



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Dpto. de Ingeniería Hidráulica y Medio Ambiente

Generación de trazados para redes de agua ramificadas,
considerando criterios energéticos, mediante técnicas de
grafos y clustering

Trabajo Fin de Máster

Máster Universitario en Ingeniería Hidráulica y Medio Ambiente

AUTOR/A: Montaña Olivares, Alvaro Raul

Tutor/a: Gómez Sellés, Elena

Cotutor/a: Del Teso March, Roberto

CURSO ACADÉMICO: 2022/2023

Trabajo Fin de Máster
GENERACIÓN DE TRAZADOS PARA
REDES DE AGUA RAMIFICADAS
CONSIDERANDO CRITERIOS
ENERGÉTICOS MEDIANTE
TÉCNICAS DE GRAFOS Y
CLUSTERING.

Intensificación: *HIDRÁULICA URBANA*

Autor:

ÁLVARO RAÚL MONTAÑA OLIVARES

Tutor:

DRA. ELENA GÓMEZ SELLÉS.

Cotutor:

DR. ROBERTO DEL TESO MARCH

Septiembre, 2023



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

máster en ingeniería
hidráulica y medio ambiente
mihma



AGRADECIMIENTOS

A los que dedicaron una parte de su tiempo y esfuerzo para que este trabajo fuera posible, en particular a quienes me han acompañado y apoyado en la distancia y a los que han estado a mi lado en esta experiencia. Una mención especial a mis tutores Elena y Roberto, por su apoyo y confianza durante este proceso.

RESUMEN

El incremento de la población global y el crecimiento económico han ocasionado un aumento significativo en el consumo de agua y energía. Este fenómeno, combinado con la disponibilidad de recursos y la gestión actual de los mismos, presenta un desafío sustancial en la búsqueda de servicios de agua sostenibles. La búsqueda de soluciones que promuevan la eficiencia es esencial para el futuro. En la actualidad, el sector del agua consume aproximadamente un 4% de la electricidad mundial, y se proyecta un aumento del 80% en el consumo para el año 2040. En adición, cerca del 50% de esta energía se emplea en la fase de distribución, según estadísticas de la Agencia Internacional de Energía (IEA).

En este contexto, identificar oportunidades para mejorar la eficiencia energética en el transporte del agua dentro de las redes de distribución es crucial, esto conlleva a una reducción en el consumo de energía y, por ende, a una disminución de los costos operativos, teniendo un impacto directo en la sostenibilidad del servicio.

Un factor condicionante del consumo energético que tradicionalmente ha sido pasado por alto es la inclusión del análisis energético en la etapa de definición del trazado de la red de distribución. En este trabajo, se propone una metodología con la cual generar trazados para redes de distribución de agua ramificadas más eficientes energéticamente, mediante el uso de técnicas de Machine Learning, en particular, el clustering y la teoría de grafos.

La metodología propuesta se desarrolla en cinco etapas, empleando la distribución e información sobre los puntos de consumo y las posibles rutas de tuberías. Se realiza una sectorización inicial de la red, a partir de la cual se define una topología base que sirve para la identificación de diversos trazados de la red principal de distribución de agua; estos trazados son evaluados en función de su eficiencia energética.

La metodología presentada permite la generación de trazados energéticamente eficientes, optimizando la energía requerida en la cabecera, adaptándose adecuadamente al terreno con el mínimo exceso de energía en los nodos y al menor costo global. Este enfoque resulta en un avance significativo hacia la consecución de servicios de agua sostenibles y eficientes desde el punto de vista energético.

Palabras claves: Redes de distribución de agua, análisis energético de una red de agua, generación de trazados, machine learning, clustering, teoría de grafos.

RESUM

L'increment de la població global i el creixement econòmic han engendrat un augment significatiu en el consum d'aigua i energia. Aquest fenomen, combinat amb la disponibilitat de recursos i la gestió actual d'aquests, presenta un desafiament substancial en la cerca de serveis d'aigua sostenibles. Pel que, la cerca de solucions que promoguen l'eficiència és essencial per al futur. En l'actualitat, el sector de l'aigua consumeix aproximadament un 4% de l'electricitat mundial, i es projecta un augment del 80% en el consum per a l'any 2040, i prop del 50% d'aquesta energia s'empra en la fase de distribució, segons estadístiques de l'Agència Internacional d'Energia (IEA).

En aquest context, identificar oportunitats per a millorar l'eficiència energètica en el transport de l'aigua dins de les xarxes de distribució és crucial, ja que això comporta a una reducció en el consum d'energia i, per tant, a una disminució dels costos operatius, tenint un impacte directe en la sostenibilitat del servei.

Un factor condicionant del consum energètic que tradicionalment ha sigut passat per alt és la inclusió de l'anàlisi energètica en l'etapa de definició del traçat de la xarxa de distribució. En aquest treball, es proposa una metodologia amb la qual generar traçats per a xarxes de distribució d'aigua ramificades més eficients energèticament, mitjançant l'ús de tècniques de Machine Learning, en particular, el clustering i la teoria de grafos.

La metodologia proposada es desenvolupa en cinc etapes, emprant la distribució i informació sobre els punts de consum i les possibles rutes de canonades. Es realitza una sectorització inicial de la xarxa, a partir de la qual es defineix una topologia base que serveix com a base per a la identificació de diversos traçats de la xarxa principal de distribució d'aigua; aquests traçats són avaluats en funció de la seua eficiència energètica.

La metodologia presentada permet la generació de traçats energèticament eficients, optimitzant l'energia requerida en la capçalera, adaptant-se adequadament al terreny amb el mínim excés d'energia en els nodes i al menor cost global. Aquest enfocament resulta en un avanç significatiu cap a la consecució de serveis d'aigua sostenibles i eficients des del punt de vista energètic.

Paraules claus: Xarxes de distribució d'aigua, anàlisi energètica d'una xarxa d'aigua, generació de traçats, machine learning, clustering, teoria de grafos.

ABSTRACT

The increase in global population and economic growth has led to a significant rise in water and energy consumption. This phenomenon, combined with resource availability and current management practices, poses a substantial challenge in the pursuit of sustainable water services. Therefore, seeking solutions that promote efficiency is crucial for the future. Currently, the water sector accounts for approximately 4% of global electricity consumption, and an 80% increase in consumption is projected by 2040. About 50% of this energy is used in the distribution phase, according to statistics from the International Energy Agency (IEA).

In this context, identifying opportunities to enhance energy efficiency in water transportation within distribution networks is crucial, as it leads to reduced energy consumption and, consequently, lower operational costs, directly impacting the sustainability of the service.

A determinant factor of energy consumption that has traditionally been overlooked is the inclusion of energy analysis in the stage of defining the layout of the distribution network. This work proposes a methodology to generate more energy-efficient layouts for branched water distribution networks using Machine Learning techniques, specifically clustering and graph theory.

The proposed methodology unfolds in five stages, utilizing distribution and information about consumption points and potential pipeline routes. An initial network sectorization is performed, from which a baseline topology is defined. This baseline serves as a foundation for identifying various layouts for the main water distribution network. These layouts are evaluated based on their energy efficiency.

The presented methodology enables the generation of energetically efficient layouts, optimizing the energy required at the source, adapting suitably to the terrain with minimal excess energy at nodes, and achieving the lowest overall cost. This approach represents a significant step towards achieving sustainable and energetically efficient water services.

Key words: Water distribution networks, energy analysis of a water network, layout generation, machine learning, clustering, graph theory.

ÍNDICE

AGRADECIMIENTOS	i
RESUMEN	ii
RESUM	iii
ABSTRACT	iv
ÍNDICE	v
ÍNDICE DE FIGURAS.....	vii
ÍNDICE DE TABLAS.....	ix
Capítulo 1. Introducción.....	1
1.1.1. Problemática.....	2
1.2. Motivación y objetivos.....	5
1.2.1. Objetivo principal.....	5
1.2.1.1. Objetivos específicos.....	6
1.3. Ámbito de aplicación.....	6
1.4. Estructura del documento	7
Capítulo 2. Fundamentos.	8
2.1. Redes de distribución de agua.	8
2.1.1. Etapas de un proyecto de red de distribución de agua.....	9
2.1.2. Modelo matemático de una red de distribución de agua.....	10
2.2. Sectorización de las redes de distribución de agua.	11
2.3. Machine Learning para el análisis de datos.	12
2.3.1. El agrupamiento en clústeres.....	13
2.3.1.1. Algoritmos para el agrupamiento en clústeres.....	15
2.3.1.2. Etapas para el agrupamiento en clústeres.....	17
2.4. Teoría de grafos.....	19
2.4.1. Conceptos clave.....	19
2.4.2. Tipos fundamentales de grafos.....	19
2.4.3. Formas de representación de un grafo	20
2.4.4. Árboles de expansión.	21
2.4.5. Teorema de Kirchhoff.....	21
2.5. La energía en redes de distribución de agua.	22
Pérdidas operacionales.	23

Pérdidas estructurales o energía topográfica.....	24
Capítulo 3. Metodología para la generación de trazados energéticamente convenientes.	27
3.1. Introducción.....	27
3.2. Propuesta metodológica.....	28
3.2.1. Generación de clústeres o sectorización inicial.....	29
3.2.2. Generación de la topología base.....	32
3.2.3. Generación de los posibles trazados.....	34
3.2.4. Cálculo de redes.....	36
3.2.5. Análisis de resultados y selección de alternativas de trazados.....	40
Capítulo 4. Caso de estudio.	42
4.1. Presentación.....	42
4.2. Etapa 1. Generación de clústeres o sectorización inicial.....	44
Estandarización de las variables.	45
Análisis de componentes principales (PCA).	46
Generación de los clústeres.	47
Evaluación de la calidad de los clústeres.....	51
4.3. Etapa 2. Generación de la topología base.....	53
4.4. Etapa 3. Generación de los posibles trazados.	54
4.5. Etapa 4. Cálculo de redes.....	56
4.6. Etapa 5. Análisis de resultados y selección de alternativas de trazados.	58
4.7. Discusión y análisis de resultados.	66
Capítulo 5. Conclusiones y desarrollos futuros.	70
5.1. Conclusiones.	70
5.1. Desarrollos futuros.....	72
Capítulo 6. Referencias.....	73

ÍNDICE DE FIGURAS

Figura 1. Consumo de electricidad en el sector del agua por proceso, 2014-2040 (Electricity Consumption in the Water Sector by Process, 2014-2040 – Charts – Data & Statistics, s. f.) [7].....	3
Figura 2. Gráfico del dendrograma empleado en la clusterización jerárquica.....	14
Figura 3. Agrupamiento de datos basado en centroides.....	15
Figura 4. Representación gráfica de un grafo.....	19
Figura 5 Gráfica balance de energía para una fuente de energía (a)rígida y (b) variable.....	23
Figura 6. Subdivisión de la energía topográfica [23].....	25
Figura 7. Metodología propuesta para la generación de redes de redes principales de distribución de agua.	29
Figura 8. Diagrama de flujo para la etapa de clusterización.....	30
Figura 9. Generación de la topología base.....	33
Figura 10. Generación de los posibles trazados.....	34
Figura 11. Cálculo de redes ramificadas.	37
Figura 12. Análisis y selección de alternativas.....	41
Figura 13. Distribución espacial de los nudos de consumo, clasificados de acuerdo con la cota.	43
Figura 14. Distribución espacial de los nudos de consumo, clasificados de acuerdo con la demanda base.....	43
Figura 15. Ejes posibles para la instalación de tuberías en la localización de la red.	44
Figura 16. Diagrama de caja y bigotes para la característica elevación.....	45
Figura 17. Histograma para la para la característica elevación.	45
Figura 18. Diagrama de caja y bigotes para la característica elevación.....	45
Figura 19. Histograma para la para la característica demanda base.	45
Figura 20. Diagrama de Pareto para los componentes principales	46
Figura 21. Gráfico de dispersión de los datos para los componentes principales 1 y 2.	47
Figura 22. Dendrograma generado mediante el clusterización jerárquica.	48
Figura 23. Codo de Jambú.....	49
Figura 24. Gráfico del coeficiente de la silueta.....	49
Figura 25. Gráfico del coeficiente de Dunn.....	49
Figura 26. Gráfico del índice de Davies - Boulding.....	49
Figura 27. clústeres generados mediante el algoritmo de k-means.....	50
Figura 28. Centroides obtenidos mediante el algoritmo de K-means.	50
Figura 29. Gráfico de silueta para los clústeres generados mediante el algoritmo de K-means.....	51

Figura 30. Gráfico de silueta para los clústeres generados mediante el algoritmo jerárquico.....	52
Figura 31. Primera selección de los ejes principales para la topología base.....	53
Figura 32. Definición de los ejes principales para la topología base.	53
Figura 33. topología base.....	54
Figura 34. topología base en Epanet.	55
Figura 35. Ejemplo de trazado generado a partir de un árbol de expansión del grafo G.....	57
Figura 36. Ejemplo de trazado generado a partir de un árbol de expansión del grafo G.....	57
Figura 37. Ejemplo RPDA dimensionada y calculada con cabecera en el nudo 13...58	
Figura 38. Ejemplo RPDA dimensionada y calculada con cabecera en el nudo 10...58	
Figura 39. Histograma de frecuencias Erc.....	59
Figura 40. Diagrama de cajas y bigotes para la Erc.	59
Figura 41. Configuración 4453 de la RPDA con mínimo valor de Erc.....	60
Figura 42. Configuración 4471 de la RPDA con mínimo valor de Erc.	60
Figura 43. Variación de la Erc de acuerdo con la pendiente de diseño.	61
Figura 44. Diagrama de cajas y bigotes para la pendiente y la Erc.	61
Figura 45. Topología base con distribución de cotas en los nudos y longitud en los arcos.....	61
Figura 46. Histograma para el nudo crítico.	62
Figura 47. Diagrama de cajas y bigotes para el nudo crítico y la energía mínima requerida en cabecera.....	62
Figura 48. Diagrama de cajas y bigotes para el nudo de entrada y la energía mínima requerida en cabecera.....	62
Figura 49. Valores mínimos y mediana de ERC para los nudos de entrada a la red.62	
Figura 50. Diagrama de cajas y bigotes para nudo de entrada a la red y Erc, para las mejores soluciones.....	63
Figura 51. Diagrama de cajas y bigotes para nudo crítico de la red y Erc, para las mejores soluciones.....	63
Figura 52. Diagrama de cajas y bigotes para nudo de entrada a la red y $\Sigma\Delta H_i$, para el grupo de las mejores soluciones.	64
Figura 53. Diagrama de cajas y bigotes para nudo de entrada a la red y C_{Total} , para el grupo de las mejores soluciones.	64
Figura 54. Diagrama de nube de puntos para $\Sigma\Delta H_i$ y C_{Total} , para el grupo de soluciones con menor Erc.....	64
Figura 55. Frente Óptimo de Pareto para $\Sigma\Delta H_i$ y C_{Total} , para el grupo de soluciones con menor Erc.....	64
Figura 56. Trazado energéticamente conveniente 2633-C16-1.....	65
Figura 57. Trazado energéticamente conveniente 2609-C16-1.....	65
Figura 58. Trazado energéticamente conveniente 2645-C16-1.....	66
Figura 59. Trazado energéticamente conveniente 2615-C16-1.....	66

