, 2006, 2005d, 2007, 2008b); Hart et al. (2008); Propato (2006); Propato et al. (2005); Watson et al. (2005)	sí	no	no	promedio	No
Aral et al. (2010); Berry et al. (2008a) Dorini et al. (2006); Eliades y Polycarpou (2006); Guan et al. (2006); Gueli (2006); Hart et al. (2008); Krause et al. (2006, 2008); Leskovec et al. (2007); Preis y Ostfeld (2006a, 2008b); Wu y Walski (2006);	sí	no	sí	promedio	no
Krause y Guestrin (2009); Watson et al. (2006, 2009)	sí	no	no	robusto	no
Berry et al. (2008a, 2009); Preis y Ostfeld (2008b); Wu et al. (2008)	sí	sí	no	promedio	no

Además de la clasificación anterior, otros autores han propuesto enfoques diferentes para la solución al problema del emplazamiento óptimo de sensores. Entre ellos, Xu et al. (2008) que presentó una formulación en la que no utilizó simulaciones de la calidad del agua, sino que analizó la topología de los flujos de una RDA para identificar los lugares más importantes para la colocación de sensores. Preis y Ostfeld (2006b) plantearon un método de optimización con el objetivo de facilitar el análisis de los datos provenientes del sensor y para identificar el lugar o la fuente de inyección de un contaminante. Más adelante, Xu et al. (2010) propusieron una solución al problema de localización de sensores con un modelo de redes Bayesianas. En su trabajo consideran sensores imperfectos en la RDA y hacen inferencias sobre los eventos de contaminación con base en las señales del sensor. Su modelo propuesto se divide en dos etapas: 1) identificación de las ubicaciones óptimas de los sensores y 2) inferencia de la probabilidad de ocurrencia de un evento de contaminación y su posible fuente de contaminación.

En la literatura se pueden encontrar trabajos relacionados con la simulación de las RDAs y el desarrollo de algoritmos para la ubicación de sensores. Una cuestión más a resolver ha sido la elección del método de solución para optimizar la red de sensores; esto se debe a la variedad de métodos se han aplicado para abordar las diferentes formulaciones de optimización. Por ejemplo, la programación entera, los algoritmos genéticos y de búsqueda local y otros métodos de optimización heurísticos. Generalmente, la elección del método se basa en factores relacionados con la seguridad de la ejecución de la solución final, la capacidad computacional y el tiempo de funcionamiento disponible para llevar a cabo la optimización. Como propuesta al problema anterior, Xu et al. (2009) desarrollaron un protocolo de ayuda para determinar cuál formulación de optimización y cuál método de solución son los mejores, según las necesidades del sistema de distribución y de las capacidades de cómputo disponibles. Por su parte, Grayman et al. (2006) desarrollaron un marco interactivo para la toma de decisiones con el que las empresas gestoras del suministro de agua podrían evaluar las fortalezas y debilidades de diferentes diseños de redes de monitorización. Isovitsch y Vanbriesen (2007, 2008) implementaron una técnica basada en un Sistema de Información Geográfica (SIG) para analizar la disposición y la sensibilidad de los diferentes diseños y configuraciones de redes de sensores. Posteriormente, simularon eventos de contaminación aleatorios en una RDA que ya disponía de una red de sensores optimizada (Isovitsch y Vanbriesen, 2009). Su trabajo se centró en simular un aumento de la desinfección para originar el instante en el que la contaminación alcanza un sensor. Además, calcularon un rango de coeficientes de decrecimiento del contaminante, para simular su reacción con el desinfectante. Con este análisis mostraron que un sistema de impulso-respuesta puede ser efectivo para reducir significativamente el volumen de agua contaminada consumida.

En la última década, a partir del año 2010 hasta la actualidad, tal como se ha podido apreciar a partir de los datos de la figura 2.3, se ha mantenido el interés en el desarrollo de trabajos sobre la mejora de la seguridad de las RDAs. De acuerdo con Rathi et al., 2015, las áreas de investigación con mayor impulso, junto con algunas referencias significativas, son:

- a) la evaluación de los impactos de eventos de contaminación (Ciaponi et al., 2019; Davis y Janke, 2011 y 2008; Perelman y Ostfeld, 2010);
- b) la ubicación de sensores para la detección de contaminantes (Antunes y Dolores, 2016; Giudicianni et al., 2020; Herrera et al., 2011b; Di Nardo et al., 2018; Nazempour et al., 2018; Perelman y Ostfeld, 2013); (Brentan et al., 2021a; Cardoso et al., 2018; Gutiérrez-Pérez et al., 2017);
- c) la identificación de la fuente de contaminación (Arad et al., 2013; Oliveira et al., 2018; Perelman et al., 2012; Perelman y Ostfeld, 2013b; Shafiee y Zechman, 2013); y
- d) la limpieza o recuperación de la RDA después de un evento de contaminación (Ostfeld y Salomons, 2014; Poulin et al., 2010).

Dada la amplitud de este tópico de investigación, en esta tesis se aborda de manera exclusiva el problema de la ubicación óptima de sensores desde una perspectiva que hibrida el estudio topológico de la red, a través de elementos de la teoría de grafos, con la optimización evolutiva multiobjetivo, mediante el algoritmo ASO que refina, mejora y produce soluciones óptimas, a partir de las cuales los gestores de un abastecimiento pueden elaborar estrategias de monitorización y diseñar mecanismos de alerta temprana para la identificación y solución de eventos de contaminación en la red.

3 Enfoque multiobjetivo para la ubicación óptima de sensores de calidad en redes de distribución de agua

Como se ha mencionado anteriormente, las RDAs son infraestructuras complejas formadas por un gran conjunto de elementos (válvulas, bombas, depósitos, tuberías y otros elementos), que pueden alcanzar hasta miles de enlaces y uniones, cuya dispersión espacial y su dinámica propia hacen que su gestión y, en general, la evaluación de los problemas de calidad del agua, en particular la detección de intrusiones de contaminantes, sea un problema complejo de resolver. Si bien el objetivo principal de una RDA es suministrar agua en cantidad y calidad aceptables desde las fuentes hasta los consumidores, también es esencial mantener la RDA lejos de posibles escenarios de fallo (Herrera et al., 2016; Ostfeld et al., 2014), con capacidad de detectar rápidamente las fuentes de intrusión de contaminación (Di Cristo y Leopardi, 2008; Guan et al., 2006; Islam et al., 2015; Nafi et al., 2018; Preis y Ostfeld, 2006c; Tryby et al., 2010), habilidad para detección y minimización de fugas (Candelieri et al., 2014; Mamlook y Al-Jayyousi, 2003; 2018; Sarrate et al., 2014), etc. Para abordar estos temas, las empresas de servicios de agua necesitan instalar, en diferentes puntos de la RDA, sensores que midan, de manera continua y directamente, diversos parámetros. Por ejemplo, para el problema de las fugas se busca la ubicación óptima de sensores de presión y caudal (Francés-Chust et al., 2020; Nam et al., 2018). En términos de calidad, otros tipos de sensores deben medir la conductividad, el pH, y la concentración de diferentes sustancias. Como ya se especificó, el diseño de una red adecuada de sensores de calidad es el objetivo principal de esta tesis.

En el campo de la ubicación óptima de sensores para monitoreo de la calidad del agua, existen diversas consideraciones que condicionan su función, como el coste de inversión y la extensión de la RDA que debe ser cubierta. Esto plantea las dos cuestiones esenciales:

- ¿cuántos sensores deben ser instalados? y,
- dado un número de sensores, ¿en dónde deben ser emplazados?

Además, dada la gran variedad de contaminantes posibles, la identificación de las fuentes de contaminación dentro de la RDA y el tiempo de reacción de las

empresas de agua para hacer frente a un evento de contaminación, son también elementos importantes que considerar.

La ubicación óptima de sensores en RDAs urbanas ha representado en la última década un campo de investigación relevante y en auge, dedicado a optimizar los procesos de seguridad, control, gestión y mantenimiento de una red. Sin embargo, existen diversas cuestiones técnicas todavía por solucionar. Las más importantes, como se acaba de mencionar, son cuántos sensores instalar y, dado el número de sensores, dónde instalarlos. En este sentido, se debe realizar un análisis de compensación para poder responder la primera pregunta. Tener más sensores en la RDA representa tener que manipular más datos y, por tanto, permite un conocimiento más completo sobre el sistema. Sin embargo, tener más sensores también significa mayor inversión económica. El argumento económico hace que la elección de una estrategia de protección sea una decisión crucial, pues la selección de puntos adecuados de monitoreo puede aportar más y mejor información sobre el sistema a menor coste. Y, recíprocamente, una elección no adecuada echará por tierra una inversión en sensores haciéndola inservible o, en el mejor de los casos, poco eficiente. Considerando la compleja tarea de definir la posición y el número de sensores en una red de agua, es razonable que existan muchos trabajos en la literatura que presentan diferentes enfoques para el problema de la ubicación óptima de sensores (Adedoja et al., 2018; Rathi y Gupta, 2014). Esta circunstancia se puede comprobar en la abundancia de referencias en el primero de estos trabajos, que presenta una reciente revisión sistemática del estado del arte en la ubicación óptima de sensores para alertas de contaminación en RDAs.

Debido a la naturaleza de los objetivos involucrados, la decisión sobre dónde colocar los sensores en la red y la cantidad que se instalarán, es un problema difícil de solucionar. Es un hecho que, para conseguir los objetivos de calidad deseados, las empresas de servicios de agua están obligadas a medir los parámetros de calidad del agua continuamente, de modo que la calidad pueda ser monitoreada adecuadamente. Con este fin, es necesario diseñar una red óptima de sensores que cubra áreas y puntos de la RDA de una manera estratégicamente óptima. Parece obvio que los sensores no puedan colocarse aleatoriamente a lo largo y ancho de la red. Colocar sensores puede parecer una

tarea simple, pero teniendo en cuenta los costes de la compra, la instalación y el mantenimiento de los sensores, así como la extensión de las tuberías que deben cubrirse, el diseño de la red de sensores más adecuada resulta un problema complejo, ya que participan diversas condiciones que compiten entre sí, como veremos en este capítulo.

La contaminación se propaga en la RDA, originada principalmente, por los flujos de agua a lo largo y ancho del sistema. Por lo tanto, un buen conocimiento de los flujos de agua conduce a una buena comprensión de los escenarios de contaminación. No cabe duda de que, como lo sugieren Di Nardo et al. (2013 y 2014), el uso de un sistema simplificado (por ejemplo, la sectorización o división del sistema en los denominados Distritos Métricos (DMAs, por *District Metered Areas*) (Campbell et al. 2014; Ferrari et al., 2014; Herrera et al., 2012; Izquierdo et al., 2009), incluida la sectorización dinámica (Wright et al., 2014), puede dar lugar a respuestas rápidas a eventos extremos, como una intrusión contaminante. Obviamente, esto reduciría la cantidad de elementos necesarios para el monitoreo y, a su vez, de los gastos operativos.

En esta tesis no se pretende cubrir todos los temas necesarios para proteger una RDA contra posibles intrusiones contaminantes. Más bien, se propone una solución para el problema del diseño de una red de sensores óptima para medir la calidad del agua que, como hemos anticipado, considera la determinación del número de sensores y sus ubicaciones en la RDA. Este problema se aborda desde una perspectiva de optimización multiobjetivo. Los objetivos que se considerarán para resolver el problema son: tiempo de detección, volumen de agua consumida, tasa de fallo de detección y coste de la red de sensores. Trabajos relacionados con este enfoque son los realizados por Antunes y Dolores (2016), Cardoso et al. (2018), Cheifetz y Mandel (2018), Rathi y Gupta (2017), Zhao et al. (2016a) y Gutiérrez-Pérez et al. (2017), siendo este último, de producción propia, el que se toma como principal referencia y que corresponde a uno de las referencias citadas en la revisión sistemática del estado del arte en la ubicación de sensores para la calidad dada en Adedoja et al. (2018).

Empezamos por revisar brevemente los fundamentos de la optimización multiobjetivo. Después, enfocaremos de manera más precisa el problema que

nos ocupa. El nivel de concreción continuará con una definición clara de los objetivos a considerar y la forma en cómo tales objetivos permiten valorar distintas redes de sensores de calidad, que permitirá la optimización buscada que conduzca al diseño de la red de sensores de calidad finalmente adoptado.

3.1 Fundamentos de la optimización multiobjetivo

La mayoría de los problemas de optimización del mundo real son multiobjetivo; en particular, lo es el problema de la ubicación óptima de sensores en una RDA. Es decir, se consideran dos o más objetivos que se deben satisfacer simultáneamente, y que pueden – es lo habitual – estar en conflicto entre sí. Además, existen diversas restricciones⁸, que se deben cumplir en todo caso.

La solución de problemas de optimización multiobjetivo ha presentado históricamente dos tendencias. Por un lado, con el fin de simplificar la obtención de la solución, se plantea y soluciona el problema como mono-objetivo. Para ello, se utiliza sólo una función donde los objetivos individuales se ponderan; y/o se consideran algunos de los objetivos como restricciones, cuyo incumplimiento penaliza adecuadamente a la función mono-objetivo a optimizar. La otra tendencia es el uso de métodos que proporcionen no una única solución, sino un conjunto de soluciones que representen determinados compromisos entre los objetivos involucrados. Para la obtención de la solución finalmente adoptada, en ambas tendencias, se requiere contar con alguna información adicional. En el caso mono-objetivo esta información permite decidir los pesos asignados a los objetivos y/o los valores de penalización en la optimización. En el caso multiobjetivo, la información se utiliza para optar por alguna de las múltiples soluciones de compromiso encontradas. En la primera tendencia, el vector de pesos y/o los factores de penalización se dan *a priori*, y no existe una idea clara de los posibles efectos que dichos pesos y/o factores puedan originar sobre la solución obtenida. En el caso multiobjetivo, la información adicional se utiliza para seleccionar la solución más conveniente entre el conjunto de soluciones

⁸ Estas restricciones pueden ser restricciones de frontera o funciones de restricción. Las de frontera se usan cuando el valor de una variable de decisión está limitada a un rango. Las funciones restrictivas son más complejas y se dividen en restricciones de desigualdad y de igualdad.

óptimas obtenidas, y sobre el cual se pueden establecer más claramente las prioridades en la selección (Montalvo, 2010).

En general, un problema de optimización multiobjetivo puede formularse como: hallar el vector de las variables de decisión $\vec{X} = [x_1^*, x_2^*, \dots, x_n^*]^T$ que satisfaga m restricciones de desigualdad:

$$g_i(\vec{X}) \leq 0, \quad i = 1, 2, \dots, m \quad (1)$$

p restricciones de igualdad:

$$h_i(\vec{X}) = 0, \quad i = 1, 2, \dots, p \quad (2)$$

y que optimice, es decir, que minimice (o maximice) varios objetivos (o funciones objetivo), expresados de manera vectorial mediante:

$$\vec{f}(\vec{X}) = [f_1(\vec{X}), f_2(\vec{X}), \dots, f_k(\vec{X})]^T, \quad (3)$$

siendo k el número de objetivos (el superíndice T denota la transposición matricial/vectorial).

El propósito de la optimización multiobjetivo es encontrar distintos vectores de decisión \vec{X} , que minimicen (o maximicen) los objetivos, estableciendo equilibrios entre ellos y cumpliendo, en todo caso, las restricciones.

En la optimización multiobjetivo, la función objetivo de la ecuación (3) toma valores en un subconjunto de un espacio multidimensional al que se denomina espacio objetivo. Para cada punto del espacio de decisión o espacio de soluciones Ω (constituido por las variables de decisión), existirá un punto en el espacio objetivo. Esta definición se conoce comúnmente como optimalidad de Pareto en honor al científico del siglo XIX, Vilfredo Pareto⁹. Las soluciones óptimas de

⁹ La noción más aceptada de "óptimo" en el contexto de la optimización multiobjetivo es la propuesta originalmente por Francis Ysidro Edgeworth en 1881 y generalizada posteriormente por Vilfredo Pareto en 1896.

Pareto constituyen el llamado conjunto óptimo de Pareto. La imagen del conjunto óptimo de Pareto en el espacio objetivo se denomina frente de Pareto. En la Figura 3.1 se puede ver la correspondencia entre puntos del espacio de soluciones o espacio de decisión (izquierda) y sus imágenes en el espacio objetivo (derecha). Encontrar soluciones óptimas de Pareto para un problema multiobjetivo, a veces, también se denomina sin más optimización de Pareto. El objetivo de la optimización de Pareto es encontrar un conjunto de soluciones cuyas imágenes en el espacio objetivo constituyan un conjunto denominado frente de Pareto con una convergencia adecuada y una distribución/diversidad, lo más uniforme posible, a lo largo de dicho frente de Pareto (Kukkonen y Coello Coello, 2017).

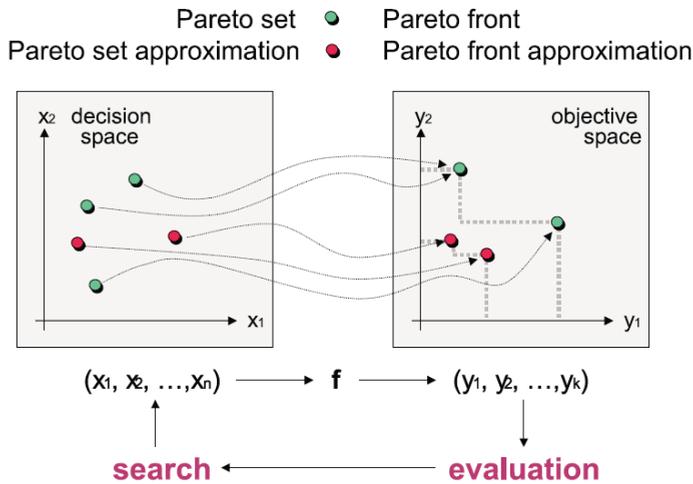


Figura 3.1: Ilustración de un problema general de optimización multiobjetivo (Zitzler et al., 2004).

Los conceptos clave de la optimización multiobjetivo son: dominancia de Pareto, conjunto de óptimos de Pareto y frente de Pareto. Consideraremos, para las definiciones siguientes, que cada componente de la función multiobjetivo debe minimizarse (obsérvese que esto puede conseguirse cambiando el signo de las funciones objetivo a maximizar).

Dados dos vectores de decisión \vec{X} e \vec{Y} , se dice que $\vec{X} = [x_1^*, x_2^*, \dots, x_n^*]^T$ **domina** a $\vec{Y} = [y_1^*, y_2^*, \dots, y_n^*]^T$, denotado como $\vec{X} < \vec{Y}$, si y sólo si \vec{X} es parcialmente menor que \vec{Y} , es decir, $\forall i \in \{1, 2, \dots, k\}, f_i(\vec{X}) \leq f_i(\vec{Y}) \wedge \exists i \in \{1, 2, \dots, k\} : f_i(\vec{X}) < f_i(\vec{Y})$.

El **conjunto de óptimos de Pareto** (\mathcal{P}^*), para un problema multiobjetivo dado, expresado por la función vectorial $\vec{f}(\vec{X})$, se define como:

$$\mathcal{P}^* := \{\vec{X} \in \Omega \mid \neg \exists \vec{X}' \in \Omega : \vec{f}(\vec{X}') \ll \vec{f}(\vec{X})\} \quad (4)$$

Para un problema multiobjetivo y un conjunto de óptimos de Pareto \mathcal{P}^* , el **frente de Pareto** (\mathcal{FP}^*) se define como:

$$\mathcal{FP}^* := \{\vec{f}(\vec{X}) = (f_1(\vec{X}), f_2(\vec{X}), \dots, f_k(\vec{X}))^T \mid \vec{X} \in \mathcal{P}^*\} \quad (5)$$

En la toma final de decisiones para la selección de la estrategia a implementar, se utilizan determinados criterios (con frecuencia no técnicos) que permiten a los tomadores de decisiones decantarse por alguna de las soluciones del frente de Pareto. No obstante, existen mecanismos que utilizan métodos de decisión multicriterio para realizar la selección de la solución final de una manera más sistemática (ver, por ejemplo, Brentan et al. (2021a; 2021b) y Carpitella et al., 2019)).

3.2 Planteamiento general de los objetivos para la ubicación de sensores de calidad en RDAs

Una vez más, recordamos que el principal propósito de una RDA es proporcionar agua segura a los consumidores en cantidad y calidad adecuadas. Sin embargo, la calidad del agua puede deteriorarse durante el transporte desde la planta de potabilización hasta el consumidor principalmente por: 1) la calidad química y biológica de la fuente de agua; 2) la efectividad y la eficiencia de los procesos de tratamiento; 3) la integridad de la planta de tratamiento, las instalaciones de almacenamiento y el sistema de distribución; 4) la antigüedad, tipo, diseño y mantenimiento de la red de distribución; 5) la calidad del agua tratada; y 6) mezcla de agua de diferentes fuentes y otras condiciones hidráulicas. Es por todo

esto que la calidad del agua debe ser monitoreada de manera rutinaria en varios puntos de una RDA con el objetivo último de lograr la detección temprana de contaminantes y el aislamiento de las áreas afectadas para que la exposición de los consumidores a la contaminación sea mínima. La determinación de la ubicación óptima de los sensores en una RDA es un desafío que requiere especial atención para proteger a la sociedad del consumo de agua de calidad deficiente o contaminada.

Teóricamente, los contaminantes se pueden introducir en una RDA en cualquier punto y en cualquier momento. Esto implicaría monitorear en cada uno de los nodos todo el tiempo para lograr una protección completa; pero esto en la práctica no es factible. Por lo anterior, se requiere cubrir el número máximo de consumidores con un número limitado de sensores considerando dos aspectos fundamentales: 1) el riesgo asociado a una cobertura no adecuada, y 2) la necesidad de maximizar la probabilidad de detección (PD) de eventos de contaminación. El riesgo se puede definir en términos de porcentaje de la población expuesta, de la longitud de tubería contaminada, o en términos del número de detecciones fallidas. Por otro lado, PD es la probabilidad de que al menos uno de los sensores detecte una contaminación. Un sensor puede detectar un evento de contaminación falso o no detectar un evento, si la concentración está por debajo de los límites de detección. Además, puede haber un retraso en la respuesta de los sensores establecidos. Por lo tanto, también se han considerado objetivos adicionales como: a) el tiempo de respuesta del sensor, que mide el tiempo transcurrido entre el registro de un evento de contaminación en un sensor y la respuesta que proporciona; b) la probabilidad de fallo de detección, que se define como la proporción de eventos no detectados por todos los sensores y, por último, c) la redundancia de detección del sensor, definida como la probabilidad de detección de un evento de contaminación por un número específico de sensores durante un tiempo específico.

De manera más específica, varios objetivos han sido propuestos para la elección de la ubicación de los sensores en las RDAs. De acuerdo con Adedoja et al. (2018) y Rathi y Gupta (2015), los principales objetivos para la ubicación de sensores o puntos de monitoreo en una RDA están relacionados con:

- a) conseguir una detección temprana de cualquier evento de contaminación, es decir, minimizar el tiempo de detección (TD) para cuya definición existen varias alternativas que se discutirán en apartados siguientes; implícitamente, este objetivo implica
- b) tratar de conseguir que ningún evento de contaminación quede sin identificar, para lo que hay que maximizar la eficiencia de detección;
- c) minimizar el impacto o las consecuencias de un evento de contaminación, que se pueden definir como:
 - el volumen de agua consumida (VC), definido como la cantidad de agua contaminada consumida por la población antes de la detección del evento de contaminación,
 - la población expuesta a la contaminación (PE) definida como el número de personas expuestas al evento de contaminación antes de su detección y, por último,
 - la extensión de la contaminación (EC) definida como la longitud de la tubería contaminada por un evento de contaminación.
- d) Finalmente, es obvio que un objetivo es la inversión necesaria para la monitorización de la red, objetivo que supone, claramente, la minimización del coste de la red de sensores y de su instalación, mantenimiento, etc.

Idealmente, estos objetivos deben alcanzarse con una solución de distribución de sensores adecuada. Sin embargo, es un hecho la naturaleza mutuamente conflictiva de los objetivos, que los hace estirar en direcciones no necesariamente coincidentes. En general, mejorar uno de ellos con certeza resultará en detrimento, total o parcial, de otro.

Por ejemplo, maximizar la cobertura de protección en la red requerirá aumentar la cantidad de sensores (esto significa el coste) o, probablemente, necesitar aceptar tiempos de detección más largos. En consecuencia, la solución final resultará de un compromiso entre objetivos, en lugar de una única "mejor alternativa". Es por ello que la resolución adecuada de problemas de esta naturaleza requiere el uso de un enfoque multiobjetivo. La idea principal no es encontrar una única solución óptima, sino encontrar (una aproximación de) el frente de Pareto cuyas soluciones representen buenas compensaciones entre todos los objetivos involucrados. La alternativa mono-objetivo de primero

deliberar sobre los objetivos y luego agregar su influencia ponderada en una sola expresión, para resolver un solo problema objetivo, probablemente resultará, en el mejor de los casos, en encontrar solo un punto del frente de Pareto, lo que parece facilitar la decisión. No obstante, este enfoque es equivalente a producir una compensación *a priori*, sin tener una idea clara de cómo la solución obtenida se relaciona con el resto de las soluciones potenciales del problema. Por ejemplo, este enfoque no puede responder preguntas de coste marginal, tales como: si es conveniente comprar un sensor adicional para obtener una mejora razonable en otro objetivo; esto es así porque en el enfoque mono-objetivo no hay forma de saber cuánta mejora en la cobertura de protección y el tiempo de detección aportaría ese sensor adicional. Esas son las preguntas que un enfoque multiobjetivo ayuda a responder. En la misma línea están las cuestiones y respuestas necesarias para encontrar una solución de colocación de sensores que represente una buena compensación entre todos los objetivos involucrados.

3.3 Definición de las funciones objetivo utilizadas

En esta tesis no se pretende cubrir todos los aspectos relacionados con la protección de la RDA contra cualquier posible intrusión de contaminantes. Más bien, la tesis se concentra en proponer una solución sólo para el problema de colocación de los sensores, es decir, determinar de manera óptima el número de sensores y sus ubicaciones. Como ya se ha dicho, este problema de optimización se aborda desde una perspectiva multiobjetivo. Y los objetivos específicos que se considerarán para resolver el problema de ubicación óptima de sensores son:

- se desea identificar los problemas de calidad del agua lo antes posible, esto es, **minimizar el tiempo de detección**;
- al menos un sensor siempre (idealmente) debería ser capaz de identificar un problema o fallo de calidad, sin importar donde se localiza la fuente de contaminación; esto significa **maximizar la cobertura de protección**;
- se debe **minimizar la cantidad o volumen de agua consumida** con deficiente o mala calidad, lo que implica atender especialmente a las áreas con densidad de población alta en comparación con aquellas áreas con menor densidad de población y una tasa de consumo menor;

- por último, **minimizar el coste de la inversión**, que es directamente proporcional al número de sensores que pueden instalarse; y tal coste debe ser el mínimo posible.

Estos objetivos se evalúan sobre una determinada red de sensores arbitraria, Λ , es decir, una solución candidata del problema, y se pretende seleccionar aquellas redes de sensores Λ que constituyan el frente de Pareto del problema. Para proceder a la evaluación, se generan en la RDA una serie de eventos de contaminación que ponen a prueba a cada una de tales redes de sensores.

La generación de los escenarios de contaminación consiste en producir eventos de contaminación en cada uno de los nodos de la RDA, esto es, contaminar (computacionalmente hablando, claro) cada uno de los nodos de la RDA por orden, y evaluar cada evento uno a uno. El planteamiento de dichos escenarios considera los siguientes supuestos:

- a) La simulación de la intrusión contaminante en cada nodo de la RDA se realiza por un período de 24 horas, es decir, realizando simulación en período extendido (SPE).
- b) La intrusión se realiza en todos los nodos de la RDA ordenadamente y uno a uno, sin realizar eventos simultáneos.
- c) Los momentos en los que se realiza la intrusión contaminante se determinan en las horas de consumo máximo, pues es el momento de mayor exposición de la población. Sin embargo, este supuesto no es fijo y es flexible a la variación horaria que se requiera en otros escenarios.
- d) Se asume que los sensores en la RDA estarán funcionando adecuadamente, por lo que no se considera redundancia en la medición.
- e) A pesar de todas posibilidades contempladas, se asume que la RDA no está protegida completamente ante cualquier tipo de agente contaminante.

Para todos estos eventos de contaminación, los objetivos que se consideran para evaluar la idoneidad de una red de sensores Λ se describen de manera precisa a continuación.

3.3.1 Tiempo de detección

Este objetivo se refiere, básicamente, a una medida adecuada del lapso de tiempo transcurrido desde que una fuente de contaminación emite el contaminante, hasta el momento en que el contaminante alcanza al menos un sensor de Λ que lo detecta.

Existen dos formas comúnmente empleadas para definir la función objetivo-tiempo de detección para la red de sensores; se considera: 1) el tiempo máximo de detección de diferentes escenarios de contaminación; o 2) el tiempo promedio de detección de todos los escenarios de contaminación (según Kumar et al., 1999). La primera definición es más adecuada para el caso en que se detectan todos los escenarios de contaminación, mientras que la segunda se utiliza cuando la detección de todos los eventos no es posible debido al número limitado de sensores de monitoreo. El trabajo de (Watson et al. 2004) consideró solo los eventos detectados, mientras que otros autores incluyeron también los eventos no detectados, considerando para estos un tiempo de detección en función del tiempo de simulación.

La optimización del tiempo de detección, ignorando los eventos no detectados, minimizará dicho tiempo para los eventos detectados, lo que tiende a ubicar sensores cerca de la fuente. Cuando también se consideran los eventos no detectados, estos harán que la tendencia sea a convertir los eventos no detectados en eventos detectados, de modo que el TD disminuya lo máximo posible. De esta manera, los sensores se colocarán lejos de la fuente para, al menos, detectar todos los eventos en cualquier período de tiempo mínimo. En resumen, el tiempo de detección es un parámetro clave en la monitorización de sensores y marca cómo se realiza la modelación de la ubicación de la red de sensores.

En esta tesis, consideraremos como tiempo de detección básico para un evento de contaminación particular la diferencia de tiempo entre el inicio de un incidente de contaminación y su detección en el primer sensor de Λ que lo detecta.

Y la función objetivo para la red de sensores Λ , que se tratará de minimizar, es el tiempo promedio para todos los eventos de contaminación, es decir:

$$TD(\Lambda) = \frac{1}{N} \sum_{i=1}^N \tau(i), \quad (6)$$

donde N es el número de nodos de la red susceptibles de producir algún evento de contaminación; y $\tau(i)$ es, para cada escenario de contaminación detectado, el tiempo transcurrido desde el inicio de la contaminación en el nodo i hasta la identificación de la primera concentración de contaminante por un sensor de Λ (es decir, el tiempo de la primera detección de contaminante), y es $\tau(i) = 0$ para los escenarios no detectados.

A partir de las simulaciones de los eventos de contaminación es posible determinar un tiempo de llegada promedio de la contaminación desde los nodos de la red hasta que un primer sensor de la red que detecta la contaminación.

Observemos que este cálculo no incluye los eventos no detectados, ya que para tales eventos el valor de $\tau(i)$ es 0, como se ha especificado más arriba. Esta situación ocurre típicamente en casos en que los sensores no detectan el contaminante porque, por ejemplo, está ubicado más “aguas arriba” de donde el contaminante se inyecta en la RDA. Para resolver este problema, existen propuestas como las de Berry et al. (2012) y Hart et al. (2008) que suponen un tiempo de detección igual al tiempo de simulación (notamos que el tiempo de simulación se introduce como dato para definir cuánto tiempo debe estar funcionando la simulación). La alternativa aquí considerada asume que el tiempo de detección es igual a cero cuando no hay detección. Esto no es un problema si el cálculo está utilizando un enfoque de objetivos múltiples, en el que otros objetivos como el fallo de detección y el volumen de agua contaminada consumida también son considerados, como ocurre en nuestro estudio y como vemos más abajo. Justamente, consideramos, a continuación, estos dos objetivos.

3.3.2 Volumen de agua contaminada consumida

El volumen consumido (VC) de agua contaminada y la población de consumidores afectada son medidas del alcance que puede tener un evento de contaminación en los consumidores. La población afectada puede basarse en la densidad poblacional de consumidores en la red, o bien en el volumen contaminado consumido, es decir, en la demanda asociada en el modelo de la RDA a cada uno de los nodos afectados por un evento de contaminación. Es claro que esta es una información directamente disponible en el modelo hidráulico y utilizada por la SPE desplegada para analizar hidráulicamente la red. Por ello, en esta tesis utilizaremos el volumen de agua consumida, que calcularemos a partir de las demandas asociadas a los nodos de consumo.

Una expresión cerrada para este objetivo es la siguiente:

$$VC(\Lambda) = \sum_{i=1}^N \sum_{j=1}^N d(j, i), \quad (7)$$

donde $d(j, i)$ es la demanda asociada al nodo j durante el evento contaminante correspondiente al nodo i . Observemos que las redes de sensores Λ con muchos eventos no detectados (y, consecuentemente, un tiempo de detección promedio bajo) presentarán valores deficientes (grandes) para este objetivo. Esto permite paliar el efecto, virtualmente adverso, de no haber considerado los eventos no detectados para el cálculo de TD , tal como hemos comentado más arriba.

Cabe notar, también, que el VC está relacionado con el tiempo de detección. Si se minimizan TD y VC resultarán diferentes ubicaciones de los sensores. La minimización de VC seleccionará ubicaciones para los sensores cerca o en los nodos de mayor demanda o población.

La cuantificación precisa de los impactos reales derivados del volumen de agua contaminada consumida es un aspecto que precisaría de mayor elaboración. Por supuesto, la estimación de la población afectada podría formularse de una manera mucho más compleja. Sin embargo, insistir en el cálculo del número efectivo de personas afectadas abrirá, entre otras, la pregunta de cuánta agua

debería haberse consumido para que los consumidores se vean afectados por la concentración de contaminante que aparece en el agua. Además, se debe considerar que las acciones tomadas directamente por la empresa de servicios públicos tras detectarse una contaminación no ocurren instantáneamente. Lleva algún tiempo informar a la población, operar válvulas cuando sea necesario y realizar cualquier acción adicional como respuesta al evento de contaminación. Esto significa que las consecuencias de la contaminación serán probablemente mayores que el cálculo descrito aquí. Sin embargo, la forma en que estamos ejecutando los cálculos hace que este cálculo sea proporcional a lo que sucederá en el caso de un evento de contaminación real. Es por ello que la expresión para $VC(A)$ puede utilizarse como un indicador de la cantidad de consumidores afectados para optimizar la ubicación de los sensores, a pesar de que se requeriría un análisis diferente para tener un cálculo preciso de cuántas personas se verían exactamente afectadas.

3.3.3 Fallo de detección

Existen varias alternativas para valorar el fallo en la detección de eventos de contaminación. Por ejemplo, Ostfeld y Salomons (2008), utilizan la probabilidad de detección, que se define como la función:

$$FD = \frac{1}{S} \sum_{r=1}^S d_r \quad (8),$$

donde $d_r = 1$ si el escenario de contaminación r es detectado, y $d_r = 0$ en otro caso; S expresa el número total de escenarios de contaminación considerados.