Figura 60. Delta Z ($Z_i - Z_{min}$) de los nudos de la RPDA.....	67
Figura 61. Demanda base en los nudos de consumo.....	67
Figura 62. Gráfico de frecuencia de barras apiladas para valores de E_{RC} y el nudo de entrada.....	67
Figura 63. Gráfico de frecuencia de barras apiladas para valores de E_{RC} y el nudo de entrada.....	67
Figura 64. Conjunto de mejores con mínimo valor de E_{RC} para cada $J_{diseño}$	68
Figura 65. Soluciones con mínimo valor de E_{RC} para cada $J_{diseño}$	68

ÍNDICE DE TABLAS

Tabla 1. Rango de intensidades energéticas para segmentos del ciclo de uso del agua[10].....	4
Tabla 2. Datos resumen de la red.....	42
Tabla 3. Rango de las variables que definen los nudos de consumo.....	46
Tabla 4. Varianza por componente principal.....	47
Tabla 5. Variables asociadas a cada sector.....	51
Tabla 6. resultados de la evaluación de la agrupación mediante clústeres.....	53
Tabla 7. Conjunto de elementos que componen el grafo base G.....	55
Tabla 8. Fragmento de los resultados que definen los árboles de expansión válidos.....	55
Tabla 9. Elementos que conforman la RPDA.....	56
Tabla 10. Pendiente óptima para la instalación.....	57
Tabla 11. Parámetros de diseño de la RPDA.....	58
Tabla 12. Rango en que oscilan las E_{RC} para las RPDA.....	59
Tabla 13. Resumen resultados para la configuración 1 y 2 con menor E_{RC}	60
Tabla 14. Resumen de resultados para los trazados de la RPDA energéticamente convenientes.....	65
Tabla 15. Resumen de resultados para la solución óptima con $J_{Diseño}$ 9 m.c.a/km....	68

Capítulo 1.

Introducción.

La vida y el agua tienen relación intrínseca y esencial, a pesar de que el agua es un recurso abundante en nuestro planeta no siempre está presente en la misma cantidad, ni con la misma calidad en todos los lugares. El hombre no es ajeno a esta realidad y nuestra historia fue, es y será en función de la disponibilidad del recurso y su aprovechamiento.

Desde la antigüedad, las poblaciones humanas se asentaron cerca a fuentes de agua con capacidad para satisfacer sus necesidades diarias de agua, no obstante, contar con agua en la cantidad y calidad requerida ha sido una preocupación dentro de las poblaciones humanas. Tener agua fresca constantemente, ha llevado a que el hombre usando su ingenio, desarrolle soluciones para el aprovisionamiento del agua, captándola en las fuentes y conduciéndola hasta donde se requiere. El crecimiento poblacional, el aumento en las necesidades de agua, y la diversidad de usos en que requerimos el agua nos ha llevado al desarrollo de obras de ingeniería de gran magnitud y complejidad relacionadas con el abastecimiento de agua.

Civilizaciones como la Romana hicieron alarde de grandes obras de ingeniería para la conducción del agua. Haciendo uso de la gravedad los ingenieros romanos, transportaron el agua mediante canales y conductos de diversos materiales en los que el agua fluía a una presión mayor a la atmosférica salvando valles y largas distancias, para llegar a las ciudades donde el agua era distribuida entre diferentes usos como termas, fuentes públicas e incluso a viviendas donde se disponía de grifos privados. Este modelo, en que el agua se conduce mediante una red de conductos interconectados a una presión mayor a la atmosférica y se distribuye a distintos puntos de forma simultánea es el antecesor de las redes de distribución de agua modernas.

Para el siglo XVII, sistemas como el sistema de distribución de agua de Londres o el sistema para el suministro de agua en los Jardines de Versalles, transportaban ya parcialmente el agua mediante una red de tuberías presurizadas, durante el siglo XVII y XVIII ciudades como Nueva York, São Paulo llevaron a cabo proyectos para la construcción de importantes obras de infraestructura para el abastecimiento de agua mediante sistemas presurizados. Con la revolución industrial, la construcción

de redes para la distribución de agua se potencio, la producción en masa de tuberías y el desarrollo de nuevas tecnologías para el bombeo y control del agua impulsaron la construcción de infraestructura en el mundo, las nuevas tuberías de hierro fundido no solo eran más resistentes, también facilitaban la construcción de la red, abaratando los costes asociados a la construcción lo que permitió que el agua llegara a más lugares.

Durante los XIX y XX creció el número y la extensión de las redes de distribución de agua en mundo, la llegada de nuevos materiales como el policloruro de vinilo PVC y el desarrollo tecnológico en la industria del agua facilitaron la creación, ampliación o mejora de las redes de distribución de agua. Con la digitalización de la industria del agua se incorporan nuevas herramientas para la mejora de los procesos de diseño, control y gestión de las redes de agua; la búsqueda de la eficiencia y la sostenibilidad es el nuevo objetivo de cara a afrontar desafíos como la escasez del recurso hídrico, el cambio climático y el crecimiento de la población.

Lograr este objetivo, requiere de estudios más minuciosos y profundos a los que tradicionalmente se realizan durante la proyección y operación de una red de distribución de agua, el avance en la ciencia y el desarrollo tecnológico han mejorado la comprensión de los fenómenos asociados con la distribución de agua a presión y las nuevas herramientas no solo han permitido realizar análisis más detallados, sino que también más complejos. En este contexto un aspecto que cobra especial relevancia es el energético, identificar oportunidades para mejorar la eficiencia desencadena en menores consumos de energía, con una consecuente reducción de los costes asociados durante la operación, lo que finalmente impacta en la sostenibilidad de los servicios de agua.

Por consiguiente, conocer energéticamente cómo se comportan las Redes de Distribución de Agua (RDA) es fundamental en la gestión eficiente de los recursos hídricos. Los estudios que involucran los aspectos hídricos y energéticos de la RDA proporcionan una comprensión holística de cómo se comporta la red, estudiando además el consumo eléctrico de los equipos de presión. Otros aspectos como la eficiencia hídrica de la red, la gestión de la demanda, su topología y su relación con el lugar donde se implanta, facilitan, entre otras, la toma de decisiones, propuestas de cambios en la operación y optimizar su rendimiento.

1.1.1. Problemática.

La población mundial ha pasado de alrededor de 6.000 millones de personas en el año 2000 a aproximadamente 7.700 millones en 2023, de acuerdo con las proyecciones de las Naciones Unidas este número podría alcanzar los 9.700 millones para el año 2050, llegando incluso a aproximadamente 11.000 millones hacia el año 2100 [1]. Este vertiginoso aumento demográfico y el paralelo desarrollo económico global traen consigo un aumento en las demandas de recursos esenciales como agua, energía y alimentos, estrechamente relacionados. Sin embargo, mantener este

crecimiento poblacional se enfrenta a una gran tensión ante la incertidumbre sobre la futura disponibilidad de estos recursos cruciales.

Actualmente el consumo mundial anual de agua se estima en alrededor de 4.600 km³ [2]. De esta cantidad, cerca del 10% es utilizado para usos domésticos, un 20% en la industria y la mayor porción, cerca del 70% se destina a la agricultura. Se anticipa que este último sector experimente un incremento adicional de aproximadamente el 30% para el año 2050[3]. Respectivamente, el consumo global de energía en el 2022 es cercano a 85,5 billones de kilovatios hora (kWh) por año, siendo la industria el mayor consumidor con cerca del 40%, mientras que el uso doméstico y el transporte compartieron un 25% respectivamente. Se espera un aumento adicional entre el 25% y el 35% para el año 2025 [4].

La interconexión existente entre el recurso agua y la energía, conocida como el nexo agua-energía, se fundamenta en la necesidad de agua en la producción de energía eléctrica, principalmente hidroeléctrica y en la necesidad de energía durante todo el ciclo integral del agua. A nivel mundial, aproximadamente el 10% del agua dulce extraída se destina a la producción de energía, una cifra que se eleva al 40% en naciones industrializadas. En un sentido inverso, el sector agua consume aproximadamente el 4% de la electricidad global, con cifras que se elevan al 9% en regiones como Oriente Medio[5]. La Agencia Internacional de Energía (IEA) pronostica un aumento del 80% en el consumo eléctrico mundial en el sector agua para 2040, impulsado por la creciente demanda de energía eléctrica requerida para el bombeo, tratamiento y transporte del agua. En este contexto, el sector se ve obligado a adaptarse a estas nuevas necesidades, que junto con la tendencia del alza en los precios de la energía, hacen imperativa la búsqueda de la eficiencia lo largo de todo el ciclo integral del agua [6].

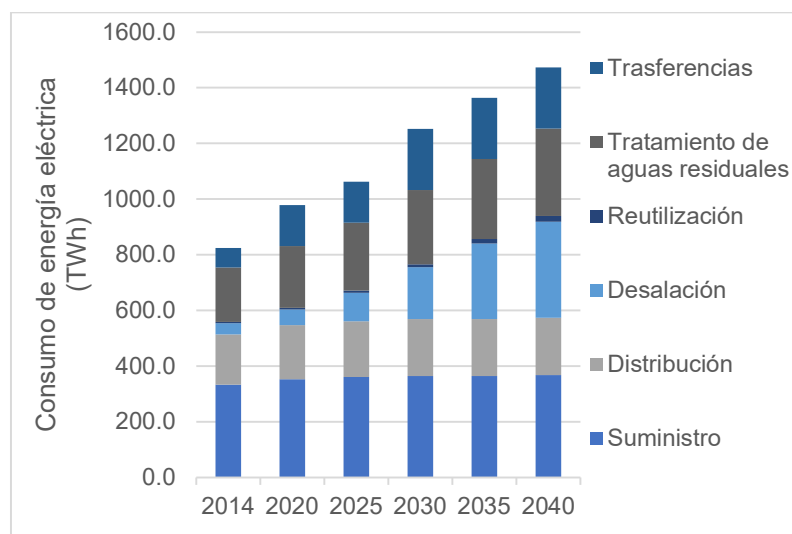


Figura 1. Consumo de electricidad en el sector del agua por proceso, 2014-2040 (Electricity Consumption in the Water Sector by Process, 2014-2040 – Charts – Data & Statistics, s. f.)[7]

En la actualidad, el sector agua consume cerca de 1000 TWh de electricidad a nivel global, principalmente en las etapas de bombeo (90%), tratamiento y transporte. Según el Programa Mundial de Evaluación de Recursos Hídricos (WWAP), el consumo eléctrico global en la fase de distribución de agua alcanza los 200 TWh anuales, lo que representa el 25% del consumo total del sector[8]. No obstante, la fase completa de distribución (suministro y transporte) presenta los valores de intensidad energética más altos (ver Tabla 1) y requiere aproximadamente la mitad del consumo total de energía. De acuerdo con las proyecciones de la Agencia Internacional de Energía se prevé un aumento de al menos 50% en el consumo eléctrico del sector hídrico en las dos próximas décadas y lograr el acceso al agua potable y saneamiento básico (ODS 6), incrementaría al menos un 1% el consumo global de energía[9] [6].

Tabla 1. Rango de intensidades energéticas para segmentos del ciclo de uso del agua[10]

Etapas	kWh/m³
Captación	0,29 - 1,22
Potabilización	0,03 - 4,23
Distribución	0,18 - 0,32
Drenaje y depuración	0,29 - 1,22
Vertido	0 - 0,11
Tratamientos avanzados	0,11 - 0,32

La mejora de la eficiencia energética durante la etapa de distribución se convierte en una necesidad de cara a conseguir servicios de agua más sostenibles, con menores consumos y que alivie la crisis energética. No obstante, este avance se encuentra condicionado por diversos factores que influyen en el gasto energético de las redes. Estos factores pueden variar en función de la estructura de la red, la infraestructura utilizada y las características específicas del entorno, entre los más relevantes están la eficiencia de los equipos, la gestión de las presiones, las pérdidas de agua, la topografía, el tamaño y longitud de la red, la demanda de agua, la eficiencia de la energía de bombeo, la gestión de la energía, el estado de los componentes y el uso de energías renovables. Una evaluación minuciosa de estas variables y la realización de un balance energético exhaustivo pueden catalizar mejoras sustanciales en esta área.

En conjunto, la combinación de estos factores puede afectar significativamente el consumo energético en una red de distribución de agua. La adopción de prácticas y tecnologías que optimicen la eficiencia de la red, una gestión eficaz, junto con diseño (topología y dimensionado) eficiente, son clave para reducir el consumo de energía y promover una operación sostenible.

En este sentido, la configuración y el trazado de la red emergen como aspectos sumamente influyentes en la eficiencia energética. Usualmente, la definición del

trazado en una Red de Distribución de Agua (RDA) queda en manos del proyectista, quien lo genera a partir de las restricciones normativas existentes y la configuración espacial del lugar de implantación, considerando el comportamiento hidráulico de la red y los costes de inversión para su definición. No obstante, aunque las soluciones sean técnicamente correctas, parten de la subjetividad y suele ignorar el análisis energético de la RDA y de sus costes asociados. Este desequilibrio resulta en redes con consumos energéticos más elevados, presiones excesivas y costos operativos elevados, llegando incluso a ser insostenibles con el tiempo.

En resumen, la consideración de los aspectos energéticos tanto en la planificación de nuevas redes como en la evaluación de alternativas de optimización de redes existentes es fundamental, especialmente en el contexto de futuras crisis y precios de energía elevados. Sin embargo, se debe abordar el desafío del costo computacional implicado en la búsqueda y cálculo de soluciones óptimas, debido a la multidimensionalidad y correlación existente entre las variables relacionadas en el problema, y la magnitud y la complejidad de los sistemas.

1.2. Motivación y objetivos.

Este trabajo tiene como principal motivación desvelar parte de la complejidad que existe a la hora de definir el trazado (layout) que energéticamente resulte más conveniente para una Red de Distribución de Agua (RDA). Definir el recorrido de las tuberías en una red es una actividad que, a pesar de su aparente sencillez, suele estar particularmente influenciada por la subjetividad del proyectista, la imposición de restricciones ajenas a los criterios técnicos y por el contexto espaciotemporal en el que se quiere realizar la implantación de la red.