En una línea similar, Rathi et al. (2015) definen la probabilidad de detección como la proporción de eventos detectados por un sensor de la red; también, como la probabilidad de que un sensor de la red detecte un evento de contaminación. La expresión que la representa es:

$$PD = \sum_{x=1}^J \sum_{p=1}^p \alpha_{xp} b_{xp} \quad (9),$$

donde $b_{xp} = 1$, si el evento de contaminación x es detectado durante el patrón p ; si no, es $b_{xp} = 0$.

Obviamente, el objetivo es maximizar la probabilidad de detección de fallo.

Otra aproximación, considera el complemento de esta probabilidad que es su diferencia a 1. Se trata de una forma recíproca de considerar el fallo, tratando de minimizar la cantidad de fallos de detección. Es la alternativa que seguimos en esta tesis y que definimos formalmente como:

$$FD(\Lambda) = \sum_{i=1}^N (\tau(i) = 0), \quad (10)$$

donde la condición lógica $(\tau(i) = 0)$ es igual a 1 si la condición es verdadera, es decir, el evento de contaminación iniciado en el nodo i no fue detectado por ningún sensor de Λ , y 0 en caso contrario. Esta expresión actúa como un contador de eventos de contaminación no detectados. Obsérvese que las redes de sensores Λ con tiempos de detección promedio bajos, derivados de los ceros asociados a eventos no detectados, presentarán una puntuación muy pobre (grande) en este objetivo.

La ubicación de sensores para maximizar la PD o minimizar la $FD(\Lambda)$, da como resultado la localización de los sensores lejos de la(s) fuente(s) de abastecimiento de modo que se reduzca el número de eventos no detectados. Asimismo, en la consideración del fallo de detección se pueden considerar sólo aquellos eventos que han sido registrados como eventos detectados y que tienen un TD menor que un tiempo tolerable determinado. Por lo tanto, cualquier evento detectado más allá de un tiempo específico no se incluye al evaluar el fallo. Es decir, el fallo puede considerarse limitado por el tiempo o con restricción de tiempo. Obviamente, la maximización con limitación de tiempo restringiría las ubicaciones de los sensores para alejarse de la fuente (Rathi et al., 2015). En esta tesis no se utiliza esta limitación de tiempo.

3.3.4 Costes de implementación

Los costes son proporcionales a la cantidad de sensores que se instalarán. Incluso si se desconoce cuántos sensores se instalarán finalmente, se puede suponer, obviamente, que más sensores implicarán más costes, y viceversa. Los costes totales reales estarán más allá del coste por sensor y deberán incluir también los costes de instalación y mantenimiento como mínimo. Una expresión sencilla para este objetivo es:

$$C(\Lambda) = \alpha N_s, \quad (11)$$

donde N_s es el número de sensores y α es el coste promedio estimado por sensor.

No será necesario un cálculo detallado de los costes para el problema de optimización que planteamos porque muchos de los detalles de los costes no tienen nada que ver directamente con las variables de decisión en el problema. Solo se puede decidir el número de sensores a instalar, que solo influirá en los costes de los sensores de manera proporcional al número seleccionado. En esta tesis, según lo expresado por la fórmula anterior, es claro que los costes se estiman como el número de sensores en la red de sensores Λ , multiplicado por un coste promedio estimado por sensor.

3.3.5 El cálculo de las matrices de contaminación y de agua contaminada consumida

El cálculo de estas funciones objetivo se realiza mediante la construcción y uso de dos matrices, a saber, la matriz de contaminación y la matriz de agua contaminada consumida. Estas matrices se construyen al realizar todas las simulaciones de eventos de contaminación necesarias y se almacenan adecuadamente en función del tamaño de la red hidráulica. Luego, los cálculos efectivos de los tres primeros objetivos se realizan mediante búsquedas adecuadas a estas matrices.

Describimos ahora estas matrices.

3.3.5.1 Matriz de contaminación

Esta **matriz se almacena**, para cada evento de contaminación, es decir, para la contaminación de un nodo susceptible de la RDA, el tiempo que el contaminante tarda, en la SPE correspondiente, en llegar al primer nodo de la red de sensores Λ que detecta la contaminación.

Al realizar las SPEs adecuadas, se va generando esta matriz, que se emplea, durante la optimización, para determinar los tiempos mínimos de detección para contaminaciones en cada uno de los nodos de la RDA.

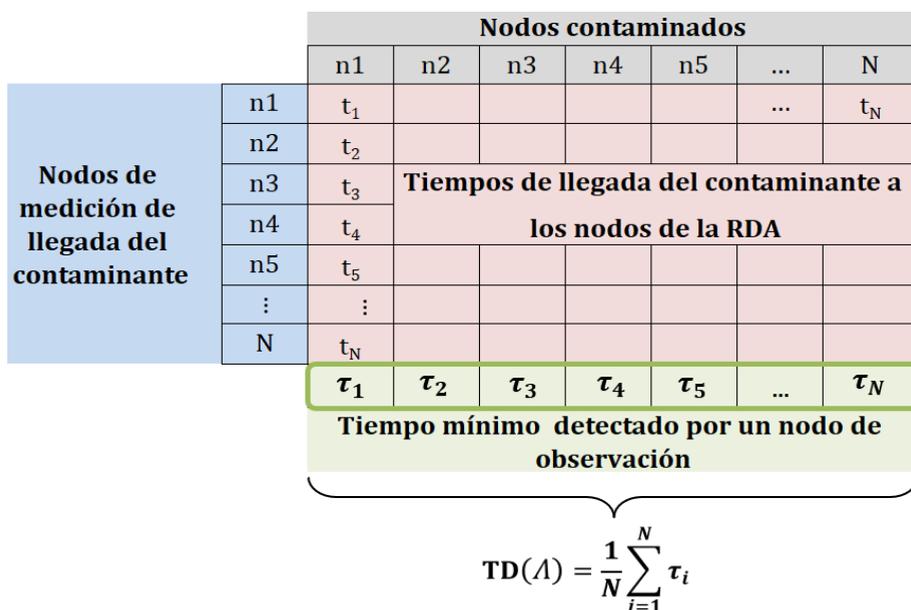


Figura 3.2. Configuración de la matriz de tiempos de contaminación para la obtención del tiempo de detección mínimo de los sensores

La Figura 3.2. muestra la configuración general de la matriz construida a partir de las simulaciones de contaminación. Esta matriz registra los tiempos de llegada del contaminante en cada nodo de medición, según el nodo contaminado. En cada caso, se elige el tiempo mínimo en el que el contaminante alcanzó un nodo de la red de sensores actual. El tiempo de detección promedio para la solución de sensores que se evalúa, corresponde al valor a minimizar en la función objetivo de la ecuación (6).

El cálculo de la matriz de tiempos de contaminación puede describirse en los pasos del siguiente pseudocódigo:

- 1) Se inicializa a cero la suma dada por (6);
- 2) se inicia una iteración sobre el número de nodos de la RDA, origen de eventos de contaminación; se hace $i = 1$;
 - a. se introduce el contaminante en el nodo i de la RDA;
 - b. se inicia la simulación hidráulica;
 - c. se calcula el tiempo hasta que algún sensor de Λ detecta la contaminación;
 - d. se acumula dicho tiempo al sumatorio general;
 - e. se elimina la contaminación del nodo i ;
- 3) finalmente, al completar la iteración 2), se calcula el promedio de todos los tiempos de detección, es decir, del sumatorio construido.

Esta matriz se utiliza para calcular las funciones objetivo $TD(\Lambda)$ y $FD(\Lambda)$ para una red arbitraria de sensores, Λ . También se utiliza, como paso previo, para calcular $VC(\Lambda)$, como se ve a continuación.

3.3.5.2 *Matriz de agua contaminada consumida*

De manera similar, y simultáneamente con la matriz de contaminación, durante las SPEs se va construyendo la matriz de agua contaminada consumida en la que se almacena precisamente la cantidad de agua contaminada consumida en cada escenario por cada nodo de consumo de la red.

El siguiente proceso ilustra cómo se implementa el cálculo del agua contaminada promedio que se ha suministrado a la población.

- 1) Se inicializa la suma de agua contaminada a 0;
- 2) se inicia una iteración sobre el número de nodos de la RDA, origen de eventos de contaminación; se hace $i = 1$;
 - a. se contamina el nodo i ;
 - b. se considera el tiempo hasta que uno de los sensores de Λ es alcanzado por el contaminante. Observemos que este cálculo

- utiliza el tiempo de detección para la red de sensores Λ que se ha calculado previamente, como se explica en el apartado 3.3.1;
 - c. para cada nodo alcanzado por el contaminante, se suma su demanda real a la suma del consumo de agua contaminada;
 - d. se elimina la contaminación del nodo i ;
- 3) se devuelve (tras completarse la iteración anterior) la suma de agua total contaminada consumida.

Esta matriz se utiliza para calcular la función objetivo $VC(A)$ para la red de sensores actual Λ . Hay que tener en cuenta que un elemento (j, i) de esta matriz representará cuánta agua contaminada se ha consumido en el nodo j desde el momento en que se inyecta un contaminante en el nodo i , hasta el momento en que un sensor en Λ detecta el contaminante; y este último cálculo utiliza la matriz de contaminación.

3.4 Conclusión

Las dos preguntas más importantes cuando se busca proteger las redes de agua contra eventos de contaminación accidental o intencional que amenazan la calidad del agua: cuántos sensores se necesitan y dónde deben colocarse. Responder a estas preguntas implica tomar decisiones preliminares sobre qué criterios y requisitos se deben considerar para lograr una buena solución, que inexorablemente debe obtenerse mediante un adecuado enfoque multiobjetivo. En este capítulo hemos considerado el tiempo de detección, el fallo de detección, el consumo de agua contaminada y el coste.

La solución final debe basarse en una compensación entre los objetivos involucrados, con especial énfasis en la tolerancia al fallo que la empresa de servicios públicos pueda permitirse en su sistema de suministro de agua. Se puede lograr una mejora en todos los objetivos analizados agregando nuevos sensores, pero esto, por supuesto, tiene la consecuencia de incrementar los costes, lo que puede ser una restricción (insalvable) para la implementación de la solución.

La información adicional sobre la utilidad bajo análisis también puede influir en el diseño final de la red de sensores adoptado. Este es el caso, por ejemplo, del tiempo medio de reacción de la empresa cuando se detecta un evento. Si el tiempo de reacción es relativamente grande, puede ser beneficioso utilizar aquellas soluciones en las que los sensores estén ubicados ligeramente aguas arriba. De esta forma, al menos una parte de la población no se vería afectada por el evento de contaminación. Hay que tener en cuenta que en este caso también podría haber un mayor número de fallos de detección. Obtener alguna mejora en ambos objetivos (menor tiempo de detección y menor número de fallos de detección) implica agregar más sensores a la solución.

Este capítulo trata de dar una descripción general simplificada sobre cómo lidiar en la práctica con la colocación de sensores de calidad del agua para proteger infraestructuras tan importantes como los sistemas de distribución de agua. Y, por supuesto, las decisiones no solo deben basarse en la experiencia práctica, sino que es preciso ejecutar cálculos con los modelos hidráulicos. En este sentido, tanto los modelos informáticos como la experiencia deben utilizarse conjuntamente para lograr mejores resultados.

4 Metodología para localizar sensores de calidad en redes de distribución de agua

En este Capítulo se expone la metodología seguida en esta Tesis para la localización óptima de una red de sensores para monitorizar la calidad del agua en una RDA. El método propone la utilización conjunta de la teoría de grafos y la plataforma evolutiva ASO (*Agent Swarm Optimization*). Esta combinación de métodos permite tener una solución inicial no aleatoria, proporcionada por el preproceso realizado con teoría de grafos, que favorece la solución del problema a través del enfoque heurístico, es decir, con el algoritmo evolutivo.

La teoría de grafos se ha utilizado para el estudio de distintos tipos de redes mediante grafos matemáticos, que permiten representar redes complejas del mundo real como, por ejemplo, redes de información, redes comerciales, redes sociales, redes biológicas y redes de sistemas de ingeniería, entre estas últimas las RDAs. En este sentido, la teoría de redes complejas emplea la teoría de grafos para clasificar redes, analizar sus estructuras, incluyendo vulnerabilidad, robustez y tolerancia a fallos o ataques. Con base en lo anterior, la RDA, vista como una red compleja, se estudia extrayendo la información de la red y se reconocen las principales medidas de su estructura, sus elementos más centrales y las medidas del espectro de su matriz de adyacencia que componen el grafo. Por otro lado, debido a que el emplazamiento de sensores para medir la calidad del agua en una RDA es un problema multiobjetivo, el algoritmo evolutivo ASO dará solución a las funciones objetivo planteadas en el apartado 3.3, partiendo de los nodos propuestos por la teoría de redes. El algoritmo tendrá un conjunto de soluciones iniciales que facilitará la convergencia hacia soluciones no dominadas del frente de Pareto. Esta inicialización, favorecerá la optimización, lo que tendrá como efecto añadido que el espacio de soluciones se vea reducido gracias al resultado obtenido en el preproceso.

Primero se hace referencia a las bases teóricas de la teoría de redes para el análisis de redes complejas y se describen los principales conceptos que se utilizarán para obtener los nodos-sensores en la RDA. Luego, se explican los fundamentos de ASO y el planteamiento para el diseño óptimo de la red de

sensores en una RDA, así como el planteamiento de la combinación de ambos enfoques para obtener una metodología que permita tomar en cuenta las características topológicas de una RDA, así como los factores hidráulicos y económicos asociados a la misma. En la Figura 4.1 se presenta la visión general de la propuesta metodológica.

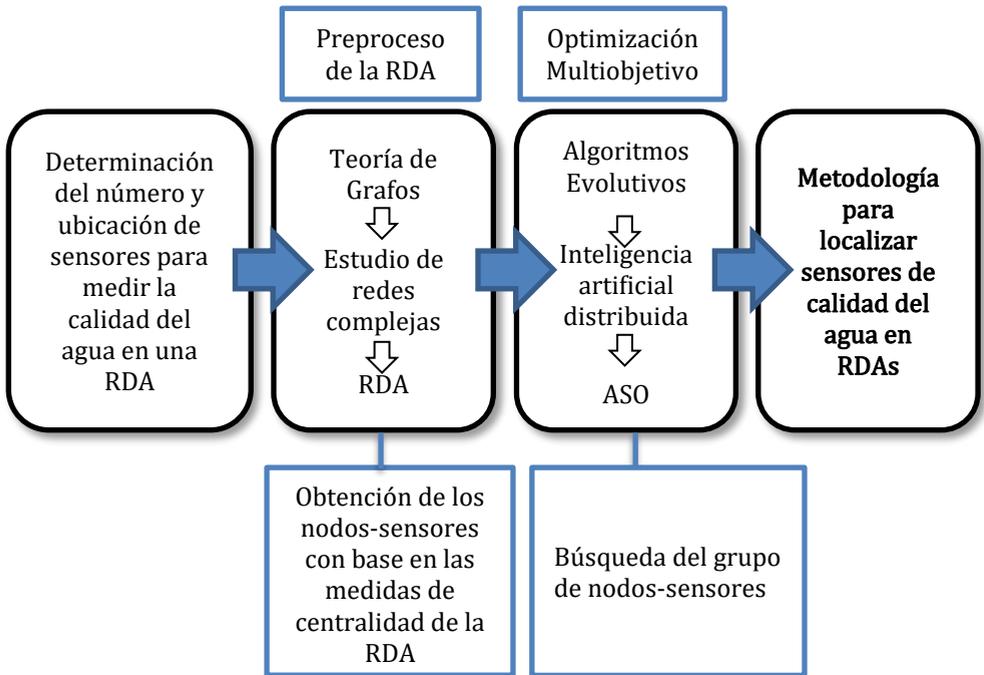


Figura 4.1: Enfoque general de la propuesta metodológica.

4.1 La Teoría de grafos como herramienta para el análisis de RDAs

Existe una gran variedad de sistemas complejos que son representados en forma de redes. Una red (matemáticamente denominada grafo) está formada por un conjunto de elementos llamados nodos que se conectan entre ellos a través de enlaces. La teoría de grafos o teoría de redes es una rama básica de las matemáticas discretas que se encarga del estudio y análisis de redes. En la literatura, a menudo se cita la solución del problema de los puentes de

Königsberg¹⁰ como la primera solución con teoría de redes y, desde entonces, la teoría de grafos es una rama esencial de las matemáticas. Además del enfoque puramente matemático, la aplicación de la teoría de redes ha tenido importantes resultados en diferentes áreas del conocimiento. Las redes complejas se pueden entender como grafos matemáticos con topologías no-triviales (no puramente aleatorias, pero tampoco completamente determinísticas) en las que sus nodos y enlaces normalmente poseen propiedades relacionadas con un sistema, físico o virtual, que el grafo matemático representa.

Debido a las posibilidades para estudiar redes de diversas características, es posible extraer la información y representar redes complejas de varios tipos como, por ejemplo: Internet, la *World Wide Web* (WWW), redes sociales¹¹ u otras conexiones entre individuos o entidades sociales, redes de infraestructura, como redes de transporte, agua, telecomunicaciones o redes de energía eléctrica, sistemas biológicos o médicos y redes de distribución y suministro de bienes y servicios, además de otras (Herrera et al., 2020). Por otra parte, se denominan redes complejas, a aquellas redes cuya estructura es irregular y están dinámicamente envueltas en el tiempo (Boccaletti et al., 2006). De forma general, las redes complejas se pueden agrupar en redes tecnológicas, biológicas, sociales o de información. Asimismo, sus estructuras topológicas pueden ser centralizadas, descentralizadas y distribuidas, dependiendo de la importancia jerárquica o redundancia de sus configuraciones (Newman, 2003). Dependiendo del tipo de red que se analice, los nodos y enlaces pueden ser definidos de diferentes maneras; por ejemplo, en una RDA, como puntos de consumo, depósitos y tuberías.

Los métodos de la teoría de redes complejas son útiles para identificar *a priori* las vulnerabilidades mediante el análisis topológico y dinámico de la estructura;

¹⁰ En 1736 el matemático suizo Leonhard Euler publicó la solución al problema de “Los Siete Puentes de Königsberg,” que consistió en la búsqueda de un viaje de ida y vuelta, que recorre cada uno de los puentes de la ciudad prusiana de Königsberg exactamente una vez.

¹¹ En la década de 1930, los sociólogos dieron importancia a los patrones de conexión entre entidades sociales para la comprensión del funcionamiento de la sociedad humana. Los estudios de redes sociales abordaron cuestiones de centralidad (qué individuos están mejor conectados con otros o tienen más influencia) y conectividad (si los individuos están conectados entre sí a través de la red y cómo).

y para guiar y enfocar con más detalle el estudio de las áreas críticas (Kröger y Zio, 2011c). Los principales propósitos de estudio de esas redes han sido:

- caracterizar los elementos de la red y sus posibles causas de vulnerabilidad,
- determinar su estabilidad y robustez frente a perturbaciones, fallos o anomalías y ataques; y
- construir diversos escenarios para el análisis de la topología de la red.

En este contexto, las RDAs son redes complejas con elementos interdependientes que pueden ser representadas y estudiadas desde la perspectiva de la teoría de redes. Los nodos de la RDA corresponden a los puntos de consumo y las fuentes de abastecimiento (ya sean depósitos o embalses) y los enlaces, a las tuberías, válvulas y bombas. Trabajos como los de Giudicianni et al. (2020); Herrera et al. (2015); Di Nardo et al. (2018); Nazempour et al. (2018); y Yazdani y Jeffrey (2010, 2011), entre otros, exploraron y demostraron los alcances de la aplicación de los fundamentos de la teoría de grafos en RDAs, considerándolas como grafos con características de redes complejas y con una importante relación topológica entre los elementos de la red. Utilizaron técnicas determinísticas de redes complejas y teoría de grafos para cuantificar la conectividad y estimar su resiliencia y su redundancia. Además, exploran una variedad de estrategias para comprender la formación, estructura, eficiencia y vulnerabilidad de las RDAs. Otro aspecto analizado en redes de distribución de agua ha sido la sectorización. Tzatchkov et al. (2006) abordan ese tema presentando algoritmos basados en teoría de grafos para la obtención de sectores independientes, así como para la obtención del conjunto de nodos que pertenecen a cada sector, el conjunto de nodos desconectados y el nodo fuente de contribución. En la misma línea Herrera (2011) y Herrera et al. (2010) desarrollaron una metodología, basada en aprendizaje semi-supervisado, para la conformación de clústeres de abastecimiento (*District Metered Areas*). Dicha metodología utiliza tanto información gráfica como vectorial. Por su parte, Campbell et al. (2014) propuso una metodología de sectorización de redes de abastecimiento de agua, basada en los conceptos de centralidad y detección de comunidades de la teoría de redes sociales. Por otro lado, el ordenamiento de los elementos de una RDA según su importancia ha tenido propuestas importantes. Entre las referencias más

relevantes se puede mencionar a Michaud y Apostolakis (2006) quienes introdujeron una metodología basada en el estudio de escenarios de riesgo para evaluar la importancia de cada nodo. En Izquierdo et al. (2008) el análisis se enfoca en la evaluación de la importancia relativa de las tuberías, con respecto al proceso de distribución del agua a través de la red. Por su parte, Grubestic et al. (2008) estudiaron aplicaciones de la teoría de grafos y flujo en redes en el análisis de la vulnerabilidad y robustez de las RDAs. En la mayoría de los trabajos se revisaron las distintas medidas existentes para la ordenación de los nodos de la red las cuales, en su mayoría, están basadas en técnicas estadísticas y de análisis espectral.

4.1.1 Marco conceptual y principales medidas de análisis de redes

En este apartado se hace una introducción a los principales conceptos del estudio de redes. Debido a que existe una gran cantidad de literatura dedicada al tema, no se pretende hacer una revisión exhaustiva, sino presentar los elementos que se consideran más importantes y que son utilizadas en esta tesis.

La teoría de grafos se basa en un conjunto de medidas que ayudan a evaluar redes complejas del mundo real. El análisis topológico se enfoca en las propiedades estructurales de los grafos a escala local y global. Las medidas globales proporcionan información sobre el grado de conectividad de cada nodo con respecto a los demás nodos, mientras que las medidas locales, evalúan en qué medida los vecinos próximos de cada nodo están conectados entre sí.

De manera formal, un grafo dirigido (o no dirigido) se define como un par $G = (V, E)$ que consta de dos conjuntos, el conjunto finito $V, V \neq \emptyset$, y el conjunto E de pares ordenados (o no ordenados) de elementos de V . Los elementos de $V = \{v_1, \dots, v_N\}$ son denominados vértices de G , mientras que los elementos de $E \equiv \{e_1, \dots, e_M\}$ son los arcos. Los números de elementos en V y E están denotados por N y M , respectivamente. Asimismo, el orden y el tamaño de un grafo $G = (V, E)$ están dados por el número de vértices N y el número de arcos M , respectivamente. Por lo tanto, $G(N, M) \equiv (V, E)$ denota un grafo finito y no vacío. En general, un nodo se representa por su enumeración $i = 1, \dots, N$, dentro del conjunto V . En un grafo no dirigido cada uno de los arcos es definido por una

pareja de nodos i y j , y se denota por (i, j) o e_{ij} . Si e_{ij} es un arco, entonces se dice que los nodos i y j son los extremos o nodos finales de e_{ij} . Un nodo y un arco son incidentes, si el nodo es uno de los extremos del arco (i, j) . Dos nodos son adyacentes o vecinos si están unidos por un arco e_{ij} . Por tanto, el arco e_{ij} representa el enlace desde el nodo i al j y, por tanto, $e_{ij} \neq e_{ji}$.

Las medidas de análisis de redes complejas son clasificadas como medidas de conectividad básica, medidas espectrales y medidas estadísticas:

- a) Las medidas de conectividad básica, que incluyen la conectividad de nodos y enlaces, representan la cohesión y adhesión de la red y su sensibilidad a la eliminación de dichos elementos.
- b) Las medidas espectrales resultan del conjunto de autovalores y autovectores (espectro) de las matrices de adyacencia de la red (o de su Laplaciana). Relacionan la topología de la red con la intensidad de la conectividad y la cohesión gráfica mediante el análisis del espectro de la matriz de adyacencia de la red. Dos medidas espectrales importantes son: la conectividad algebraica, que cuantifica la robustez de la red frente a fallos en nodos o enlaces; y la brecha espectral, que cuantifica el grado de conectividad de la red y proporciona información de la robustez del grafo.
- c) Las medidas estadísticas cuantifican las propiedades subyacentes de la red. Esto se realiza a través de la construcción de patrones de frecuencia y la conformación de subredes. Algunas medidas de esta clase son la distribución del grado de los nodos, el coeficiente de agrupamiento y el promedio de la longitud de los caminos.

Estos conceptos se explicitan de manera progresiva a continuación, conforme se presentan otros elementos de interés.

La estructura topológica de una RDA se describe matemáticamente en términos de un grafo; es decir, como un conjunto de vértices (nodos de consumo y fuentes de abastecimiento) y arcos (tuberías, válvulas y bombas). El análisis de una RDA basado en la teoría de grafos proporciona información sobre las propiedades

más significativas de la estructura, a escala global y local, y su relación con el análisis de vulnerabilidad en la distribución de agua. Primero, destacando el papel que desempeñan sus componentes; segundo, haciendo una evaluación preliminar de su vulnerabilidad basada en la simulación de fallos (representados principalmente por la eliminación de nodos y enlaces) y, por último, revalorando las propiedades topológicas de la red.

En lo sucesivo, se indicará $G(N, M)$ cuando se hable de una red con un número determinado de nodos y enlaces, específicamente de una RDA con N nodos y M enlaces (con el sentido físico descrito anteriormente). Con el fin de visualizar ciertas características de una red, en la Figura 4.2 se representan tres tipos de redes que ejemplifican una RDA mallada con $N = 7$ nodos y $M = 14$ enlaces. La red (A) es no dirigida o sin dirección de los flujos y no ponderada, la red (B) tiene indicadas las direcciones de los flujos y, por último, la red (C) es no dirigida y muestra los pesos $w_{i,j}$ en cada enlace que indican algún parámetro relacionado con la tubería, como la longitud, el diámetro o la rugosidad.

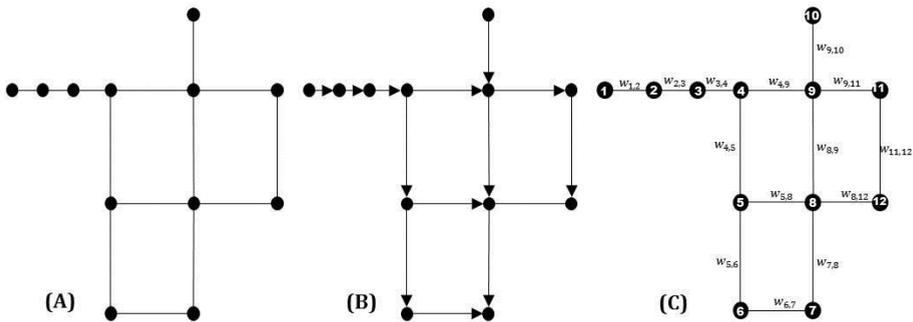


Figura 4.2: Representación de RDAs: sin dirección de los flujos (A), con dirección de los flujos (B) y no dirigida con pesos de las tuberías (C).

En los siguientes puntos revisamos las medidas básicas para el análisis topológico de una red.

- En el contexto de la teoría de redes, un concepto esencial es el de accesibilidad o alcance de dos nodos diferentes, pues dos nodos que no son adyacentes pueden alcanzarse el uno desde el otro. Así, se dice que una red está conectada si, para cada par de nodos i y j , existe una ruta o

camino desde i hasta j ; en caso contrario, se dice que es una red no conectada o desconectada.

- Un concepto muy importante para la caracterización de la estructura interna de un grafo es el camino más corto, que representa la comunicación y circulación dentro del grafo. Por esto, resulta útil obtener todas las longitudes de las rutas más cortas de un grafo y representarlas como una matriz, en donde cada una de sus entradas es la distancia geodésica desde el nodo origen al nodo destino. Así, se define un camino o recorrido desde el nodo i al j como una secuencia alterna de nodos (nodos adyacentes) y enlaces. La longitud del camino está dada por el número de enlaces en la sucesión. Entonces, un camino de longitud M es también una red $G = (N, M)$ con nodos pertenecientes al conjunto $V \equiv \{v_1, v_2, \dots, v_N\}$ y enlaces del conjunto $E \equiv \{v_1 v_2 \dots v_N\}$. Donde los nodos v_1 y v_N son los extremos del recorrido. Una red con un sólo nodo es un camino de longitud cero. Los caminos se representan dando la sucesión de los nodos que lo componen. En otras palabras, un recorrido $i - j$ en G es una secuencia finita que alterna nodos y enlaces, que se inicia en el nodo i y termina en el nodo j , incluyendo los M enlaces $e_i = \{v_{i-1} v_i\}$, donde $1 \leq i \leq N$. Lo anterior se define mediante la expresión $i = v_1, e_1, \dots, e_{M-1}, v_{N-1}, e_M, v_N = j$. Con base en el concepto de camino, una ruta se define como un camino en el cual un nodo no es visitado más de una vez. La ruta con la mínima longitud se conoce como la ruta o el camino más corto. Los conceptos anteriores, referidos a una RDA, pueden ser útiles para el análisis de los caudales o flujos, en referencia a su recorrido según la longitud de las tuberías, la velocidad y/o el tiempo de recorrido del caudal desde un nodo a otro. De lo anterior, se definen el diámetro de la red y la longitud promedio del camino más corto. El diámetro se expresa como:

$$Diam(G) = \max(d_{ij}), \quad (12)$$

donde d_{ij} es la distancia geodésica desde el nodo i al nodo j . El valor máximo de d_{ij} se llama diámetro del grafo.

- La longitud promedio del camino más corto está dada por:

$$L = \frac{1}{N(N-1)} \sum_{i,j \in \mathcal{V}, i \neq j} d_{ij} \quad (13)$$

Dónde N es el número de nodos de G . Si en G existen componentes desconectados, entonces L diverge.

- La densidad de tuberías (q) se considera el indicador más básico de la vinculación de los enlaces o escasez general de la estructura de una red. Esta medida varía según la escala y, por lo tanto, su magnitud puede cambiar al cambiar el tamaño de la red. Con esta medida se puede caracterizar a una red dispersa¹² o escasa por su valor de q , en el sentido de que el número de tuberías está por debajo del máximo posible. Una RDA dispersa se define como una red con el número de tuberías linealmente proporcional al número de nodos, es decir, $M = O(|N|^k)$ para $1 < k < 2$ (Diestel, 2005 en Yazdani et al., 2011).
- El grado o conectividad k_i de un nodo i de un grafo G , es el número de arcos o enlaces incidentes al nodo y es definido en términos de la matriz de adyacencia A (Boccaletti et al., 2006):

$$k_i = \sum_{j \in \mathcal{V}} a_{ij}. \quad (14)$$

Si $k_i = 0$, se dice que i es un nodo aislado. Un grafo $G = (V, E)$ es regular si todos sus nodos tienen el mismo grado. Si el grado común es k se dice que el grafo es k -regular. Los grafos 3-regular son llamados también grafos cúbicos. Si el grafo es dirigido, entonces el grado del nodo tiene dos componentes: el número de arcos o enlaces salientes $k_i^{out} = \sum_j a_{ij}$ es decir, el grado de salida del nodo. Por el contrario, el número de enlaces entrantes $k_i^{in} = \sum_j a_{ji}$ se refiere al grado de entrada del nodo. Por lo tanto, el grado total de un grafo dirigido es $k_i = k_i^{out} + k_i^{in}$. Por último, se denomina sucesión de grados a la lista ordenada y no

¹² Una red dispersa es aquella en que sus nodos son de grado de valor bajo, es decir con una cantidad de enlaces significativamente menor que la red con todos los nodos conectados entre sí.

decreciente del grado de todos los nodos del grafo. La mayoría de las redes complejas tienen una topología heterogénea; es decir, algunos nodos tienen un gran número de enlaces conectados y/o muchos nodos con pocos enlaces conectados. Además del grado de los nodos, es interesante conocer el grado promedio $\langle k \rangle$, ya que así se puede saber la conectividad global de la red. El $\langle k \rangle$ de una red con N nodos y M enlaces se denota con la expresión:

$$\langle k \rangle = \frac{2M}{N}. \quad (15)$$

El grado medio del nodo $\langle k \rangle$ es una medida básica de la conectividad. Refleja la similitud topológica general de la red a las redes perfectas o estructuras en forma de rejilla o cuadrícula, lo que es importante para la distribución igualada del flujo y la presión bajo diferentes demandas. La relación de enlace o tubería por nodo (e) y el grado medio de los nodos ($\langle k \rangle$) ayudan a determinar la dispersión de la RDA hacia una estructura de árbol o de malla. El valor de e en redes organizadas espacialmente se encuentra entre $1 < e < 2$, $e = 1$ representa a una red planar¹³ en forma de árbol y $e = 2$ una red con retículas o mallas regulares bidimensionales. Se sabe que, en general, las estructuras en forma de cuadrícula o mallado facilitan la distribución equilibrada del flujo y la presión bajo diferentes tasas de demanda y ubicaciones en las RDAs. Por lo anterior, e es una media simple que, hasta cierto punto, también muestra la eficiencia hidráulica de una RDA.

- El concepto de agrupamiento (conocido también como *clustering* o transitividad), es una propiedad típica de las redes sociales o de individuos conocidos en las que dos elementos, con un elemento “amigo” en común, es probable que se conozcan o se relacionen. En términos de un grafo G genérico, la transitividad significa la presencia de un gran número de triángulos en él.

¹³ Una red es planar si se puede incrustar en el plano de modo que sus bordes se intersecan sólo en un nodo que incida mutuamente con ellos.

Esto se puede cuantificar mediante la determinación de la transitividad T del grafo G , definida como el número relativo de triples transitivos, es decir, la fracción de tríos de nodos conectados que también forman triángulos (Newman, 2003):

$$T = \frac{3 \times \text{número de triángulos en } G}{\text{número de triples de nodos conectados en } G} \quad (16)$$

El factor 3 en el numerador de la ecuación anterior, compensa el hecho de que cada triángulo completo de tres nodos aporta tres triples conectados, uno centrado en cada uno de los tres nodos y asegura que $0 \leq T \leq 1$ con $T = 1$ para M_N . Cabe mencionar que T es una media clásica utilizada en ciencias sociales para indicar, localmente, qué tan agrupada está una red.

- Una alternativa a la medida anterior, es el coeficiente de agrupamiento C introducido por Watts y Strogatz (1998) para medir el grado local de agrupamiento. Debido a que muchas redes complejas presentan una fuerte tendencia a la propiedad de agrupamiento, el coeficiente de agrupamiento permite saber la densidad de triángulos en la red, es decir dos nodos conectados a un tercero, que a su vez está directamente conectado a otro. Por tanto, el coeficiente de agrupamiento de todo el grafo está dado por el promedio de c_i de todos los nodos en el grafo G :

$$C(G) = \langle c \rangle = \frac{1}{N} \sum_{i \in V} c_i. \quad (17)$$

Por definición, $0 \leq c_i \leq 1$ y $0 \leq C \leq 1$. En Latora y Marchiori (2003) se pueden observar algunas diferencias entre los coeficientes C y T .

- La conectividad algebraica (λ_2) se define como el segundo autovalor más pequeño de la matriz Laplaciana normalizada del grafo. La matriz de adyacencia o conexión de un grafo G con N nodos, es una matriz $N \times N$ definida como $A = (a_{ij})$, donde $a_{ij} = 1$ en caso de que exista un enlace entre los nodos i y j y $a_{ij} = 0$ para el caso contrario. Por lo tanto, la diagonal de A está formada por ceros. Un grafo G puede describirse

completamente por su matriz de adyacencia A , por lo que la representación matricial de una red resulta muy útil en la práctica. Por otro lado, la matriz Laplaciana de G se define como $L = D - A$, donde $D = \text{diag}(k_i)$, siendo k_i el grado del nodo i . El autovalor más pequeño de la matriz Laplaciana es cero con multiplicidad igual al número de los componentes conectados del grafo. La conectividad algebraica es un valor positivo cuya magnitud indica la robustez y buena conexión del grafo. Por otro lado, valores altos de conectividad representan una robustez muy alta frente a desconexiones o desacoplamientos provocados de los elementos del grafo.

- La brecha espectral ($\Delta\lambda$) corresponde a la diferencia entre el primer y segundo autovalor de la matriz de adyacencia A . Si el resultado es un valor muy alto, representa una condición necesaria de las propiedades de “buena expansión” y la similitud relativa entre la estructura del grafo y los “expansores” (que se definen como una familia de redes regulares dispersas conectadas de manera óptima). La falta de una buena expansión, representada por un valor pequeño de brecha espectral, generalmente se observa a través de la baja conectividad simultánea, la escasez y la presencia de puentes y puntos de articulación cuya eliminación da como resultado la división del grafo en dos o más grafos grandes. Por otro lado, la existencia de una buena expansión junto con una distribución uniforme de los grados de los nodos da como resultado una mayor robustez estructural contra fallas de nodos y enlaces.

De las medidas descritas anteriormente, se considera que los tres conceptos de mayor importancia son la distribución del grado de los nodos de la red, el coeficiente de agrupamiento y el promedio de la longitud de los caminos. Su importancia se debe a que estas medidas cuantifican las propiedades de organización de la red y, a su vez, relacionan los patrones estructurales más frecuentes con su redundancia.

Asimismo, son importantes las medidas de conectividad de nodos y arcos, ya que estas pueden cuantificar el número mínimo de ataques o fallos necesarios para eliminar o inutilizar a un grupo de nodos desconectados. En general, el valor de

la conectividad de los nodos es menor que el de la conectividad de los arcos. Lo que podría indicar que la existencia de nodos poco fiables en la red, en comparación con la existencia de arcos no fiables, puede generar consecuencias más graves en cuanto a fallos y recuperación se trata.

Sin embargo, estas medidas pueden ser triviales para el caso de algunas redes, como es el caso de las RDAs, debido a su baja redundancia, a la dispersión que presentan y a las conexiones individuales de los usuarios finales de la red. Por ello, se emplean medidas alternativas, tales como la brecha espectral y la conectividad algebraica, que cuantifiquen de manera más precisa la conectividad y la robustez de los enlaces. La brecha espectral es utilizada para identificar las propiedades de buena expansión de los grafos. Las redes con buena expansibilidad son aquellas que presentan una estructura topológica de las conexiones, tal que, cualquier conjunto de nodos está conectado de forma robusta a otros nodos, incluso si la red es dispersa. Por el contrario, las redes sin propiedades de buena expansión son las que en sus conexiones presentan estructuras de cuello de botella y puntos de articulación o puentes, lo que facilita su separación en dos o más subredes mediante la eliminación de esos elementos. En Di Nardo et al. (2018a) se demuestra la aplicación de los métodos espectrales para el estudio y gestión de las RDAs.

En la Tabla 4.1 se resumen las medidas descritas anteriormente.

Tabla 4.1. Medidas básicas para el análisis de una red.

Medida	Ecuación	Definición
Longitud promedio del camino más corto	$L = \frac{1}{N(N-1)} \sum_{i,j \in V; i \neq j} d_{ij}$	El número promedio de pasos a lo largo de los caminos más cortos entre todos los posibles pares de nodos de red.
Densidad de conexión	$q = \frac{2M}{N(N-1)}$	Fracción entre el total y número máximo de enlaces.
Promedio del grado de los nodos	$\langle k \rangle = \frac{2M}{N}$	Valor promedio de la distribución del grado de los nodos.

Medida	Ecuación	Definición
Diámetro.	$Diam(G) = \max(d_{ij})$	La máxima distancia geodésica entre dos nodos.
Coefficiente de agrupamiento.	$C = \frac{3n_{\Delta}}{n_3} = C(G) = \frac{1}{N} \sum_i c_i$	La fracción entre el número total de triángulos y el número total de triples conectados.
Coefficiente de Mallado.	$R_m = \frac{M - N + 1}{2N - 5}$	La fracción entre el total y el número máximo de bucles.
Brecha espectral.	$\Delta\lambda$	La diferencia entre el primero y segundo autovalor de la matriz de adyacencia de la red.
Conectividad algebraica.	λ_2	El segundo autovalor más pequeño de la matriz Laplaciana normalizada de la red.

En los siguientes apartados se hace referencia a tres medidas que son utilizadas en el proceso de selección de los nodos – sensores, dentro de la metodología propuesta en esta tesis. Estas medidas proporcionan la centralidad y conectividad de los nodos de un grafo o red, pero con diferentes enfoques conceptuales.

4.1.2 Medidas de centralidad en redes

Este apartado presenta tres de las principales medidas de centralidad de los nodos en redes complejas. Estas medidas son de especial interés, pues cuantifican la importancia de los nodos de la red respecto de su conectividad y rol en la distribución de tráfico en la red (caudal de agua en caso de red de tuberías, o flujo de datos en caso de red de telecomunicaciones).

4.1.2.1 Medida de centralidad del autovector (*eigenvector centrality*)

La medida de centralidad del autovector o vector propio [$e_c(v)$] proporciona el nodo más importante en cada clúster, desde un punto de vista topológico y de

conectividad. La expresión matemática que lo define es (Bonacich, 1972; Giudicianni et al., 2020):

$$A_S \mathbf{x} = \lambda_{\max} \mathbf{x}, \quad (18)$$

donde A_S es la matriz de adyacencia del subgrafo G_S (recordemos que los elementos de la matriz de adyacencia $a_{i,j}$ son iguales a 1 si el nodo i y el nodo j están conectados, y son igual 0 en otro caso); y λ_{\max} es el autovalor más grande. El autovector de centralidad calcula la centralidad para un nodo con base en la centralidad de sus nodos vecinos, de acuerdo con las coordenadas del autovector $e_c(v)$, correspondientes al autovector $\mathbf{x} = e_c$, asociado con el autovalor más alto λ_{\max} de la matriz A_S .

4.1.2.2 *Medida de ordenación de los nodos (pagerank)*

Como se mencionó al inicio de este capítulo, el ordenamiento de los elementos de un grafo o red proporciona información relevante de su estructura y conectividad. Los algoritmos de *ranking* en teoría de grafos son utilizados para decidir la importancia de un elemento de un grafo tomando información global a partir de la estructura del mismo grafo. En este grupo de algoritmos, el valor de *PageRank* es una medida muy útil de la importancia de los nodos que ayuda a entender la estructura de conectividad y funcionamiento de cualquier red y también, permite identificar elementos críticos. El algoritmo *PageRank* es el método de cálculo que usaron los fundadores de *Google* (Brin y Page, 1998). Fue diseñado como un método de análisis de los enlaces entre páginas *web* para clasificarlas cuantitativamente, según su importancia o calidad. El *PageRank* es una medida objetiva que puede corresponderse con la idea subjetiva de los usuarios sobre la importancia. Por esta correspondencia, es posible obtener excelentes resultados de búsquedas de palabras clave en la *Web*¹⁴. En general, contar las citas o los vínculos de retroceso a una página *web* determinada, proporciona una aproximación de su importancia o calidad. Sin embargo, el

¹⁴ Se entiende como una página *web* un documento o fuente de información, generalmente en formato HTML, que puede contener hiperenlaces a otras páginas *web*. Dicha página *web*, podrá ser accesible desde un dispositivo físico, una intranet, o *Internet*. Por otro lado, un sitio *Web* es un conjunto de páginas *web*, típicamente comunes a un dominio o subdominio en la *World Wide Web*.

algoritmo *PageRank* extiende esta idea al no contar por igual los enlaces de todas las páginas. Asimismo, hace una normalización en función del número de enlaces que existen en una página *web*.

Otra concepción del algoritmo *PageRank* es que puede ser interpretado como un proceso de Markov, donde los estados son las páginas y las transiciones son los enlaces entre ellas, todos igualmente probables. De esta manera, se capta la idea de modelo de navegación aleatoria. Es decir, que un usuario *web* puede ir de una página a otra con una cierta probabilidad al ingresar a alguno de los enlaces que contenga el primer lugar donde se encuentra. El algoritmo da mayor importancia a la normalización de los pesos de los vínculos entre páginas, así como a la navegación aleatoria.

La formulación que define al *PageRank* es la siguiente:

$$PR(A) = (1 - d) + d \left[\frac{PR(T_1)}{C(T_1)} + \dots + \frac{PR(T_n)}{C(T_n)} \right] \quad (19)$$

Se asume que la página *A* tiene $T_1 \dots T_n$ páginas que la direccionan (es decir, son las citas). El parámetro *d* es un parámetro o factor de amortiguamiento que se establece en el rango $[0,1]$, pero, por lo general, se recomienda el valor de 0,85. Así, el conjunto de todos los valores forma una distribución de probabilidad sobre el conjunto de todos los nodos de la red, por lo que la suma de los valores de *PageRank* debe ser igual a 1. Por último, $C(A)$ se define como el número de enlaces que salen de la página *A*.

A diferencia de otros algoritmos, *PageRank* integra el impacto tanto de los enlaces entrantes, como de los enlaces salientes de cada página *web*. El resultado es un vector de valores, también llamado *PageRank*¹⁵. El conjunto de todos los valores forma una distribución de probabilidad sobre el conjunto de todas las páginas *web* de la red, por lo que la suma de los valores de *PageRank* debe ser igual a 1. Asumiendo lo anterior, el valor de *PageRank* para cada nodo o vértice de una red cualquiera también estaría dado por la ecuación (19).

¹⁵ Se utilizará *PageRank* para hacer referencia al algoritmo y *pagerank* para el valor o *vector*.

De manera formal, dado un grafo $G = (\mathcal{V}, \mathcal{E})$ el vector *pagerank* se define como \mathbf{f} tal que satisface:

$$\mathbf{f} = \alpha \frac{\mathbf{1}}{n} + (1 - \alpha) \mathbf{f} \mathcal{P}, \quad (20)$$

con $\mathbf{f} \mathbf{1}^t = 1$ y donde α (factor de amortiguación) es una constante con valores de entre 0 y 1, pero que usualmente, como se ha dicho, toma el valor de 0,85 (Brin y Page, 1998). Además, n representa el número de páginas en el índice de páginas *web* de *Google* y \mathcal{P} es la matriz de probabilidades de transición del grafo.

\mathbf{f} es un autovector correspondiente al autovalor 1 de la matriz \mathcal{G} que está dada por:

$$\mathcal{G} = \alpha \frac{\mathbf{1} \mathbf{1}^t}{n} + (1 - \alpha) \mathcal{P}. \quad (21)$$

\mathcal{G} es la denominada matriz de *Google*. El teorema de Perron-Frobenius garantiza su existencia, indicando que \mathcal{P} estaría fuertemente conectada si todos los demás autovalores tienen un módulo de valor estrictamente menor que 1 (Ostfeld 2006).

Algunas propiedades de la matriz \mathcal{G} son (Langville y Meyer, 2006):

- a) Estocástica, porque es una combinación lineal de dos matrices estocásticas $\frac{\mathbf{1} \mathbf{1}^t}{n}$ y \mathcal{P} .
- b) Irreducible, porque cada página está directamente conectada a cada una de las otras páginas, por lo tanto, la irreductibilidad se cumple trivialmente.
- c) Aperiódica, debido a los auto-bucles, $\mathcal{G}_{ii} > 0 \forall i$.
- d) Primitiva, porque $\mathcal{G}^k > 0$ para algún k (por ejemplo, válido para $k = 1$). Esto implica que existe un único \mathbf{f}^t positivo.

En el cálculo del vector *pagerank*, para resolver la ecuación (21), se tiene la siguiente ecuación recurrente (Brin y Page, 1998):

$$\mathbf{f}_{t+1} = \alpha \frac{1}{n} + (1 - \alpha) \mathbf{f}_t \mathcal{P}, \quad (22)$$

a partir de una distribución inicial razonable del vector \mathbf{f}_0 . Por lo tanto, el *PageRank* de una página se define de manera recursiva y depende del número y *PageRank* de las páginas enlazadas. Asimismo, los dos resultados importantes que se obtienen son: primero, para las matrices estocásticas como \mathcal{G} , $\lambda_1 = 1$, por lo tanto λ_2 obtiene la convergencia. Segundo, si los espectros de \mathcal{P} y \mathcal{G} , son $\sigma(\mathcal{P}) = 1, \mu_2, \dots, \mu_m$ y $\sigma(\mathcal{G}) = 1, \lambda_2, \dots, \lambda_m$, respectivamente, se tiene que:

$$\lambda_k = \alpha_k \mu_k \quad \text{para } k = 2, 3, \dots, n. \quad (23)$$

Finalmente, la ecuación (23) garantiza convergencia, con una precisión apropiada. Existen otras alternativas de cálculo basadas en el espectro de la matriz de transición, \mathcal{P} , basados, en su mayoría, en el importante papel que juegan los autovalores y autovectores en la matriz, (Bryan y Leise, 2006; Langville y Meyer, 2006 y Pedroche 2007).

4.1.2.3 *Medida de intermediación (betweenness)*

Una medida de la relevancia de un nodo se puede obtener mediante el conteo de los caminos geodésicos que pasan a través del nodo, que se define como intermediación (*betweenness*). Junto con el grado y la cercanía (*closeness*) de un nodo, que se define como el inverso de la distancia promedio desde otros nodos, el parámetro *betweenness* es una media de la centralidad de un nodo, introducida originalmente para cuantificar la importancia de un individuo en una red social (Freeman, 1977; Latora y Marchiori, 2007). De manera más precisa, la centralidad o intermediación de un nodo i , se define como:

$$b_i = \sum_{j,k \in \mathcal{V}, j \neq k} \frac{n_{jk}(i)}{n_{jk}}, \quad (24)$$

donde n_{jk} es el número de caminos cortos que conectan a los nodos j y k , mientras $n_{jk}(i)$ es el número de caminos cortos que conectan a los nodos j y k pasando a través del nodo i . Aquí cabe hacer referencia a los algoritmos más

comunes para encontrar caminos cortos como son, el algoritmo de Dijkstra, el método de búsqueda de primera amplitud o anchura (*breadth-first*) y otros algoritmos más eficaces para calcular la intermediación (Crucitti et al., 2006a; Girvan y Newman, 2002). Además del concepto de intermediación asociado a los nodos de un grafo, éste también puede extenderse a los enlaces o arcos. La intermediación de un arco se define como el número de caminos cortos entre pares de nodos que pasan a través de un arco particular.

4.1.3 Detección de comunidades en redes

Este apartado presenta los rudimentos de los algoritmos de detección de comunidades en redes complejas. La identificación de comunidades o clústeres en la red puede ayudar a comprender cómo la topología puede ser no homogénea en toda la red. A menudo, las redes asociadas con infraestructura urbana, como redes de abastecimiento de agua, se distribuyen formando áreas discretas de alta conectividad dentro de ellas y baja conectividad entre ellas. Por tanto, el sistema, representado por la red compleja, está formado por subredes conectadas entre sí, donde el tráfico en la red fluye más fácilmente a nivel interno de subred que en la conexión existente entre una subred y otra. *Clustering de grafos y clustering de grafos espectral* son alternativas a la definición de subredes mediante la detección de comunidades (Brentan et al., 2018; Campbell et al., 2014; Campbell González, 2017; Martins Alves, 2020).

Los algoritmos de detección de comunidades se basan en similitudes entre cualquier red compleja (como la red que representa a una infraestructura urbana, red de agua potable) y una red social. Los algoritmos de agrupamiento en redes sociales se basan en la detección de áreas en las que los individuos están fuertemente relacionados (vinculados) y la información se difunde fácilmente. El método *Walktrap* (Pons y Latapy, 2005; Yang et al., 2016) es un algoritmo de agrupamiento jerárquico que identifica las estructuras de la comunidad a diferentes escalas. Se basa en realizar paseos aleatorios (*random walks*) en la red. Es más probable que estos paseos permanezcan dentro de un área determinada si hay pocos enlaces que conduzcan a otra parte de la red. Esto significa que *Walktrap* subdivide la red en áreas o comunidades discretas de tal manera que se maximiza la densidad de conectividad entre nodos (modularidad)

por comunidad. El algoritmo *Walktrap* primero calcula las distancias entre todos los vértices adyacentes. El algoritmo ejecuta paseos aleatorios de forma iterativa y utiliza sus resultados para fusionar comunidades separadas de forma ascendente. El proceso de partición evoluciona fusionando iterativamente los 2 nodos o clústeres más cercanos en una nueva comunidad, para crear una nueva partición, teniendo en cuenta las nuevas distancias recalculadas. El algoritmo continúa actualizando las distancias entre comunidades. El proceso se repite hasta tener un grupo único. En cada iteración se mide la calidad de la partición, por ejemplo, evaluando su modularidad. Esto permite elegir la mejor partición *i*-ésima, la de máxima modularidad, como resultado del algoritmo. Un algoritmo equivalente a *Walktrap*, pero computacionalmente más eficiente y escalable, es el algoritmo de *Louvain* (Blondel et al., 2008; Traag et al., 2019) que es adecuado para redes de gran tamaño. *Louvain* utiliza un proceso iterativo para mejorar la escalabilidad de la detección general de la comunidad basada en la optimización de la modularidad calculada durante el proceso de detección de comunidades. En el proceso que sigue *Louvain* primero se encuentran comunidades de pequeño tamaño, a nivel local, optimizando la modularidad en cada nodo. Después, las comunidades se extienden agregando más nodos mediante la iteración del proceso de agregación de nodos con que se inicia el algoritmo, pero entendiendo cada comunidad del paso anterior como un hiper-nodo al que se le pueden agregar, o no, otros nodos, o hiper-nodos; esto es otras pequeñas comunidades calculadas en iteraciones anteriores del algoritmo.

4.2 Implementación de los fundamentos de la teoría de redes para seleccionar nodos de monitoreo en una RDA.

En este apartado se describe la propuesta basada en teoría de grafos para la selección de nodos – sensores en una RDA. La Figura 4.3 esquematiza el proceso general.

Para la obtención y el tratamiento de la información se hicieron algunas consideraciones. Primero, que la red de abastecimiento es un grafo particular que incluye tanto información geográfica como de conectividad. Segundo, que los nodos y arcos del grafo corresponden a los puntos de consumo de la red de abastecimiento y a las tuberías de la red, respectivamente.

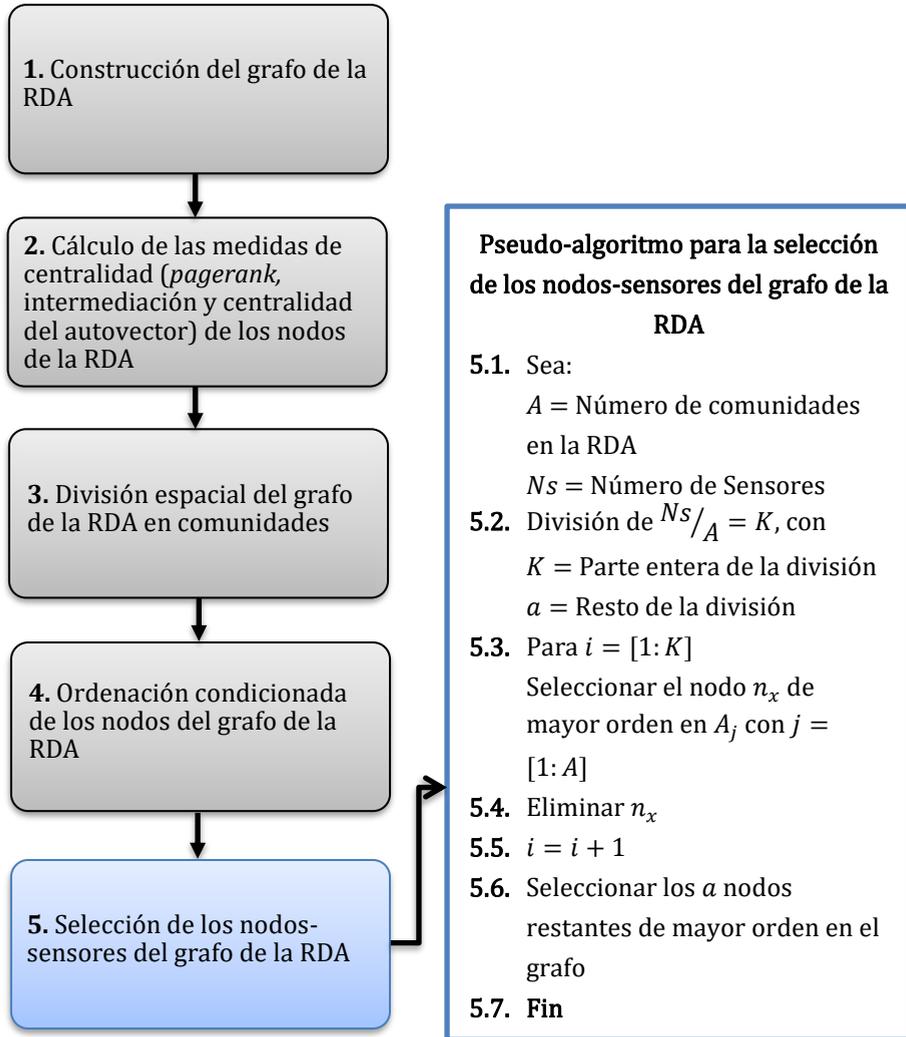


Figura 4.3: Proceso para la selección de los nodos sensores de una RDA

Como se observa en la Figura 4.3 el primer paso del proceso es extraer todos los elementos de la RDA junto con sus propiedades y construir su grafo. Los nodos de la red compleja representando una RDA representan puntos de consumo, tanques, reservorios; mientras que los enlaces de la red compleja son tuberías, uniones, válvulas y estaciones de bombeo. Los elementos de la red compleja pueden heredar atributos de la RDA, provisionando al modelo de heterogeneidad

en sus propiedades, o atributos, que es necesaria para una mejor adaptación de la RDA al modelo de red compleja. Algunos de estos atributos pueden también resumirse como pesos de la red, jugando un papel importante en el futuro análisis dicha red compleja. Los principales atributos de los nodos son las coordenadas (X, Y) , la elevación y la demanda mientras que, para las tuberías, se considera la longitud y el diámetro de las tuberías de la RDA. Por lo tanto, el grafo de la RDA queda configurado de la siguiente manera: $\text{grafo (RDA)} = [\text{nodos (RDA)}, \text{tuberías (RDA)}, \text{atributos nodos, (RDA)}, \text{atributos tuberías (RDA)}]$.

Una vez configurado el grafo de la red, el siguiente paso es calcular las medidas de centralidad de los nodos. Para los fines de este trabajo se han elegido las medidas *pagerank*, intermediación (*betweenness*) y centralidad del autovector (*eigenvector centrality*), las cuales han sido abordadas en la sección 4.1.2.

En este punto, la idea principal es ubicar los sensores en los nodos con mayor importancia de conectividad y mayor valor de centralidad. Con esto, se pretende favorecer procedimientos de optimización heurística posteriores o, en algunos casos, se puede adoptar como herramienta única en caso de ausencia de información hidráulica de la RDA. El objetivo es tener una mejor aproximación del diseño óptimo de una red de sensores para la protección de la RDA contra ataques accidentales o intencionales y así ubicar los sensores de calidad del agua que correspondan con los nodos con mayor valor de conectividad o *pagerank*, de centralidad *betweenness* y del vector propio o *eigenvector centrality*.

Por un lado, la teoría detrás del concepto de centralidad de un grafo es identificar qué nodos atraviesan el mayor número de caminos. La centralidad del autovector proporciona el nodo más importante para cada comunidad o *cluster*, desde un punto de vista topológico y de conectividad. Así, los sensores para la RDA se ubicarán de acuerdo con esta importancia de nodo. Un aspecto crucial de este enfoque es que los nodos centrales pueden cambiar su ubicación según el peso atribuido al gráfico. En este sentido, es posible elegir un conjunto de pesos en función del objetivo de monitorización y el funcionamiento del sensor. La centralidad del autovector para un grafo no ponderado solo tiene en cuenta el grado de conectividad de los nodos y cierta información topológica. En el caso de los grafos ponderados, la información que proviene de la centralidad del vector

propio se mejora considerando también las características geométricas del grafo u otra información de las características de la tubería, y teniendo en cuenta la variabilidad espacial de la RDA. En este último caso, podría ocurrir que los nodos con menor número de conexiones estén vinculados a las tuberías más importantes, según el peso asignado al gráfico, y las conexiones de red más grandes estén conectadas a las tuberías de menor importancia (Giudicianni et al., 2020).

El valor o vector resultante *pagerank* es una medida muy útil de la importancia de los nodos que ayuda a entender la estructura de conectividad y funcionamiento de cualquier red y también, permite identificar elementos críticos. Esto puede ayudar en la gestión de puntos críticos e investigar cómo pueden afectar a otros elementos conectados en caso de un fallo o daño intencional. Trabajos como los de Gutiérrez-Pérez et al. (2012), Herrera et al. (2011a), Mihalcea (2004) y Chung y Zhao (2008) propusieron el carácter general que puede tener el algoritmo *PageRank* a la hora de establecer la ordenación por importancia de los nodos de cualquier tipo de red, lo que permite plantear su aplicación en el entorno de las redes complejas y, particularmente, las RDAs. En relación con esto último, entendiendo una RDA como un grafo de características especiales, se puede abstraer la noción de página *web* para considerarla como un nodo de consumo de la RDA. Mientras que los enlaces entre páginas se entienden como las tuberías que unen diferentes nodos.

Con el post-procesamiento de los valores de los vectores resultantes de las medidas de centralidad (*PageRank*, *betwenness* y *eigenvector centrality*) se conforma un grupo de nodos con posibilidad de funcionar como sensores en la RDA.

Para la división espacial del grafo se propone la división en comunidades utilizando el algoritmo *Walktrap*. La división del grafo de la RDA en comunidades resulta útil para considerarlas como áreas de muestreo y para la localización de los sensores, porque se busca que los sensores estén distribuidos homogéneamente en la RDA. Además, determinar áreas de mayor relevancia de la RDA resulta útil para fines de gestión de la seguridad de una RDA.

El cuarto paso del proceso metodológico se basa en un ordenamiento de los nodos condicionado por dos criterios: el valor obtenido de cada medida de centralidad y la comunidad de pertenencia. Con esto, se logra mantener la distribución homogénea de la importancia de los nodos tanto en las comunidades, como en todo el grado de la RDA.

El último paso de la metodología representa la obtención de los nodos candidatos a sensores de la RDA para el caso de cada medida, es decir, se obtiene un conjunto de nodos con base en su *pagerank*, otro con base en su *betweenness* y otro grupo con base en *eigenvector* centrality. El procedimiento para la selección de los nodos consiste en tomar el primer valor más alto ($v_{i_{A_i}}$) de cada comunidad (A) y compararlo con un umbral (u) de cada medida de centralidad. Si cumple con la condición $v_{i_{A_i}} \geq u$, entonces se elige el nodo sensor. Lo anterior se realiza para el segundo valor más alto de cada comunidad. A partir del tercer valor, se tomará el valor más alto de todas las comunidades y se comparará con u ; y si también cumple con la condición de ser mayor o igual que u , se elige el nodo sensor. Todo este desarrollo se realiza hasta completar el número de sensores (N_s) determinado *a priori*, con base en criterios principalmente económicos, del tamaño de la RDA, de la necesidad de cobertura de sitios específicos o de prioridad por su seguridad y de la capacidad de almacenamiento y gestión de datos en tiempo real necesarios.

4.3 Optimización multiobjetivo: aplicación del algoritmo evolutivo ASO para obtener una red de sensores en una RDA

En diversos campos de la ciencia y la ingeniería, las técnicas de optimización empleadas han condicionado la forma en que se abordan los problemas de optimización. Sin embargo, a pesar de que los problemas se han adaptado a las técnicas de optimización existentes, se ha demostrado que esas técnicas no se adecúan bien a muchos problemas del mundo real y son incapaces de afrontar problemas altamente dimensionales, multimodales y fuertemente no lineales.

Como consecuencia, se requieren métodos sólidos de optimización para generar resultados adecuados. Según Kukkonen y Coello Coello (2017), en el año 2002 había aproximadamente 30 técnicas de optimización multiobjetivo reportadas

en la literatura. La mayoría de ellas dirigidas a frentes de Pareto con ciertas características que requieren un punto inicial de búsqueda y la generación de una sola solución por ejecución.

En el contexto de las propuestas para la solución de problemas multiobjetivo aparecen con fuerza los algoritmos evolutivos. Los algoritmos evolutivos emplean como base a un “agente individual” que, debido a su comportamiento, es capaz de encontrar por sí mismo una posible solución al problema. Aquí, cabe comentar que un agente es un sistema informático que se encuentra en algún ambiente, y que es capaz de actuar de forma autónoma en este entorno con el fin de cumplir sus objetivos de diseño (Wooldridge, 2002). Esta definición es extendida por Montalvo (2010), que considera que un humano también puede ser considerado como un agente. Por tanto, propone la incorporación de humanos y agentes artificiales como partes activas dentro de un mismo sistema para facilitar el desarrollo de la inteligencia colectiva.

Muchos investigadores se han dado a la tarea de implementar diversos algoritmos evolutivos, entre los que destacan: algoritmos genéticos (Goldberg, 1989), colonia de hormigas (Dorigo et al., 1996), recocido simulado (Kirkpatrick et al., 1983), algoritmos meméticos (Moscato, 1989), *shuffled complex evolution algorithm* (Duan et al., 1993), *search harmony algorithm* (Geem et al., 2001), optimización de enjambre de partículas (por *Particle Swarm Optimization – PSO*) (Kennedy y Eberhart, 1995), entre otros. En relación con los algoritmos antes mencionados, el conocimiento del espacio de solución que tienen sus “agentes individuales” es muy limitado. La búsqueda eficaz de soluciones óptimas sólo es posible si existe interacción entre múltiples agentes.

Montalvo et al. (2014) señalan algunas ventajas de utilizar algoritmos evolutivos para la solución de problemas de optimización, tales como:

- 1) Pueden enfrentar problemas de variables continuas, así como de variables naturalmente discretas y variables binarias (en las decisiones de sí/no), que son frecuentes en muchos problemas del mundo real.

- 2) Sólo funcionan con la información de la función objetivo, lo que evita complicaciones asociadas con la determinación de las derivadas y otra información auxiliar.
- 3) Son procedimientos de optimización genéricos y pueden adaptarse directamente a cualesquiera funciones objetivo, incluso si éstas no se describen mediante expresiones cerradas, sino mediante procedimientos, que pueden ser bastante complejos.
- 4) Debido a que los algoritmos evolutivos funcionan con una población de soluciones, se pueden obtener muchas soluciones que pueden ser de gran interés desde el punto de vista de la ingeniería.
- 5) Son versátiles y flexibles: el análisis de sistemas bajo diversas condiciones de carga o términos de forzamiento se puede realizar dentro del proceso de optimización.

Sin embargo, a pesar de dichas ventajas, cada algoritmo tiene sus propios inconvenientes y se adapta mejor a unos problemas que a otros. Las heurísticas detrás de un determinado algoritmo evolutivo dotan a sus agentes de capacidades específicas para resolver eficazmente algunos tipos de problemas, pero son ineficaces con problemas de distinta naturaleza. Por ello, una búsqueda eficaz de soluciones óptimas sólo es posible como resultado de la interacción entre muchos agentes. Por ejemplo, en el caso de PSO las asociaciones de agentes que interactúan dan como resultado estructuras colectivas llamadas *swarms* (o enjambres) que representan el comportamiento emergente de grupos de agentes. A su vez, los enjambres pueden actuar con otros enjambres ya existentes (Montalvo et al., 2012).

Una partícula de un enjambre de PSO, una hormiga de un sistema de colonia de hormigas o un cromosoma de una estructura del algoritmo genético muestran comportamientos diferentes. Sin embargo, todos esos agentes comparten una característica común, y es que cada uno representa una solución potencial para el mismo problema en cuestión; es decir, que todos los algoritmos comparten las soluciones incorporadas por sus poblaciones. Este intercambio de soluciones y de información permite que un algoritmo pueda beneficiarse de otro. Además, los agentes están dotados de conjuntos de reglas específicas dependientes del problema encaminadas a dos objetivos claros: 1) mejorar el comportamiento del

agente para especializarse en el problema de optimización que se aborda y, al mismo tiempo, 2) reducir el espacio de búsqueda, permitiendo que las decisiones se produzcan con mayor fiabilidad y en menor tiempo de cálculo (computacionalmente hablando).

Finalmente, en un entorno combinado, una partícula de PSO podría ayudar a reforzar la feromona en los caminos de las hormigas, una hormiga podría reproducirse con un cromosoma; un cromosoma podría ser el líder de un enjambre de partículas, y así sucesivamente. Lo anterior, hace que sea posible el éxito en la solución de una gama más amplia de problemas,

En este punto, el algoritmo ASO es una apuesta en la dirección de esa integración. ASO parte del algoritmo PSO, pero con mejoras incorporadas para la solución de problemas de optimización multiobjetivo. En las secciones siguientes se explicita el funcionamiento de ASO y se presenta su aplicación en el ámbito de la ingeniería, específicamente en la conformación óptima de una red de sensores de calidad en una RDA.

4.3.1 Algoritmo evolutivo ASO

En el contexto descrito anteriormente se enmarca el algoritmo de optimización ASO desarrollado por Montalvo et al. (2010). ASO es un término abstracto que se refiere a un tipo de algoritmos extensivos, flexibles y robustos para resolver problemas de optimización complejos, en los que las técnicas convencionales y otros algoritmos evolutivos no son suficientes.

ASO es una generalización del algoritmo PSO con el enfoque de la inteligencia artificial distribuida, basado en el paradigma multiagente (de ahí el término *Agent*). Asimismo, ASO se establece en una plataforma común donde integra de manera extensible diferentes tipos de agentes (que pueden estar dotados de reglas específicas para cada problema), incluidos los humanos. En tal plataforma coexisten, interactúan y cooperan para solucionar un problema de optimización. En la medida en que se mejore la interacción de la experiencia que el factor humano puede aportar, junto con la capacidad computacional, mayores serán las posibilidades. Siempre y cuando los agentes busquen el equilibrio justo y no la

suplantación total de unos por otros. Los agentes pueden integrarse en *swarms* que a su vez pueden ser considerados como agentes en un grado de abstracción superior.