Otro aspecto que se pasa por alto es la evaluación de un número significativo de soluciones para la elección del mejor trazado, al ser un ejercicio dispendioso en tiempo y recursos, no es común que se realice, o de realizarse se hace con pocas soluciones, enfocándose en el estudio económico de la inversión, aun cuando es conocida la conveniencia de soluciones con menor consumo energético, más eficientes y sostenibles.

1.2.1. Objetivo principal.

Desde lo mencionado, este trabajo tiene como objetivo principal desarrollar una metodología integral para la generación de trazados para Redes Principales de Distribución de Agua (RPDA) que resulten energéticamente convenientes, requiriendo de una menor energía en cabecera, acoplándose mejor a las condiciones del terreno donde se implanta para tener menores excesos de energía y con un menor coste total. La RPDA se refiere a la porción de la RDA que desde la cabecera transporta el agua hasta cada uno de los sectores de forma exclusiva.

Las redes que se obtengan a partir de estos deberán hacer un uso más eficiente de los recursos. Una menor energía en la red, con presiones más cercanas a la de

servicio en los nudos de consumo, se relaciona con un consumo de agua más controlado, es decir, equivale a un menor volumen de agua fugada. Una reducción en la energía requerida en la red se traducirá en menores costes durante la operación, consiguiendo sistemas más sostenibles.

La reducción del consumo energético de la RDA desde la fase de proyección no solo impulsa su eficiencia, sino que también fortalece la sostenibilidad del servicio y atenúa su potencial impacto ambiental, disminuyendo su huella de carbono. Esta estrategia se alinea con un enfoque más amplio, la consecución de los Objetivos de Desarrollo Sostenible (ODS), específicamente los Objetivos 7 (energía asequible y no contaminante), 6 (agua limpia y saneamiento) y 13 (acción por el clima).

1.2.1.1. Objetivos específicos

Para alcanzar el objetivo general propuesto, se prevé aplicar la metodología a un caso de estudio que permita el análisis de la propuesta metodológica con los siguientes objetivos específicos:

1. Estudiar la influencia de la configuración de la red en la eficiencia energética.
2. Utilizar técnicas para el análisis de datos y algoritmos para la generación de trazados de redes de distribución de agua.
3. Aplicar la técnica de clusterización para la sectorización de RDA.
4. Analizar la eficiencia energética a partir del balance de energía en la RDA.
5. Análisis de sensibilidad de la pendiente y el nudo de entrada en la definición de un trazado energéticamente conveniente.
6. Efectuar el cálculo del coste total (coste capital y coste energía) y su relación con la definición con RDA.

1.3. Ámbito de aplicación

La metodología que se propone se aplica a redes de distribución de agua (RDA), o WDN por sus siglas en inglés (Water Distribution Network), que transporten agua mediante conductos a una presión mayor a la atmosférica, y orientada a sistemas que requieran equipos de presión. En particular, el estudio de nuevas redes o la mejora de redes existentes donde se conozcan las demandas de agua y su localización, y donde además sea posible evaluar distintas opciones de trazado o alternativas para mejorar el sistema para el suministro de agua potable y otros usos.

La metodología propuesta es una combinación de herramientas recursivas, al integrar el análisis de datos para la generación mediante clusterización de sectores y la teoría de grafos para la generación de las alternativas de trazados de redes principales de distribución de agua (RPDA) ramificadas; por lo que resulta novedosa, si se considera que su objetivo es la definición del trazado que se considere como conveniente energéticamente y no el estudio para la optimización de una red existente.

1.4. Estructura del documento

El presente documento se ha estructurado en 6 capítulos. El capítulo 1 es una introducción donde se presenta de forma general la problemática, la motivación, objetivos y el ámbito de aplicación para la metodología.

El segundo capítulo, contiene las temáticas y conceptos que son base de la metodología, y que permiten entender cada una de las etapas y procesos que se realizan. Están presentados de forma concisa, yendo de lo general a lo específico. En orden son: redes de distribución de agua, sectorización de redes de agua, machine learning, teoría de grafos y la energía en las redes de distribución de agua.

En el tercer capítulo se desarrolla la metodología para la generación de trazados para redes de agua ramificadas considerando criterios energéticos, mediante técnicas de grafos y clustering. Está compuesta por 5 etapas, clusterización, generación de la topología base, generación de los posibles trazados, cálculo de redes y el análisis de resultados; cada una esquematizada mediante diagramas de flujo que muestran la secuencia del proceso y los elementos de entrada y salida.

El cuarto capítulo, presenta la aplicación de la metodología en un caso de estudio, una fracción del área que cubre una comunidad de regantes, donde se identificó la RPDA más conveniente, a partir de la distribución de los nudos de consumo y los ejes existentes en el lugar de la implantación donde es posible construir la red de tuberías.

En el capítulo 5 se describen las conclusiones y futuras líneas de investigación y finalmente, el capítulo 6 contiene las referencias bibliográficas empleadas en la elaboración de este documento.

Capítulo 2.

Fundamentos.

Este capítulo recoge los conceptos fundamentales y necesarios adoptados en la propuesta metodológica para la generación de trazados para redes de agua ramificadas considerando criterios energéticos mediante técnicas de grafos y clusterización.

2.1. Redes de distribución de agua.

La RDA hace parte de la infraestructura necesaria para el abastecimiento de agua. De acuerdo con Cabrera et al. (2009) [11], se denomina red de distribución al conjunto de elementos encargados del transporte de agua desde los puntos de producción y almacenamiento hasta los puntos de consumo, que bien pueden ser viviendas, comercios, industrias, hidrantes de riego, o de incendio, etc.

La RDA se encuentra conformada por los conductos, accesorios y elementos especiales (válvulas, accesorios de medición, elementos de unión, etc.) con características específicas y unidos de una determinada forma, que se dimensionan para poder suministrar los caudales demandados, cumpliendo con las condiciones de servicio exigidas. Otros elementos significativos que hacen parte de la red de distribución y que condicionan su funcionamiento son los depósitos, las estaciones de bombeo y mecanismos para la regulación; la sinergia entre todos los elementos hace posible el servicio de agua, pero hacen complejo el cálculo de la red.

Las RDA se pueden clasificar de acuerdo con diversos criterios, que van desde la forma como funcionan, el tipo de servicio que prestan, la topología del sistema, el agua que transportan, etc. A continuación, se presentan algunas tipologías de redes.

Funcionamiento

- Por gravedad
- Aumento de la presión natural

Importancia

- Red principal
- Red menor de distribución

Topología de la red

- Redes ramificadas
- Redes malladas
- Redes mixtas

Servicio que prestan

- Redes de agua urbanas
- Redes de riego
- Red de extinción de incendios

El tipo de agua que trasportan

- Redes de agua potable
- Redes de agua cruda
- Redes de aguas grises

2.1.1. Etapas de un proyecto de red de distribución de agua.

El proyecto de RDA pasa por varias etapas, es un proceso en el que, definidas las necesidades de los usuarios se estudia la solución que permita satisfacerlas y brindar de forma aceptable el servicio de agua. Este proceso no es inmediato y pasa por etapas que suelen ser iterativas.

A continuación, se presenta de forma general las etapas en un proyecto de RDA.

- **Identificación de las necesidades:** Es el punto de partida de cualquier proyecto, en esta etapa se identifican y estudian todas aquellas necesidades, que hacen necesaria a la RDA.
- **Estudio de la demanda:** Se realiza un análisis de cuáles serían las demandas de agua esperadas en los puntos de consumo, de acuerdo con el uso, la cantidad de usuarios, la magnitud del proyecto, sus necesidades y cualquier otro factor que pueda interferir en la definición de la cantidad de agua que se requiere de forma inmediata, como a futuro.
- **Definición de la fuente de suministro:** Se determina la localización y capacidad de suministro de agua de la fuente o fuentes de agua de las cuales se aprovisionará la red, dependiendo del tipo de instalación podría tratarse de depósitos, embalses o algún tipo de almacenamiento con la capacidad suficiente, o directamente fuentes de agua superficiales o profundas, que podrían o no requerir de un tratamiento.
- **Definición del trazado de la red:** Esta etapa define el recorrido de los conductos, el tipo de red, la localización de los diferentes componentes que definen la red, la conectividad entre los elementos y como se pueden llegar a suplir los diferentes puntos de consumo.

- Dimensionado de los elementos que conforman la red: De acuerdo con las demandas de agua, los requisitos de servicio y la condición topográfica, se realiza el cálculo de los diferentes elementos que componen la red de suministro, de acuerdo con el trazado definido. El proceso es iterativo y pasa por la definición de materiales y parámetros de diseño que tendrá la red.
- Cálculo hidráulico de la red: En esta etapa se realiza el análisis hidráulico de la red, verificando que se cumplan en todo momento las condiciones hidráulicas para la prestación del servicio de agua.
- Evaluación de costes: Se realiza el estudio del coste de la construcción de la red, incluyendo tanto material, mano de obra y cualquier otro coste que pudiera incrementar el precio final de la instalación. También, resulta conveniente estudiar los costes asociados a la operación de la red a lo largo de su vida útil del proyecto de red.
- Puesta en marcha: comprende todas las actividades necesarias para dar inicio al servicio de la red, tales como la puesta en marcha de los equipos, pruebas, conexiones a otros servicios públicos, entre otros.
- Operación: Es el fin del proyecto de RDA, comprende todas las actividades propias de la explotación, incluido el mantenimiento de esta.

2.1.2. Modelo matemático de una red de distribución de agua.

Un modelo hidráulico de una RDA es una representación matemática simplificada de una red real, el cual generalmente se basa en líneas y nudos que representan la conectividad entre los diferentes elementos que componen la red física, como las salidas y entradas de caudal, derivaciones, tuberías, válvulas, bombas, depósitos, etc. Además, incorporan otros componentes que a pesar de no ser físicos condicionan el funcionamiento de la red como los patrones de consumo, curvas de comportamiento y controles de la red.

Son herramientas básicas que permiten simular el posible comportamiento que tendría la red en términos de caudal y presión, mediante una serie de ecuaciones de la mecánica de fluidos que relacionan la topología de la red, los parámetros de cálculo hidráulico y las características de los elementos.

La definición de un modelo parte de la topología de la red, conocida la localización de los elementos y su conectividad, es posible asignar a cada uno las características que lo distinguen y definir los parámetros del modelo en general para realizar el cálculo de las variables de salida mediante programas de análisis y simulación como EPANET. Dependiendo del tipo de análisis que se intente realizar el número de variables de entrada y salida es distintito.

A continuación, se nombran algunos tipos de modelos hidráulicos de redes.

- Modelos estáticos.

- Modelos cuasi-estático.
- Modelos en régimen transitorio.
- Modelos de calidad del agua.

Uno de los programas más populares para modelación de RDA es EPANET. Este software lleva a cabo simulaciones estáticas y cuasi-estáticas del comportamiento hidráulico y de la evolución de la calidad del agua a lo largo del tiempo en sistemas de distribución de agua a presión.

Los modelos matemáticos tienen múltiples aplicaciones y son útiles en cualquier etapa de un proyecto de red, facilitan el diseño de los diferentes elementos que la componen o el estudio de modificaciones a la red. En fase de operación permiten estudiar los problemas que se pueden presentar, escenarios de funcionamiento y planificar cambios y la expansión de la infraestructura.

Otra aplicación es el estudio de la eficiencia hídrica y energética de la red, analizar cómo es la distribución del agua que se inyecta en la red y el consumo energético del sistema es necesario para emprender acciones en la red, tendientes a mejorar el rendimiento de la red, bien sea reduciendo el volumen de agua pérdida en fugas o generando cambios a en el funcionamiento de la red que disminuyan los excesos de energía o el consumo energético.

2.2. Sectorización de las redes de distribución de agua.

El servicio que presta una RDA es el de transportar el agua desde la fuente hasta cada uno de los puntos de consumo, permitiendo el aprovisionamiento del recurso a un grupo de usuarios y usos, inscritos en un determinado espacio geográfico de forma que se garantice la asequibilidad, calidad, sostenibilidad y demás condiciones mínimas de servicio, con el menor de los costes posibles. La naturaleza del servicio y las particularidades de cada sistema de abastecimiento inducen a que, dentro de cada red, se agrupen los elementos con características similares y se creen grupos que faciliten el control de la operación, la gestión del sistema, el análisis económico de la operación, etc.

Una de las estrategias más útiles en la gestión de los sistemas de distribución de agua, es la división de la red en sectores o zonas de menor tamaño o DMA por sus siglas en inglés (District Meter Area) hidráulicamente independientes. El reparto de los nudos y otros elementos que conforman la red, cercanos y que están conectados, entre un número definido de grupos, facilita a los operadores implementar acciones para el control activo de fugas, la gestión de la presión y el control de la calidad del agua en la red, entre otras acciones.

Cada sector se suele dotar de elementos de medición y control, que permiten conocer y modificar el comportamiento de variables hidráulicas como la presión, el

caudal y la calidad del agua. Asimismo, la sectorización es también útil para realizar el balance hídrico y energético de la red, ya que simplifica el seguimiento y control de las entradas y salidas del sistema [12] [13], lo que es útil en redes con una dimensión importante, mejorando la eficiencia física y comercial del servicio de agua potable, así como reduciendo las pérdidas de agua y energía.

De acuerdo con Cabrera et al. (2019), la sectorización podría clasificarse en tres categorías de acuerdo el objetivo es la gestión: cuando el objetivo es el control de la presión, se define como áreas de gestión de presión o PMA por si siglas en inglés (Pressure Management Areas); cuando la sectorización pretende tener controlado el caudal, se conoce como áreas con medición de distrito o DMA (District Metered Areas); mientras que cuando el objetivo es la energía se conoce como áreas de Gestión Energética o EMA (Energy Management Areas) [14] [15].

2.3. Machine Learning para el análisis de datos.

La parte de la ciencia que se encarga de que máquinas y computadoras imiten la mente humana en la resolución de problemas se conoce como inteligencia artificial (IA), dentro de esta ciencia el Machine Learning o aprendizaje autónomo agrupa el conocimiento, técnicas y algoritmos que, mediante el uso de datos, pretenden imitar la forma en la que los seres humanos aprenden, de forma que a partir de datos y experiencias previas gradualmente mejoren en precisión.

El Machine Learning es un componente esencial en la ciencia de datos, ya que emplea métodos estadísticos para entrenar a los algoritmos en la realización de clasificaciones o predicciones, además de descubrir información relevante dentro del conjunto de datos. Esta información, a su vez, es utilizada en la toma de decisiones en proyectos, aplicaciones o empresas, teniendo un impacto directo en las métricas clave de crecimiento.