El objetivo principal de ASO es apoyar la toma de decisiones de ingeniería resolviendo problemas de optimización de uno o varios objetivos. En ASO, un agente puede encontrar por sí mismo una posible solución al problema. Sin embargo, como ya se mencionó, el conocimiento relacionado con el espacio de solución que tienen los agentes es muy limitado por lo que la búsqueda efectiva de soluciones óptimas solo es posible como resultado de la interacción entre varios agentes. En este sentido, el marco conceptual de ASO no es una metaheurística fija sino un entorno dinámico que combina las fortalezas de múltiples metaheurísticas donde se pueden agregar nuevos algoritmos (agentes o enjambres) en tiempo real para contribuir a la solución del problema dado. Se necesita, por tanto, una estructura jerárquica definida con precisión para coordinar los esfuerzos de manera eficiente (Montalvo et al., 2014).

Como ya se ha mencionado en párrafos anteriores, el funcionamiento del algoritmo ASO parte de la base del algoritmo PSO, con modificaciones agregadas. Al respecto, Montalvo (2010) dice que, a pesar de que PSO puede ser incluido fácilmente sobre el entorno de ASO, las partículas (o agentes) por sí solas no estarían preparadas para la solución de un problema de optimización multiobjetivo. Por ello, inicialmente se toma la hipótesis de que el frente de Pareto es totalmente desconocido y que las partículas deben ir aproximándose a ella lo más posible. Esto significa que en cada momento se puede seleccionar un frente de Pareto aproximado, constituido por las partículas que contienen soluciones no dominadas. A partir del frente de Pareto aproximado que se tenga en cada instante, se puede obtener un vector que representa un punto en el espacio objetivo, denominado punto singular, cuyas componentes estén formadas por los mejores valores hasta el momento conocidos para cada uno de los objetivos individuales. Por otro lado, la partícula o agente líder será aquella que en espacio objetivo esté más cerca del punto singular. La Figura 4.4 representa las posiciones de varias partículas, donde se han marcado aquellas que pertenecen al frente de Pareto más aproximado según lo que hasta el momento era conocido.

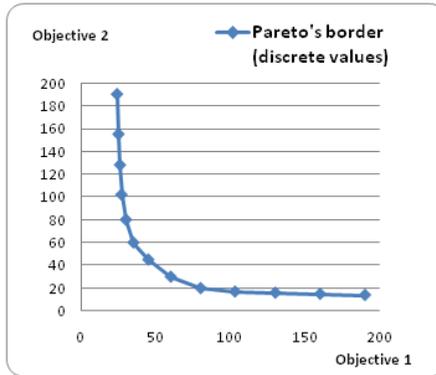


Figura 4.4: Frente de Pareto (tomada de Montalvo 2010).

En la Figura 4.5, se observan dos ejemplos de un punto singular en un frente de Pareto aproximado.

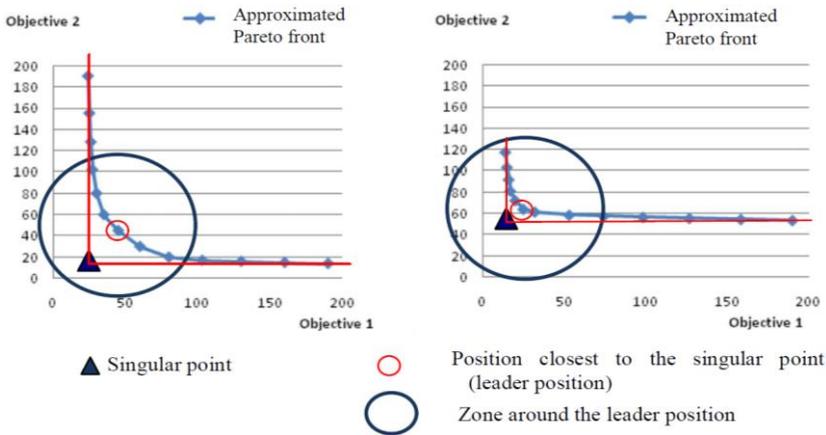


Figura 4.5: Dos ejemplos de punto singular en un frente de Pareto aproximado (tomada de Montalvo et al., 2014).

Llegar a buenas soluciones finales no sólo depende de la implementación del algoritmo, sino también del usuario (entendiéndose que pueda haber más de un usuario intercambiando información), debido al uso de varios ordenadores conectados. Decidir cuál solución elegir del frente de Pareto, cuando hay más de un elemento, requiere de la participación del usuario, desde el inicio hasta el final de la ejecución del algoritmo. Asimismo, el usuario participa de manera muy

activa decidiendo, además de la solución final, las zonas del frente de Pareto que son de mayor interés. También, el usuario puede solicitar información a los agentes sobre las soluciones encontradas, modificarlas o elaborar nuevas y, posteriormente, proponerle el resultado al algoritmo para que lo evalúe.

4.3.2 Consideraciones para la aplicación del algoritmo evolutivo ASO en una RDA.

Los enfoques que pueden ser aplicados para encontrar el frente de Pareto en un problema de optimización multiobjetivo son muy diversos (Ferreira et al., 2017; Jiang y Yang, 2017; Kukkonen y Coello Coello, 2017). En esta tesis, como se ha especificado, el algoritmo empleado en el proceso de búsqueda de la solución es ASO. El algoritmo ASO, a su combinación de varios algoritmos evolutivos multiobjetivo, añade agentes basados en reglas y análisis de datos. Asimismo, permite integrar inteligentemente el conocimiento desarrollado por el ingeniero en su tratamiento de problemas reales, con el proceso de optimización que aprende de dicho conocimiento humano para mejorar su acción y obtener mejores resultados. El algoritmo se ha integrado en WATERing © (2020)¹⁶, un paquete de software para análisis y soporte de decisiones en sistemas de distribución de agua. WATERing se conecta con el *toolkit* de EPANET (2020)¹⁷ para realizar las simulaciones hidráulicas necesarias que generan la matriz de contaminación y la matriz de agua contaminada consumida, conceptos completamente definidos en el Capítulo 3. Utilizando esta información, ASO analiza múltiples alternativas para ubicar sensores y selecciona aquellas alternativas que representan soluciones no dominadas, que integran el frente de Pareto.

A pesar de que se pueden utilizar diferentes técnicas evolutivas en ASO, para esta investigación se utilizó principalmente una versión multiobjetivo de PSO. Se introdujo una población de agentes con un comportamiento similar al de las partículas en PSO para resolver el problema. Además, se considera otra población de agentes basados en reglas, que participan en el proceso de

¹⁶ En <https://ingeniousware.com/software/watering#>

¹⁷ En <https://www.epa.gov/water-research/epanet>

búsqueda de soluciones. Los agentes basados en reglas tienen como finalidad principal guiar a los agentes convencionales de PSO en el proceso de búsqueda, consiguiendo de manera colateral reducir el espacio de decisión del problema, como consecuencia de la aplicación de reglas adecuadamente diseñadas. Esta técnica de utilización de agentes basados en reglas, se ha utilizado, aplicando reglas diversas, para reducir el espacio de decisión del problema en el diseño de un sistema de distribución de agua (Izquierdo et al., 2016b; Izquierdo et al., 2016a; Montalvo et al., 2014). En esta tesis, elaboramos reglas básicas con base en las medidas de centralidad de la teoría de redes para reducir el espacio de decisión del problema que nos ocupa, es decir, la ubicación óptima de sensores de calidad del agua en una RDA. Explicamos, a continuación, estas ideas.

Un agente "normal" basado en el comportamiento clásico de PSO, cuyo comportamiento, como es sabido, tiene una componente aleatoria, puede localizar un sensor en prácticamente cualquier nodo de la red. Sin embargo, el ingeniero, basado en su experiencia, es conocedor o intuye algunas situaciones claras, a la vez que sencillas:

- La ubicación de sensores demasiado aguas abajo de la red probablemente garantizará una buena cobertura de la red, pero, desafortunadamente, resultará en un gran tiempo de detección. Básicamente, esto se traduce en que la probabilidad de detectar un evento es relativamente alta. Sin embargo, la detección será definitivamente demasiado tarde porque la mayor parte de la red ya estaría contaminada cuando ocurra la detección, y no hay mucho más que hacer para evitar que las personas consuman agua contaminada.
- La ubicación de sensores demasiado aguas arriba de la red ayudará a detectar eventos de contaminación de manera más rápida, pero la cobertura de la red se verá comprometida. Muchos eventos de contaminación pueden ocurrir aguas abajo de la ubicación de los sensores y no serán detectados.

Clarificamos a continuación estos aspectos y elaboramos sobre la implicación que tienen en la obtención de reglas fundamentales para acelerar y mejorar el proceso de optimización.

- a) Aunque los conceptos “aguas abajo” y “aguas arriba” no están claramente definidos para una red, los tiempos de contaminación medios previamente calculados a través del objetivo $TD(\Lambda)$ dado por la ecuación (6) del Capítulo 3, y almacenados convenientemente en la matriz de contaminación son una guía perfecta para moverse aguas arriba y aguas abajo a través de la red y pueden ser utilizados como sucedáneos de los conceptos clásicos de aguas arriba y aguas abajo. Además, se observa que, como el proceso de optimización se realiza después de haber realizado un número adecuado de simulaciones que conducen a la obtención de la matriz de contaminación, no hay necesidad de simulaciones hidráulicas adicionales y, como resultado, esos conceptos sustitutos de aguas arriba y aguas abajo son adecuados.
- b) Las dos ideas mencionadas anteriormente sugieren que se deben trazar límites para definir áreas de mayor interés para la colocación de los sensores. Por razones obvias, los nodos en esta área de interés no deben estar demasiado cerca de las fuentes de agua, ni en las zonas finales de la red de tuberías.
- c) Se debe tener en cuenta otra consideración: después de que un evento de contaminación es detectado hay un tiempo de reacción por parte de la compañía de agua para tomar las acciones adecuadas.

Estas ideas tienen consecuencias importantes para la mejora del proceso de búsqueda:

- a) Se debe poner un primer límite para excluir los nodos demasiado aguas arriba como candidatos a albergar algún sensor. Los puntos anteriores a este límite serán aquellos ubicados a una cierta distancia previamente establecida de las fuentes de agua. Los técnicos deben definir esa distancia antes de ejecutar el algoritmo, y los agentes PSO deberán ser instruidos basados en reglas, teniendo así el encargo de no seleccionar nodos ubicados

demasiado cerca de una fuente de agua, es decir, a una distancia menos de la establecida.

- b) Un segundo límite para la ubicación de sensores debe considerar el tiempo de reacción de la empresa para decidir las acciones de operación que se ejecutarán y el tiempo para ejecutar esas acciones de operación en la red física. Los agentes basados en reglas utilizan ese tiempo de reacción, que también deben establecer los técnicos, para excluir ciertas ubicaciones candidatas para la colocación de sensores. En general, cualquier nodo aguas abajo de ese límite no será elegible para albergar un sensor de calidad del agua.

La idea de establecer fronteras aguas arriba y aguas abajo para la no ubicación de sensores de calidad del agua ayuda a reducir el espacio de búsqueda del problema. Sin embargo, esto no influye en el tamaño de la matriz de contaminación asociada con el problema. Recordemos que esta matriz de contaminación es el resultado del tiempo que se tarda en contaminar cada nodo después de contaminar cada punto susceptible de contaminación de la red de agua. Para los modelos de redes de agua “grandes”, no es posible mantener toda la matriz de contaminación en RAM a la vez. Esa es la razón por la que la implementación de software asociada a esta investigación utiliza una base de datos MS SQL para guardar las matrices de contaminación y de consumo de agua contaminada. En otras palabras, la evaluación de las funciones objetivo se ejecuta directamente en el servidor MS SQL y nunca es necesario que estas matrices se carguen en su totalidad. Sin este importante detalle de implementación, la solución para modelos de redes de agua “grandes” no sería asequible con el hardware que se usa normalmente en las computadoras de escritorio.

4.4 La hibridación entre la teoría de redes y el algoritmo evolutivo ASO

Además de las características mencionadas en el Capítulo 3, ASO tiene una importante habilidad: permite aceptar, dentro del enjambre (la población) que evoluciona, a otros agentes que se sugieren desde el exterior. Específicamente, ASO es capaz de utilizar soluciones sugeridas por el ingeniero, derivadas de su

experiencia y know-how, como candidatas a solución. El algoritmo recibe tales soluciones y las utiliza para intentar mejorar el frente de Pareto.

Esta inyección externa de diversidad, puede no ser de interés. Es difícil para el humano concebir soluciones como las que se manejan en este trabajo, dada la complejidad del problema. No obstante, en otras ocasiones, tales soluciones pueden ser de gran interés para la ejecución del proceso. En efecto, a veces, el algoritmo, a pesar de su componente aleatoria, incide demasiado, debido a su componente inercial, en zonas específicas del frente de Pareto, dejando despobladas otras zonas que pueden ser igualmente fértiles para albergar soluciones de interés. El ingeniero a cargo del proceso puede observar tal circunstancia y sugerir al algoritmo soluciones-semilla que le inviten a explorar otra parte del espacio de búsqueda que mejore la zona despoblada.

Como acabamos de decir, concebir buenas soluciones desde la experiencia puede resultar inviable, dada la complejidad del problema que se aborda.

Aquí es donde interviene la teoría de grafos y de redes complejas que se ha estudiado en este capítulo. En efecto, ese trabajo previo permite utilizar el propio fundamento de carácter topológico de la teoría de grafos para generar soluciones interesantes que, en todo caso, van a ser evaluadas, como cualesquiera otras, con las funciones objetivo dadas en el Capítulo 3. De tal manera, que se podrá evaluar qué tanto se acercan las aproximaciones previas obtenidas con las diferentes medidas de centralidad y conectividad basadas en la estructura del grafo de la RDA.

La idea de aportar soluciones próximas a las soluciones óptimas de Pareto, al igual que suele suceder con los algoritmos iterativos clásicos, permite acelerar la convergencia. También empuja al algoritmo a dedicarse más eficientemente a zonas específicas elegidas con tales criterios topológicos. Hay que recordar que los algoritmos evolutivos tienen también mecanismos de explotación de buenas zonas previamente identificadas. Tal explotación se realiza mediante comportamientos (no aleatorios) de los agentes que tratan de no olvidar la mejor posición ocupada por ellos mismos, al tiempo que otra componente del algoritmo les hace seguir al líder (mejor solución) del proceso. Así que, una vez más, la

adición de diversidad dada por estas soluciones exógenas acelera el proceso, reduce el espacio de búsqueda, perfila mejor y de manera más completa el frente de Pareto y, en definitiva, permite obtener soluciones de mejor calidad.

5 Caso de estudio y resultados: localización de una red de sensores en una RDA

Con el objetivo de aplicar los métodos descritos anteriormente, en este capítulo se hace un análisis de un caso de estudio de una RDA, de acuerdo con la metodología planteada de la hibridación de la teoría de redes y el algoritmo ASO. Primero, se hace referencia a las características topológicas y de estructura de la RDA con el cálculo de las medidas básicas de teoría de redes. Después, se presentan los resultados obtenidos con el proceso de selección de los nodos-sensores obtenidos a través de la importancia relativa de los nodos, dada por el algoritmo *PageRank* y las medidas de intermediación (*betweenness*) y centralidad del autovector (*eigenvector centrality*). Por último, se presenta la solución de la configuración de la red sensores, resultado de la optimización multiobjetivo del algoritmo ASO.

5.1 Caso de estudio

El caso de estudio utilizado en esta tesis corresponde a la RDA simplificada de la zona norte de la ciudad de Santa Ana de Coro, capital del Estado Falcón, Venezuela. La ciudad de Santa Ana es una ciudad costera semi árida, con gran importancia histórica que la ha llevado a ser declarada patrimonio mundial de la humanidad por la UNESCO en 1993¹⁸. Con una superficie de 1.805 km² y una población al de 211.537 habitantes (en el año 2016), la ciudad es un centro urbano con uso de suelo residencial casi en su totalidad. De ahí la importancia de analizar esta RDA, además de que a lo largo de los años ha requerido de ampliaciones continuas. La Figura 5.1 muestra el esquema de RDA, que para efectos de esta tesis la denominaremos “RDA de Coro”. La RDA de Coro ofrece un suministro por gravedad y está formada por 2889 nodos (N), incluyendo 2 fuentes de abastecimiento y 3713 tuberías (M) que suman un total de 340 km, con diferentes diámetros y materiales.

¹⁸ Fundada en 1577, fue una de las primeras ciudades coloniales de América. Posee aproximadamente 600 edificios históricos y con sus construcciones en tierra únicas en toda la región del Caribe, la ciudad de Santa Ana es el único ejemplo subsistente de una fusión lograda de las técnicas y estilos arquitectónicos autóctonos, mudéjares españoles y holandeses. (<http://whc.unesco.org/en/list/658>).



Figura 5.1. Representación de la RDA de Coro.

Para conocer las propiedades de la RDA de Coro desde el enfoque de la teoría de redes, se obtuvieron medidas básicas de la estructura de la red. El cálculo de las diferentes medidas se realizó con el lenguaje R (R Development Core Team, 2008), utilizando la librería *igraph* (Csárdi et al., 2016) que es apropiada para el análisis de grafos.

En la Tabla 5.1 se muestran los indicadores obtenidos de la RDA de Coro. La interpretación de los indicadores obtenidos de la RDA de Coro está basada en la diversa literatura consultada y en los trabajos de diferentes autores que han abordado el análisis de diferentes redes complejas.

Tabla 5.1. Indicadores de la RDA de Coro, basados en la teoría de redes.

Indicador	Valor obtenido
Número de nodos	$N = 2889$
Número de tuberías	$M = 3713$
Relación de tubería por nodo	$(e) = 1,28$
Densidad de conexión de las tuberías	$q = 8,89 \times 10^{-4}$

Indicador	Valor obtenido
Grado promedio de los nodos	$\langle k \rangle = 2,57$
Grado máximo de los nodos	$k_{max} = 5$
Coficiente de agrupamiento	$C = 5,84 \times 10^{-3}$
Longitud promedio del camino más corto	$L = 4823,35$
Díámetro	$d = 114$
Coficiente de agrupamiento	$C = 5,84 \times 10^{-3}$
Mallado	$R_m = 0,14$
Brecha espectral	$\Delta\lambda = 0,0328301$
Conectividad algebraica	$\lambda_2 = 7,707044 \times 10^{-5}$

Estas medidas estructurales mencionadas en el Capítulo 4 cuantifican los patrones de conectividad entre los componentes de la red, así como la redundancia y la robustez.

- La densidad de tuberías (q), como se ha visto, es el indicador más básico de la vinculación de los enlaces o escasez general de la estructura de una red. Los valores de los atributos de la RDA de Coro muestran que es una red dispersa: el número de tuberías está muy por debajo del máximo posible; expresado en porcentaje, sería el 0,0889% de densidad de tuberías.
- Respecto de la relación de enlace o tubería por nodo (e) y el grado medio de los nodos ($\langle k \rangle$), para la RDA de Coro sus valores son $e = 1,28$ y $\langle k \rangle = 2,57$, que la describen como una red conectada y mallada; asimismo, el grado máximo de los nodos $k_{max} = 5$ refleja la mayor conectividad de los nodos.
- En referencia al coeficiente de agrupamiento o transitividad C , que representa la redundancia de una red, el valor de $C = 5,84 \times 10^{-3}$ de la RDA de Coro, refleja que es una red con forma de cuadrícula y con bucles

no triangulares en su estructura. Esto se debe a que redes con esas características suele tener un valor pequeño de C .

- Por último, el valor obtenido del coeficiente de mallado de la RDA de Coro es $R_m = 0,14$. Este indicador también proporciona una estimación de la redundancia topológica de las RDAs respecto al número de bucles independientes (no necesariamente triangulares) presentes como un porcentaje de los bucles máximos posibles. Es decir, la RDA de Coro tiene un 14% de redundancia.
- Con referencia al diámetro (d) y la longitud promedio del camino más corto (L) de la RDA de Coro, estos indicadores resultan del cálculo de la matriz de las distancias geodésicas entre todos los nodos de la RDA. El diámetro de la red es la máxima excentricidad de la red, es decir, el valor máximo de los caminos o trayectorias más cortas entre dos nodos. Por lo tanto, $d = 114$ es el número mínimo de tuberías atravesadas o recorridas para llegar de un nodo a otro de la RDA de Coro. Por otro lado, la longitud del camino característico de la RDA de Coro $L = 4823,35$, es el promedio de las longitudes del camino más cortas y está calculada en función de la longitud de sus tuberías.
- Por último, las medidas conectividad algebraica y brecha espectral son $\lambda_2 = 7,707044 \times 10^{-5}$ y $\Delta\lambda = 0,0328301$, respectivamente.

5.2 Obtención de los nodos-sensores de la RDA de Coro con el proceso de teoría de redes

La determinación de los nodos-sensores de la RDA de Coro se describe a continuación de acuerdo con el esquema de la Figura 5.2.

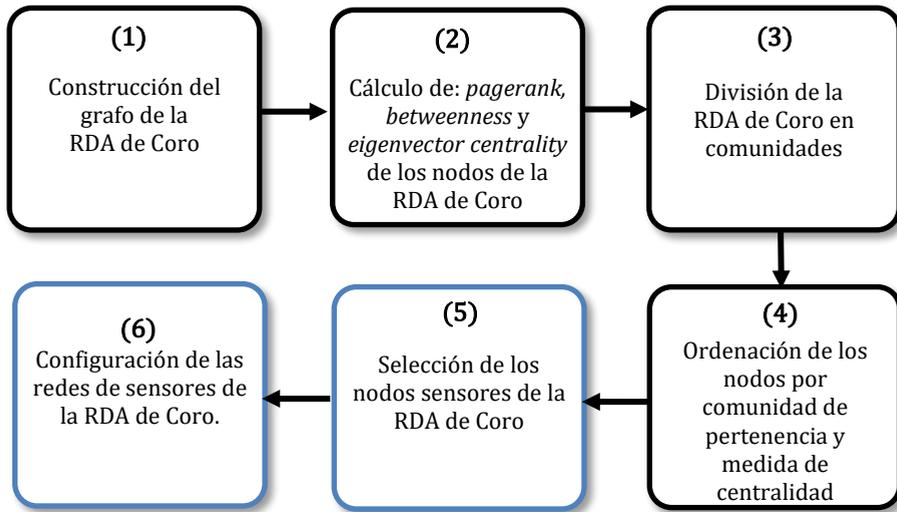


Figura 5.2: Esquema general del proceso para la obtención de los grupos de nodos sensores de la RDA de Coro.

(1) Construcción del grafo de la RDA de Coro

El primer paso del proceso consistió en abstraer el grafo de la RDA de Coro, esto se realizó conservando los atributos de los nodos y tuberías: coordenadas geográficas, elevación y demanda de los nodos; y longitud, diámetro y rugosidad de las tuberías. El resultado de la construcción del grafo de la RDA de Coro se muestra en la Figura 5.3.

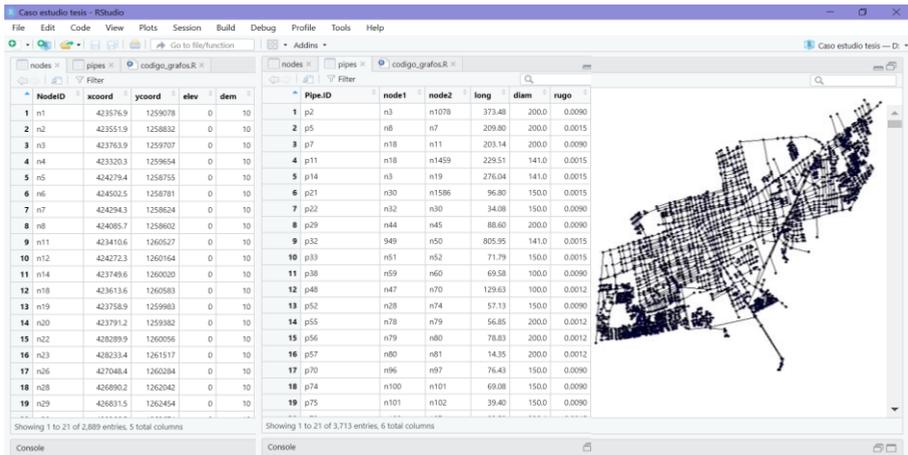


Figura 5.3: Grafo de la RDA de Coro en la interfaz del lenguaje R.

(2) Obtención de las medidas de centralidad de la RDA de Coro: centralidad del autovector, pagerank, e intermediación.

Como se explicó en el apartado 4.1.2, las medidas de centralidad de las redes nos permiten identificar aquellos nodos que tienen importancia en la cohesión de toda la red, pues su eliminación o desactivación y, en el caso de una RDA, un ataque o un fallo en la operación, podría causar fallos en otros nodos (fallos encadenados) o comprometer la operatividad de toda la RDA. Además, los nodos con mayor centralidad representan puntos de mayor tránsito en un camino de un nodo a otro.

El ordenamiento de los nodos de una RDA es una de las aplicaciones de teoría de redes que se ha utilizado para el análisis de la vulnerabilidad y robustez. En el caso de las medidas espectrales provenientes de la matriz de autovectores asociada a la matriz de adyacencia de la red, van a más allá y revelan información como la conectividad de la red o sus niveles de tolerancia a los fallos en el tránsito del flujo. Así, hemos considerado que las tres métricas propuestas para la selección de los nodos de una RDA: centralidad del autovector, *pagerank* y la intermediación, abren posibilidades en la tarea de ordenación y selección de los nodos. Esta ordenación no sólo estará relacionada con la evaluación de las vulnerabilidades, sino que es útil para análisis de calidad del agua, localización de sensores y el estudio de la viabilidad de planes de rehabilitación, entre otros.

Además, el establecer un *ranking* en los nodos es una herramienta de gran utilidad, por ejemplo, a la hora de buscar criterios eficientes de sectorización.

En el caso de la centralidad del autovector y de la intermediación de los nodos de la RDA de Coro, se calcularon considerando las longitudes de sus tuberías como pesos, para ponderar los caminos. En el cálculo del *pagerank* de los nodos de la RDA se estableció un factor de amortiguamiento (*damping factor*) igual a 0,85, que es valor más aceptado y utilizado para el análisis de redes complejas (ver ecuación (19)). El resultado de las tres medidas son los vectores asociados a la importancia relativa de los nodos y su nivel de centralidad en la RDA.

(3) División de la RDA de Coro en comunidades.

La aplicación del algoritmo *cluster walktrap* (descrito en apartado 4.1.3) para la división de la RDA de Coro, busca subgrafos densamente conectados, también llamados comunidades, a través de paseos o caminos aleatorios. La idea es que las caminatas cortas aleatorias tienden a permanecer en la misma comunidad, de tal manera que la división de la RDA está basada en la proximidad o alcance entre los nodos. Esto favorece que la conformación de las comunidades sea homogénea espacialmente. En la Figura 5.4 se muestra la división de la RDA de Coro en 41 comunidades. El número de comunidades corresponde a la máxima modularidad que es igual a 0,9142056. Al respecto, la modularidad de una red con respecto a la división mide qué tan buena es la división o qué tan separados están los nodos entre sí. Se define como:

$$Q = \frac{1}{M} \sum_{i,j} \left[A_{ij} - \frac{k_i k_j}{2M} \right] \delta(c_i, c_j), \quad (25)$$

donde M es el número de tuberías, A_{ij} es el elemento de la matriz de adyacencia A en la fila i y la columna j ; k_i es el grado del nodo i y k_j el grado del nodo j (es decir la probabilidad de un enlace entre dos nodos es proporcional a sus grados); c_i es el componente del nodo i y c_j de j . La sumatoria se extiende a todos los pares de nodos i, j y $\delta(c_i, c_j) = 1$ y $\delta(c_i, c_j) = 0$, en otro caso. Si las tuberías de la RDA (enlaces) están ponderadas, entonces serán considerados como un elemento de la matriz de adyacencia y k_i es la suma de los pesos de las tuberías adyacentes al nodo i . En el caso de la RDA de Coro, los pesos de las tuberías están dados por las

longitudes de estas. El resultado para la modularidad es un escalar numérico de la configuración dada de la RDA. El rango de valores de $Q \in [-1,1]$, cuanto mayor es su valor mejor es la partición, es decir, las comunidades encontradas están densamente conectadas internamente (esto es que hay más enlaces de los esperados aleatoriamente) y dispersamente conectadas entre sí. Con base en lo anterior, la división de la RDA de Coro en 41 comunidades corresponde a la mejor distribución espacial.

En la Figura 5.4 se puede observar que la distribución espacial de las comunidades es relativamente homogénea. En general, las comunidades están conformadas entre 25 y 200 nodos, excepto siete comunidades que contienen un mínimo número de nodos, entre 4 y 15. Información más específica de las comunidades se presenta en los Anexos.



Figura 5.4: División de la RDA de Coro en 41 comunidades.

(4) Ordenación y selección de los nodos de la RDA de Coro por comunidad de pertenencia y medida de centralidad.

Una vez que se obtuvieron las medidas de centralidad y la división en comunidades de la RDA, se ha propuesto una selección de nodos condicionada

por dos criterios: el valor más alto de la medida de centralidad y la comunidad de pertenencia. Lo anterior consistió en una búsqueda iterativa hasta encontrar los nodos que cumplieran los dos criterios de ordenación.

(5) Configuración de las redes de sensores en la RDA de Coro para cada medida de centralidad.

Con base en el proceso de la selección de los nodos que representan el valor más alto de cada medida de centralidad en cada comunidad, se obtuvieron las configuraciones espaciales de los nodos – sensores de la RDA de Coro, es decir, un total de 41 sensores que corresponden a las 41 comunidades. Adicionalmente se obtuvieron las configuraciones de 20 sensores, bajo las mismas medidas de centralidad y división de la red en comunidades. Estos resultados posteriormente serán analizados en conjunto con las soluciones obtenidas por el algoritmo ASO. Los resultados de los 20 sensores se pueden encontrar en los Anexos.

La Figura 5.5 muestra la ubicación de los 41 nodos – sensores obtenidos con el valor del *pagerank* de los nodos y distribuidos en función de la división de las comunidades. Se puede observar, en general, una distribución que abarca diversas áreas de la RDA, lo que posteriormente puede representar grandes ventajas.

La Tabla 5.2 muestra los datos de 10 nodos – sensores más relevantes, extraídos de entre el resultado de 41, por su valor alto de *pagerank*. Esta selección no excluye al resto de sensores en el proceso de optimización.

En la sección de Anexos, se indica en repositorio en el que se incluye la información completa de los 41 sensores referente a su valor de *pagerank*, intermediación y centralidad del autovector. Asimismo, se encuentra la misma información para las configuraciones de 20 nodos. Las distribuciones espaciales de estos sensores se muestran en el Anexo 1.



Figura 5.5. Ubicación de 41 sensores en la RDA de Coro según el valor de *pagerank* de los nodos.

Tabla 5.2. Nodos – sensores con el valor más alto de *pagerank* y comunidad de pertenencia.

ID del nodo sensor	Comunidad de pertenencia	Valor de <i>pagerank</i> de los nodos
897	C25	0,00068865
1034	C37	0,00064964
713	C6	0,00064253
n456	C14	0,00062035
n1362	C33	0,0006007
n1402	C8	0,00059559
n353	C9	0,00058795
427	C7	0,0005857
879	C12	0,00058227
n656	C1	0,00058008

Asimismo, la Figura 5.6 se refiere a la ubicación de los sensores para la medida de *betweenness* o intermediación. En este caso, se observa una distribución menos homogénea de los sensores, ya que el valor de la intermediación de los nodos representa la mayor conectividad o grado de los nodos. De ahí que la mayor cantidad de sensores estén ubicados en dónde hay una mayor densidad de tuberías entrantes y salientes a los nodos de la RDA. También a modo de resumen, la Tabla 5.3 presenta la información de los diez sensores con el valor de intermediación más alto.

Por último, la Figura 5.7 muestra la ubicación de los sensores obtenidos de acuerdo con el valor de centralidad del autovector. La Tabla 5.4 presenta la información correspondiente a los sensores con los diez valores más altos de la centralidad del autovector.



Figura 5.6. Ubicación de 41 sensores en la RDA de Coro según el valor de intermediación de los nodos.

Tabla 5.3. Nodos – sensores con el valor más alto de intermediación y comunidad de pertenencia

ID del nodo sensor	Comunidad de pertenencia	Valor de intermediación de los nodos
514	C30	1718800
n2467	C4	1566874
n330	C17	1495619
n633	C21	1447160
n36	C16	1213942
n851	C31	1195332
n149	C10	1069001
n964	C18	1027040
n643	C15	1017918
n944	C9	1015622



Figura 5.7. Ubicación de 41 sensores en la RDA de Coro según el valor de centralidad del autovector.

Tabla 5.4. Nodos – sensores con el valor más alto de centralidad del autovector y comunidad de pertenencia

ID del nodo sensor	Comunidad de pertenencia	Valor de centralidad del autovector de los nodos
28	C4	2,19E+09
n1927	C28	2,29E+08
n2461	C23	1,02E+07
n936	C9	4,98E+06
n2165	C41	9,98E+05
n186	C20	1,17E+02
n2256	C13	3,21E+01
n2023	C5	0,70325794
71	C18	7,01E-02
n159	C10	6,96E-02

Hay que tener en cuenta que los nodos de la RDA con valores altos de *pagerank* o centralidad del autovector, no sólo son nodos con mayor centralidad, sino que también tiene una buena conectividad con aquellos nodos en los que hay más tuberías entrantes y/o salientes, por lo que serían nodos en los que hay más caudal circulante. Por otro lado, a pesar de que las propiedades de escalabilidad que tiene el *pagerank* son para redes muy grandes, en el caso de la RDA de Coro se ha logrado una buena aproximación espacial. De igual manera, la configuración de sensores de la RDA de Coro basada en la centralidad del autovector, muestra su buen funcionamiento en redes de tamaño medio, como se puede ver en la Figura 5.7.

La importancia de los resultados obtenidos es que la ubicación de los nodos – sensores corresponde a una distribución espacial adecuada que toma en cuenta nodos o áreas de mayor conectividad, en cuanto a tuberías. Esto se traduce en puntos en los que el tránsito del caudal y la demanda puede ser mayor que en el resto de la RDA. En el caso de los nodos con mayor *pagerank* o centralidad del autovector podrían ser puntos con mayor sensibilidad a la detección de contaminantes, debido a su interacción con otros nodos de gran conectividad.

Esta propuesta de ubicación de sensores también da opción a que el tomador de decisiones analice y elija aquellos nodos de una medida u otra y hacer una combinación a conveniencia. Posteriormente, esta selección, será optimizada a

través del enfoque multiobjetivo para incorporar otros criterios no considerados como los hidráulicos, económicos o sociales.

La obtención de las diferentes configuraciones de sensores, prueban los alcances de las medidas de centralidad en combinación con la división espacial de la RDA. Adicionalmente, como ya se ha mencionado antes, estos resultados serán información de entrada para el modelo multiobjetivo.

5.3 Red de sensores aplicando la metodología: Teoría de redes – Optimización multiobjetivo

Las aproximaciones a la solución de la ubicación de sensores, obtenidas con las medidas de centralidad vistas en la teoría de redes en el apartado 5.2 serán empleadas como información de entrada para que el algoritmo multiobjetivo las considere y permita una convergencia más rápida hacia la solución. Algunos de los frentes de Pareto que se van a obtener (cuando corresponda) considerarán, dentro del espacio de soluciones, las obtenidas con cada medida de centralidad.