Los modelos de Machine Learning se dividen en tres categorías principales de acuerdo con el tipo de entrenamiento y el tipo de problema a resolver:

- **Machine Learning supervisado:** Emplea conjuntos de datos etiquetados para entrenar a los algoritmos en el análisis y agrupamiento en clústeres de conjuntos de datos sin etiquetar. Es decir que mediante un conjunto de entrada junto con las salidas deseadas correspondientes (etiquetas) el modelo aprende. Algunos de los métodos empleados son las redes neuronales, la regresión lineal y la regresión logística [16].
- **Machine Learning no supervisado:** Son algoritmos capaces de analizar y agrupar en clústeres conjuntos de datos sin etiquetar, descubriendo patrones y agrupaciones de datos ocultos sin requerir de la intervención humana. Son empleados en tres tareas principales el agrupamiento, la asociación y la

reducción de dimensionalidad, con capacidad de descubrir similitudes y diferencias en la información, algunos de los más conocidos son el análisis de componentes principales (PCA), las redes neuronales, la agrupación en clúster mediante K-means y agrupamiento jerárquico [16].

- **Machine Learning semi supervisado:** Es un punto intermedio entre el Machine Learning supervisado y no supervisado, empleando un conjunto de datos etiquetados pequeño para guiar la clasificación y la extracción de características de un conjunto de datos sin etiquetar de mayor tamaño.

Otra categoría que suele incluirse es el aprendizaje por reforzamiento, en la que, mediante un proceso de prueba y error, donde se recompensan los aciertos y se penaliza los errores el sistema aprende. Es un enfoque basado en la experiencia en el que el objetivo es maximizar la recompensa a lo largo del tiempo.

El machine learning es una herramienta poderosa, es empleado en una gran variedad de problemas como la clasificación, regresión, agrupamiento de datos, la detección de anomalías, entre otros; lo que lo hace aplicable en distintos campos. En RDA es empleado en análisis para la detección de fugas, la predicción de la demanda, la optimización de presiones, el análisis de consumos, la sectorización de la red, entre otras aplicaciones.

2.3.1. El agrupamiento en clústeres.

El agrupamiento en clústeres, clusterización o clustering, también conocido como análisis de agrupaciones, es una técnica del aprendizaje no supervisado, que en función de las similitudes o diferencias agrupa a datos no etiquetados, de forma que se puedan identificar patrones o estructuras. Los datos contenidos en un grupo o clúster guardan similitud entre ellos y características suficientes para diferenciarse de los elementos que componen otros grupos, de forma que cada dato se encuentre mejor representado en su grupo que en otros.

La clusterización se emplea frecuentemente para identificar patrones, características distintivas o segmentos en conjuntos de datos voluminosos, sin necesidad de contar con información previa sobre categorías o etiquetas. Esta técnica permite descubrir estructuras subyacentes y relaciones entre los datos, facilitando así la exploración y comprensión de conjuntos de datos complejos.

Los métodos de clusterización pueden clasificarse en función de diferentes criterios, como la naturaleza del proceso, el enfoque algorítmico y el tipo de datos que pueden manejar, en general se podrían subdividir en dos categorías principales, los métodos jerarquizados y no jerarquizados.

Los métodos jerárquicos: Se basa en el agrupamiento de los datos en una estructura jerárquica de árbol o dendrograma, utilizando las distancias entre observaciones para construir un árbol. La estrategia aglomerativa se inicia con la formación de clústeres como los propios elementos aislados, iterativamente se fusionan entre los

más cercanos conformando nuevos clústeres, hasta conformar un clúster que agrupa todos los elementos del conjunto de datos, la estrategia divisiva parte del conjunto completo separando los grupos hasta llegar a los elementos individuales.

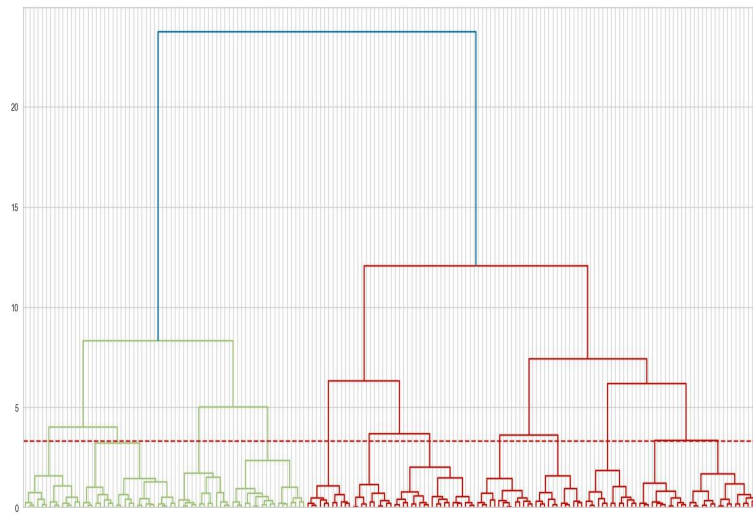


Figura 2. Gráfico del dendrograma empleado en la clusterización jerárquica.

Existen diversos métodos que se pueden emplear para el agrupamiento de los datos. El método por emplear depende del problema que se intenta resolver, algunos de los métodos de agrupamiento usados:

- Método vecino más próximo (enlace simple).
- Método vecino más lejano (enlace completo).
- Método del centroide.
- Método de vinculación promedio (intergrupo).
- Método de varianza mínima (Ward).
- Método de la mediana.

Los métodos no jerárquicos: Se basan en establecimiento de los clústeres en función de un número de grupos previamente dado, dividiendo directamente el conjunto de datos de forma que los clústeres formados que no se encuentren anidados y sean mutuamente excluyentes. Estos algoritmos agrupan los datos mediante iteraciones, asignando cada elemento al clúster más cercano o determinando centros de clústeres representativos (centroides) en el espacio de características, actualizando iterativamente los centroides para mejorar la agrupación.

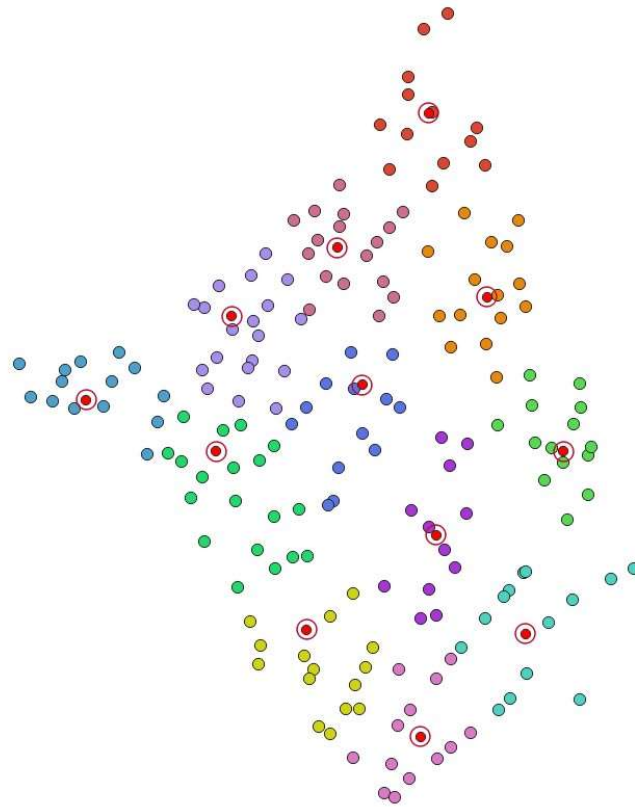


Figura 3. Agrupamiento de datos basado en centroides.

Estos métodos se pueden subdividir en cuatro categorías.

- Métodos de reasignaci3n
- Métodos de b3squeda de densidad
- Métodos directos
- Métodos reductivos

2.3.1.1. Algoritmos para el agrupamiento en cl3steres.

Dependiendo del tipo de problema que se intente resolver, el objetivo que se persigue y la naturaleza, forma y cantidad de los datos, se selecciona el tipo de algoritmo a utilizar. Existen diversas alternativas de algoritmos y variantes de estos, desarrollados bajo distintos enfoques y aplicables distintos problemas.

A continuaci3n, se listan algunos algoritmos y se profundiza en los algoritmos de algoritmo de clusterizaci3n jerárquica y clusterizaci3n K-means empleados en la metodología.

- K-means.
- Agrupamiento jerárquico (Hierarchical clustering).
- DBSCAN (Density-Based Spatial Clustering of Applications with Noise).
- OPTICS (Ordering Points to Identify the Clustering Structure).
- CLARA (Clustering Large Applications).
- Cambio medio (Mean Shift)

- Birch (Balanced Iterative Reducing and Clustering using Hierarchies).
- Gaussian Mixture Model (GMM).
- Algoritmo de Clustering Espectral (Spectral Clustering).
- Algoritmo de Propagación de Afinidad (Affinity Propagation).

Algoritmo de clusterización jerárquica.

Este algoritmo genera grupos en cada una de las fases de acuerdo con la similitud entre los clústeres, construyendo una estructura jerárquica en forma de árbol, a partir de las distancias entre datos[17].

Una de las mayores ventajas de este algoritmo es su capacidad para definir un número de clústeres por cada fase, lo que permite estudiar varias alternativas de agrupamiento óptimas y se suele utilizar de forma exploratoria previo al uso de otras técnicas de clusterización.

1. Se inicia definiendo tantos clústeres como datos existan, es decir que cada punto se considera como un clúster individual.
2. Se calcula la similitud entre los clústeres, generalmente es la medida de la distancia euclidiana entre cada par de puntos.
3. La agrupación de los datos se realiza de acuerdo con la similitud entre los puntos, agrupando los más cercanos y formando un nuevo clúster. Este agrupamiento puede ser entre clústeres individuales o existentes, lo que genera una jerarquía de clústeres.
4. Creados los nuevos clústeres se actualiza la matriz de similitud, que refleja la similitud entre el nuevo clúster y los demás puntos o clústeres.
5. Se repiten los pasos 3 y 4 hasta obtener un único clúster que contenga a todos los datos. El resultado final es el dendrograma que representa gráficamente de la estructura jerárquica de clústeres.
6. La definición del número de clústeres se realiza de acuerdo con el dendrograma, en el cual la distancia vertical entre los puntos o clústeres representa la similitud entre ellos.

La selección del número de clústeres se suele hacer al trazar una línea horizontal que atraviese las uniones entre los clústeres, el número estará en función de cuántas veces la línea horizontal corta las ramas del dendrograma.

Algoritmo clusterización K-means.

Este algoritmo utiliza una medida de la distancia entre los puntos y los centroides para agrupar los datos en clústeres. Cada clúster es representado mediante el centroide, que es el punto promedio de todos los puntos asignados a ese clúster[18]. El algoritmo es un proceso iterativo que busca minimizar la suma de los cuadrados de las distancias (SSE) dentro de los clústeres, con cada iteración, los centroides se actualizan recalculando sus posiciones en función de los puntos asignados al clúster. El proceso continúa hasta que los centroides ya no cambian significativamente o se alcanza un número máximo de iteraciones.

7. Se inicializa el número de K agrupaciones en los que se segmentará el conjunto de datos
8. Aleatoriamente se definen los K puntos como centroides iniciales de acuerdo con el número de clústeres que se definió.
9. La asignación de los datos se realiza de acuerdo con la similitud entre el centroide y los puntos (representación de los datos), generalmente la distancia euclidiana entre el dato y el centroide más cercano, aunque podría aplicarse con otras medidas de distancia.
10. Asignados todos los puntos a un clúster, se recalculan los centroides (centro de gravedad) como el promedio de las características (variables o dimensiones que describen cada punto) de los puntos asignados al clúster, desplazando el centroide al centro del clúster.
11. De acuerdo con la nueva posición de los centroides, es posible que se deban reasignar los puntos a un nuevo centroide, repitiendo el paso 2 y 3 hasta que no existan resignaciones o se alcance el número máximo de iteraciones.
12. El algoritmo converge cuando los centroides no cambian significativamente, quedando los datos asignados a un clúster que se representa mediante un centroide.

2.3.1.2. Etapas para el agrupamiento en clústeres.

A continuación, se presenta un resumen del paso a paso que se realiza para el agrupamiento de los datos en clústeres.

1. Pretratamiento de los datos: Es la revisión de la calidad de los datos y comprende el completar datos faltantes, la codificación de variables categóricas, revisión de outliers y si es necesario su eliminación, ya que podrían afectar el agrupamiento de los datos en clústeres.
2. Estandarización de las variables: Este proceso es necesario para que las variables sean comparables, ya que si existen variables con valores significativamente más altos estos prevalecerían en el estudio y perderían relevancia respecto variables con valores menos significativos. El proceso generalmente consiste en transformar las variables o características para que tengan una media igual 0 y una desviación estándar de 1.
3. Reducción de la dimensionalidad: Si los datos presentan una alta dimensionalidad (Contienen un número alto de variables) se dificulta su visualización y análisis, por lo que es conveniente reducir el número de dimensiones. Lo que mejora la eficiencia de los algoritmos de clusterización y ayuda a eliminar el ruido o la redundancia en los datos.

Uno de los métodos más utilizados para la reducción de la dimensionalidad es el Análisis de Componentes Principales (PCA), el cual busca encontrar un conjunto de nuevas variables que representen la mayor parte de la varianza de los datos originales.

4. Selección del algoritmo de clusterización: La selección del tipo de algoritmo dependerá del tipo de problema que se intente resolver y la naturaleza de los datos.
5. Aplicación y selección del número de clústeres: Se aplica el algoritmo para el agrupamiento de los datos en clústeres, dependiendo del método empleado podría ser necesario un proceso iterativo para encontrar el número óptimo de agrupaciones.
Existen diferentes formas para definir el número óptimo de clústeres de acuerdo con el método de clusterización utilizado, como método del codo o el índice de silueta.
6. Cálculo de los clústeres mediante diferentes técnicas: Se repite el cálculo de las agrupaciones con uno o más métodos que sean aplicables
7. Evaluación y validación de los resultados: Se debe evaluar la calidad de los clústeres obtenidos y su significado, además es conveniente comparar con otros métodos de clusterización.

De acuerdo con Rocío Chávez (2020), el número de clústeres óptimos se puede estimar de acuerdo con las siguientes técnicas [19].