El algoritmo multiobjetivo utilizado en esta tesis, ASO, según se ha dicho en el Capítulo 4, se ha integrado en WATERing ©, un paquete de software para análisis y soporte de decisiones en SDAs. WATERing se conecta con el kit de herramientas de EPANET para realizar las simulaciones hidráulicas necesarias que generan la matriz de contaminación y la matriz de agua contaminada consumida. Con esta información, tal como se ha descrito en el Capítulo 3, ASO analiza varias alternativas para localizar sensores y selecciona aquellas alternativas que representan soluciones no dominadas. El software WATERing está dotado de un entorno de visualización avanzado para ayudar a análisis detallados de las soluciones de optimización obtenidas. Prácticamente se pueden crear tantos gráficos Pareto 2D como se desee en este entorno. Varias relaciones entre el tiempo de detección, la cobertura de protección y el impacto de la mala calidad del agua se pueden representar simultáneamente de una manera comprensible. Las soluciones elegidas en el gráfico se seleccionarán automáticamente en el resto de los gráficos que indican cómo se comportan con respecto a todos los criterios involucrados en el problema. Además, ASO permite exportar fácilmente ficheros de resultados en hojas de cálculo, lo que permite

una customización más adecuada por parte del usuario. En esta tesis, se emplean tales ficheros de resultados en formato csv para obtener los gráficos que aparecen en ese capítulo.

Para la obtención de los resultados, se utilizan los números de sensores sugeridos por la teoría de grafos, 20 y 41. Además, en la optimización se obtienen resultados para otros dos números de sensores 3 y 10, de modo que pueda observarse con facilidad cómo el aumento en el número de sensores conduce hacia soluciones más efectivas.

Para todos estos casos se ha realizado la simulación utilizando ASO, tal como se describe en el Capítulo 3, es decir, primero realizando las simulaciones mediante el uso de la SPE en EPANET, para construir las matrices de contaminación y de agua contaminada consumida que se almacenan adecuadamente. Posteriormente, con base en esas matrices, se realiza la optimización.

Hay que hacer observar aquí que el coste computacional de estas simulaciones y de la optimización final es realmente importante, por lo que se ha realizado mediante la distribución de cálculos en varios *cores* controlados por un servidor central. Los números básicos al respecto son los siguientes:

- Número de evaluaciones de la función objetivo realizadas para cada caso de estudio: 450 000
- Tiempo de cálculo aproximado: 27 horas de cálculo
- Los cálculos fueron ejecutados en paralelo en un ordenador con las siguientes características:
 - Tipo de procesador: Intel(R) Xeon(R) W-2123 CPU @ 3.60GHz
 - Procesador de 4 *cores* físicos con capacidad para ejecutar 2 *threads* diferentes en cada *core*.
 - Memoria RAM: 16,0 GB
 - Sistema operativo: Windows 10 Professional 64bit

A continuación, se presentan los resultados obtenidos y se añaden comentarios adecuados que explican las gráficas que se han obtenido.

La Figura 5.8 muestra una proyección 2D del frente de Pareto que representa el tiempo de detección promedio frente a un indicador del consumo de agua contaminada (el indicador normalizado es el consumo promedio de agua contaminada en l/s). Se pueden observar varias soluciones con tiempo de detección igual a cero. Hay que tener en cuenta que se está suponiendo, como se explicó en el Capítulo 3, que el tiempo de detección es igual a cero para eventos no detectados. Evidentemente, para la mayoría de estas soluciones, el consumo de agua contaminada es elevado. Obviamente, estas no son soluciones interesantes para el diseño final de la red de sensores que se tenga que implementar. Por otra parte, las soluciones de la parte superior del conjunto representado, aun siendo soluciones con tiempo de detección no nulo y, por tanto, soluciones que sí detectan el evento de contaminación, tampoco resultan interesantes porque corresponden a tiempos promedio elevados y a consumos grandes de agua contaminada de nuevo. Esta representación bidimensional de la frontera de Pareto apunta con bastante claridad hacia las soluciones próximas al vértice izquierdo de la figura con aspecto parabólico, que son soluciones que representarían (respecto de estos dos objetivos) mejores compromisos que las demás.

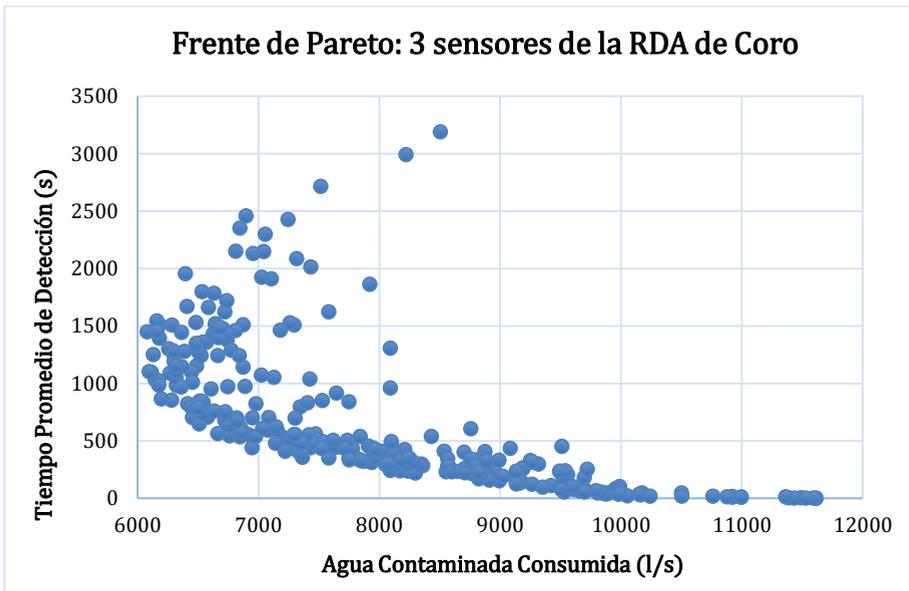


Figura 5.8. Solución del Frente de Pareto para la ubicación de 3 sensores en la RDA de Coro.

Las Figuras Figura 5.9, Figura 5.10 y Figura 5.11 son análogas a la anterior, pero ahora para los casos de 10, 20 y 41 sensores. Cabe realizar comentarios análogos a los expresados para el caso de 3 sensores. No obstante, además, ahora se puede observar cómo al aumentar el número de sensores, los valores del agua contaminada consumida mejoran (se reducen), como cabía esperar. Basta observar los valores de las escalas de los ejes. Cabe observar que no ocurre lo mismo (de manera directa) con el tiempo de detección promedio. La razón es que, como se ha indicado en el Capítulo 3, los eventos no detectados contribuyen con tiempos de detección nulos. Y es obvio, que cuantos menos sensores tiene la red, menos eventos es capaz de detectar y, por tanto, ese tiempo es menor. En este caso, como se ve más adelante, la compensación se realiza con los otros objetivos.

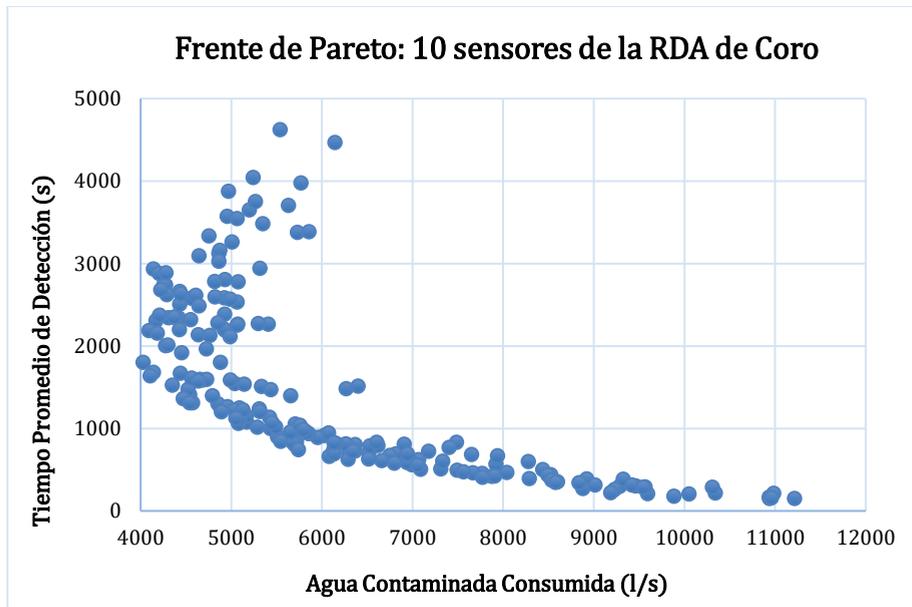


Figura 5.9. Solución del Frente de Pareto para la ubicación de 10 sensores en la RDA de Coro.

Además, en las Figuras Figura 5.10 y Figura 5.11 se han añadido las tres soluciones encontradas utilizando la teoría de grafos, representadas con símbolos y colores distintos a los puntos azules, que corresponden a las

soluciones proporcionadas por ASO. Puede observarse cómo estas soluciones, aun quedando finalmente dominadas por otras soluciones del frente de Pareto, no obstante, quedan relativamente próximas a tal frente, lo que significa que proporcionan, dentro de la idea de esta tesis, buenas aproximaciones que, sin duda, permiten al algoritmo de optimización iniciar su búsqueda desde buenas posiciones iniciales. Sin embargo, también resulta claro que un modelo puramente topológico no proporciona, en general, las mejores soluciones, como cabría esperar de un modelo que no considera la hidráulica del problema. La topología ayuda a describir mejor la conectividad entre los elementos de la RDA. La hidráulica, por otra parte, toma en cuenta la topología y define, además, los sentidos y velocidades de circulación del agua. La forma en que un contaminante se distribuirá a través de la RDA estará condicionada no sólo por la topología, sino también por los sentidos y velocidades del agua circulante en la red.

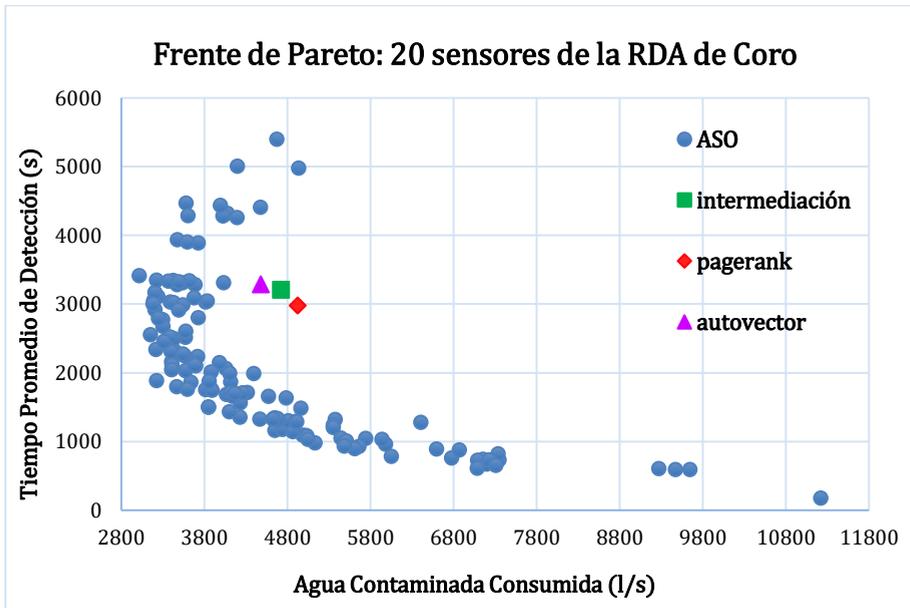


Figura 5.10. Solución del Frente de Pareto para la ubicación de 20 sensores en la RDA de Coro, con las soluciones de centralidad.

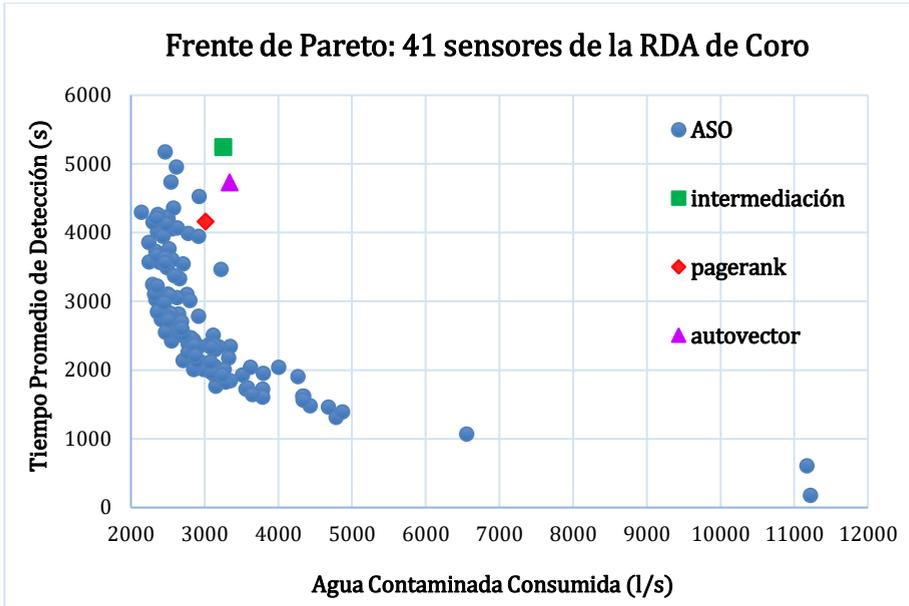


Figura 5.11. Solución del Frente de Pareto para la ubicación de 41 sensores en la RDA de Coro con las soluciones de centralidad.

La Figura 5.12 relaciona la cantidad de fallos de detección con el tiempo promedio de detección de nuevo para el caso de 3 sensores. Las soluciones con tiempos de detección muy altos representan diseños de redes de sensores con sensores ubicados en nodos muy aguas abajo de la red. En estos casos, se tarda más en detectar un contaminante (como promedio, considerando todos los posibles eventos de contaminación) pero el fallo de detección es mucho menor. Por el contrario, muchas soluciones con tiempos de detección demasiado pequeños, presentan grandes fallos de detección, y corresponden, principalmente, a diseños con sensores demasiado aguas arriba o, como se ha comentado más arriba, a casos con tiempo de detección igual a cero, es decir, eventos no detectados. Las soluciones que representan el mejor compromiso (correspondientes a las próximas al vértice de la parábola de las figuras anteriores) se ubican en la zona media de línea de puntos presentada, en las proximidades y por encima y a la derecha del punto (1000,2600), para este frente de Pareto correspondiente al caso de 3 sensores.

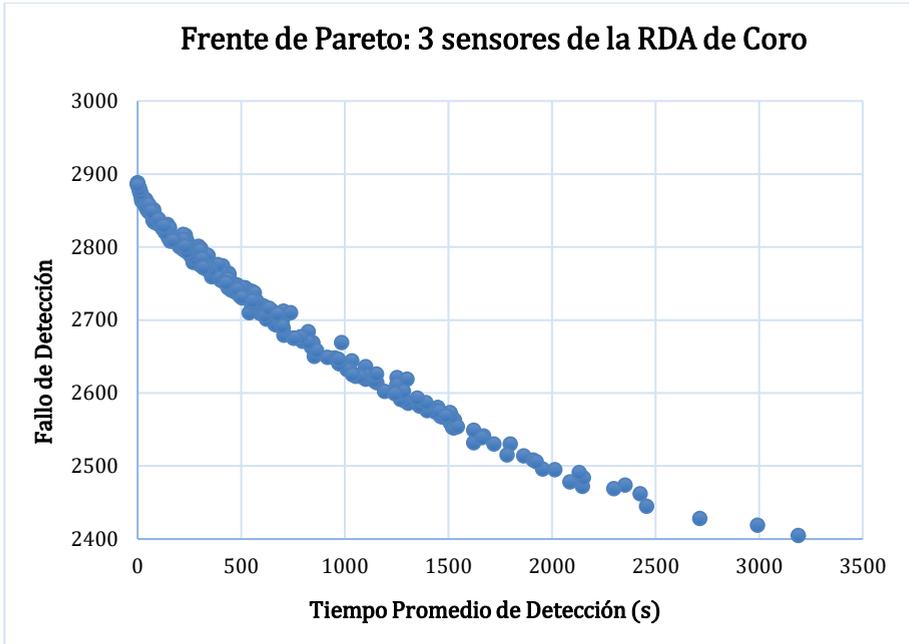


Figura 5.12.. Tiempo de detección promedio vs número de fallos de detección, para el caso de 3 sensores.

También aquí, al presentar las proyecciones 2D con las mismas variables para los casos de mayor número de sensores ubicados (de nueve, 10, 20 y 41), se puede observar en las correspondientes Figuras, Figura 5.13 a Figura 5.15, cómo los valores de los objetivos varían conforme aumenta el número de sensores. En estos casos, sin embargo, los mejores puntos se van desplazando hacia la zona inferior derecha, mostrando la mejora progresiva en el fallo de detección, cuyos valores, se reducen hasta alrededor de 2300 para 10 sensores, de 2100 para 20 sensores y de 1800 para 41 sensores. Para el tiempo promedio, recordamos lo señalado más arriba: parece empeorar, pero ello revela que el número de eventos detectados es mayor, conforme hay más sensores. Es decir, estas mejoras, lo recordamos, se producen a costa de incrementar el coste de la inversión en sensores, junto con el de su instalación y mantenimiento.

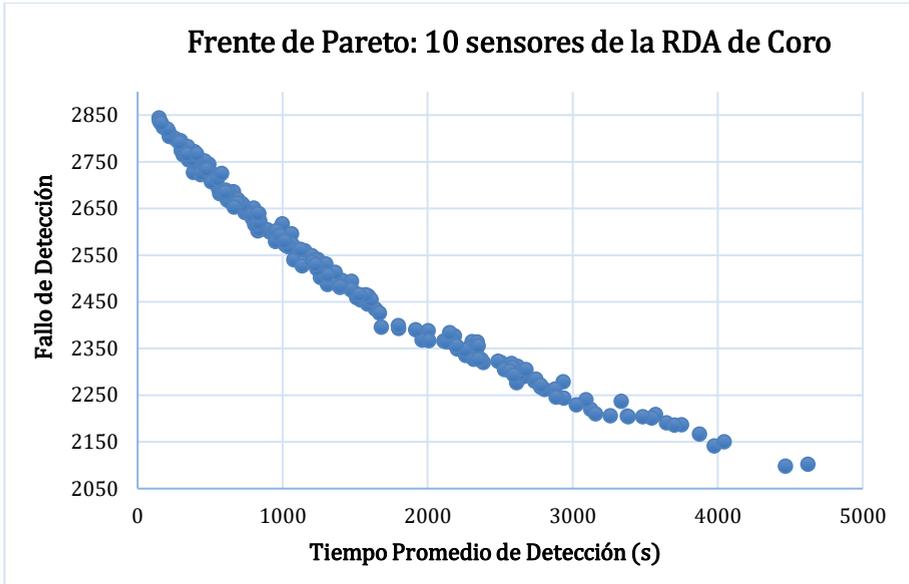


Figura 5.13. Tiempo de detección promedio vs número de fallos de detección, para el caso de 10 sensores.

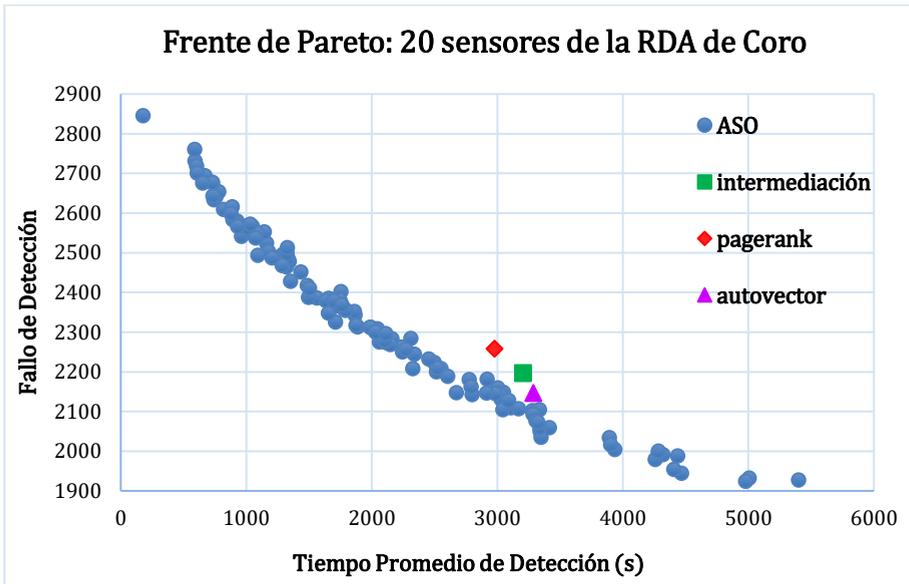


Figura 5.14. Tiempo de detección promedio vs número de fallos de detección, para el caso de 20 sensores con las soluciones de centralidad.

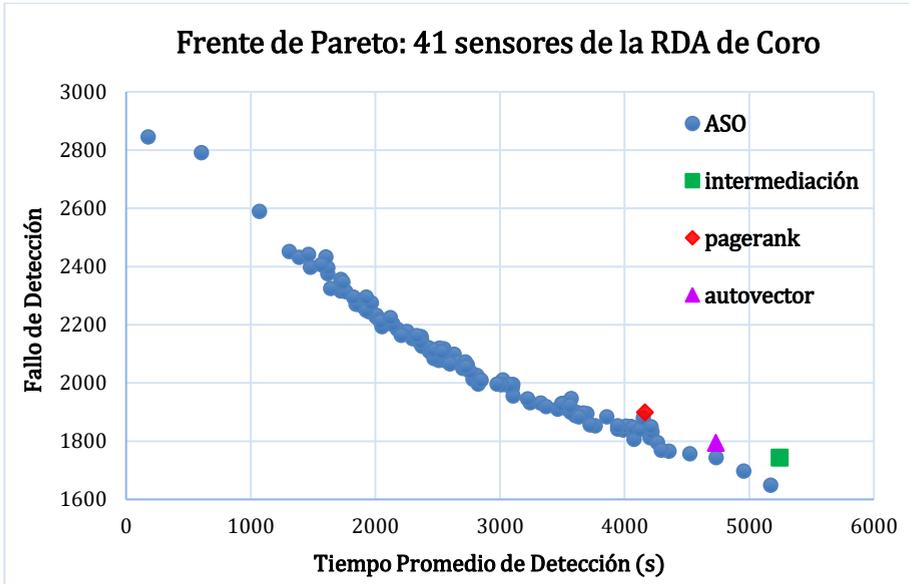


Figura 5.15. Tiempo de detección promedio vs número de fallos de detección, para el caso de 41 sensores con las soluciones de centralidad.

La Figura 5.16 muestra que el indicador de agua contaminada consumida puede incrementarse debido a dos razones principales: o corresponde a soluciones con fallos de detección promedio altos (sensores ubicados demasiado cerca de las fuentes que no pueden detectar contaminación aguas abajo) o a soluciones en que los sensores quedan ubicados en nodos en posiciones muy aguas abajo, y se requiere más tiempo en promedio para detectar los efectos de la contaminación. La relación entre el tiempo de detección y el fallo de detección se ha mencionado anteriormente y se puede ver en la Figura 5.12. De nuevo, las soluciones situadas hacia el vértice de esta otra forma parabólica representan los mejores compromisos entre los objetivos. Se trata de las soluciones con Fallo de detección próximo a 2600 (ver más arriba) y situadas más a la izquierda, con consumos bajos de agua contaminada.

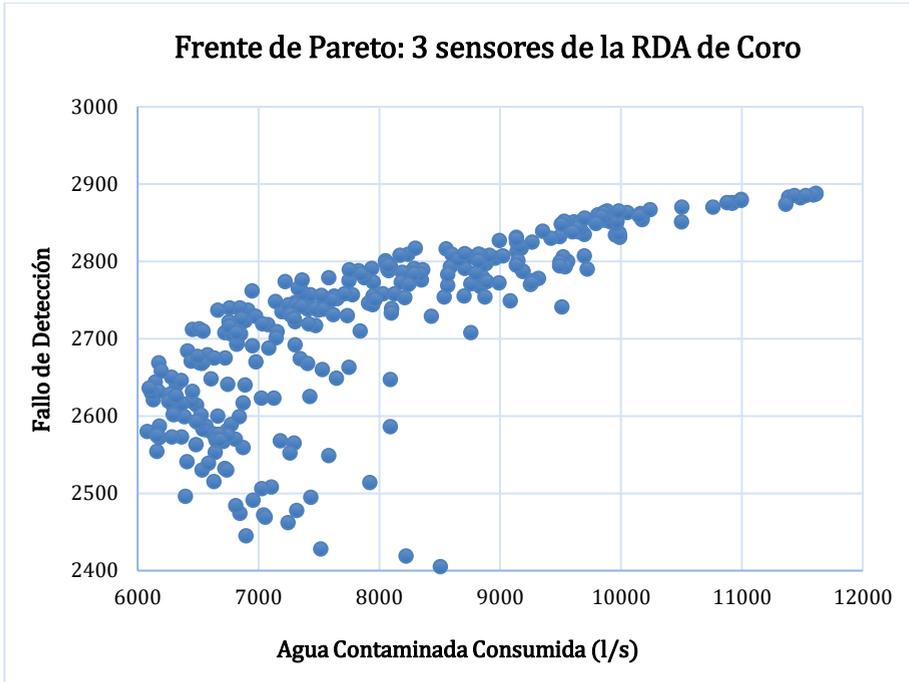


Figura 5.16. Fallo de detección promedio frente al consumo promedio de agua contaminada para la red de 3 sensores.

De manera análoga a como se ha hecho antes, las Figuras 5.17 a 5.19 presentan la misma proyección referida a los casos de 10, 20 y 41 sensores. Comentarios análogos pueden hacer observar cómo los valores de los objetivos son, naturalmente, mejores conforme aumenta el número de sensores. En este caso, los dos objetivos representados muestran una clara mejoría conforme aumenta el número de sensores.

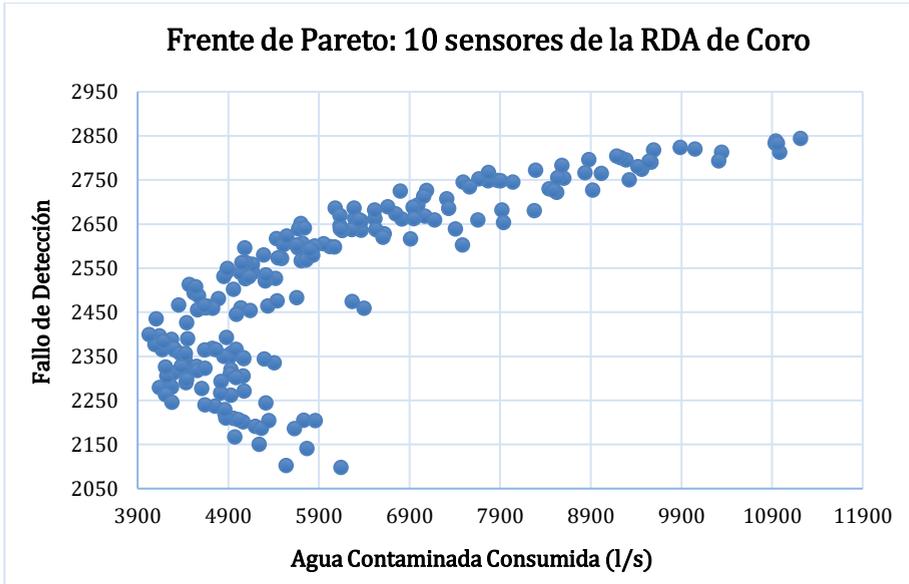


Figura 5.17. Fallo de detección promedio frente al consumo promedio de agua contaminada para la red de 10 sensores.

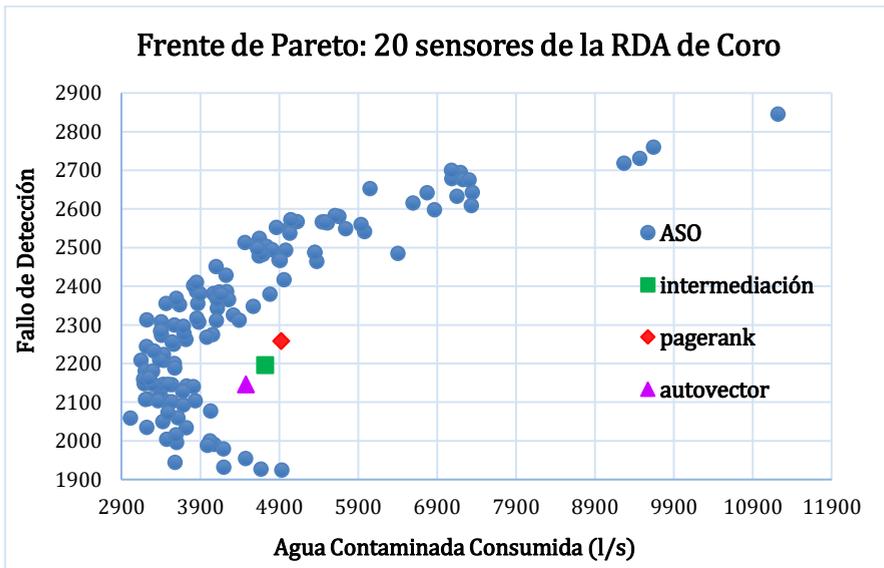


Figura 5.18. Fallo de detección promedio frente al consumo promedio de agua contaminada para la red de 20 sensores con las soluciones de centralidad.

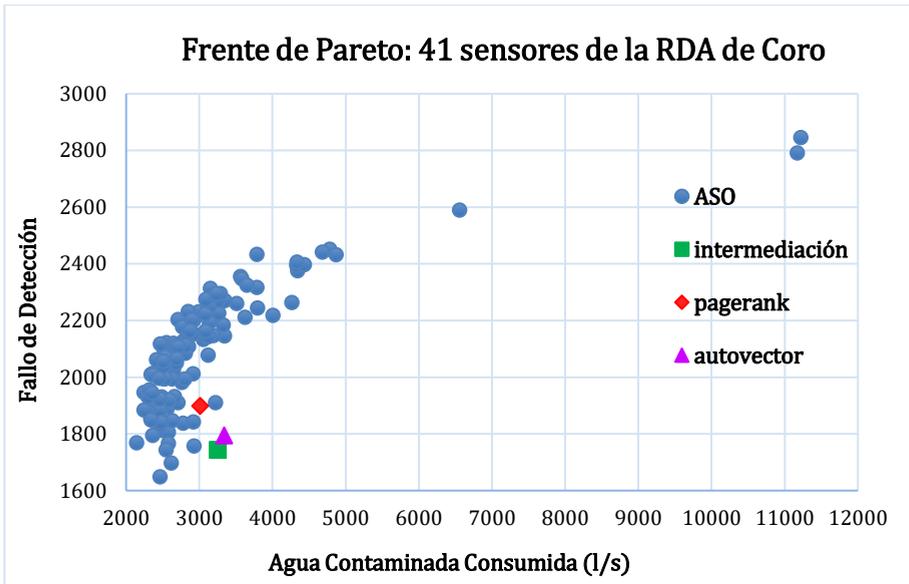


Figura 5.19. Fallo de detección promedio frente al consumo promedio de agua contaminada para la red con 41 sensores con las soluciones de centralidad.

De nuevo, observamos, en las dos últimas figuras, las soluciones proporcionadas por las medidas de centralidad de la teoría de grafos.

Las figuras anteriores ayudan a ubicar un cierto número de sensores en la red, concretamente 3, 10, 20 y 41 sensores, para el caso de las figuras mencionadas. Sin embargo, se requieren acciones adicionales para evaluar cuántos sensores deben instalarse en una red. Agregar sensores implica una inversión y esto tiene una clara limitación. Dependerá del presupuesto de la empresa y las recompensas por instalarlos.

Un posible enfoque consiste en establecer primero un presupuesto objetivo y luego calcular el número máximo de sensores a instalar sobre la base de ese presupuesto. El enfoque se ejecutará eventualmente para decidir un diseño óptimo único. Sin embargo, a pesar de ser un enfoque prácticamente realista, cabe preguntarse qué pasaría si, por ejemplo, la cobertura de la red que detecta la intrusión de contaminantes se incrementara un 25% por una inversión adicional del 10%. ¿La empresa de servicios públicos agregaría este 10% a su presupuesto en tal caso? Además, ¿qué pasa si la cobertura de la red se incrementó en un 30%, o si el tiempo de detección se redujo en un cierto

porcentaje? Siguiendo esta línea de pensamiento, la siguiente pregunta sería ¿cuánta mejora / beneficio se obtendría al aumentar el presupuesto para adquirir un sensor más? La misma idea puede hacer que los gerentes de servicios públicos se pregunten cuánta mejora / beneficio se perdería si deja de instalarse uno de los sensores para reducir el presupuesto. Responder a este tipo de preguntas requiere una representación de las mejoras / beneficios obtenidos en función del número de sensores instalados.

Decidir un presupuesto a priori sin considerar la mejora / beneficio recibido en función del número de sensores instalados no conducirá, en general, a una buena decisión. A pesar de que se puede decidir un presupuesto tentativo inicial, ambos aspectos deben combinarse antes de tomar una decisión final. Las mejoras / beneficios obtenidos se expresarán en términos de los objetivos considerados para la colocación de sensores: reducir el tiempo de detección de intrusión de contaminantes o problemas de calidad del agua; proteger a la población del consumo de agua bajo requisitos mínimos de calidad; y maximizando la cobertura de protección en la red. Estos tres objetivos deben combinarse con el objetivo adicional de mantener el coste dentro del marco de un presupuesto que pueda afrontar la empresa.

Sin entrar aquí en las cuestiones anteriores, que precisarían del conocimiento experto de los ingenieros a cargo de la red, presentamos a título ilustrativo una de las soluciones del frente de Pareto para 41 sensores. En la Figura 5.20 se presenta este frente de Pareto en el que se ha marcado con un círculo naranja la solución seleccionada (aquí se ha elegido una de las que menor cantidad de agua contaminada consumida presentan), que bien podría proporcionar un buen compromiso para la ubicación óptima de esa cantidad de sensores. En esta gráfica 3D aparecen, además las soluciones de centralidad obtenidas mediante la teoría de grafos.

La solución correspondiente se muestra en la Figura 5.21, en la que se presenta la red de Coro, y en la que se han destacado los nodos-sensores correspondientes a esa solución.

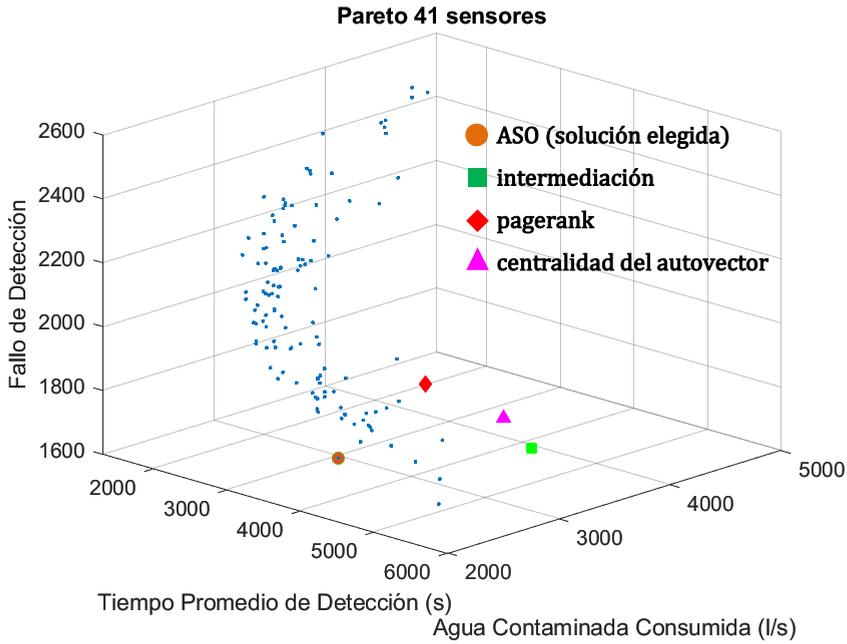


Figura 5.20. Frente de Pareto para 41 sensores, con solución elegida (círculo naranja) y soluciones de centralidad (otras figuras de colores).



Figura 5.21. Red de Coro con los sensores correspondientes a la solución elegida

Para finalizar, añadimos algunas figuras que permiten aportar más luz al estudio. Presentamos primero proyecciones 2D simultáneas de los frentes de Pareto estudiados. La Figura 5.22 presenta dichos frentes, para 3, 10, 20 y 41 sensores diferenciados por colores, para el caso de Tiempo promedio de detección frente a volumen de agua contaminada consumida. Siguiendo el mismo orden de la exposición anterior, la Figura 5.23 presenta la misma idea para el Fallo de detección frente al Tiempo de detección promedio. Y la Figura 5.24, hace lo propio respecto de los objetivos Fallo de detección frente a Agua contaminada consumida. Puede observarse con facilidad, cómo al incrementar el número de sensores todos los objetivos mejoran, como cabía esperar.

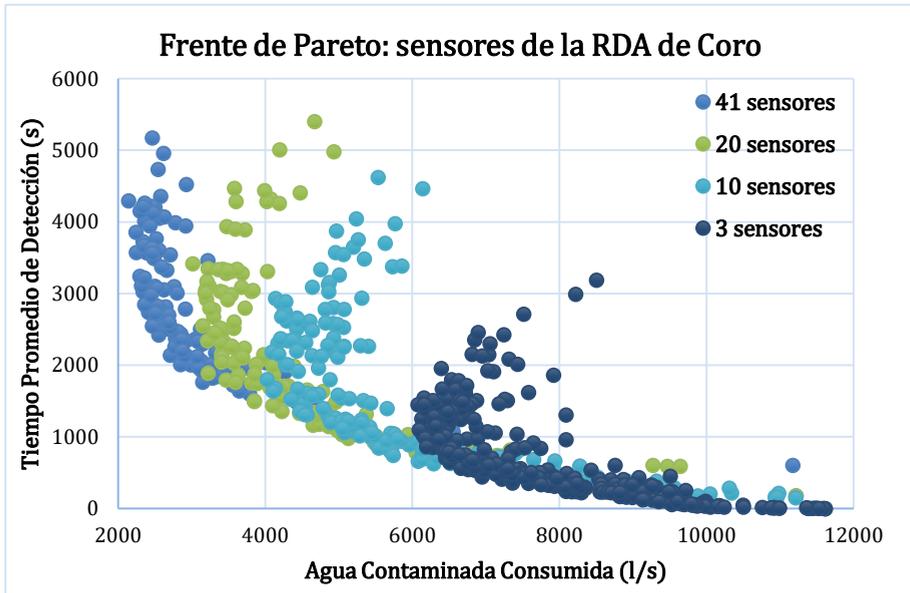


Figura 5.22. Consumo de agua contaminada vs Fallo de detección. Comparación de soluciones para 3, 10, 20 y 41 sensores.

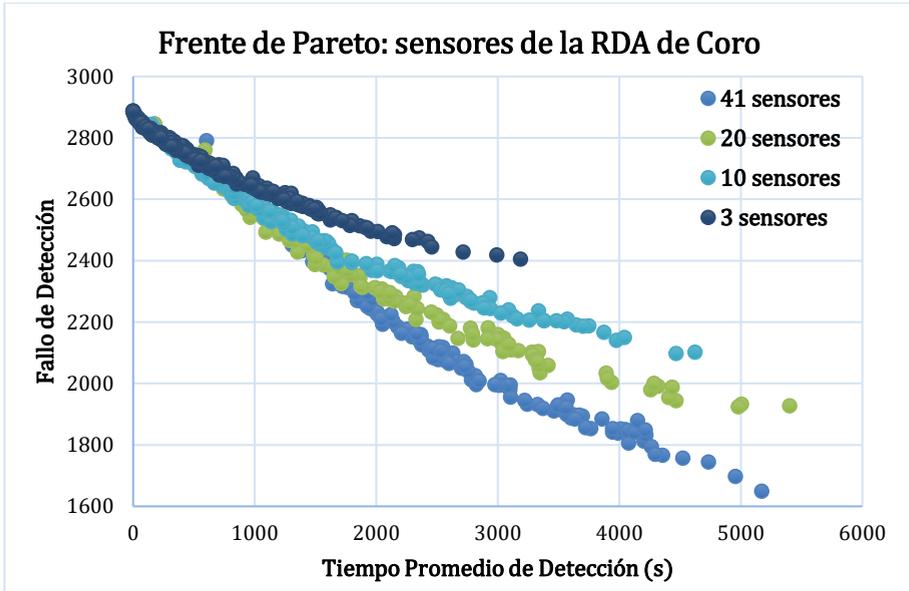


Figura 5.23. Tiempo promedio de detección vs Fallo de detección. Comparación de soluciones para 3, 10, 20 y 41 sensores.

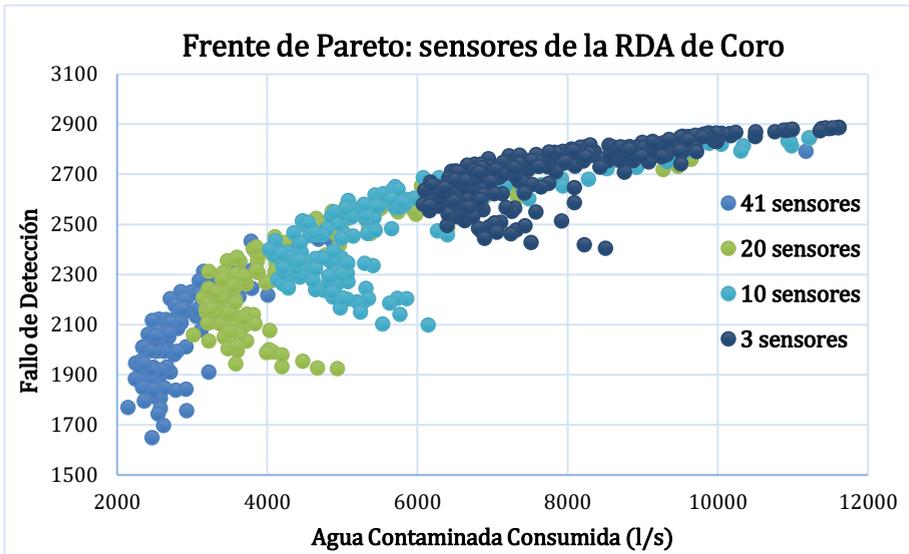


Figura 5.24. Tiempo promedio de detección vs Fallo de detección. Comparación de soluciones para 3, 10, 20 y 41 sensores.

Finalmente, puede ser ilustrativo añadir, además, otros gráficos tridimensionales de los frentes de Pareto estudiados. Con el fin de no sobrecargar, se presentan:

- el frente para 10 sensores (Figura 5.25); observamos que este gráfico incluye también las proyecciones sobre los tres planos coordenados que corresponden a las Figura 5.9, a la Figura 5.15 y a la Figura 5.19), y el conjunto de las cuatro fronteras estudiadas (Figura 5.26).
- Una vez más, en esta figura puede observarse la mejora que se produce, de manera natural, al aumentar el número de sensores.

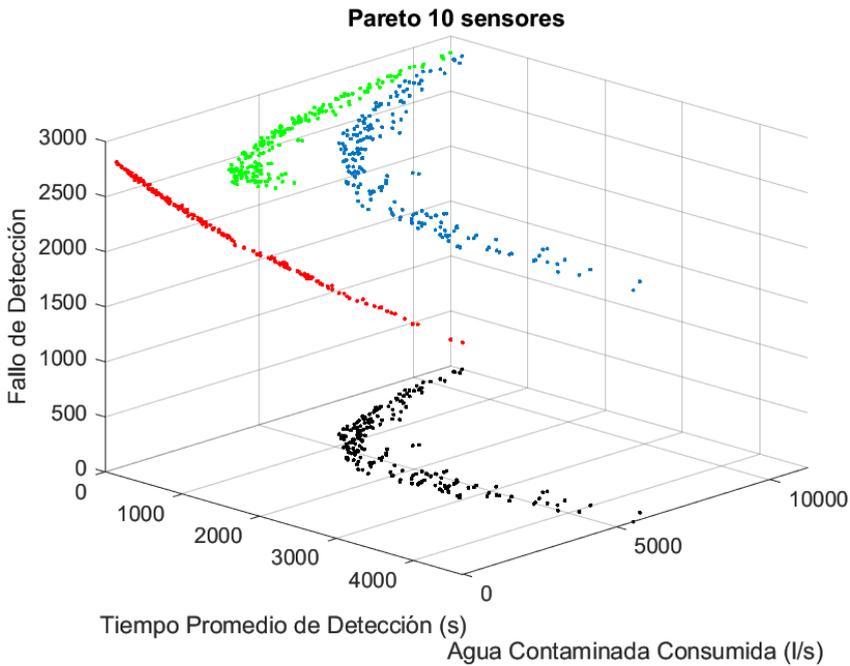


Figura 5.25. Frente de Pareto 3D para el caso de redes con 10 sensores y sus proyecciones sobre los planos coordenados (en azul los puntos tridimensionales; en verde, rojo y negro, sus proyecciones).

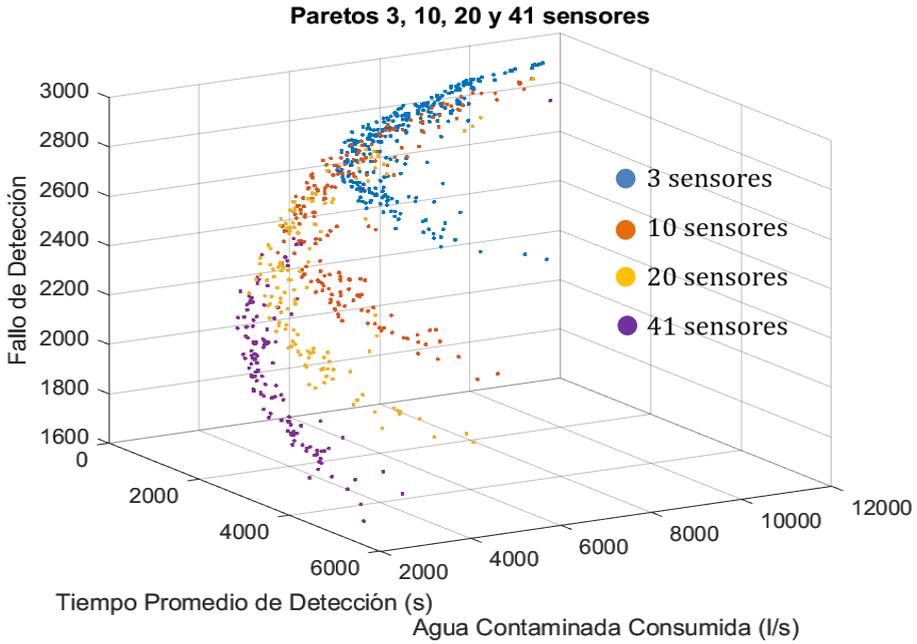


Figura 5.26. Frentes de Pareto 3D para los números de sensores considerados.

Y, para concluir, en la Figura 5.27 se presenta el frente de Pareto 3D para 20 sensores donde se han destacado las soluciones aportadas por la teoría de grafos. Una vez más, siguen viéndose cerca del frente de Pareto obtenido con la optimización multiobjetivo y, en la vista presentada, es perfectamente visible cómo las soluciones obtenidas con solo la teoría de grafos (sin consideraciones hidráulicas) quedan dominadas por las soluciones del frente de Pareto.

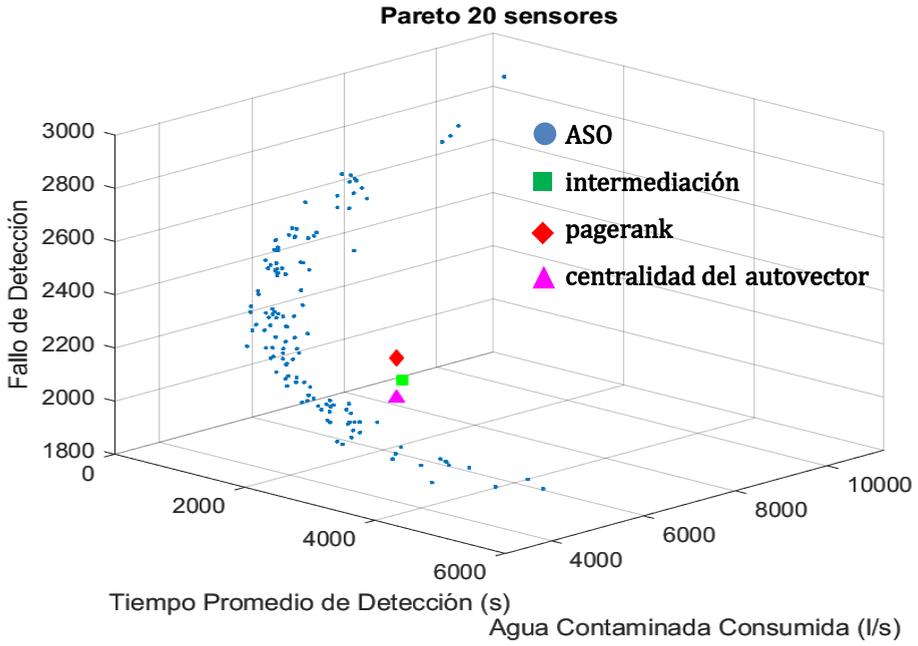


Figura 5.27. Frente de Pareto 3D para el caso de 20 sensores, mostrando las soluciones (dominadas) obtenidas mediante la teoría de grafos.

6 Conclusiones y trabajos futuros

6.1 El problema

El agua potable es un recurso fundamental para la salud y el bienestar de los ciudadanos y para la vitalidad económica de las ciudades. Sin embargo, la calidad del agua se deteriora fácilmente debido a múltiples causas. En esta tesis se ha puesto especial interés en la intrusión de contaminantes originada por un incidente casual o por un ataque malintencionado.

Como se ha observado en la tesis, las redes de suministro de agua potable tienen, en comparación con otras redes, características únicas, incluyendo su complejidad de escala, su susceptibilidad para ser atacadas, su dificultad de monitorización y, también, su alto riesgo al que están expuestas, ya que la salud de los ciudadanos depende fuertemente de ellas.

Debido a estas razones, monitorizar adecuadamente la calidad del agua de una RDA representa un gran desafío. Por ello, la investigación en este campo es muy activa y contempla aspectos muy diversos. Uno de tales aspectos, que es el objetivo de esta tesis, es la ubicación de sensores de monitorización de la calidad del agua en tales sistemas.

A raíz de ciertos sucesos de gravedad acaecidos y documentados en la literatura, las empresas de agua y los gobiernos han tenido cada vez mayor preocupación por la posibilidad de dañar la calidad del agua debido a incidentes de contaminación accidental o intencional. Y, en particular, la Agencia de Protección Ambiental de Estados Unidos (EPA) proporcionó recomendaciones sobre las acciones que las empresas de agua pueden tomar para minimizar los impactos potenciales, que han servido de directriz en este campo de trabajo.

6.2 Hacia la solución

Una RDA, como se ha visto, puede describirse mediante un grafo en el que los nodos representan fuentes de agua, tanques y uniones, y los arcos denotan las líneas (tuberías, bombas y válvulas). La contaminación puede ocurrir en cualquier nodo del sistema y extenderse a través de las tuberías a todo el sistema. Para proteger a la población de estos peligros, se colocan sensores de control de la calidad del agua en el sistema.

La utilización de redes de sensores de calidad del agua es, como consecuencia, un enfoque fundamental para detectar incidentes de contaminación en los sistemas de distribución de agua potable.

Se puede colocar sensores en cualquier nodo para detectar contaminación o polución. Sin embargo, debido a la estructura a gran escala de una RDA y al alto coste (compra, instalación y mantenimiento) de los sensores de monitoreo de la calidad del agua, no es posible colocar sensores en todos los nodos del sistema. Por ello, se han desarrollado diversos métodos de optimización para dotar a las empresas de distribución de agua de herramientas para elegir número de sensores (mínimo, con el objetivo de reducir la inversión), y sus mejores ubicaciones (para conseguir una detección de máxima eficacia). Resulta, pues, fundamental investigar cómo implementar sensores en las RDAs.

En esta tesis se han abordado cuestiones fundamentales en el ámbito de la protección contra eventos de contaminación intencional o accidental y problemas de la calidad del agua de una RDA. Para lograr dicha protección es básico el control y supervisión de la RDA mediante una red de sensores de calidad. Como se ha explicado a lo largo de este trabajo, hay que resolver principalmente dos preguntas: cuántos sensores se requieren y dónde colocarlos. La solución a estas cuestiones requiere una decisión sobre los criterios y objetivos que se deben considerar para lograr una solución adecuada.

En el Capítulo 3 se ha razonado que son cuatro los objetivos considerados de manera estándar en la evaluación del desempeño para diferentes algoritmos: alguna medida del tiempo de detección, la probabilidad de fallo en la detección,

el volumen consumido de agua contaminada, y el coste del despliegue de sensores. Algunos de estos objetivos son mutuamente conflictivos. Por ejemplo, el tiempo de detección entra en conflicto con la probabilidad de detección. En términos generales, el primero estimula el despliegue de sensores aguas arriba, mientras que el segundo se decanta por el despliegue de sensores aguas abajo. Y, ni qué decir tiene, que el coste es un objetivo, típicamente contrario a cualquier mejora.

Así que la ubicación de sensores se considera, de manera general, un problema de optimización, en general, multiobjetivo. Además, el modelo de optimización, a menudo, se combina con software hidráulico y de calidad del agua para estimar los riesgos de contaminación.

6.3 Sobre la implementación de la solución

Recientemente, la mayoría de las investigaciones de optimización han utilizado simulaciones de transporte de contaminantes dinámicas, fundamentalmente EPANET, para realizar la simulación del comportamiento hidráulico y de la calidad del agua dentro de la red de tuberías presurizadas.

El primer paso habitualmente consiste en generar eventos contaminantes y usar el simulador para calcular determinada matriz o matrices que evalúan el impacto en el sistema. El segundo paso utiliza un algoritmo de optimización para seleccionar un diseño óptimo de sensores del conjunto de soluciones potenciales.

Los retos en este trabajo son diversos. Fundamentalmente, el problema de la ubicación óptima de sensores en una RDA es un problema discreto de optimización combinatoria. No obstante, esta área de investigación es claramente multidisciplinar ya que involucra a ciencias ambientales, ciencias de la computación y tecnología de control. La solución para tales problemas está limitada por varios desafíos clave, entre los que hay que nombrar la complejidad computacional, la incertidumbre y la complejidad de la evaluación de la calidad de detección.

Entre las soluciones propuestas en la literatura, además del juicio del ingeniero, podemos citar técnicas basadas en reglas o medidas topológicas, técnicas de optimización computacional y técnicas de tipo teórico (basadas en el cálculo de la sensibilidad y de la entropía).

En esta tesis hemos considerado una hibridación de dos de tales técnicas:

1. Las técnicas basadas en reglas, que tratan de clasificar la ubicación potencial de los sensores según algunas métricas de tipo topológico.
2. Las técnicas de optimización, que adoptan un modelo computacional para estimar el rendimiento del diseño correspondiente a una determinada red de sensores.

Los métodos basados en reglas definen mecanismos para clasificar las posibles ubicaciones de los sensores sin utilizar, en general, un modelo hidráulico y de calidad de la RDA, por lo que suponen, en general, poco coste de cálculo y almacenamiento. Estos métodos se pueden utilizar para desplegar sensores en una red a gran escala, y el diseño se realiza sin ningún modelo de transporte hidráulico.

Varios de estos métodos están basados en el concepto básico aportado por la teoría de grafos que utiliza conceptos como estructura de comunidad y varios tipos de métricas del grafo, tales como centralidad, conectividad, modularidad, importancia de los nodos, etc., para identificar un conjunto de nodos clave para colocar sensores. Los resultados muestran que el método produce resultados muy similares en comparación con los métodos basados en la optimización.

Por un lado, los métodos basados en redes son efectivos especialmente para RDAs a gran escala, ya que no necesitan ni la simulación de la calidad hidráulica del agua, ni el cálculo de optimización, por lo que puede generar una solución casi óptima con menos consumo computacional. Sin embargo, estos métodos pueden ser básicos para ubicar sensores porque, generalmente, no tienen en cuenta las demandas nodales o la hidráulica de la red. Además, reglas de tipo general no pueden, con frecuencia, ser satisfactorias para diferentes tipos de

RDA's (por ejemplo, con estructuras básicamente arbórea, o, por el contrario, muy malladas).

Esta es una razón por la que muchos estudios abogan por un método basado en la optimización para obtener el diseño óptimo de la red de sensores.

Los métodos basados en la optimización necesitan, en general, incorporar tales técnicas de optimización con modelos hidráulicos y de calidad del agua como EPANET. Con EPANET, es posible tener un conocimiento bastante preciso de cómo fluye el agua y cómo se comportan los contaminantes en la red, antes de colocar los sensores. Esta es la razón por la que los métodos basados en optimización también se denominan de optimización por simulación.

Combinando ambas metodologías, en esta tesis hemos diseñado un método híbrido que combina la facilidad para ubicar sensores de manera óptima desde el punto de vista topológico con la eficacia de la optimización multiobjetivo para refinar y mejorar los diseños sugeridos por la metodología basada en la teoría de grafos.

6.4 La propuesta híbrida planteada en la tesis

Es difícil encontrar una solución óptima para el problema planteado ya que se trata de uno de los problemas denominados como NP-*hard* o NP-completo, pues no pueden ser resueltos (utilizando técnicas computacionales habituales) en tiempo polinomial.

Por otra parte, los algoritmos deterministas han mostrado grandes inconvenientes, al no poder manejar diversos tipos de datos (discretos, continuos y booleanos) y al mostrar, por tanto, eficiencia tan solo para problemas muy bien condicionados y de tamaño muy pequeño. Es decir, son métodos bien conocidos por tener dificultades en los casos de redes a gran escala.

Por el contrario, los algoritmos heurísticos, aunque, generalmente no pueden garantizar la solución óptima, se sabe que encuentran rápidamente una solución

casi óptima en la práctica, al tiempo que muestran una enorme facilidad para adaptarse a datos cualesquiera y a trabajar cooperativamente con simuladores hidráulicos como EPANET. Los algoritmos heurísticos tienen, en general, una buena capacidad de búsqueda global. Y, la combinación de un método basado en optimización con un método basado en reglas, específicamente la teoría de redes, es una idea alternativa para resolver el problema de la ubicación óptima de sensores de manera efectiva. El método basado en optimización puede utilizar los resultados del método basado en teoría de redes como base para la optimización posterior. Por ejemplo, en esta tesis hemos utilizado un algoritmo heurístico de optimización de dos etapas, donde la primera fase usa el método basado en redes para generar algunas de las soluciones iniciales, y la segunda fase adopta la solución como población inicial para hacer una búsqueda de exploración adicional, complementaria.

La solución propuesta en esta tesis para la solución de dicho problema ha sido la hibridación de dos enfoques: la teoría de redes y la optimización multiobjetivo con el entorno evolutivo ASO. Esta conjunción, por un lado, ha permitido realizar un preproceso de la RDA y conocer sus características desde una perspectiva de red compleja, y entender *a priori* la posible respuesta a determinados sucesos o fallos.

Con las medidas básicas y de centralidad que se obtuvieron en el caso de estudio, se ha podido demostrar que se puede obtener conocimiento de una RDA sin la previa modelación hidráulica. En este sentido, es posible evaluar el funcionamiento de una RDA cuando no se cuenta con información extensa sobre los parámetros y características hidráulicas de la red. Saber qué medidas o indicadores muestran la redundancia, robustez y conectividad de una RDA permite tener un *background* para llegar a soluciones más complejas. Los resultados obtenidos de la RDA de Coro demuestran el alcance de las técnicas de la teoría de redes, así como ya lo han demostrado una gran cantidad de trabajos. En el caso de la RDA de estudio, se ha encontrado que al igual que otras RDAs es una red dispersa y poco redundante. Por otro lado, las medidas de centralidad propuestas han sido adecuadas para encontrar los nodos-sensores, pues, a pesar de que, en general, son aplicadas en redes con una gran cantidad de conexiones

y densidad de nodos (en especial *pagerank*), es viable su utilización en redes de menor tamaño.

Una ventaja de la propuesta metodológica de este trabajo es que, con el enfoque basado en la teoría de redes, se considera la dimensión espacial de la estructura de la RDA. Para la selección y ubicación de los sensores, se toman en cuenta criterios básicos para la división de la RDA, más las características de conectividad. Esto permite tomar decisiones sobre qué usuarios deben ser priorizados o protegidos, qué área puede representar mayor vulnerabilidad, ya sea por una gran cantidad de conexiones o densidad de nodos o, porque es un área de la RDA alejada, que requiere protección. Lo anterior puede estar relacionado con políticas o acciones orientadas a la equidad en el suministro, a la regulación de la demanda, o costes del servicio.

La solución multiobjetivo final se ha basado en una compensación entre los objetivos involucrados y la tolerancia que se pueda tener a que no sean cumplidos. Se puede lograr una mejora en todos los objetivos analizados agregando nuevos sensores, pero esto por supuesto, tendría la consecuencia de incrementar los costes. lo que puede ser una limitante para la implementación de las soluciones.

Para la red de estudio de esta tesis, una red de tamaño mediano se han analizado fundamentalmente 4 escenarios con distintos números de sensores. Los escenarios más básicos se han llevado a cabo con el fin de comprobar cómo al aumentar el número de sensores, los demás objetivos obtienen mejores comportamientos. Los escenarios para 20 y 41 nodos han utilizado las soluciones propuestas por la teoría de grafos como parte de la población inicial del algoritmo evolutivo. Y en la optimización, dicho algoritmo ha conseguido refinar las soluciones obteniendo aproximaciones de sus frentes de Pareto que, claramente mejoran (dominan) a las soluciones de la teoría de grafos. Sin embargo, cabe hacer notar que tales soluciones, a pesar de no contar con la componente hidráulica del transporte de contaminante, consiguen comportamientos muy cercanos a algunas de las soluciones del frente de Pareto que las dominan. De hecho, nos hemos preguntado por la calidad de las soluciones con las medidas de centralidad para una cantidad mayor de nodos

que las soluciones obtenidas con la optimización. En el caso concreto estudiado, se puede ver que las soluciones espectrales para 41 nodos son mejores que las soluciones multiobjetivo para 20 nodos, como muestra la figura 6.1, que presenta la frontera de Pareto obtenida para la solución de 20 sensores, junto con las soluciones espectrales para 41 nodos. Se puede apreciar, fácilmente que estas últimas no son dominadas por aquellas.

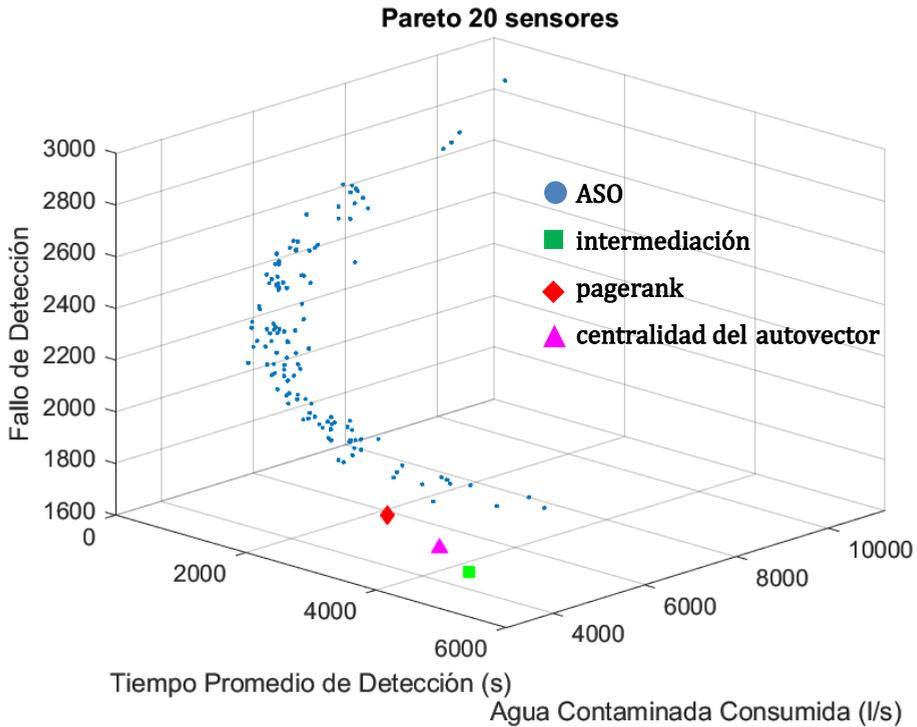


Figura 6.1 Las soluciones de centralidad para 41 nodos mejoran a las soluciones multiobjetivo para 20 nodos.

La información adicional sobre la utilidad también puede influir en la ubicación final de los sensores adoptados. Este es el caso, por ejemplo, del tiempo medio de reacción de la empresa gestora del suministro cuando se detecta un evento. Si el tiempo de reacción es relativamente grande entonces puede ser conveniente utilizar soluciones con los sensores ubicados un poco más aguas arriba si se desea evitar que al menos una parte de la población reciba agua contaminada. Hay que tener en cuenta que en este caso también podría haber un mayor

número de fallos de detección. Obtener alguna mejora en ambos lados (tiempo de detección mínimo y fallos de detección reducidos) implica agregar más sensores a la solución.

Esta tesis muestra una descripción general simplificada sobre cómo enfrentar en la práctica el problema de la ubicación de sensores de calidad del agua para proteger a los SDAPs. Se recomienda no basar las decisiones solo en la experiencia práctica, sino ejecutar cálculos de modelos hidráulicos. Tanto los modelos computacionales como la experiencia, deben combinarse para lograr los mejores resultados posibles.

6.5 Direcciones futuras

Como hemos observado en la tesis, ha habido muchos estudios dedicados a métodos basados en optimización y basados en la teoría de redes, para un diseño óptimo o casi óptimo de sensores durante las últimas décadas. Sin embargo, la investigación sobre la colocación de sensores en SDAPs es aún un problema incipiente. Se necesita mucho más esfuerzo para comprender completamente el problema a fin de proponer métodos más avanzados. Entre la inmensidad de futuras direcciones posibles de investigación sobre el tema podemos citar:

- Dentro de las posibilidades que abre la división de una RDA en comunidades, se tiene que, dado un número de áreas, estas se pueden reagrupar por medio de técnicas sencillas de *clustering* para obtener áreas de mayor tamaño y permitir el manejo del número de sensores a colocar. Esto permite considerar la distribución espacial de la red.
- En relación con lo anterior, el problema inverso sería aumentar o dividir cada comunidad, como una red, y hacer un proceso iterativo por cada área hasta lograr el número de comunidades por sensores.
- Hacer un ranking combinando las soluciones con algoritmos multicriterio, añadir criterios de factibilidad, es decir, por motivos de accesibilidad se pueden elegir soluciones cercanas factibles que permitan la instalación alternativa al punto exacto de monitoreo.

- La mejora en la elaboración del modelo matemático para el problema en una RDA que permita mejorar la descripción de la dinámica del transporte de contaminante bajo condiciones topológicas diversas de la red y bajo diferentes dinámicas de las demandas.
- El análisis estocástico del sistema que permitiría un tratamiento más acertado de la incertidumbre, la consideración de los denominados sensores imperfectos que pueden proporcionar falsos positivos y negativos en su información, y, lo que son elementos clásicos en las RDAs, la consideración del envejecimiento de las tuberías en el problema, las dinámicas de las demandas que pueden alterar la edad del agua en las tuberías, etc., aspectos que pueden implicar el tener que adaptar y/o renovar las redes de sensores ya desplegadas.
- La computación paralela y/o distribuida. Dada la naturaleza NP-*hard* del problema, es obvio que se hacen imprescindibles técnicas más eficientes y con menor coste computacional que las técnicas actualmente utilizadas. En este sentido, la eficiencia en la solución del problema se vería enormemente mejorada con el diseño de algoritmos distribuidos basados en el denominado *cloud-computing*, que proporcionarían métodos más flexibles, eficaces y tolerantes a fallos. Por supuesto, la computación cuántica podría significar, como en otros muchos campos, el salto de calidad imprescindible para la solución de un problema NP-*hard*, como el planteado.
- La incorporación de las ventajas de las redes de sensores móviles. Aunque de utilización realmente muy escasa en la actualidad para la monitorización de la calidad del agua en RDAs, ya que los sensores móviles están en una etapa muy incipiente, no cabe duda de que cualquier avance en esta línea sería de un enorme interés. Y, la acción conjunta de sensores fijos y móviles deberá mejorar aún más la capacidad de monitorización de tal red mixta de sensores.

- Las mejoras en el diseño, implementación y prueba del sistema. En esta línea queremos indicar el desarrollo de sistemas potentes de alerta temprana para la solución a los problemas de contaminación. Estas mejoras, además de una red óptima, implicarán la identificación precisa de la fuente de contaminación y la definición de una respuesta adecuada en términos de protocolos de información para los consumidores, y acciones a desarrollar por el propio abastecimiento para llevar a cabo una evacuación óptima del contaminante de la red, así como para una puesta en funcionamiento del sistema lo más rápida posible. Estas componentes deben ser investigadas de la manera más exhaustiva posible. Por ejemplo, la evacuación de agua contaminada radioactivamente debe ser considerada con sumo detalle, ya que un simple vertido sobre el terreno podría, a su vez, tener consecuencias desastrosas.