- Codo de Jambú: Es una representación de la inercia intraclases (WSS) que es una medida de la similitud entre los puntos de acuerdo con el número de clústeres.
A medida que disminuye el valor WSS aumenta el número de clústeres, lo que es conveniente. El número de clústeres se suele definir como el punto en que el valor del WSS disminuye de forma menos significativa.
- Coeficiente de silueta: Representa el promedio de los coeficientes de silueta, mide la similitud que existe entre un individuo y los demás individuos agrupados en el clúster, comparado con los individuos ubicados en el clúster más cercano. Es deseable que el coeficiente de silueta sea cercano a 1, ya que indican que los clústeres se encuentran bien separados y son más homogéneos.
- Índice de Dunn: Su objetivo es identificar clústeres densos y bien separados, clústeres bien definidos tendrán un diámetro pequeño y estarán bien separados por lo que un mayor valor del índice representa una mejor partición.
- Davies-Bouldin index: Esta basado en la distancia aproximada entre los clústeres y su dispersión, su objetivo también es identificar clústeres densos y bien separados. A menor valor del índice es mejor la agrupación de los clústeres.

2.4. Teoría de grafos

Forma parte de las matemáticas discretas y se encarga del estudio de los grafos, que son estructuras matemáticas que representan de forma abstracta las relaciones existentes en un conjunto de objetos. Se componen de dos elementos fundamentales, los vértices o nudos que representan los objetos y las aristas o arcos (enlaces) que representa las relaciones entre los vértices.

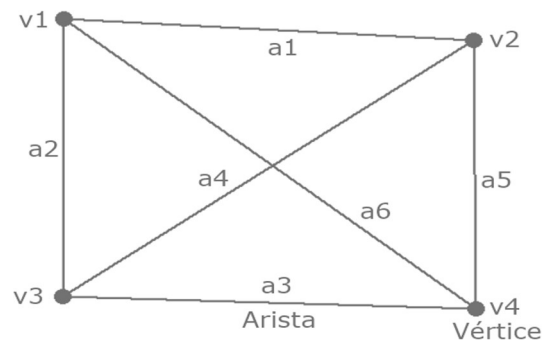


Figura 4. Representación gráfica de un grafo.

La Figura 4, representa gráficamente un grafo donde $V = \{v1, v2, v3, v4\}$ es el conjunto de vértices y $E = \{(v1, v2), (v1, v3), (v1, v4), (v2, v4), (v2, v3), (v3, v4)\}$ es el conjunto de pares no ordenados de vértices (aristas).

Los grafos se utilizan para modelar y analizar una gran variedad de fenómenos en distintas áreas del conocimiento, como la informática, biología, física, economía y la ingeniería, resolviendo problemas que involucran redes, caminos, flujos, conexiones, agrupaciones como por ejemplo la modelización de relaciones sociales, la interacción entre genes, la programación de tareas, la optimización de redes en ingeniería y la resolución de problemas de enrutamiento, entre otros.

2.4.1. Conceptos clave

Grado: corresponde al número de aristas que están conectadas a un vértice o aristas incidentes al vértice.

Camino: es una sucesión de vértices distintos conectados por aristas, de manera que el vértice final de una arista sirve de inicial a la siguiente arista, sin que se repitan vértices ni aristas en la secuencia.

Ciclo: Es un camino cerrado, es decir que inicia y acaba en el mismo vértice sin repetir ningún otro vértice ni arista.

2.4.2. Tipos fundamentales de grafos

Existen diversos criterios empleados en la clasificación de los grafos como la direccionalidad o el peso de las aristas, la existencia de ciclos o la conexión.

- **Direccionalidad:** los grafos se pueden clasificar como dirigidos o no dirigidos dependiendo si las aristas tienen o no una dirección específica.

- **Peso:** los grafos en los que las aristas tienen asociado algún tipo de valor se denominan como ponderados, si no tienen valores asociados serán no ponderados.
- **Conexión:** si existe al menos un camino entre cada par de vértices en el grafo este se clasifica el grafo como conexo, si existen componentes aislados será un grafo des conexo.
- **Ciclicidad:** los grafos en los que es posible trazar un ciclo cerrado de aristas que comienza y termina en el mismo vértice se denomina como cíclico, en el caso de no contener ciclos será un grafo acíclico.

Otras clasificaciones para los grafos son:

- **Regular:** un grafo se considera como regular cuando todos los vértices tienen el mismo grado.
- **Simple:** es aquel en el que no hay dos vértices que estén conectados solo por una arista, y no hay ningún vértice que esté conectado a sí mismo.
- **Subgrafos:** es un grafo que se compone de un subconjunto de los vértices y arista de un grafo mayor.
- **Planar:** es un grafo que es posible de dibujar en un plano, sin que exista cruces de aristas.
- **Bipartido:** un grafo es bipartido cuando los vértices se pueden dividir en dos conjuntos, en donde las aristas conectan vértices de conjuntos diferentes y no hay aristas que conecten vértices dentro del mismo conjunto.
- **Completo:** si cada par de vértices está conectado por una arista se clasifica como completo, es decir que cada vértice está conectado a los demás.

2.4.3. Formas de representación de un grafo

Un grafo G se define como el par ordenado $G = (V, E)$, donde V es el conjunto de vértices y E es el conjunto de aristas, cada arista es un par de vértices (v_1, v_2) donde v_1 y v_2 son vértices incluidos en el conjunto V . Gráficamente un grafo se puede representar como puntos (vértices) y líneas (aristas) en cualquier plano.

Aparte de su forma gráfica los grafos se pueden representar de forma matemática, a continuación, se enuncia las principales formas en las que se pueden representar:

Matriz de adyacencias: La matriz $A(G)$ es una matriz cuadrada con tamaño igual al número de vértices $n = |V|$, que representa la relación que existe entre vértices. El valor en la celda (i, j) es 1 si existe una arista entre los vértices i y j y 0 si no están conectados.

$$A = (a_{ij}), \text{ donde } a_{ij} = \begin{cases} 1, & \text{si } (v_i, v_j) \in E \\ 0, & \text{si } (v_i, v_j) \notin E \end{cases}$$

Lista de adyacencia: es una lista con tamaño igual al número de vértices, en el que para cada vértice se listan los vértices que están conectados a él por una arista.

Lista de aristas: es una lista de pares de vértices (arista) que conforman un grafo.

En adición existen otras formas de representar los grafos como la función de adyacencias, secuencias de grados y la matriz de incidencias.

2.4.4. Árboles de expansión.

Los árboles de expansión, generadores o recubridores representan un conjunto de grafos que se forman a partir de los vértices y aristas de un grafo $G = (V, E)$. Un árbol de expansión es un subgrafo $T = (V, E')$, que contiene la totalidad de los vértices de G y un subconjunto de sus aristas de forma que sea conexo[20].

Para que un subgrafo se considere como un árbol de expansión debe cumplir las siguientes condiciones:

- T es un árbol de expansión si es conexo y no contiene ciclos.
- T es un árbol de expansión si contiene todos los vértices de G , $V(T) = V(G)$.
- T es un subgrafo de G , o sea $E'(T) \subseteq E(G)$, lo que equivale a decir que conjunto de aristas de T es un subconjunto de las aristas de G .

Algunas de las propiedades de un árbol de expansión son:

- No es único, pueden existir distintos árboles de expansión en un grafo.
- Agregar cualquier arista de G que no está en T , crea un ciclo.
- Remover cualquier arista de T , lo desconecta.

Los árboles de expansión tienen aplicabilidad en varias áreas que involucran la optimización de redes y la distribución de recursos, para resolver problemas que impliquen conectar de forma eficiente y optimizada un conjunto de nudos para reducir la distancia o los costes.

2.4.5. Teorema de Kirchhoff.

Este teorema permite determinar el número posible de árboles de expansión de un grafo conexo, mediante la relación entre las propiedades de la matriz de adyacencia (A) y la matriz de Laplace de un grafo (L). De forma general establece que, si se calcula el cofactor de la matriz de Laplace del grafo, el resultado será igual a el número de árboles de expansión.

La matriz de Laplace $L(G)$ del grafo se forma partir de la matriz de adyacencias y la diagonal (D) con los grados de los vértices, por lo que es también una matriz cuadrada de $n \times n$, donde $n = |V|$.

- En la diagonal principal de la matriz $L(G)$, el valor en la celda (i, j) es igual al grado del vértice i , cuando $i=j$.
- El valor en la celda (i, j) es igual a -1 , si existe una arista entre los vértices i y j .
- El valor en la celda (i, j) es igual 0 , si no existe una arista entre los vértices i y j , es decir que no están conectados.

A continuación, se resume el algoritmo que permite obtener el número de árboles de expansión de un grafo G [21].

1. Construir la matriz de Laplace $L(G)$ del grafo como $L=D-A$.
2. Se elimina cualquier fila y columna de la matriz $L(G)$.
3. Se calcula el determinante de la nueva matriz obtenida en el paso anterior.

2.5. La energía en redes de distribución de agua.

La energía suministrada a la RDA depende de la diferencia de altura entre el nivel del agua en el depósito o fuente, y el nivel donde se ubican los usuarios. Esta energía se conoce como energía natural o energía potencial gravitacional. Cuando la ubicación de la fuente no garantiza la energía mínima requerida para brindar el servicio, es necesario agregar energía desde una fuente externa, generalmente en forma de energía de presión o cinética. En estos casos se requiere incorporar elementos mecánicos a la red, bombas que suelen convertir la energía eléctrica en energía mecánica y posteriormente en energía de presión.

El transporte de agua en la red lo condicionan los valores de caudal (Q) y la presión (P); transportar, elevar y mantener el flujo de agua a lo largo de la red con las condiciones que requieren los usuarios, implica un consumo energético. Este gasto o requisito de energía está definido por el valor de la potencia requerida evaluada en un periodo de tiempo específico de análisis[22].

$$E = \gamma Q H \Delta t \quad (\text{Eq. 1})$$

Donde E = energía suministrada/requerida, Q = caudal transportado, H = altura hidráulica, Δt = delta de tiempo evaluado, γ = peso específico del agua.

La carga hidráulica, altura de energía o cabeza es una medida de la cantidad de energía total que es necesaria para que el agua circule y es función de la cabeza de presión, la cabeza de velocidad, la cabeza de posición o elevación. En el cálculo de tuberías, la altura de energía se utiliza para el cálculo del gradiente hidráulico, como la diferencia de carga entre dos puntos, asociada con las pérdidas de carga.

$$\frac{V^2}{2g} + \frac{P}{\gamma} + Z = H \quad (\text{Eq. 2})$$

Donde H_r = altura de energía, $\frac{V^2}{2g}$ = altura cinética, $\frac{P}{\gamma}$ = altura de presión, Z = altura de posición.

El valor de la energía con la que mínimamente debería suministrarse a un usuario para abastecerse está definido por la presión de servicio y el caudal demandado, la

suma de esta energía entregada a cada usuario (j) equivale a la energía mínima requerida por el sistema E_{uo} .

$$E_{uo} = \gamma \sum_{j=1}^n v_j \left[(Z_j - Z_l) + \frac{P_{0,j}}{\gamma} \right] \quad (Eq. 3)$$

Donde v_j es el volumen de agua consumido en el nodo j durante el período considerado, Z_l es la altura del nudo más bajo de la RDA, Z_j es la altura del nodo j, $P_{0,j}/\gamma$ es la presión de suministro mínima en el nodo j.

En el proceso de transporte se producen una serie de consumos de energía, conocidos como pérdidas que son debidas a factores como la fricción, los cambios de presión, las fugas de agua y otras ineficiencias que generan que la cantidad de energía que se inyecta o requiere E_{SR} sea mayor a E_{uo} . Las pérdidas de energía pueden clasificarse como operaciones o estructurales de acuerdo con su origen [14].

La ecuación (Eq. 4) desarrollada por Cabrera et al. (2019) relaciona los valores de energía mínima requerida por los usuarios y las pérdidas estructurales y operaciones con el valor de la energía inyectada o requerida por el sistema, lo que permite realizar el balance energético de la red [23].

$$E_{SR} = E_{uo} + E_{pr} + E_{fr} + E_{lr} + E_{er} + E_{tr} \quad (Eq. 4)$$

Donde E_{SR} energía suministrada real, E_{uo} energía mínima requerida, E_{pr} pérdidas de energía asociadas a la estación de bombeo, E_{fr} pérdidas de energía asociadas a la fricción, E_{lr} pérdidas de energía asociadas a las fugas de agua, E_{er} energía de excesos y E_{tr} energía topográfica. Los valores de E_{pr} y E_{er} son iguales a 0 en el caso de fuentes rígidas de suministro.

De acuerdo con Del Teso et al. (2019), a continuación se explica y presenta el cálculo de las diferentes componentes del balance energético en una RDA [23].

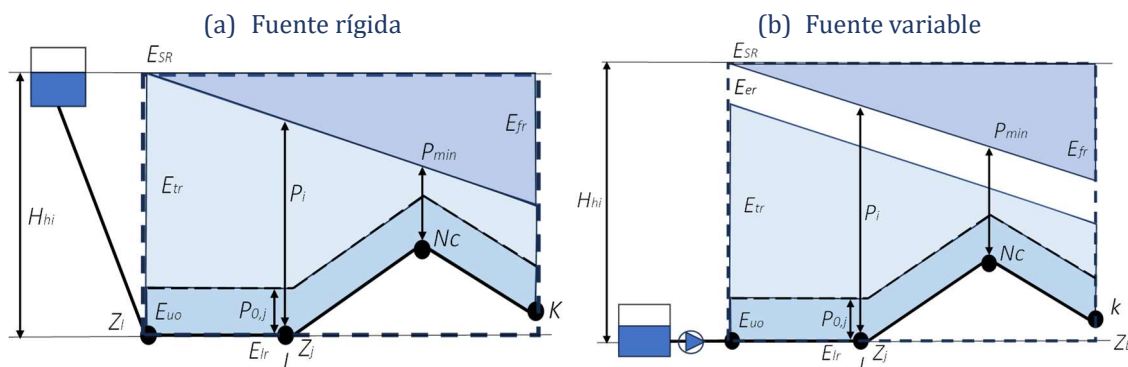


Figura 5 Gráfica balance de energía para una fuente de energía (a)rígida y (b) variable.

Pérdidas operacionales.

Las pérdidas operaciones se relacionan directamente con la operación de la red y se generan por ineficiencias en la estación de bombeo, la fricción en las tuberías, las

fugas de agua, los excesos de energía y despresurizaciones como las que se generan en los tanques.

Pérdidas de energía asociadas a la estación de bombeo: están asociadas con la eficiencia del conjunto bomba, motor y si corresponde, la conexión eléctrica. Dependen directamente del estado y de la selección de los elementos.

Pérdidas de energía asociadas a la fricción: es la fracción de la energía que se disipa por el rozamiento del agua con las tuberías del sistema.

$$E_{fr} = \gamma \sum_{i=1}^n q_i \Delta h_i \Delta t = \gamma \sum_{i=1}^n v_i \Delta h_i \quad (\text{Eq. 5})$$

Donde m corresponde al número de tuberías, q_i al caudal conducido en cada tubería, Δh_i las pérdidas de carga en la tubería, v_i el volumen transportado y Δt el periodo de tiempo. En términos de nudos las pérdidas por fricción están dadas por la siguiente ecuación (Eq. 6).

$$E_{fr} = \gamma \sum_{j=1}^n v_{g,j} \left[H_{hi} - \left((Z_j - Z_l) + \frac{P_j}{\gamma} \right) \right] \quad (\text{Eq. 6})$$

Donde $v_{g,j}$ es el volumen total en el nudo (demanda + fugas de agua), H_{hi} la mayor altura piezométrica en los nudos y $\frac{P_j}{\gamma}$ la presión el nudo j .

Pérdidas de energía asociadas a las fugas de agua: corresponden con el volumen de agua fugado por el valor de la altura piezométrica en el nudo donde se asocia la fuga.

$$E_{lr} = \gamma \sum_{j=1}^n v_{l,j} \left[(Z_j - Z_l) + \frac{P_j}{\gamma} \right] \quad (\text{Eq. 7})$$

Donde $v_{l,j}$ es el volumen de fugas de agua en el nudo.

Excesos de energía: Se presenta cuando existe una diferencia positiva entre la energía del nudo crítico y el valor mínimo requerido.

$$E_{er} = \gamma \sum_{j=1}^n v_{g,j} \left[\frac{P_{NC}}{\gamma} - \frac{P_{0,NC}}{\gamma} \right] \quad (\text{Eq. 8})$$

Donde $\frac{P_{NC}}{\gamma}$ es el valor de la energía en el nudo crítico y $\frac{P_{0,NC}}{\gamma}$ el valor de la presión de servicio.

Pérdidas estructurales o energía topográfica.

Es básicamente un excedente de energía vinculada a la topografía, disposición espacial del lugar de implantación, el trazado y configuración de la red. Debido a la diferencia de altura entre el nudo crítico y los demás nudos que conforman la red se suministra más energía de la que se requiere en los nodos del sistema de distribución. La energía topográfica es una ineficiencia que debe minimizarse en la

medida de lo posible, con el objetivo de que se inyecte a la red solo la cantidad necesaria [22].

$$E_{tr} = \gamma \sum_{j=1}^n v_{g,j} \left(\frac{P_j}{\gamma} - \frac{P_{0,j}}{\gamma} \right) - E_{er,j} \quad (Eq. 9)$$

Donde $\frac{P_{0,j}}{\gamma}$ es la presión mínima requerida en el nudo j y $E_{er,j}$ el exceso de energía existente en cada nudo. La Figura 5 ilustra la diferencia existente entre una fuente de energía rígida y una fuente variable. En fuentes rígidas, la altura es estimada para evitar excesos en el nodo crítico durante el momento punta, pero no pueden adaptarse a las necesidades exactas de energía a lo largo del tiempo. Como consecuencia se generan ineficiencias que son inevitables. A diferencia de los sistemas con bombas donde la altura requerida puede ser ajustada a medida que cambia la demanda.

La energía topográfica se puede subdividir en tres componentes de acuerdo con sus características, inevitable E_{tr}^u ligada al flujo, E_{tr}^f y manejable E_{tr}^m , como se muestra en la Figura 6.

$$E_{tr} = E_{tr}^u + E_{tr}^f + E_{tr}^m \quad (Eq. 10)$$

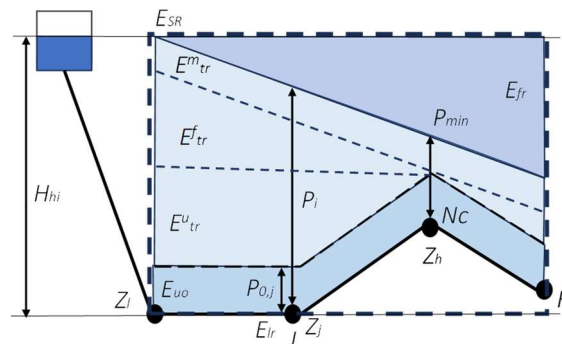


Figura 6. Subdivisión de la energía topográfica [23].

Energía topográfica inevitable o ideal: Es la energía mínima que se debe suministrar a un sistema de distribución de agua para que el nudo crítico pueda recibir el servicio en una situación ideal (sin pérdidas por fricción). Afecta al resto de los nudos de consumo de menor cota, ya que para que el nudo más alto tenga presión, el resto de los nudos del sistema tendrán sobrepresión y no se puede evitar, salvo modificando el diseño ya que se relaciona con las características del terreno y la ruta de la tubería.

$$E_{tr}^u = \gamma \sum_{j=1}^n v_{g,j} (Z_h - Z_j) \quad (Eq. 11)$$

Donde Z_h es la altura del nudo más desfavorable.

La energía topográfica inevitable debido al flujo: Es necesaria para cumplir con la presión de servicio. Por lo tanto, depende de las pérdidas por fricción existentes y es una función de la pendiente hidráulica con la cual se dimensiona la red. Sin embargo, reducirla solo es posible hasta el punto en el que deja de ser viable la garantía de suministro.

$$E_{tr}^f = \gamma \sum_{j=1}^n v_{g,j} \left(Z_j + \frac{P_j}{\gamma} - \frac{P_{min}}{\gamma} - Z_h \right) \quad (Eq. 12)$$

Donde $\frac{P_{min}}{\gamma}$ es la presión de nodo menos favorable.

La energía topográfica gestionable: Corresponde a la parte de la energía topográfica que se puede gestionar y se genera debido a la configuración de la red, ya que las subredes que se derivan de tramos intermedios pueden presentar excesos de energía. Se puede recuperar (utilizando PAT) o disipar (utilizando PRV).

$$E_{tr}^m = \gamma \sum_{j=1}^n v_{g,j} \left(\frac{P_{min}}{\gamma} - \frac{P_{0,j}}{\gamma} \right) - E_{er,j} \quad (Eq. 13)$$

Donde $E_{er,j}$ es la energía de exceso en el nudo, $\frac{P_{min}}{\gamma}$ la presión mínima del nudo menos favorable y $\frac{P_{0,j}}{\gamma}$ la presión de servicio en el nudo j.

Capítulo 3.

Metodología para la generación de trazados energéticamente convenientes.

3.1. Introducción.

Una RDA surge de la necesidad de suplir el servicio de agua a un determinado grupo de usos o usuarios, que bien porque no cuenta con agua o no les es suministrada en condiciones, requieren implementar una solución que les garantice su abastecimiento de agua. No obstante, cada solución es singular; al encontrarse inscrita a un espacio y tiempo, por lo que es necesario evaluar las condiciones particulares que condicionan la red, como el lugar de implantación y los requerimientos de los usuarios que dependerán de ella.

Como ya se ha comentado, el estudio de un proyecto de red para el suministro de agua, parte de la distribución espacial de los posibles usuarios del sistema y sus características. Este espacio, resulta uno o quizás el mayor de los condicionantes para el trazado de la red, al estar definido por el desnivel existente y por los elementos que limitan la ubicación de los componentes de la red de distribución.

El objetivo del trazado de la RDA es garantizar que el agua llegue de manera eficiente y segura a todos los puntos donde pueda existir un consumo dentro del área de influencia de la red. Su definición intenta seguir los ejes viales, caminos o alineamientos existentes, de forma que los usuarios queden cubiertos por alguna de las tuberías que conforman la red y puedan conectarse para recibir el servicio de agua.

Algunos de los factores que condicionan el trazo de la red son:

- Condiciones geográficas y topográficas del terreno de implantación de la red, como las pendientes del terreno, accidentes geográficos, cuerpos de agua, etc.

- Distribución de los elementos que conforman el espacio donde se implanta la RDA, como manzanas, parcelas, caminos y cualquier otro elemento catastral o que pueda condicionar el trazado.
- Normas y regulaciones relacionadas con la construcción de infraestructuras de distribución de agua.
- Restricciones relacionadas con el uso del suelo, propiedad del terreno, conveniencia o restricciones en el uso del suelo.
- La ubicación de la fuente de agua y aspectos relacionados como la distancia, la calidad del agua, la energía requerida, etc.
- Condiciones técnicas del servicio, presiones mínimas y máximas, calidad del agua suministrada, velocidades, etc.

La naturaleza que tienen las RDA hace que para su definición y estudio sea necesario que tanto la distribución espacial de los usuarios, como la configuración del lugar de implantación y cualquier otro insumo que intervenga, sean presentados e incorporados mediante algún tipo de representación cartográfica que sirva de base para el proyecto.

El trazado de la RDA es el principal elemento que define la topología de la red. Además, condiciona su funcionamiento, operación y por consiguiente su coste, de ahí que estudiar su definición resulta conveniente para obtener sistemas más eficientes. Con este contexto y como se ha comentado previamente, el objetivo de la metodología es la generación de trazados que resulten energéticamente más convenientes, mediante el uso de técnicas de machine learning, específicamente de la clusterización y la teoría de grafos.

3.2. Propuesta metodológica

En consideración de lo expuesto, la generación de trazados energéticamente convenientes se propone a partir de dos entradas de información: la capa de ejes en los que sería viable la instalación de tubería y la capa con la distribución e información de los usuarios o de nudos de consumo.

Capa de posibles ejes para tuberías: contiene todos los ejes en los que podría ser viable la implantación de la red, como ejes viales existentes, caminos, servidumbres y cualquier otro espacio en el que no exista restricciones y sea factible la construcción, operación y mantenimiento de la red. Se genera con base en la información catastral, planimetría, ortofotos o cualquier otro recurso que permita identificar los ejes.

La selección de los ejes deberá alinearse a las normativas locales y restricciones que pueden existir a nivel legal y espacial en el sitio de implantación de la RDA.

Capa de usuarios o nudos de consumo: contiene la distribución espacial de los usuarios o nudos de consumo de la red y todas aquellas características particulares que identifican a cada individuo. Es indispensable que cada elemento cuente como

mínimo con coordenadas espaciales, cota la que se encuentra y con la demanda base o consumo de agua.

La generación de las capas de entrada requiere de un trabajo previo de búsqueda, edición y depurado de la información, aunque no hace parte de la metodología, si es necesaria para la obtención de un buen resultado. En esta etapa preliminar es vital el uso de un sistema de información geográfica, con el cual poder localizar espacialmente, complementar y analizar las diferentes capas de información de la zona de implantación, que se requieren para generar los “inputs” o entradas de la metodología.



Figura 7. Metodología propuesta para la generación de redes de redes principales de distribución de agua.

La metodología para la generación de trazados energéticamente convenientes en RPDA se desarrolla en 5 etapas. La Figura 7 presenta la secuencia de etapas y un resumen con las entradas y salidas de cada una.

3.2.1. Generación de clústeres o sectorización inicial.

El uso de técnicas derivadas del aprendizaje autónomo y en particular de la clusterización para la sectorización de las RDA y en la generación de los DMAs, es recientemente analizado en diversos estudios y aplicaciones [24] [25] [26]. Estas técnicas permiten identificar de manera objetiva y basada en datos, grupos de nudos

estrechamente interconectados en la red y delimitar áreas geográficas con características hidráulicas similares [27] [28] [29] [30].

La generación de clústeres a partir de la localización y características de los usuarios o nodos de consumo es una herramienta de análisis con muchas posibilidades entre ellas se destacan el balance hídrico de la red, o los estudios de demanda. Además, permite definir la red principal de distribución de agua dentro del sistema de suministro de agua potable. Esta infraestructura que, partiendo desde la cabecera, sirve para conducción del agua de forma exclusiva a los sectores de la red, sin incluir otras derivaciones, facilita el análisis de soluciones particulares a cada sector que pueden considerar las particularidades heredadas de los nodos de consumo que lo conforman.

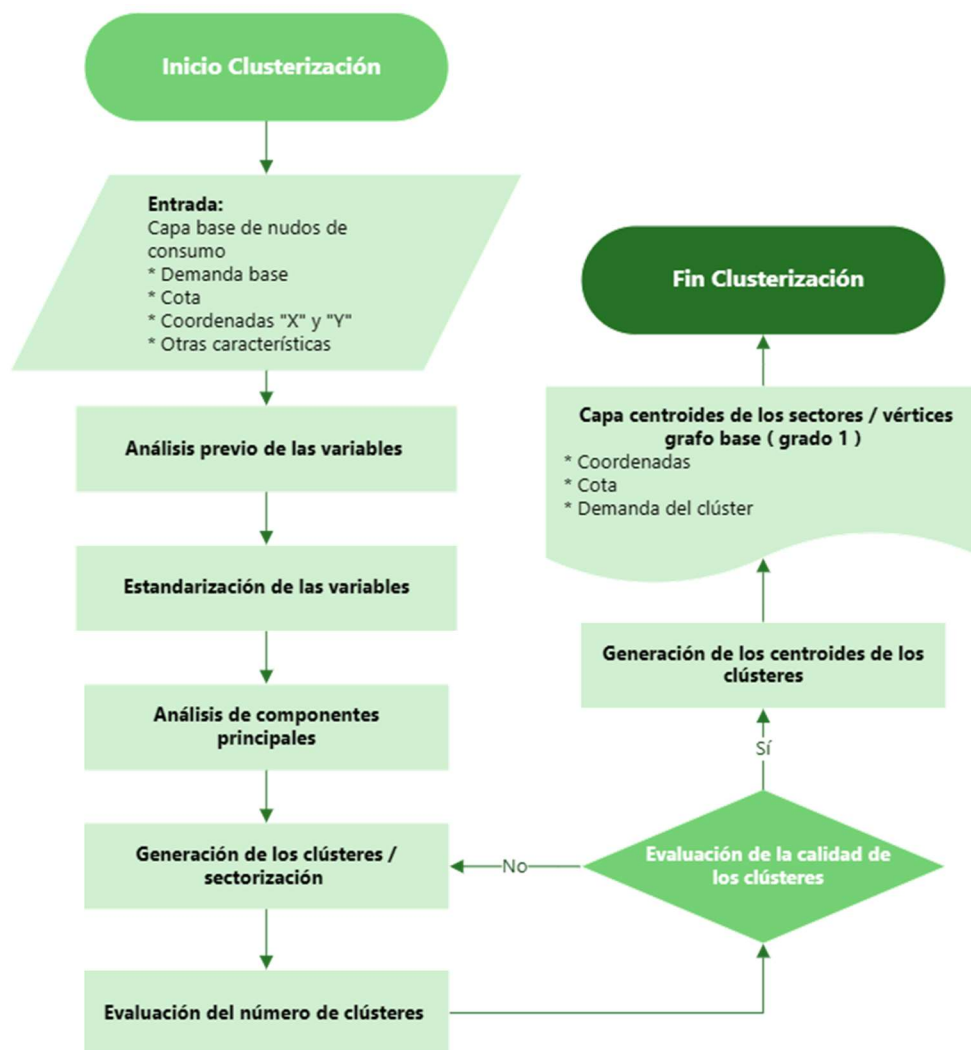


Figura 8. Diagrama de flujo para la etapa de clusterización.