Referencias bibliográficas

- Adedoja, O. S., Hamam, Y., Khalaf, B., and Sadiku, R. (2018). "A state-of-the-art review of an optimal sensor placement for contaminant warning system in a water distribution network." *Urban Water Journal*, Taylor & Francis, 15(10), 985–1000.
- Adger, N., Aggarwal, P., Agrawala, S., Alcamo, J., Allali, A., Cruz, R. V., Alcaraz, E. D. A., Easterling, W., Field, C., Fischlin, A., Fitzharris, B. B., García, C. G., Hanson, C., Harasawa, H., Huq, S., Jones, R., Bogataj, L. K., Karoly, D., Klein, R., Mortsch, L., Niang-diop, I., Nicholls, R., Nováky, B., Nurse, L., Nyong, A., Oppenheimer, M., Palutikof, J., Parry, M., Patwardhan, A., Lankao, R., Rosenzweig, C., Schneider, S., Semenov, S., Smith, J., and Stone, J. (2007). "Intergovernmental Panel on Climate Change 2007 : Impacts , Adaptation and Vulnerability Working Group II Contribution to the Intergovernmental Panel on Climate Change Fourth Assessment Report Summary for Policymakers." (2300).
- Akgun, I., Kandakoglu, A., and Fahri, A. (2010). "Expert Systems with Applications Fuzzy integrated vulnerability assessment model for critical facilities in combating the terrorism." *Expert Systems With Applications*, Elsevier Ltd, 37(5), 3561–3573.
- Al-Zahrani, M. A., and Moied, K. (2001). "Locating Optimum Water Quality Monitoring Stations in Water Distribution System." *Bridging the Gap*, American Society of Civil Engineers (ASCE), 1–9.
- Al-Zahrani, M. A., and Moied, K. (2003). "Optimizing Water Quality Monitoring Stations Using Genetic Algorithms." *The Arabian Journal for Science and Engineering*, 28(1B), 57–75.
- Alzamora, F., and Ayala, H. B. (2006). "Optimal sensor location for detecting contamination events in water distribution systems using topological algorithms." *8th Annual Water Distribution System Analysis Symposium (WDSA)*, American Society of Civil Engineers (ASCE), Cincinnati, Ohio, United States.
- Antunes, C. H., and Dolores, M. (2016). "Sensor Location in Water Distribution Networks to Detect Contamination Events - A Multiobjective Approach Based on NSGA-II." *2016 IEEE Congress on Evolutionary Computation, CEC 2016*, IEEE, 1093–1099.
- Apostolakis, G. E., and Lemon, D. M. (2005). "A screening methodology for the identification and ranking of infrastructure vulnerabilities due to terrorism." *Risk Analysis*, 25(2), 361–376.
- Arad, J., Housh, M., Perelman, L., and Ostfeld, A. (2013). "A dynamic thresholds scheme for contaminant event detection in water distribution systems." *Water Research*, 47(5), 1899–1908.
- Aral, M. M., Guan, J., and Maslia, M. L. (2010). "Optimal Design of Sensor

- Placement in Water Distribution Networks." *Journal of Water Resources Planning and Management*, 136(1), 5–18.
- Bahadur, R., Samuels, W. B., Grayman, W. M., Amstutz, D., and Pickus, J. (2003). "PipelineNet: A Model for Monitoring Introduced Contaminants in a Distribution System." *World Water and Environmental Resources Congress 2003*, American Society of Civil Engineers (ASCE), 1–6.
- Baker, G. H. (2005). "A Vulnerability Assessment Methodology for Critical Infrastructure Facilities." *DHS Symposium: R&D Partnerships in Homeland Security*, Boston, Massachusetts, 1–13.
- Berry, J., Boman, E., Riesen, L. A., Hart, W. E., Phillips, C. A., and Watson, J.-P. (2012). "User's Manual TEVA-SPOT Toolkit.Version 2.5.2." United States Environmental Protection Agency (EPA).
- Berry, J., Carr, R. D., Hart, W. E., Leung, V. J., Phillips, C. A., and Watson, J.-P. (2008a). "On the Placement of Imperfect Sensors in Municipal Water Networks." *Water Distribution Systems Analysis Symposium 2006*, American Society of Civil Engineers, Reston, VA, Cincinnati, Ohio, United States, 1–13.
- Berry, J., Carr, R. D., Hart, W. E., Leung, V. J., Phillips, C. A., and Watson, J.-P. (2009). "Designing Contamination Warning Systems for Municipal Water Networks Using Imperfect Sensors." *Journal of Water Resources Planning and Management*, 135(4), 253–263.
- Berry, J., Hart, W. E., Phillips, C. A., and Uber, J. (2004). "A General Integer-Programming-Based Framework for Sensor Placement in Municipal Water Networks." *Critical Transitions in Water and Environmental Resources Management*, American Society of Civil Engineers, Reston, VA, 1–10.
- Berry, J., Hart, W. E., Phillips, C. A., Uber, J. G., and Watson, J.-P. (2006). "Sensor Placement in Municipal Water Networks with Temporal Integer Programming Models." *Journal of Water Resources Planning and Management*, 132(4), 218–224.
- Berry, J. W., Boman, E., Phillips, C. A., and Riesen, L. A. (2008b). "Low-memory Lagrangian Relaxation Methods for Sensor Placement in Municipal Water Networks." *World Environmental and Water Resources Congress 2008*, 10.
- Berry, J. W., Carr, R. D., Hart, W. E., and Phillips, C. A. (2007). "Scalable Water Network Sensor Placement via Aggregation." *World Environmental and Water Resources Congress 2007*, American Society of Civil Engineers, Reston, VA, 1–11.
- Berry, J. W., Fleischer, L., Hart, W. E., Phillips, C. A., and Watson, J.-P. (2005a). "Sensor Placement in Municipal Water Networks." *Journal of Water Resources Planning and Management*, 131(3), 237–243.
- Berry, J. W., Hart, W. E., Phillips, C. A., Uber, J. G., and Walski, T. M. (2005b). "Water Quality Sensor Placement in Water Networks with Budget Constraints." *Impacts of Global Climate Change*, American Society of Civil Engineers, Reston, VA, 1–11.
- Berry, J. W., Hart, W. E., Phillips, C. A., Uber, J. G., and Watson, J.-P. (2005c).

- “Validation and Assessment of Integer Programming Sensor Placement Models.” *Impacts of Global Climate Change*, American Society of Civil Engineers, Reston, VA, 1–11.
- Berry, J. W., Hart, W. E., Phillips, C. A., and Watson, J.-P. (2005d). “Scalability of Integer Programming Computations for Sensor Placement in Water Networks.” *Impacts of Global Climate Change*, American Society of Civil Engineers, Reston, VA, 1–13.
- Blondel, V. D., Guillaume, J. L., Lambiotte, R., and Lefebvre, E. (2008). “Fast unfolding of communities in large networks.” *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), 1–12.
- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D. U. (2006). “Complex networks: Structure and dynamics.” *Physics Reports*, 424(4–5), 175–308.
- Bonacich, P. (1972). “Technique for Analyzing Overlapping Memberships.” *Sociological Methodology*, [American Sociological Association, Wiley, Sage Publications, Inc.], 4, 176–185.
- Bouchon, S. (2006). “The Vulnerability of interdependent Critical Infrastructures Systems : Epistemological and Conceptual State-.”
- Boudou, R. (2006). “Évaluation de la vulnérabilité d’une ville face à son réseau d’eau potable: un approche par conséquences.” École Polytechnique de Montréal.
- Brentan, B., Campbell, E., Goulart, T., Manzi, D., Meirelles, G., Herrera, M., Izquierdo, J., and Luvizotto, E. (2018). “Social Network Community Detection and Hybrid Optimization for Dividing Water Supply into District Metered Areas.” *Journal of Water Resources Planning and Management*, 144(5), 04018020.
- Brentan, B., Carpitella, S., Barros, D., Meirelles, G., Certa, A., and Izquierdo, J. (2021a). “Water Quality Sensor Placement: A Multi-Objective and Multi-Criteria Approach.” *Water Resources Management*, 35(1), 225–241.
- Brentan, B. M., Carpitella, S., Izquierdo, J., Luvizotto, E., and Meirelles, G. (2021b). “District metered area design through multicriteria and multiobjective optimization.” *Mathematical Methods in the Applied Sciences*, (April 2020), mma.7090.
- Brin, S., and Page, L. (1998). “The anatomy of a large-scale hypertextual Web search engine.” *Computer Networks and ISDN Systems*, 30(1–7), 107–117.
- Bryan, K., and Leise, T. (2006). “The \$25,000,000,000 Eigenvector: The linear Algebra Behind Google.” *SIAM Review*, 48, 569–581.
- Campbell, E., Ayala-Cabrera, D., Izquierdo, J., Pérez-García, R., and Tavera, M. (2014). “Water supply network sectorization based on social networks community detection algorithms.” *Procedia Engineering*, 89(May 2016), 1208–1215.
- Campbell González, E. (2017). “Sectorización de Redes de Abastecimiento de Agua Potable basada en detección de comunidades en redes sociales y optimización heurística.” Universitat Politècnica de València.

- Candelieri, A., Conti, D., and Archetti, F. (2014). "A Graph based Analysis of Leak Localization in Urban Water Networks." *Procedia Engineering*, 70, 228–237.
- Cardoso, S. M., Barros, D. B., Oliveira, E. C. M., Brentan, B. M., Junior, E. L., and Ribeiro, L. C. L. J. (2018). "Multiobjective Optimization and Rule Extraction for Optimum Location of Quality Sensors in Water Distribution Network." *1st International WDSA / CCWI 2018 Joint Conference*, Kingston, Ontario, Canada, 8.
- Carpitella, S., Brentan, B., Montalvo, I., Izquierdo, J., and Certa, A. (2019). "Multi-criteria analysis applied to multi-objective optimal pump scheduling in water systems." *Water Supply*, 19(8), 2338–2346.
- Carr, R. D., Greenberg, H. J., Hart, W. E., Konjevod, G., Lauer, E., Lin, H., Morrison, T., and Phillips, C. A. (2006). "Robust optimization of contaminant sensor placement for community water systems." *Mathematical Programming. Series B*, 107(1–2), 337–356.
- Carr, R. D., Greenberg, H. J., Hart, W. E., and Phillips, C. A. (2004). "Addressing Modeling Uncertainties in Sensor Placement for Community Water Systems." *Critical Transitions in Water and Environmental Resources Management*, American Society of Civil Engineers (ASCE), 1–10.
- CDC Centers for Disease Control and Prevention CDC. (2014). "Emergency Preparedness and Response. Bioterrorism Agents Diseases."
- Chastain, J. R. J. (2004). "A Heuristic Methodology for Locating Monitoring Stations to Detect contamination Events in Potable Water Distribution Systems." University of South Florida.
- Chastain, J. R. J. (2006). "Methodology for locating monitoring stations to detect contamination in potable water distribution systems." *Journal of infrastructure systems*, 12(4), 252–259.
- Cheifetz, N., and Mandel, P. (2018). "A comparative study of multi-objective methods for sensor placement optimization applied on realistic WDN MULTI-OBJECTIVE SENSOR PLACEMENT." *1st International WDSA / CCWI 2018 Joint Conference*, Kingston, Ontario, Canada.
- Chung, F., and Zhao, W. (2008). "PageRank and Random Walks on Graphs." *In Fete of Combinatorics and Computer Science Conference*, Keszthely, Hungary., 131–132.
- Ciaponi, C., Creaco, E., Di Nardo, A., Di Natale, M., Giudicianni, C., Musmarra, D., and Santonastaso, G. F. (2019). "Reducing impacts of contamination in water distribution networks: A combined strategy based on network partitioning and installation of water quality sensors." *Water (Switzerland)*, 11(6), 1–16.
- Clark, R. M., and Deininger, R. A. (2000). "Protecting the Nation's Critical Infrastructure: The Vulnerability of U.S. Water Supply Systems." *Journal of Contingencies and Crisis Management*, 8(2), 73–80.
- Clark, R. M., Nabil, A., Atluri, V., Halem, M., Vowinkel, E., Tao, P. C., Cummings, L., and Ibrahim, E. A. (2004). "Developing an Early Warning System for

- Drinking Water Security and Safety." *Water Supply Systems Security*, L. W. Mays, ed., McGraw-Hill, New York, London.
- Coburn, A., Spence, R., and Pomonis, A. (1994). "Vulnerability and Risk Assessment." UNDP/UNDRO Disaster Management Training Programme. Cambridge Architectural Research Limited, Cambridge, United Kingdom.
- Cozzolino, L., Mucherino, C., Pianese, D., and Pirozzi, F. (2006). "Positioning, Within Water Distribution Networks, of Monitoring Stations Aiming at an Early Detection of Intentional Contamination." *Civil Engineering and Environmental Systems*, 23(3), 161–174.
- Di Cristo, C., and Leopardi, A. (2008). "Pollution Source Identification of Accidental Contamination in Water Distribution Networks." *Journal of Water Resources Planning and Management*, 134(2), 197–202.
- Crucitti, P., Latora, V., and Porta, S. (2006). "Centrality in Networks of Urban Streets." *Chaos*, 16(1), 015113–1--015113–9.
- Csárdi, G., Nepusz, T., and Airoldi, E. M. (2016). *Statistical Network Analysis with igraph (R)*. Springer.
- Davis, M. J., and Janke, R. (2011). "Patterns in Potential Impacts Associated with Contamination Events in Water Distribution Systems." *Journal of Water Resources Planning and Management*, 137(1), 1–9.
- Davis, M. J., and Janke, R. . (2008). "Importance of exposure model in estimating impacts when a water distribution system is contaminated." *Journal of Water Resources Planning and Management*, 134(5), 449–456.
- Diestel, R. (2005). *Graph theory. Graph Theory*, Springer-Verlag Heidelberg, New York 1997, 2000, 2005 This, Berlin, Heidelberg.
- Dorigo, M., Maniezzo, V., and Colorni, A. (1996). "Ant system: Optimization by a colony of cooperating agents." *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 26(1), 29–41.
- Dorini, G., Jonkergouw, P., Kapelan, Z., di Pierro, F., Khu, S.-T., and Savic, D. (2006). "An Efficient Algorithm for Sensor Placement in Water Distribution Systems." *8th Annual Water Distribution System Analysis Symposium*, American Society of Civil Engineers (ASCE), Cincinnati, Ohio, United States, 1–13.
- Duan, Q. Y., Gupta, V. K., and Sorooshian, S. (1993). "Shuffled complex evolution approach for effective and efficient global minimization." *Journal of Optimization Theory and Applications*, 76(3), 501–521.
- Einarsson, S., and Rausand, M. (1998). "An Approach to Vulnerability Analysis of Complex Industrial Systems." *Risk Analysis*, 18(5), 535–546.
- Eliades, D., and Polycarpou, M. (2006). "Iterative Deepening of Pareto Solutions in Water Sensor Networks." *8th Annual Water Distribution System Analysis Symposium*, American Society of Civil Engineers, Cincinnati, Ohio, United States, 1–19.
- Estrategia Internacional para la Reducción de Desastres (EIDR). (2009). "2009 UNISDR Terminología sobre Reducción del Riesgo de Desastres." *Estrategia Internacional para la Reducción de Desastres de las Naciones*

- Unidas (UNISDR)*, Estrategia Internacional para la Reducción de Desastres de las Naciones Unidas (UNISDR), Ginebra, Suiza.
- Ezell, B. C. (2007). "Infrastructure Vulnerability Assessment Model (I-VAM)." *Risk analysis : an official publication of the Society for Risk Analysis*, 27(3), 571–83.
- Ferrari, G., Savic, D., and Becciu, G. (2014). "Graph-Theoretic Approach and Sound Engineering Principles for Design of District Metered Areas." *Journal of Water Resources Planning and Management*, American Society of Civil Engineers, 140(12), 4014036.
- Ferreira, J. C., Fonseca, C. M., Denysiuk, R., and Gaspar-Cunha, A. (2017). "Methodology to select solutions for multiobjective optimization problems: Weighted stress function method." *Journal of Multi-Criteria Decision Analysis*, 24(3–4), 103–120.
- Francés-Chust, J., Brentan, B. M., Carpitella, S., Izquierdo, J., and Montalvo, I. (2020). "Optimal Placement of Pressure Sensors Using Fuzzy DEMATEL-Based Sensor Influence." *Water (Switzerland)*, 12(493), 1–18.
- Freeman, L. C. (1977). "A Set of Measures of Centrality Based on Betweenness." *Sociometry*, 40(1), 35.
- Ghimire, S. R., and Barkdoll, B. D. (2006a). "Heuristic Method for the Battle of the Water Network Sensors: Demand Based Approach." *8th Annual Water Distribution System Analysis Symposium*, American Society of Civil Engineers (ASCE), Cincinnati, Ohio, United States, 1–10.
- Ghimire, S. R., and Barkdoll, B. D. (2006b). "A Heuristic Method for Water Quality Sensor Location in a Municipal Water Distribution System: Mass-Released Based Approach." *8th Annual Water Distribution System Analysis Symposium*, American Society of Civil Engineers (ASCE), Cincinnati, Ohio, United States, 1–11.
- Girvan, M., and Newman, M. E. J. (2002). "Community structure in social and biological networks." *Proceedings of the National Academy of Sciences*, 99(12), 7821–7826.
- Giudicianni, C., Herrera, M., Di Nardo, A., Greco, R., Creaco, E., and Scala, A. (2020). "Topological Placement of Quality Sensors in Water-Distribution Networks without the Recourse to Hydraulic Modeling." *Journal of Water Resources Planning and Management*, 146(6), 6.
- Gleick, P. H., and Heberger, M. (2011). "Water Conflict Chronology." *The World's Water*, P. H. Gleick, ed., Island Press/Center for Resource Economics, 175–214.
- Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Longman Publishing Co., Inc., USA.
- Gordon, K., and Dion, M. (2008). *Protection of "critical infrastructure" and the role of investment policies relating to national security. OECD (Organisation for Economic Co-operation and Development)*.
- Grayman, W. M., Deininger, R. A., Males, R. M., and Gullick, R. W. (2004). "Source Water Early Warning Systems." *Water Supply Systems Security*, L. W. Mays,

- ed., McGraw-Hill, New York, London, 11.1-11.30.
- Grayman, W. M., Ostfeld, A., and Salomons, E. (2006). "Locating Monitors in Water Distribution Systems: Red Team-Blue Team Exercise." *Journal of Water Resources Planning and Management*, 132(August), 300-304.
- Grigg, N. S. (2003). "Water Utility Security: Multiple Hazards and Multiple Barriers." *Journal of Infrastructure Systems*, 9(2), 81-88.
- Grubestic, T. H., Matisziw, T. C., Murray, a. T., and Snediker, D. (2008). "Comparative Approaches for Assessing Network Vulnerability." *International Regional Science Review*, 31(1), 88-112.
- Guan, J., Aral, M. M., Maslia, M. L., and Grayman, W. M. (2006). "Identification of Contaminant Sources in Water Distribution Systems Using Simulation - Optimization Method : Case Study." *Journal of Water Resources Planning and Management*, 132(4), 252-262.
- Gueli, R. (2006). "Predator-Prey Model for Discrete Sensor Placement." *8th Annual Water Distribution System Analysis Symposium*, American Society of Civil Engineers (ASCE), Cincinnati, Ohio, United States, 1-9.
- Gutiérrez-Pérez, J. A., Herrera, M., Izquierdo, J., and Pérez-García, R. (2012). "An approach based on ranking elements to form supply clusters in water supply networks as a support to vulnerability assessment." *International Environmental Modelling and Software Society (iEMSS) 2012 International Congress on Environmental Modelling and Software Managing Resources of a Limited Planet, Sixth Biennial Meeting*, D. B. R. Seppelt, A.A. Voinov, S. Lange, ed., Leipzig, Germany.
- Gutiérrez-Pérez, J. A., Herrera, M., Pérez-García, R., and Ramos-Martínez, E. (2013). "Application of Graph-Spectral Methods in the Vulnerability Assessment of Water Supply Networks." *Mathematical and Computer Modelling*, Elsevier Ltd, 57, 1853-1859.
- Gutiérrez-Pérez, J., Montalvo, I., Francés-Chust, J., Herrera, M., and Izquierdo, J. (2017). "Water Quality Sensor Placement with a Multiobjective Approach." *CMN 2017 Congress on Numerical Methods in Engineering*, I. Arias, J. Blanco, S. Clain, P. Flores, P. Lourenço, J. J. Ródenas, and M. Tur, eds., Valencia, España.
- Haines, Y. Y. (2006). "On the Definition of Vulnerabilities in Measuring Risks to Infrastructures." 26(2), 293-296.
- Haines, Y. Y., Matalas, N. C., Lambert, J. H., Bronwyn, J. A., and Fellows, J. F. R. (1998). "Reducing vulnerability of water supply systems to attack." *Journal of Infrastructure Systems*, 4(4), 164-177.
- Hart, W. E., Berry, J. W., Boman, E. G., Murray, R., Phillips, C. A., Riesen, L. A., and Watson, J.-P. (2008). "The TEVA-SPOT Toolkit for Drinking Water Contaminant Warning System Design." *World Environmental and Water Resources Congress 2008*, American Society of Civil Engineers, Reston, VA, 1-12.
- Hart, W. E., Mckenna, S. A., Klise, K., Cruz, V., and Wilson, M. (2007). "CANARY: A Water Quality Event Detection Algorithm Development Tool." *World*

- Environmental and Water Resources Congress 2007*, 1–9.
- Hart, W. E., and Murray, R. (2010). "Review of Sensor Placement Strategies for Contamination Warning Systems in Drinking Water Distribution Systems." *Journal of Water Resources Planning and Management*, 136(6), 611–619.
- Hasan, J., States, S., and Deininger, R. (2004). "Safeguarding The Security Of Public Water Supplies Using Early Warning Systems: A Brief Review." *Journal of Contemporary Water Research & Education*, (129), 27–33.
- Herrera Fernández, A. M. (2011). "Improving water network management by efficient division into supply clusters." Universitat Politècnica de València.
- Herrera, M., Abraham, E., and Stoianov, I. (2015). "Graph-theoretic surrogate measures for analysing the resilience of water distribution networks." *Procedia Engineering*, Elsevier B.V., 119(1), 1241–1248.
- Herrera, M., Abraham, E., and 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., Canu, S., Karatzoglou, A., Pérez-García, R., and Izquierdo, J. (2010). "An approach to water supply clusters by semi-supervised learning." *International Environmental Modelling and Software Society (iEMSs). Modelling for Environment's Sake, Fifth Biennial Meeting*, T. F. David A. Swayne, Wanhong Yang, A. A. Voinov, A. Rizzoli, ed., Ottawa, Canadá, pp-1-8.
- Herrera, M., Gutiérrez-Pérez, J. A., Izquierdo, J., and Pérez-García, R. (2011a). "Ajustes en el Modelo Page-Rank de Google para el Estudio de la Importancia Relativa de los Nodos de la Red de Abastecimiento." *X Seminario Iberoamericano de Planificación, Proyecto y Operación de Sistemas de Abastecimiento de Agua (SEREA)*, Morelia, México.
- Herrera, M., Gutiérrez-Pérez, J. A., Izquierdo, J., and Pérez-García, R. (2012). "A multi-level spectral clustering process to ascertain sensor location for mitigating effects of a potential contamination in a water supply network." *International Environmental Modelling and Software Society (iEMSs). Managing Resources of a Limited Planet, Sixth Biennial Meeting*, Leipzig, Germany.
- Herrera, M., Gutiérrez-Pérez, J., Izquierdo, J., and Pérez-García, R. (2011b). "Ubicación de Sensores en un Sistema de Abastecimiento de Agua Basada en un Criterio de Zonas Hidráulicas." *VII Seminario Euro-Latioamericano de Sistemas de Ingeniería (SELASI)*, Universitat Politècnica de València, Fluing-IMM, Valenc, 85–90.
- Herrera, M., Pérez-Hernández, M., Parlikad, A. K., and Izquierdo, J. (2020). "Multi-agent systems and complex networks: Review and applications in systems engineering." *Processes*, 8(3).
- Hrudey, S. E., and Krewski, D. (1995). "Is there a safe level of exposure to a carcinogen?" *Environmental Science and Technology*, 29(8), 370A-375A.
- Huang, J. J., McBean, E. A., and James, W. (2006). "Multi-objective Optimization for Monitoring Sensor Placement in Water Distribution Systems." *8th*

- Annual Water Distribution System Analysis Symposium*, American Society of Civil Engineers (ASCE), Cincinnati, Ohio, United States, 1–14.
- Iglesias Rey, P. L., García Mares, F. J., López Jiménez, P. A., and Pérez-García, R. (2003). “Análisis y modelación de redes de distribución de agua. Introducción a las redes de abastecimiento.” Grupo Multidisciplinar de Modelación de Fluidos (GMMF), Universidad Politécnica de Valencia.
- Islam, N., Farahat, A., Al-Zahrani, M. A. M., Rodriguez, M. J., and Sadiq, R. (2015). “Contaminant intrusion in water distribution networks: review and proposal of an integrated model for decision making.” *Environmental Reviews*, NRC Research Press, 23(3), 337–352.
- Isovitich, S. L., and Vanbriesen, J. M. (2007). “Spatial Analysis of Optimized Sensor Locations using GIS.” *World Environmental and Water Resources Congress 2007*, American Society of Civil Engineers (ASCE), 1–9.
- Isovitich, S. L., and VanBriesen, J. M. (2008). “Sensor Placement and Optimization Criteria Dependencies in a Water Distribution System.” *Journal of Water Resources Planning and Management*, 134(2), 186–196.
- Izquierdo, J., Campbell, E., Montalvo, I., and Pérez-García, R. (2016a). “Injecting problem-dependent knowledge to improve evolutionary optimization search ability.” *Journal of Computational and Applied Mathematics*, Elsevier B.V., 291, 281–292.
- Izquierdo, J., Herrera, M., Montalvo, I., and Pérez-García, R. (2009). “Division of Water Supply Systems into District Metered Areas Using a Multi-agent Based Approach.” *Software en Data Technologies. 4th International Conference, ICSoft 2009*, J. Cordeiro, A. Ranchordas, and B. Shishkov, eds., Springer Berlin Heidelberg, Sofia, Bulgaria, 167–180.
- Izquierdo, J., Montalvo, I., Campbell, E., and Pérez-García, R. (2016b). “A hybrid, auto-adaptive and rule-based multi-agent approach using evolutionary algorithms for improved searching.” *Engineering Optimization*, 48(8), 1365–1377.
- Izquierdo, J., Montalvo, I., Pérez, R., and 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.
- Jiang, S., and Yang, S. (2017). “Evolutionary Dynamic Multiobjective Optimization: Benchmarks and Algorithm Comparisons.” *IEEE Transactions on Cybernetics*, IEEE, 47(1), 198–211.
- Kennedy, J., and Eberhart, R. (1995). “Particle swarm optimization.” *Proceedings of ICNN'95 - International Conference on Neural Networks*, IEEE, London, UK, 1942–1948.
- Kessler, A., Ostfeld, A., and Sinai, G. (1998). “Detecting Accidental Contaminations in Municipal Water Networks.” *Journal of Water Resources Planning and Management*, 124(4), 192–198.
- Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983). “Optimization by Simulated Annealing.” *Science*, American Association for the Advancement of Science, 220(4598), 671–680.

- Krause, A., and Guestrin, C. (2009). "Robust Sensor Placement Detecting Adversarial Contaminations in Water Distribution Systems." *World Environmental and Water Resources Congress 2009*, American Society of Civil Engineers (ASCE), 1–10.
- Krause, A., Leskovec, J., Guestrin, C., VanBriesen, J., and Faloutsos, C. (2008). "Efficient Sensor Placement Optimization for Securing Large Water Distribution Networks." *Journal of Water Resources Planning and Management*, 134(6), 516–526.
- Krause, A., Leskovec, J., Isovitsch, S., Xu, J., Guestrin, C., Vanbriesen, J., Small, M., and Fischbeck, P. (2006). "Optimizing Sensor Placements in Water Distribution Systems Using Submodular Function Maximization." *8th Annual Water Distribution System Analysis Symposium*, American Society of Civil Engineers (ASCE), Cincinnati, Ohio, United States, 1–17.
- Kröger, W., and Zio, E. (2011a). "Introduction and Definition of Key Terms." *Vulnerable systems*, Springer-Verlag London Limited, London, 1–7.
- Kröger, W., and Zio, E. (2011b). "Challenges to Methods for the Vulnerability Analysis of Critical Infrastructures." *Vulnerable systems*, Springer-Verlag London Limited, London, 33–39.
- Kröger, W., and Zio, E. (2011c). "Basic Approaches." *Vulnerable systems*, Springer-Verlag London Limited, 41–54.
- Kukkonen, S., and Coello Coello, C. A. (2017). "Generalized Differential Evolution for Numerical and Evolutionary Optimization." *Studies in Computational Intelligence*, 253–279.
- Kumar, A., Kansal, M. L., and Arora, G. (1999). "Detecting accidental contaminations in municipal water networks." *Journal of Water Resources Planning and Management*, 125(5), 308–310.
- Langville, A. N., and Meyer, C. D. (2006). *Google's PageRank and Beyond: The Science of Search Engine Rankings*. Princeton University Press.
- Latora, V., and Marchiori, M. (2003). "Economic small-world behavior in weighted networks." *European Physical Journal B*, 32(2), 249–263.
- Latora, V., and Marchiori, M. (2007). "A measure of centrality based on network efficiency." *New Journal of Physics*, 9.
- Lee, B. H., and Deininger, R. A. (1992). "Optimal Locations of Monitoring Stations in Water Distribution System." *Journal of Environmental Engineering*, 118(1), 4–16.
- Lee, B. H., Deininger, R. A., and Clark, R. M. (1991). "Locating Monitoring Stations in Water Distribution Systems." *Journal American Water Works Association*, 83(7), 60–66.
- Leskovec, J., Krause, A., Guestrin, C., Faloutsos, C., and Vanbriesen, J. (2007). "Cost-effective Outbreak Detection in Networks."
- Li, H. (2007). "Hierarchical Risk Assessment of Water Supply Systems." (March).
- Mamlook, R., and Al-Jayyousi, O. (2003). "Fuzzy sets analysis for leak detection in infrastructure systems : a proposed methodology." *Clean Techn Environ Policy*, 6, 26–31.

- Martins Alves, C. (2020). "La capacidad hidráulica en las redes de agua potable y su influencia en el proceso de sectorización." Universitat Politècnica de València.
- Matalas, N. (2005). "Acts of nature and potential acts of terrorists: Contrast relative to water resource systems." *Journal of Water Resources Planning and Management*, (April), 79–80.
- Mays, L. W. (2004a). "Water Supply Security: An Introduction." *Water Supply Systems Security*, (1863), 1.1-1.12.
- Mays, L. W. (2004b). "Vulnerability Assessment, Emergency Response Planning: Summary of What's Available." *Water Supply Systems Security*, L. W. Mays, ed., McGraw-Hill, New York, London, 3.1-3.42.
- Mays, L. W. (Ed.). (2004c). *Water Supply Systems Security*. McGraw-Hill, New York, New York, USA.
- Michaud, D., and Apostolakis, G. E. (2006). "Methodology for Ranking the Elements of Water-Supply Networks." *Journal of Infrastructure Systems*, 12(4), 230–242.
- Mihalcea, R. (2004). "Graph-based ranking algorithms for sentence extraction, applied to text summarization." *Proceedings of the ACL 2004 on Interactive poster and demonstration sessions* -, (4), 20-es.
- Moglia, M., Technology, I., Burn, S., Leader, T., Infrastructure, U. W., Technology, I., Meddings, S., and Technology, I. (2006). "DECISION SUPPORT SYSTEM FOR WATER PIPELINE RENEWAL." 11(January), 237–256.
- Montalvo Arango, I. (2010). "Diseño óptimo de sistemas de distribución de agua mediante Agent Swarm Optimization." Universidad Politècnica de Valencia.
- Montalvo, I. (2008). "Diseño Óptimo de Sistemas de Distribución de Agua Mediante Particle Swarm." Universitat Politècnica de València.
- Montalvo, I., Izquierdo, J., Pérez-García, R., and Herrera, M. (2014). "Water distribution system computer-aided design by agent swarm optimization." *Computer-Aided Civil and Infrastructure Engineering*, 29(6), 433–448.
- Montalvo, I., Izquierdo, J., Schwarze, S., and Pérez-García, R. (2010). "Agent Swarm Optimization: A paradigm to tackle complex problems. Application to water distribution system design." *Modelling for Environment's Sake: Proceedings of the 5th Biennial Conference of the International Environmental Modelling and Software Society, iEMSs 2010*, 2035–2043.
- Montalvo, I., Izquierdo, J., Schwarze, S., and Pérez-García, R. (2012). "Agent swarm optimisation, a novel approach in swarm intelligence." *WIT Transactions on the Built Environment*, 35–47.
- Mora Rodríguez, J. de J. (2011). "Modelación de la Intrusión Patógena a través de Fallos y su Implicación en la Red de Distribución de Agua." Universitat Politècnica de València.
- Moscato, P. (1989). *On Evolution, Search, Optimization, Genetic Algorithms and Martial Arts: Towards Memetic Algorithms*. Pasadena, CA 91125, U.S.A.
- Moteff, J., Copeland, C., and Fischer, J. (2003). *Critical Infrastructures: What Makes an Infrastructure Critical?* Library of Congress Washington DC

- Congressional Research Service.*
- Murray, A. T., and Grubestic, T. H. (Eds.). (2007). *Critical Infrastructure. Reliability and Vulnerability. Advances in Spatial Science*, Springer Berlin Heidelberg New York.
- Murray, R., Haxton, T., Janke, R., Hart, W. E., Berry, J., and Phillips, C. (2010). "Sensor Network Design for Drinking Water Contamination Warning Systems. A compendium of Research Results and Case Studies using the TEVA-SPOT Software." (April), 79.
- Nafi, A., Crastes, E., Sadiq, R., Gilbert, D., and Piller, O. (2018). "Intentional contamination of water distribution networks: developing indicators for sensitivity and vulnerability assessments." *Stochastic Environmental Research and Risk Assessment*, 32(2), 527–544.
- Nam, Y., Arai, Y., Koizumi, A., and Kunizane, T. (2018). "Planning Models for Optimal Deployment of Water Leak Sensors and Their Effectiveness." *1st International WDSA / CCWI 2018 Joint Conference*, Kingston, Ontario, Canada.
- Di Nardo, A., Giudicianni, C., Greco, R., Herrera, M., and Santonastaso, G. F. (2018a). "Applications of graph spectral techniques to water distribution network management." *Water (Switzerland)*, 10(1), 1–16.
- Di Nardo, A., Giudicianni, C., Greco, R., Herrera, M., Santonastaso, G. F., and Scala, A. (2018b). "Sensor Placement in Water Distribution Networks based on Spectral Algorithms." *HIC 2018. 13th International Conference on Hydroinformatics*, G. La Loggia, G. Freni, V. Puleo, and M. De Marchis, eds., EPiC Series in Engineering, 593–600.
- Di Nardo, A., Di Natale, M., Guida, M., and Musmarra, D. (2013). "Water Network Protection from Intentional Contamination by Sectorization." *Water Resources Management*, 27, 1837–1850.
- Di Nardo, A., Di Natale, M., Musmarra, D., Santonastaso, G. F., Tzatchkov, V., and Alcocer-Yamanaka, V. H. (2014). "A District Sectorization for Water Network Protection from Intentional Contamination." *Procedia Engineering*, Elsevier B.V., 70, 515–524.
- National Oceanic and Atmospheric Administration (NOAA). (2010). *Adapting to Climate Change : A Planning Guide for State Coastal Managers.*
- National Research Council. (2004). *Preventing earthquake disasters: The grand challenge in earthquake engineering: A research agenda for the Network for Earthquake Engineering (NESS).*
- Nazempour, R., Monfared, M. A. S., and Zio, E. (2018). "A Complex Network Theory Approach for Optimizing Contamination Warning Sensor Location in Water Distribution Networks." *International Journal of Disaster Risk Reduction*, Elsevier Ltd, 30(April), 225–234.
- Newman, M. E. J. (2003). "The Structure and Function of Complex Networks." *SIAM Review*, 45(2), 167–256.
- Oliveira, E. C. M., Melo Brentan, B., Danta, R. F., dos Santos Macedo, L., Luvizotto Junior, E., and Ribeiro, L. C. L. J. (2018). "Detection of Chemical Intrusion

- Compounds in Water Distribution Networks by Quality Sensors Data Mining." *1st International WDSA / CCWI 2018 Joint Conference*, Kingston, Ontario, Canada.
- Ostfeld, A. (2006). "Enhancing Water-Distribution System Security through Modeling." *Journal of Water Resources Planning and Management*, 132(4), 209–210.
- Ostfeld, A., Oliker, N., and Salomons, E. (2014). "Multiobjective optimization for least cost design and resiliency of water distribution systems." *Journal of Water Resources Planning and Management*, 140(12), 1–12.
- Ostfeld, A., and Salomons, E. (2003). "An Early Warning Detection System (EWDS) for Drinking Water Distribution Systems Security." *World Water and Environmental Resources Congress 2003*, American Society of Civil Engineers (ASCE), 1–7.
- Ostfeld, A., and Salomons, E. (2004). "Optimal Layout of Early Warning Detection Stations for Water Distribution Systems Security." *Journal of Water Resources Planning and Management*, 130(5), 377–385.
- Ostfeld, A., and Salomons, E. (2005a). "Securing Water Distribution Systems Using Online." *Journal of Water Resources Planning and Management*, 131(5), 402–405.
- Ostfeld, A., and Salomons, E. (2005b). "Optimal early warning monitoring system layout for water networks security: Inclusion of sensors sensitivities and response delays." *Civil Engineering and Environmental Science*, 22(3), 151–169.
- Ostfeld, A., and Salomons, E. (2008). "Sensor Network Design Proposal for the Battle of the Water Sensor Networks (BWSN)." *Water Distribution Systems Analysis Symposium 2006*, American Society of Civil Engineers, Reston, VA, Ohio, United States, 1–16.
- Ostfeld, A., and Salomons, E. (2014). "Optimal disinfection of water distribution networks following a contamination event." *Procedia Engineering*, Elsevier B.V., 89(d), 168–172.
- Ostfeld, A., Uber, J. G., Salomons, E., Berry, J. W., Hart, W. E., Phillips, C. A., Watson, J., Dorini, G., Jonkergouw, P., Kapelan, Z., di Pierro, F., Khu, S.-T., Savic, D., Eliades, D., Polycarpou, M., Ghimire, S. R., Barkdoll, B. D., Gueli, R., Huang, J. J., McBean, E. A., James, W., Krause, A., Leskovec, J., Isovitsch, S., Xu, J., Guestrin, C., VanBriesen, J., Small, M., Fischbeck, P., Preis, A., Propato, M., Piller, O., Trachtman, G. B., Wu, Z. Y., and Walski, T. (2008). "The Battle of the Water Sensor Networks (BWSN): A Design Challenge for Engineers and Algorithms." *Journal of Water Resources Planning and Management*, 134(6), 556–568.
- Panguluri, S., Meiners, G., Hall, J., and Szabo, J. (2009). "Distribution System Water Quality Monitoring: Sensor Technology Evaluation Methodology and Results." *Environmental Protection Agency*, 1–60.
- Panguluri, S., Phillips, W. R. J., and Clark, R. M. (2004). "Cyber threats and IT/SCADA system vulnerability." *Water Supply Systems Security*, L. W. Mays,

- ed., McGraw-Hill, New York, London.
- Parks, S. L. I., and VanBriesen, J. M. (2009). "Booster Disinfection for Response to Contamination in a Drinking Water Distribution System." *Journal of Water Resources Planning and Management*, 135(6), 502–511.
- Pedroche, F. (2007). "Métodos de Cálculo del Vector PageRank." *SeMA Journal: Boletín de la Sociedad Española de Matemática Aplicada*, 39, 7–30.
- Perelman, L., Arad, J., Housh, M., and Ostfeld, A. (2012). "Event detection in water distribution systems from multivariate water quality time series." *Environmental Science and Technology*, 46(15), 8212–8219.
- Perelman, L., and Ostfeld, A. (2010). "Extreme Impact Contamination Events Sampling for Water Distribution Systems Security." *Journal of Water Resources Planning and Management*, 136(1), 80–87.
- Perelman, L., and Ostfeld, A. (2013a). "Application of graph theory to sensor placement in water distribution systems." *World Environmental and Water Resources Congress 2013: Showcasing the Future - Proceedings of the 2013 Congress*, 617–625.
- Perelman, L., and Ostfeld, A. (2013b). "Bayesian Networks for Source Intrusion Detection." *Water Resources Planning and Management*, 139(4), 426–432.
- Pérez-García, R., Izquierdo, J., Herrera, M., and Gutiérrez-Pérez, J. A. (2009). "La vulnerabilidad en los sistemas de abastecimiento de agua: prevención, alerta temprana y recuperación." *IX SEREA. Seminario Iberoamericano sobre Planificación, Proyecto y Operación de Sistemas de Abastecimiento de Agua*, 35–50.
- Pons, P., and Latapy, M. (2005). "Computing Communities in Large Networks Using Random Walks." *Computer and Information Sciences - ISCIS 2005*, pInar Yolum, T. Güngör, F. Gürgen, and C. Özturan, eds., Springer Berlin Heidelberg, Berlin, Heidelberg, 284–293.
- Poulin, A., Mailhot, A., Periche, N., Delorme, L., and Villeneuve, J. (2010). "Planning Unidirectional Flushing Operations as a Response to Drinking Water Distribution System Contamination." *Journal of Water Resources Planning and Management*, 136(6), 647–657.
- Preis, A., and Ostfeld, A. (2006a). "Multiobjective Sensor Design for Water Distribution Systems Security." *8th Annual Water Distribution System Analysis Symposium*, American Society of Civil Engineers (ASCE), Cincinnati, Ohio, United States, 1–17.
- Preis, A., and Ostfeld, A. (2006b). "Optimal Sensors Layout for Contamination Source Identification in Water Distribution Systems." *Water Distribution Systems Analysis Symposium 2006*, American Society of Civil Engineers (ASCE), 1–12.
- Preis, A., and Ostfeld, A. (2006c). "Contamination Source Identification in Water Systems: A Hybrid Model Trees–Linear Programming Scheme." *Journal of Water Resources Planning and Management*, 132(4), 263–273.
- Preis, A., and Ostfeld, A. (2008a). "Multiobjective Contaminant Sensor Network Design for Water Distribution Systems." *Journal of Water Resources*

- Planning and Management*, 134(4), 366–377.
- Preis, A., and Ostfeld, A. (2008b). "Genetic algorithm for contaminant source characterization using imperfect sensors." *Civil Engineering and Environmental Systems*, 25(1), 29–39.
- Propato, M. (2006). "Contamination Warning in Water Networks: General Mixed-Integer Linear Models for Sensor Location Design." *Journal of Water Resources Planning and Management*, 132(4), 225–233.
- Propato, M., and Piller, O. (2008). "Battle of the Water Sensor Networks." *Water Distribution Systems Analysis Symposium 2006*, American Society of Civil Engineers, Reston, VA, Ohio, United States, 1–8.
- Propato, M., Piller, O., and Uber, J. G. (2005). "A Sensor Location Model to Detect Contaminations in Water Distribution Networks." *Impacts of Global Climate Change*, American Society of Civil Engineers, Reston, VA, 1–12.
- Qiao, J. (2005). "Security Enhancement and Consequence Mitigation Strategies for Water Infrastructure Against Physical Destruction." Purdue University.
- R Development Core Team. (2008). "R: A Language and Environment for Statistical Computing."
- Ramos-Martínez, E., Herrera, M., Gutiérrez-Pérez, J., Izquierdo, J., and Pérez-García, R. (2014). "Rehabilitation actions in water supply systems: Effects on biofilm susceptibility." *Procedia Engineering*, 89, 225–231.
- Rathi, S., and Gupta, R. (2014). "Sensor placement methods for contamination detection in water distribution networks: A review." *Procedia Engineering*, Elsevier B.V., 89, 181–188.
- Rathi, S., and Gupta, R. (2015). "A critical review of sensor location methods for contamination detection in water distribution networks." *Water Quality Research Journal of Canada*, 50(2), 95–108.
- Rathi, S., and Gupta, R. (2017). "Optimal sensor locations for contamination detection in pressure-deficient water distribution networks using genetic algorithm." *Urban Water Journal*, Taylor & Francis, 14(2), 160–172.
- Rathi, S., Gupta, R., and Ormsbee, L. (2015). "A review of sensor placement objective metrics for contamination detection in water distribution networks." *Water Science and Technology: Water Supply*, 15(5), 898–917.
- Sarrate, R., Blesa, J., and Nejari, F. (2014). "Clustering Techniques Applied to Sensor Placement for Leak Detection and Location in Water Distribution Networks." *2014 22nd Mediterranean Conference on Control and Automation, MED*, Palermo, Italy, 109–114.
- Shafiee, M. E., and Zechman, E. M. (2013). "Integrating Genetic Programming and Agent-based Modeling to Identify Sensor-based Rules for Flushing Contaminated Water from a Pipe Network." *World Environmental and Water Resources Congress 2013*, 784–790.
- Shastri, Y., and Diwekar, U. (2006). "Sensor Placement in Water Networks: A Stochastic Programming Approach." *Journal of Water Resources Planning and Management*, 132(3), 192–203.
- Traag, V. A., Waltman, L., and van Eck, N. J. (2019). "From Louvain to Leiden:

- guaranteeing well-connected communities." *Scientific Reports*, 9(1), 1–12.
- Trachtman, G. B. (2008). "A "Strawman" Common Sense Approach for Water Quality Sensor Site Selection." *Water Distribution Systems Analysis Symposium 2006*, American Society of Civil Engineers, Reston, VA, Ohio, United States, 1–13.
- Tryby, M. E., Propato, M., and Ranjithan, S. R. (2010). "Monitoring Design for Source Identification in Water Distribution Systems." *Journal of Water Resources Planning and Management*, 136(6), 637–646.
- Tzatchkov, V., Alcocer-Yamanaka, V., and Bourguett Ortíz, V. (2006). "Graph Theory Based Algorithms for Water Distribution Network Sectorization Projects." *Water Distribution Systems Analysis (WDSA)*, 1–15.
- Uber, J., Janke, R., Murray, R., and Meyer, P. (2004). "Greedy Heuristic Methods for Locating Water Quality Sensors in Distribution Systems." *Critical Transitions in Water and Environmental Resources Management*, G. Sehlke, D. F. Hayes, and D. K. Stevens, eds., American Society of Civil Engineers (ASCE), 1–9.
- Watson, J.-P., Greenberg, H. J., and Hart, W. E. (2004). "A Multiple Objective Analysis of Sensor Placement Optimization in Water Networks." *Critical Transitions in Water and Environmental Resources Management*, American Society of Civil Engineers (ASCE), 1–10.
- Watson, J.-P., Hart, W. E., and Berry, J. W. (2005). "Scalable High-Performance Heuristics for Sensor Placement in Water Distribution Networks." *Impacts of Global Climate Change*, American Society of Civil Engineers (ASCE), 1–12.
- Watson, J.-P., Hart, W. E., and Murray, R. (2006). "Formulation and Optimization of Robust Sensor Placement Problems for Contaminant Warning Systems." *Water Distribution Systems Analysis Symposium 2006*, American Society of Civil Engineers (ASCE), 1–13.
- Watson, J.-P., Murray, R., and Hart, W. E. (2009). "Formulation and Optimization of Robust Sensor Placement Problems for Drinking Water Contamination Warning Systems." *Journal of infrastructure systems*, 15(December), 330–339.
- Watts, D. J., and Strogatz, S. H. (1998). "Collective Dynamics of 'Small-World' Networks." *Letters to Nature*, 393, PP. 440-442.
- Wooldridge, M. (2002). "Intelligent Agents: The Key Concepts BT - Multi-Agent Systems and Applications II." V. Mařík, O. Štěpánková, H. Krautwurmová, and M. Luck, eds., Springer Berlin Heidelberg, Berlin, Heidelberg, 3–43.
- Wright, R., Stoianov, I., Parpas, P., Henderson, K., and King, J. (2014). "Adaptive water distribution networks with dynamically reconfigurable topology." *Journal of Hydroinformatics*, 16(6), 1280–1301.
- Wu, X.-G., Zhang, T.-Q., and Huang, Y.-D. (2008). "Optimal Algorithm for Determining Locations of Water Quality Sensors in Water Supply Networks under Multi-Objective Constraints." *Journal of Hydraulic Engineering*, 39(4), 433–439.

- Wu, Z. Y., and Walski, T. (2006). "Multi-Objective Optimization of Sensor Placement in Water Distribution Systems." *8th Annual Water Distribution System Analysis Symposium*, American Society of Civil Engineers (ASCE), Cincinnati, Ohio, United States, 1–11.
- Xu, J., Fischbeck, P. S., Small, M. J., VanBriesen, J. M., and Casman, E. (2008). "Identifying Sets of Key Nodes for Placing Sensors in Dynamic Water Distribution Networks." *Journal of Water Resources Planning and Management*, 134(4), 378–385.
- Xu, J., Small, M., Fischbeck, P., and Vanbriesen, J. (2010). "Integrating Location Models with Bayesian Analysis to Inform Decision Making." *Journal of Water Resources Planning and Management*, 136(2), 209–216.
- Xu, J., Vanbriesen, J. M., Small, M. J., and Fischbeck, P. S. (2009). "Decision Making under Information Constraints." *World Environmental and Water Resources Congress 2009*, American Society of Civil Engineers (ASCE), 1–9.
- Yang, Z., Algesheimer, R., and Tessone, C. J. (2016). "A comparative analysis of community detection algorithms on artificial networks." *Scientific Reports*, Nature Publishing Group, 6(July).
- Yazdani, A., and Jeffrey, P. (2010). "A note on measurement of network vulnerability under random and intentional attacks." 1–9.
- Yazdani, A., and Jeffrey, P. (2011). "Complex network analysis of water distribution systems." *Chaos*, 21(1).
- Yazdani, A., Otoo, R. A., and Jeffrey, P. (2011). "Resilience enhancing expansion strategies for water distribution systems: A network theory approach." *Environmental Modelling and Software*, Elsevier Ltd, 26(12), 1574–1582.
- Zhao, Y., Schwartz, R., Salomons, E., Ostfeld, A., and Poor, H. V. (2016). "New formulation and optimization methods for water sensor placement." *Environmental Modelling and Software*, 76, 128–136.
- Zitzler, E., Laumanns, M., and Bleuler, S. (2004). "A Tutorial on Evolutionary Multiobjective Optimization." *Metaheuristics for Multiobjective Optimisation*, X. Gandibleux, M. Sevaux, K. Sörensen, and V. T'kindt, eds., Springer Berlin Heidelberg, Berlin, Heidelberg, 3–37.

Publicaciones propias

Artículos en Revistas

- 1) **Gutiérrez-Pérez, J. A.** Herrera, M., Pérez-García, R., & Ramos-Martínez, E. (2013). Application of graph-spectral methods in the vulnerability assessment of water supply networks. *Mathematical and Computer Modelling*, 57(7-8), 1853-1859.
- 2) Herrera Fernández, A. M., **Gutiérrez-Pérez, J. A.**, Izquierdo Sebastián, J., & Pérez García, R. (2012). Combining multiple perspectives on clustering. Node-pipe case in hydraulic sectorization. *International Journal of Complex Systems in Science*, 2(1), 17-20.
- 3) Benítez, J., Delgado-Galván, X., **Gutiérrez-Pérez, J. A.**, & Izquierdo, J. (2011). Balancing consistency and expert judgment in AHP. *Mathematical and Computer Modelling*, 54(7-8), 1785-1790.
- 4) Ramos-Martínez, E., Herrera, M., **Gutiérrez-Pérez, J. A.**, Izquierdo, J., & Pérez-García, R. (2014). Rehabilitation actions in water supply systems: Effects on biofilm susceptibility. *Procedia Engineering*, 89, 225-231.

Participación en Congresos

- 5) **Gutiérrez-Pérez, J. A.** Idel Montalvo, Jorge Francés-Chust, Manuel Herrera & Joaquín Izquierdo. "Water quality sensor placement with a multiobjective approach". Congreso de Métodos Numéricos en Ingeniería (CMN), Valencia, España. Julio, 2017.
- 6) Campbell, E., Pérez-García, R., Izquierdo, J., Ayala-Cabrera, D., Tavera, M. & **Gutiérrez-Pérez, J. A.** "Sectorización multinivel de redes de abastecimiento de agua potable basada en métodos de detección de comunidades en redes

- sociales”. XXVI Congreso Latinoamericano de Hidrogeología e Hidráulica, Santiago de Chile, Chile, 2014.
- 7) **Gutiérrez-Pérez, J.A.** Ramos-Martínez, E., Herrera, M., Izquierdo, J. & Pérez-García, R. “Kernel spectral clustering for identify vulnerable areas of biofilm development in drinking water distribution systems”. Mathematical Modelling in Engineering & Human Behaviour 2013 Conference, Valencia, España.
 - 8) **Gutiérrez-Pérez, J.A.** Herrera, M., Izquierdo, J. & Pérez-García, R. (2013). “Sectorización mediante el agrupamiento espacial de la información completa de la red de abastecimiento”. XII Simposio Iberoamericano sobre sistemas de abastecimiento y drenaje urbano (SEREA). Buenos Aires, Argentina, 2013.
 - 9) Ramos-Martínez, E., **Gutiérrez-Pérez, J.A.** Herrera, M., Izquierdo, J. & Pérez-García, R. (2013). “Metadata on biofilm development in drinking water distribution systems”. Artificial Intelligence Research and Development. Proceedings of the 16th International Conference of the Catalan Association of Artificial Intelligence.
 - 10) **Gutiérrez-Pérez, J. A.** Herrera, M., Izquierdo, J., Pérez-García & R., Ramos-Martínez, E. (2012). “Enfoque multi-agente para la identificación de elementos vulnerables de una red de abastecimiento de agua”. VIII Seminario Euro-Latinoamericano de Sistemas de Ingeniería (SELASI). La Habana, Cuba.
 - 11) **Gutiérrez-Pérez, J. A.** Herrera, M., Izquierdo, J., Pérez-García R. & Ramos-Martínez, E. (2012). “Identificación de elementos vulnerables de una red de abastecimiento de agua mediante sistemas multi-agente”. XI Seminário Ibero-Americano sobre Sistemas de Abastecimento e Drenagem (SEREA). Coimbra, Portugal. ISBN: 978-989-26-0298-1
 - 12) **Gutiérrez-Pérez, J. A.** Herrera, M., Izquierdo J. & Pérez-García, R. (2012). “An approach based on ranking elements to form supply clusters in water supply networks as a support to vulnerability assessment”. International Environmental Modelling and Software Society (iEMSS). International

- Congress on Environmental Modelling and Software Managing Resources of a Limited Planet, Sixth Biennial Meeting, Leipzig, Alemania.
- 13) Herrera, M., **Gutiérrez-Pérez, J. A.**, Izquierdo, J. & Pérez-García, R. (2012). “A multi-level spectral clustering process to ascertain sensor location for mitigating effects of a potential contamination in a water supply network”. International Environmental Modelling and Software Society (iEMSs). International Congress on Environmental Modelling and Software Managing Resources of a Limited Planet, Sixth Biennial Meeting, Leipzig, Alemania.
 - 14) Herrera, M., **Gutiérrez-Pérez, J. A.**, Izquierdo, J. & Pérez-García, R. (2012). “Métodos de análisis espectral para la sectorización de la red de abastecimiento y su posterior gestión”. XXV Congreso Latinoamericano de Hidráulica. San José, Costa Rica.
 - 15) **Gutiérrez-Pérez, J. A.**, Herrera, M., Pérez-García, R. & Ramos-Martínez, E. (2011). “Application of graph-spectral methods in the vulnerability assessment of water supply networks”. Modelling for engineering and human behaviour 2011. ISBN 978-84-695-2143-4.
 - 16) **Gutiérrez-Pérez, J. A.**, Herrera, M., Pérez-García, R. & Izquierdo, J. (2011). “Aplicación de técnicas de teoría de grafos en el análisis de la vulnerabilidad de redes abastecimiento de agua”. X Seminario Iberoamericano sobre planificación, proyecto y operación de sistemas de abastecimiento de agua (SEREA). Morelia, México. ISBN 978-84-89487-34-5.
 - 17) **Gutiérrez-Pérez, J. A.**, Herrera, M., Izquierdo, J. & Pérez-García, R. (2011). “Método basado en grafos para evaluar la vulnerabilidad de redes de abastecimiento de agua”. VII Seminario Euro-Latinoamericano de Sistemas de Ingeniería (SELASI). Valencia, España. ISBN 978-84-89487-35-2.
 - 18) Herrera, M., **Gutiérrez-Pérez, J. A.**, Izquierdo, J. & Pérez-García, R. (2011). “Ajustes en el modelo *pagerank* de *Google* para el estudio de la importancia relativa de los nodos de la red de abastecimiento”. X

- Seminario Iberoamericano sobre planificación, proyecto y operación de sistemas de abastecimiento de agua (SEREA). Morelia, México. ISBN 978-84-89487-34-5.
- 19) Pérez-García, R., Izquierdo, J., Herrera, M., **Gutiérrez-Pérez, J. A.** & Ramos-Martínez, E. (2011). "La vulnerabilidad de los sistemas de abastecimiento de agua". X Seminario Iberoamericano sobre planificación, proyecto y operación de sistemas de abastecimiento de agua (SEREA). Morelia, México. ISBN 978-84-89487-34-5.
- 20) Herrera, M., **Gutiérrez-Pérez, J. A.**, Izquierdo, J. & Rafael Pérez-García. (2011). "Ubicación de sensores en un sistema de abastecimiento de agua basada en un criterio de zonas hidráulicas". VII Seminario Euro-Latinoamericano de Sistemas de Ingeniería (SELASI). Valencia, España. ISBN 978-84-89487-35-2.
- 21) Izquierdo, J., Montalvo, I., Pérez-García, R. & **Gutiérrez-Pérez, J. A.** (2010). "A multi-agent framework for an IEDSS in urban water management". International Congress on Environmental Modelling and Software, Modelling for Environment's Sake, Fifth Biennial Meeting. Ottawa, Canada.
- 22) **Gutiérrez-Pérez, J. A.**, Pérez-García, R., Izquierdo, J. & Herrera, M. (2009). "Reducción de la vulnerabilidad en redes de abastecimiento mediante la instalación de una red de sensores de calidad". IX Seminario Iberoamericano sobre planificación, proyecto y operación de sistemas de abastecimiento de agua y saneamiento (SEREA). Valencia, España. ISBN: 978-84-89487-30-7
- 23) Pérez-García, R., Izquierdo, J., Herrera, M. & **Gutiérrez Pérez, J. A.** (2009). "La vulnerabilidad en los sistemas de abastecimiento de agua: prevención, alerta temprana y recuperación". IX Seminario Iberoamericano sobre planificación, proyecto y operación de sistemas de abastecimiento de agua y saneamiento (SEREA). Valencia, España. ISBN: 978-84-89487-30-7
- 24) Pérez-Gracia, R., Izquierdo, J., Herrera, M., Tavera, M. & **J. A. Gutiérrez-Pérez.** (2008). "Aplicación de los sistemas de información y decisión en la predicción de escenarios de riesgo en sistemas de abastecimiento de

agua". VIII SEREA Seminario Ibero-Americano sobre Sistemas de Abastecimiento e Drenagem, Lisboa, Portugal. ISBN 978-989-95838-0-1.

Capítulos de Libro

- 25) Ramos-Martínez, E., Gutiérrez-Pérez, J. A., Herrera, M., Izquierdo, J. & Pérez-García, R. (2014). "Pipe Database Analysis Transduction to Assess the Spatial Vulnerability to Biofilm Development in Drinking Water Distribution Systems". *Mathematical Modeling in Social Sciences and Engineering*. Eds.: J. C. Cortés, L. A. Jódar and R. J. Villanueva. Nova Science Publishers, New York, pp. 71-80. ISBN 978-1-63117-339-4.
- 26) Herrera, M., Ramos-Martínez, E., Gutiérrez-Pérez, J. A., Izquierdo, J. & Pérez-García, R. (2014). "On Kernel Spectral Clustering for Identifying Areas of Biofilm Development in Water Distribution Systems". *Mathematical Modeling in Social Sciences and Engineering*. Eds.: J. C. Cortés, L. A. Jódar and R. J. Villanueva. Nova Science Publishers, New York, pp. 81-89. ISBN 978-1-63117-339-4.
- 27) Ramos-Martínez, E., Gutiérrez-Pérez J. A., Herrera, M., Izquierdo, J. & Pérez García, R. (2013). "Metadata on biofilm development in drinking water distribution systems". *Artificial Intelligence Research and Development. Proceedings of the 16th International Conference of Catalan Association for Artificial Intelligence*. IOS Press. pp. 307-310. DOI: 10.3233/978-1-61499-320-9-307.
- 28) Gutiérrez-Pérez, J. A., Herrera, M., Pérez-García, R., & Ramos-Martínez, E. (2011). "Application of graph spectral methods in the vulnerability assessment of water supply networks". *Modelling for engineering and human behaviour 2011*. Eds.: L. Jódar, L. Acedo and J.C. Cortés, Instituto de Matemática Multidisciplinar IMM-UPV. ISBN 978-84-695-2143-4 Vol. 1, pp. 154-157.
- 29) Pérez-García, R., Gutiérrez-Pérez, J. A., Izquierdo, J. & Herrera, M. (2011). "Forecasting potential risk scenarios in water supply systems: application

- of information and decision support systems”. Applications of intelligent data analysis in urban water systems. Eds.: Joaquín Izquierdo y Rafael Pérez-García. Grupo FluIng–IMM–UPV, pp. 169-191. ISBN 978-84-89487-33-8.
- 30) Benítez, J., Delgado-Galván, X. V., **Gutiérrez-Pérez, J. A.** & Izquierdo, J. (2011). “Balancing Consistency and Expert Judgment in AHP”. Applications of intelligent data analysis in urban water systems. Eds.: Joaquín Izquierdo y Rafael Pérez-García. Grupo FluIng–IMM–UPV, pp. 141-151. ISBN 978-84-89487-33-8.
- 31) **Gutiérrez-Pérez, J. A.**, Pérez-García, R., Izquierdo, J. & Herrera, M. (2010). “Análisis de eventos de contaminación en redes de abastecimiento de agua con técnicas de control estadístico de procesos”. Retos tecnológicos y metodológicos en la gestión técnica de los sistemas urbanos de agua. I Seminario de la Red Iberoamericana IT-AGUA. Editores: Rafael Pérez-García y Joaquín Izquierdo. Grupo FluIng–IMM–UPV, pp. 3-58. ISBN 978-84-89487-32-1.
- 32) **Gutiérrez-Pérez, J. A.**, Pérez-García, R., Izquierdo, J. & Herrera, M. (2010). “Monitorización de intrusiones contaminantes deliberadas en redes de agua potable mediante técnicas de control estadístico de procesos”. Planificación, proyecto y operación de sistemas de abastecimiento de agua. Editores: Rafael Pérez-García y Joaquín Izquierdo. Grupo FluIng–IMM–UPV, pp. 307-318. ISBN 978-84-89487-31-4.
- 33) Pérez-García, R., Izquierdo, J., Herrera, M. & **Gutiérrez-Pérez, J. A.** (2010). “Análisis y cuantificación de la vulnerabilidad en los sistemas de abastecimiento de agua”. Planificación, proyecto y operación de sistemas de abastecimiento de agua. Editores: Rafael Pérez-García y Joaquín Izquierdo. Grupo FluIng–IMM–UPV, pp. 195-206. ISBN 978-84-89487-31-4.

Anexos

Anexo 1. Representación espacial de los 20 sensores obtenidos por medio de los valores de las medidas de centralidad de los nodos de la RDA de Coro.

- **Anexo 1.A.** Sensores obtenidos por el valor de *pagerank*.



- **Anexo 1.B.** Sensores obtenidos por el valor de intermediación (*betweenness*).



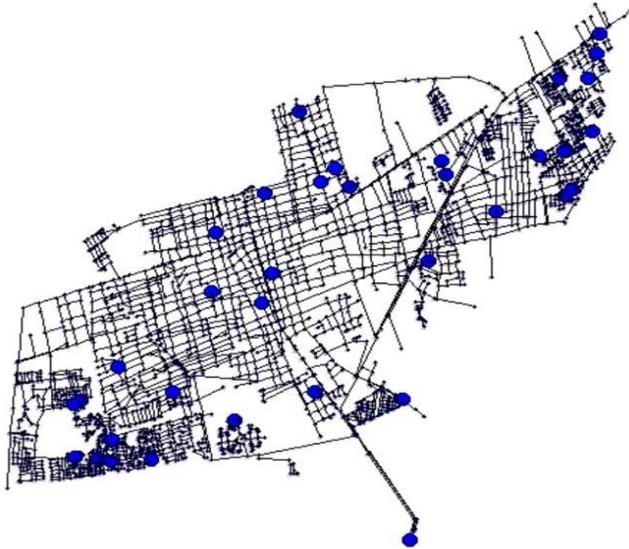
- **Anexo 1.C.** Sensores obtenidos por el valor de centralidad del autovector (*eigenvector centrality*).



Anexo 2. Representación espacial de cinco soluciones seleccionadas del frente de Pareto para 41 sensores del algoritmo ASO de la red de Coro. Son, al igual que la solución representada en la Figura 5.21, algunas de las soluciones elegidas en función de su mejor comportamiento respecto al agua contaminada consumida.

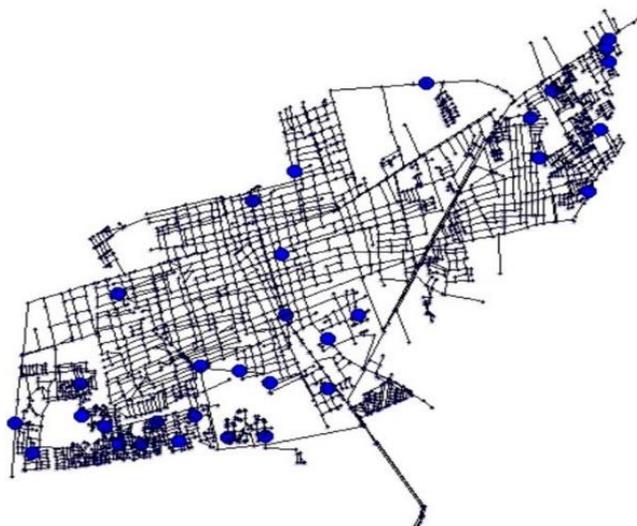
- **Anexo 2.A.**

Solución de sensores ASO: bird21



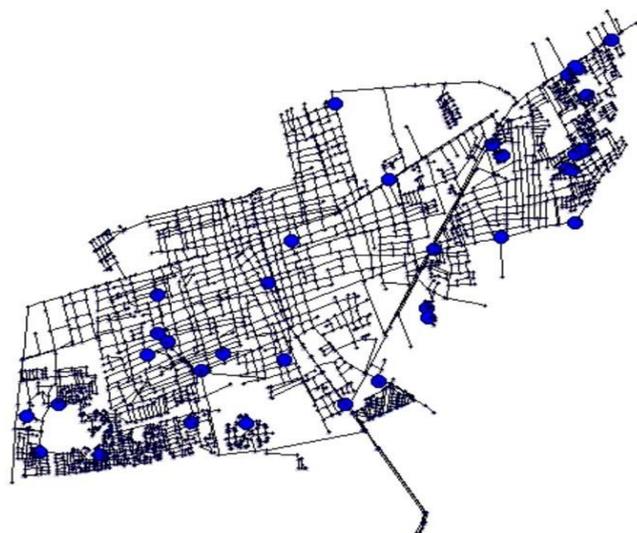
- Anexo 2.B.

Solución de sensores ASO: bird39



- Anexo 2.C.

Solución de sensores ASO: bird70 Op2



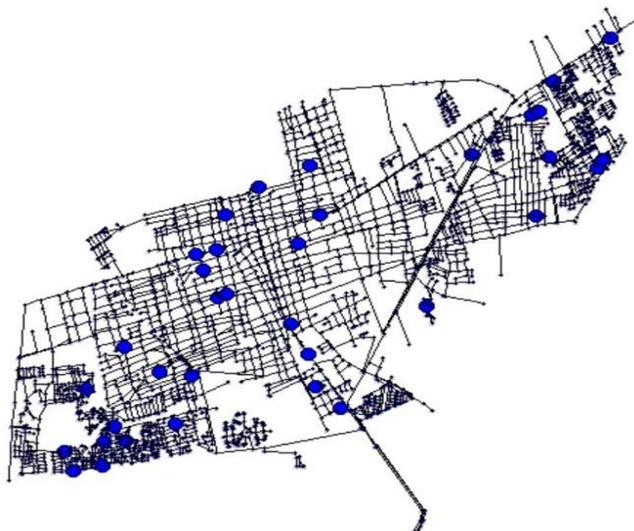
- Anexo 2.D.

Solución de sensores ASO: bird89



- Anexo 2.E.

Solución de sensores ASO: bird90



A continuación, se hace una lista de los anexos correspondientes a los resultados obtenidos a través de los diferentes métodos utilizados en esta tesis. Esta información se ha conformado en ficheros de formato csv y está contenida y disponible para su consulta en:

https://github.com/joagupre/Anexos_Tesis

Anexo 3. Resultados de los cálculos de la RDA de Coro, obtenidos con el método de teoría de redes.

- **Anexo 3.A.** Medidas de centralidad (*pagerank*, *betweenness* y *eigenvector centrality*) y grado de los nodos de la RDA de Coro.
- **Anexo 3.B.** Autovalores de la Matriz de Adyacencia del grafo de la RDA de Coro.
- **Anexo 3.C.** Autovalores de la Matriz Laplaciana del grafo de la RDA de Coro.

Anexo 4. Información de las comunidades de la RDA de Coro: número de nodos agrupados y promedio de las medidas de centralidad de los nodos de la RDA de Coro por comunidad.

Anexo 5. Información de los 20 sensores obtenidos por medio de los valores de las medidas de centralidad de los nodos de la RDA de Coro.

- **Anexo 5.A.** Sensores obtenidos por el valor de *pagerank*.
- **Anexo 5.B.** Sensores obtenidos por el valor de intermediación. (*betweenness*).
- **Anexo 5.C.** Sensores obtenidos por el valor de centralidad del autovector (*eigenvector centrality*).

Anexo 6. Información de los 41 sensores obtenidos por medio de los valores de las medidas de centralidad de los nodos de la RDA de Coro.

- **Anexo 6.A.** Sensores obtenidos por el valor de *pagerank*.
- **Anexo 6.B.** Sensores obtenidos por el valor de intermediación (*betweenness*).
- **Anexo 6.C.** Sensores obtenidos por el valor de centralidad del autovector (*eigenvector centrality*).

Anexo 7. Soluciones del algoritmo ASO para la ubicación de sensores en la RDA de Coro

- **Anexo 7.A.** Resultados de 3 sensores.
- **Anexo 7.B.** Resultados de 10 sensores.
- **Anexo 7.C.** Resultados de 20 sensores.
- **Anexo 7.D.** Resultados de 41 sensores.

Anexo 8. Soluciones del algoritmo ASO para la ubicación de sensores con base en las medidas de centralidad de los nodos de la RDa de Coro.

- **Anexo 8.A.** Resultados de 20 sensores.
- **Anexo 8.B.** Resultados de 41 sensores.