La definición de sectores no solo es conveniente en etapas en las que se conoce la topología de la red previamente, sino que también es potencialmente aplicable en etapas tempranas de conceptualización. Esto posibilita la definición de la infraestructura principal que conformará la red y su análisis energético, con el

objetivo de conseguir soluciones que se adapten mejor a las exigencias de los usuarios y a las condiciones del lugar donde se implantan.

En este contexto, en el que aún no se conoce la conexión entre los elementos de la red, son aplicables técnicas de clusterización no supervisado para agrupar los nudos de la red en sectores, mediante métodos como el de K-Means o clusterización jerárquica, es posible determinar de manera eficiente y objetiva los grupos. La aplicación de estas técnicas desde la definición del propio trazado de la red permite llegar a soluciones con un mejor rendimiento energético, lo que las hace más sostenibles.

La Figura 8 representa el algoritmo para la etapa de clusterización, el cual es una simplificación del código desarrollado para realizar esta tarea de forma automática. La rutina se implementó directamente en la consola de Python de la herramienta software Qgis, desde la cual es posible la generación de nuevas aplicaciones y análisis avanzados no disponibles en el programa. Qgis es un sistema de información geográfica (SIG) de licencia libre y código abierto, que sirve para visualizar, editar y analizar datos geoespaciales.

Para realizar el agrupamiento, es necesaria la capa con la distribución espacial de los nudos o usuarios con los atributos más relevantes que los describen, como por ejemplo la cota, la demanda base, el patrón de consumo y cualquier otro que permita categorizarlos. Inicialmente se analizan las variables características de los nodos de consumo, identificando aquellas comunes a todos los nudos. Luego, se evalúa la integridad de la información y se toman decisiones sobre la necesidad de completar o descartar variables.

Debido a que la escala de los datos puede introducir sesgos, como segundo paso es necesario que se realice un proceso de estandarización, para así asegurar que las variables de menor escala no pierdan importancia en el análisis y puedan impactar en la definición de los clústeres. Esto equilibra la influencia de las variables y garantiza la convergencia del algoritmo de clusterización.

El tercer paso del algoritmo es el análisis de componentes principales o PCA. Con ella se reduce la cantidad de variables consideradas durante el proceso. Además, el PCA elimina la multicolinealidad al crear nuevas variables no correlacionadas, evitando la redundancia de información y capturando la variabilidad relevante de los datos.

Tras la definición de las nuevas variables sin correlación, se procede a la determinación del número óptimo de agrupaciones o clústeres. Este proceso se fundamenta en la aplicación y comparación de los resultados obtenidos mediante los algoritmos de clusterización jerárquica y K-Means. Conforme se detalló en el capítulo de fundamentos, el número de clústeres puede ser determinado a través de una variedad de métodos, los cuales varían según el algoritmo empleado.

En el caso de la técnica de clusterización jerárquica, la elección del número de clústeres suele basarse en la interpretación visual del dendrograma. Esto conlleva considerar aspectos como el tamaño deseado de los clústeres, la distancia (similitud) entre los grupos de datos o umbrales predefinidos. Paralelamente, en la técnica K-Means se recurre también a métodos gráficos, entre ellos el conocido codo de Jambú. Este método establece el valor de K (número de clústeres) en el punto de inflexión del gráfico, donde la disminución en la suma de las distancias al cuadrado experimenta una significativa desaceleración.

Dada la naturaleza particular del problema, la determinación del número "óptimo" de clústeres se presenta como un proceso iterativo, cuyo objetivo teórico reside en hallar un mínimo número de agrupaciones. No obstante, en el contexto de la sectorización de RDA, se podría considerar valores superiores, siempre y cuando los clústeres generados sean visualmente distinguibles, se ajusten a las condiciones del problema y reflejen agrupaciones lógicas de nudos con buena cohesión.

Una vez se haya definido el número de clústeres, los resultados obtenidos tras la aplicación de las técnicas podrían presentar diferencias considerables. En este contexto, es conveniente llevar a cabo una revisión de la coherencia entre las distintas soluciones, tomando en consideración tanto la calidad de los clústeres generados como las características deseables en los sectores, en conjunción con el cumplimiento de los parámetros preestablecidos.

Finalmente, se procede a la generación de la capa de centroides, que representará a los sectores de la RDA a partir de los clústeres conformados. A cada sector se le asignarán las características heredadas de los elementos agrupados, como la suma de las demandas base, la cota representativa del sector y cualquier otra propiedad que resulte importante en el cálculo de la red.

3.2.2. Generación de la topología base

La generación de la topología base requiere como entrada la capa con la distribución espacial de caminos posibles para la instalación de la red y atributos de los centroides que representan la sectorización realizada para la RPDA, estos centroides que actuarán como vértices extremos con un grado igual a 1, conformarán un grafo destinado a representar la estructura de la red. Las aristas de este grafo serán definidas a partir de los ejes donde sea viable de la instalación de tuberías dentro del área de implantación de la red.

La Figura 9, ilustra el diagrama de flujo propuesto para esta fase de la metodología en la cual es esencial nuevamente el uso de una herramienta SIG. Inicialmente, se realiza la superposición de las capas de entrada con los centroides y los posibles ejes. A partir de esta capa resultante, se seleccionan aquellos ejes que posibilitan una conexión óptima a los sectores previamente definidos.



Figura 9. Generación de la topología base.

El objetivo de esta etapa es la conformación de una malla. Esto asegura la existencia de alternativas de trazado, dando prioridad a aristas redundantes con capacidad de conexión a vértices extremos (centroides), manteniendo a su vez, estos vértices como los únicos con grado 1. Posteriormente, se completa la topología al agregar los segmentos faltantes que conecten todos los vértices, asegurando la presencia de un nudo en la intersección entre dichas aristas. El resultado podría asociarse con una red de distribución mixta.

La generación de las posibles configuraciones de trazado para la RPDA se fundamenta en la equivalencia intrínseca entre los grafos y los componentes de un modelo de red, donde los nudos y tuberías se asocian a los vértices y aristas de un grafo, respectivamente. La topología base será entonces un grafo G compuesto por el conjunto V de vértices (nudos) y el conjunto E de aristas (tuberías), cuyas propiedades son conexo, no dirigido y tener la presencia de ciclos. Esta configuración simplifica el análisis de diversas alternativas de conectividad para cada sector y facilita una evaluación más precisa de los trazados que resultan más convenientes en función de los objetivos planteados.

3.2.3. Generación de los posibles trazados.

El estudio de RDA que sean energéticamente más eficientes, requiere de una evaluación exhaustiva de diversas alternativas de trazado, con el propósito de identificar aquellas configuraciones con un menor requerimiento de energía en cabecera, menor exceso de energía y que sean económicamente rentable en términos de su construcción y operación, de forma que se obtengan soluciones más sostenibles.

Las alternativas de trazado para la RPDA resultan de las combinaciones de aristas que conforman el grafo base $G = (V, E)$. El nuevo subgrafo resultante T será un árbol de expansión que se conforma por el conjunto V de vértices del grafo base y por el subconjunto E' , con el mínimo número de aristas que son necesarias para conectar todos los vértices. Este nuevo grafo seguirá teniendo las propiedades de ser conexo y no dirigido, pero será acíclico. Además, cumplirá con el criterio de que número de aristas que lo definen será igual al número de vértices - 1, característico de una red ramificada.

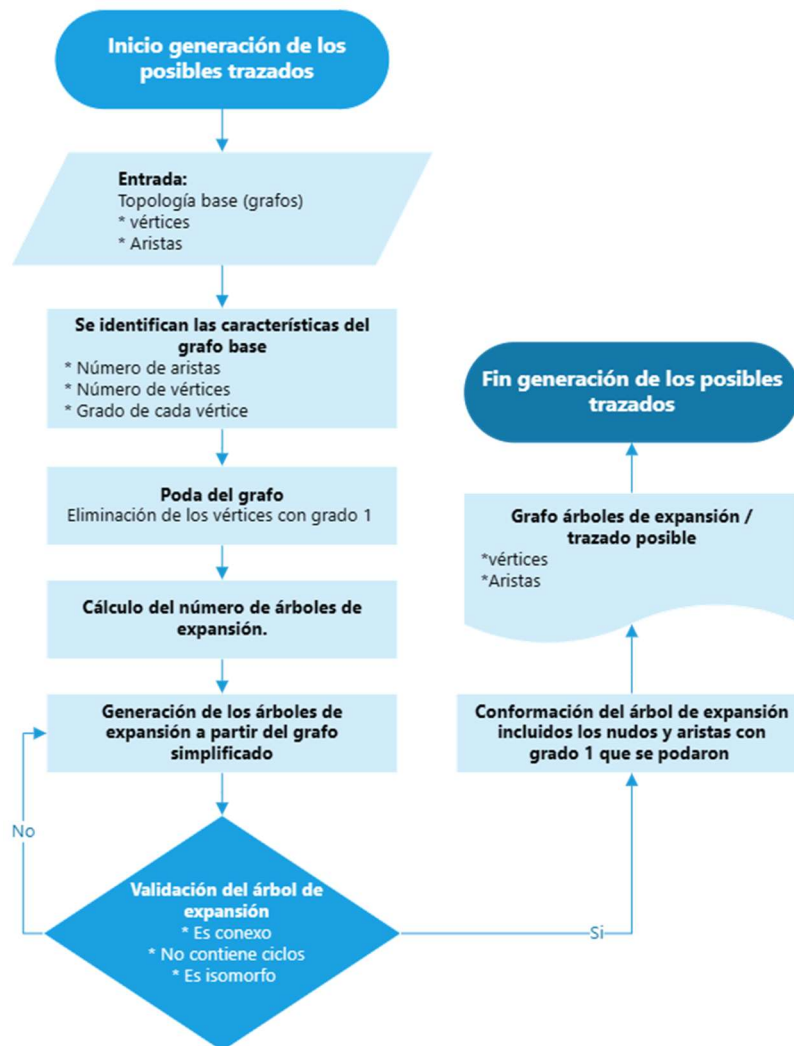


Figura 10. Generación de los posibles trazados.

El diagrama de flujo presentado en la Figura 10, describe los pasos para la generación de árboles de expansión a partir de la topología base (grafo base). Este algoritmo, es una simplificación del código que se elaboró para listar todos los posibles de árboles de expansión de un grafo de forma automática y que hace parte una aplicación más grande desarrollada en Visual Basic for Applications (VBA). El código, aplica un algoritmo de Búsqueda en Profundidad (DFS), para generar todas las combinaciones de aristas. Esta combinación de aristas será válida si cumple con los criterios que definen un árbol de expansión, por lo que se implementan nuevamente funciones de DFS que revisan que sea conexo y acíclico.

El algoritmo DFS se emplea para recorrer y explorar un grafo de manera recursiva. Comenzando desde un vértice inicial (raíz), se selecciona una arista incidente a dicho vértice y avanza en profundidad hacia los vértices no visitados. El algoritmo utiliza una pila (stack) para almacenar los vértices visitados o no visitados según se defina, cuando se alcanza un vértice extremo o se cumple un criterio de terminación, el algoritmo retrocede al vértice anterior y continúa explorando vértices no visitados hasta cubrir todos los nudos posibles del árbol.

La entrada para esta etapa de la metodología es un archivo plano de texto en formato INP que define la topología del grafo base, del cual se extrae la cantidad de elementos que lo conforman y se calcula el grado de cada vértice. A partir de esta información y de acuerdo con la conexión existen entre los vértices se generan las matrices, listas y demás elementos que permitan representar y entender el grafo presentados en el numeral 2.4.

La tarea de enumerar todos los árboles de expansión representa un desafío significativo en términos computacionales, debido a la complejidad exponencial inherente al problema. A medida que el tamaño y la complejidad del grafo aumentan, la cantidad de árboles de expansión se incrementa de manera significativa. Por consiguiente, se hace necesario llevar a cabo una "poda" del grafo base, que es un proceso que implica la eliminación de todos aquellos nudos con grado igual a uno y sus aristas adyacentes, lo que resulta en una simplificación del grafo base, y, en consecuencia, en una reducción de la complejidad del problema y del tiempo de cálculo computacional requerido.

Para el nuevo grafo base simplificado G' , se identifican tanto los vértices como las aristas que lo componen. A continuación, se procede a generar nuevas listas, matrices y representaciones matemáticas específicas necesarias para los procesos de generación de los árboles de expansión posteriores. El algoritmo DFS encargado de generar todas las combinaciones de aristas, utiliza por el ejemplo la lista de vértices incidentes por nudo como entrada. Por otro lado, el cálculo de la cantidad de árboles de expansión se lleva a cabo mediante el Teorema de Kirchhoff, utilizando la matriz de adyacencia. Este teorema establece que cualquier cofactor de la matriz $AG-DG$, donde AG es la matriz de adyacencia y DG es la matriz diagonal con valores

correspondientes a los grados de los vértices del grafo, es equivalente al número de árboles de expansión del grafo.

Como se mencionó previamente, una vez que el algoritmo ha generado la combinación de aristas, se verifica si estas forman un árbol de expansión válido, en cuyo caso al árbol de expansión resultante se le incorporan los vértices y aristas previamente eliminados durante el proceso de poda en el grafo base. Los árboles de expansión generados representan los posibles trazados que podría tener la RPDA a partir de la sectorización realizada. El producto de esta fase se traducirá en la creación de un nuevo archivo de texto en el que se registren las aristas excluidas del conjunto E' , representando así cada combinación de aristas válidas.

3.2.4. Cálculo de redes

El cálculo y dimensionado de las redes RPDA ramificadas que se generan a partir de los trazados definidos por los árboles de expansión, se desarrolla de acuerdo con el diagrama de flujo presentado en la Figura 11. Este algoritmo, es una simplificación del código que se elaboró para realizar esta tarea de forma automática y que junto con el código de la etapa anterior hace parte una aplicación más grande desarrollada en VBA.

El código desarrollado emplea la herramienta Toolkit de Epanet en su versión 2.2, que es una librería de enlaces dinámicos (DLL) de funciones de Epanet que junto con los demás ficheros que conforman la herramienta, permite a los desarrolladores emplear las funcionalidades de Epanet en sus propias aplicaciones. De forma general, el módulo elaborado para el dimensionado y cálculo de las redes utiliza las funciones, para obtener y editar información de la red, la generación del proyecto, así como las funciones del motor hidráulico para el cálculo.

Las entradas para este proceso son las redes ramificadas generadas en el numeral 3.2.3 que contiene el conjunto de aristas excluidas (E') y el archivo INP que contiene a la topología base. De la topología base se eliminan las aristas que no forman parte del árbol de expansión analizado, resultando un archivo INP que contiene la propuesta de RPDA a evaluar y en el que aún no se definen los parámetros y propiedades de los elementos que conforman la red. En adición se incorpora en el análisis, la variación para cada propuesta del nudo de entrada a la red, ya que este punto influye en la obtención de soluciones energéticamente más convenientes. Este análisis implica la generación, cálculo y dimensionado de posibilidades igual al número de nudos por cada propuesta de trazado.

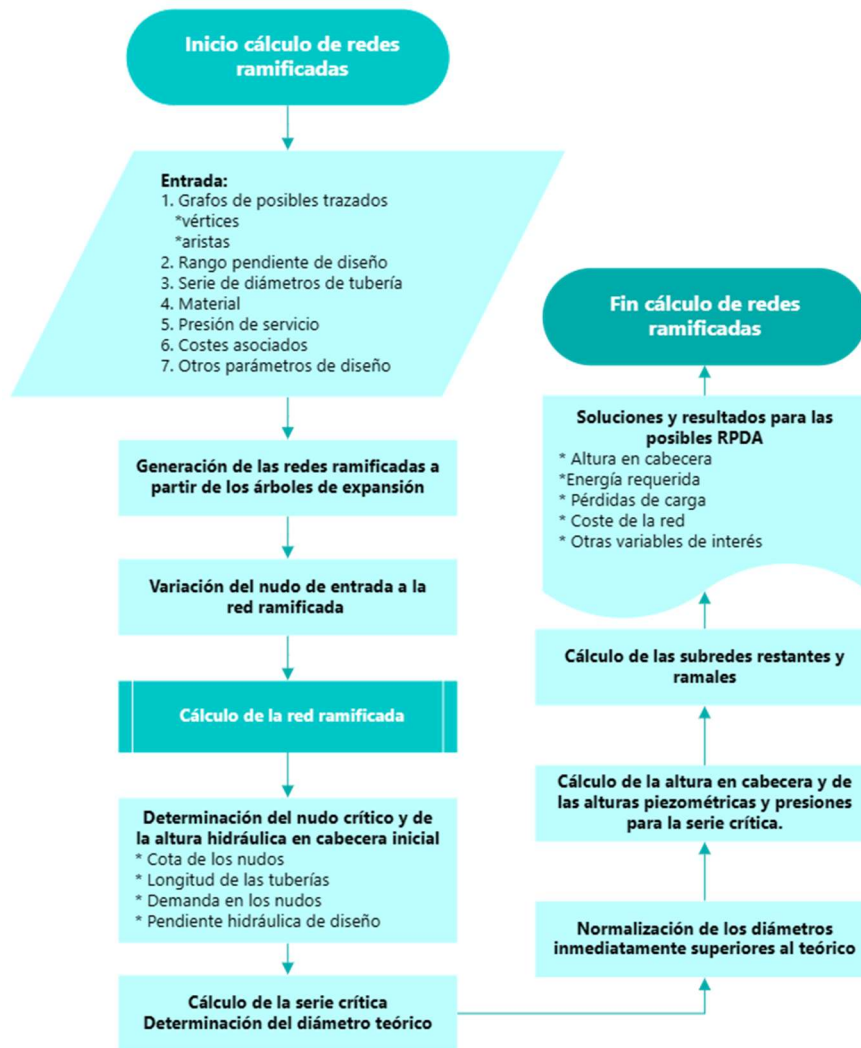


Figura 11. Cálculo de redes ramificadas.

Para el cálculo de las RPDA se optó por el método de la pendiente hidráulica de diseño, en específico se incorporó al código el algoritmo para la solución de problemas con altura cuya cabecera es desconocida, donde la variable de entrada es la pendiente de diseño para la ruta crítica. Para la aplicación de la metodología se propone realizar el dimensionado para un rango de pendientes cercanas a la óptima. Esto permite evaluar su influencia en la definición de este parámetro considerando que pendientes bajas implican soluciones que requieren menos energía en la cabecera, pero tienen un coste mayor de construcción; y que pendientes altas tendrán un mayor consumo energético pero menores costes de construcción asociado a diámetros menores, con mayor pérdida de carga.

$$H_f = f \frac{8LQ^2}{g \pi^2 D^5} \quad (\text{Eq. 14})$$

$$J = \left(\frac{4}{\pi^2}\right) \frac{f Q^2}{D^5 2g} \quad (2) \quad (\text{Eq. 15})$$

Donde f = factor de fricción, que se supone constante en este análisis, H_f = pérdida de carga debida a la fricción, L = longitud de la tubería, Q = caudal que circula por la tubería, D = el diámetro del tubo, J = pendiente hidráulica.

A partir del método de la pendiente óptima descrito por Vasconcelos et al. (2022), que considera los costes asociados con las tuberías, costes energéticos a lo largo de la vida útil del proyecto, la tasa de crecimiento de la demanda, tiempo de operación y tasa de interés, es posible calcular este parámetro. Debido a que, en la etapa de conceptualización de la red es probable que no se conozcan la totalidad de estos parámetros necesarios para el cálculo propuesto, la ecuación (Eq. 16) se puede adaptar utilizando los parámetros conocidos.

$$J_{opt} = \left(\frac{8f}{\pi^2 g} \right)^{\frac{m}{5+m}} \left(\frac{\eta}{\gamma} \right)^{\frac{5}{5+m}} Q_j^{\frac{2m-5}{5+m}} \left[\frac{0.2mK_{mat}}{(0.365TC_{kW,h,j} + 0.012C_{kW,j})K_2} \right]^{\frac{5}{5+m}} \quad (Eq. 16)$$

Donde η = es la eficiencia de la bomba, γ = el peso específico del agua, K_{mat} = coeficiente en la función de costes de capital, m = exponente en la función de costes del material, T = periodo de operación de la diaria de la bomba, $C_{kW,h,j}$ = Coste unitario de energía consumida, $C_{kW,j}$ = Coste unitario por potencia contratada, K_2 = coeficiente en términos de la tasa de interés, la vida útil del proyecto y del crecimiento en la demanda de agua.

En términos generales, el dimensionado de la red se realiza siguiendo el siguiente proceso:

- Definición de la serie crítica y la altura hidráulica requerida en cabecera a partir de la longitud desde el nudo crítico al punto de cabecera de la red, la diferencia de nivel entre estos dos puntos, la presión de servicio requerida y la pendiente de diseño.
- Dimensionado de cada uno de los tramos que compone la serie crítica a partir de la ecuación de Darcy-Weisbach, considerando la rugosidad del material seleccionado para la red, el caudal a transportar, la longitud del tramo y la pendiente de diseño.
- La normalización de los diámetros calculados se realiza hacia el diámetro inmediatamente superior de acuerdo con la serie de diámetros seleccionada para el proyecto, lo que favorece tanto el cálculo como la obtención de soluciones energéticamente eficientes.
- A partir de la serie crítica normalizada se calcula nuevamente la altura de cabecera, presiones, alturas piezométricas en los nudos intermedios, velocidades y otras variables de interés.
- Finalmente, a partir de los nudos de altura conocida se dimensionan y calculan las subredes y tramos aislados de la red. Para obtener la pendiente de diseño, ahora se considerará como un problema de altura en cabecera

conocida, siendo esta la diferencia entre la altura del nudo conocido y la presión mínima requerida en el extremo de la serie o tramo calculado.

El cálculo y dimensionado de las redes se realiza en régimen permanente (estático), considerando que las condiciones hidráulicas (caudales y altura piezométrica) no fluctúan en el tiempo.

Las variables empleadas en el cálculo de los criterios para el análisis energético de la red se fundamentan en los resultados obtenidos de su cálculo, que se realiza luego de dimensionar el último tramo de la RDA analizada.

Los criterios que se proponen para evaluar si el trazado propuesto es energéticamente conveniente son:

1. La energía requerida en cabecera.
2. El valor de energía de exceso en los nodos.
3. El coste de la red, incluyendo los gastos de materiales, construcción y operación.

La energía requerida en la cabecera: es el mínimo valor de energía que se necesita para que la red pueda operar, se obtiene a partir del balance energético de la red, presentado por Del Teso et al. (2019). En el que la energía suministrada real (E_{sr}) equivalente a la energía requerida en la cabecera (E_{rc}) en una etapa de proyección de la red, es función de la energía mínima requerida por los usuarios, la energía asociada con las pérdidas por fricción y las pérdidas estructurales del sistema vinculadas con el desnivel del terreno.

$$E_{sr} = E_{uo} + E_{fr} + E_{tr} \quad (Eq. 17)$$

Donde E_{sr} = energía suministrada real, E_{uo} = energía mínima requerida por los usuarios, E_{fr} = energía perdida en fricción y E_{tr} = energía topográfica o pérdidas energéticas estructurales.

Energía de exceso en los nudos: hace referencia a la diferencia entre la energía entregada en un nudo y el valor mínimo requerido para cumplir las condiciones de servicio. Este exceso de energía se refleja como una presión por encima del valor establecido como presión de servicio.

En un escenario ideal, toda la energía inyectada en la red se destinará por completo a cubrir la energía mínima requerida por los nudos. Sin embargo, las pérdidas energéticas tanto estructurales como operacionales no pueden eliminarse. Teniendo en cuenta que la energía en cabecera es la mínima que puede tener la red y que no existen gastos energéticos distintos a los ocasionados por la fricción (pérdidas estructurales), los excesos de energía que se generen serán consecuencia de la configuración de la red. De acuerdo con Del Teso March (2020), la energía topográfica en la red equivale a las pérdidas estructurales relacionadas con el perfil

del terreno, es decir, con las diferencias de altura presentes en el área donde se implanta la red y su recorrido.

El coste de la red, incluyendo los gastos de operación: los costes de la red no solo consideran aquellos costes relacionados con la construcción de la red, además se incluyen los asociados a la operación del sistema a lo largo de su vida útil. Para la metodología expuesta, los costes de operación serán los costes debidos al consumo energético de la red.

$$C_{total} = C_{capital} + C_{energía} \quad (Eq. 18)$$

Los costes energéticos durante la operación del proyecto pueden diferenciarse en dos, el coste asociado con la energía consumida o demanda (€/kWh) y el coste unitario debido a la potencia contratada o al uso (€/kW).

$$C_{energía} = C_{E.demanda} + C_{uso} \quad (Eq. 19)$$

La salida de esta etapa será un archivo CSV, donde se almacenan los resultados obtenidos para cada red incluyendo los valores obtenidos con los criterios energéticos adoptados. Además, se guardan otras variables relevantes para determinar la configuración óptima de la RPDA.

3.2.5. Análisis de resultados y selección de alternativas de trazados.

Como última etapa se realiza el análisis de los resultados, para así identificar aquellas alternativas de RPDA que por su trazado presentan mejores criterios energéticos y pueden considerarse como convenientes. La Figura 12 presenta el diagrama de flujo empleado para la elección de los trazados óptimos desde una perspectiva energética.

La cantidad de resultados que se obtienen para cada red y la cantidad de redes posibles generan un volumen de datos significativo, haciendo necesario el uso del software RStudio, cuyo entorno de desarrollo integrado (IDE) permite escribir, depurar y ejecutar código en R, que es un lenguaje de programación enfocado al análisis estadístico.

Con el total de soluciones se crea una base de datos, y se realiza el primer análisis identificando los rangos en que oscilan los criterios energéticos de las soluciones. Además, se analizan aquellas variables que permitan identificar tendencias como el nudo de entrada a la red, el nudo crítico y la influencia de la pendiente de diseño en la generación de soluciones convenientes.



Figura 12. Análisis y selección de alternativas.

Con base en el menor valor de la energía requerida en cabecera, se descartan aquellas soluciones que, bajo criterio de mínimos, están por encima hasta un delta del 5%. No obstante, este delta de energía requerirá un ajuste de acuerdo con las soluciones obtenidas cuidando conservar un número de soluciones representativo.

A partir de los resultados seleccionados por mínimo valor de energía requerida en la cabecera, se filtran las soluciones con un menor valor de energía de exceso en los nudos, escogiendo aquellas soluciones que se encuentran más cercanas al valor mínimo. De igual forma, se filtran y escogen las soluciones con un menor coste total.

Mediante un cruce entre las dos selecciones, se eligen aquellos resultados que destacan en ambos criterios, por lo que se consideran que son energéticamente favorables. La cantidad de soluciones a considerar dependerá del criterio particular; no obstante, deberán considerarse aspectos como el tamaño y complejidad de la red, la conveniencia de las soluciones, los costes y complejidad del trazado propuesto.

Capítulo 4.

Caso de estudio.

La metodología propuesta se empleó para la definición de la red principal de distribución de agua en un caso de estudio en el que, a partir de la distribución espacial de los nudos de consumo, sus características y la información catastral de la zona se generaron todos los trazados posibles y se identificaron aquellos que resultaron ser energéticamente más convenientes.

Como ya se comentó, la metodología se desarrolla en 5 etapas en las que se propone el uso de una herramienta de software de acuerdo con el problema que se intenta resolver en cada etapa. Además, se apoya en el uso de códigos que agilizan la generación de los resultados y facilitan la implementación de la metodología.

4.1. Presentación.

El caso escogido para implementar la metodología es una fracción del área que cubre la comunidad de regantes de Villa Real, en la provincia de Castellón, España. En la zona escogida se ubican un grupo de hidrantes para el suministro de agua para el riego a una serie de parcelas, de los cuales se conoce su demanda base y localización, los hidrantes se encuentran agrupados en nudos de consumo de acuerdo con su cercanía.

La estimación de la demanda base no se contempla como parte de la metodología, al considerarse un insumo que debe cuantificarse en el caso de redes nuevas o es conocido en el caso de redes existentes. No obstante, para el caso de estudio depende de la dotación por tipo de cultivo, que es constante al tratarse de un riego programado.

Previo al inicio de la metodología y conocidas las coordenadas de los nudos y sus atributos (características), es necesario crear la capa o shape con la distribución espacial de los nudos que es la primera entrada de la metodología. La Figura 13 y la Figura 14, muestran la distribución en altimetría de los nudos y la demanda base asociada a cada uno, respectivamente, así como la distribución espacial de los 178 nudos que representan los puntos en los que se prevé que exista un consumo.

Tabla 2. Datos resumen de la red.

Número de nudos de consumo	178
Caudal total demandado	242,33 l/s
Área de implantación de la red	5.304.238m ²

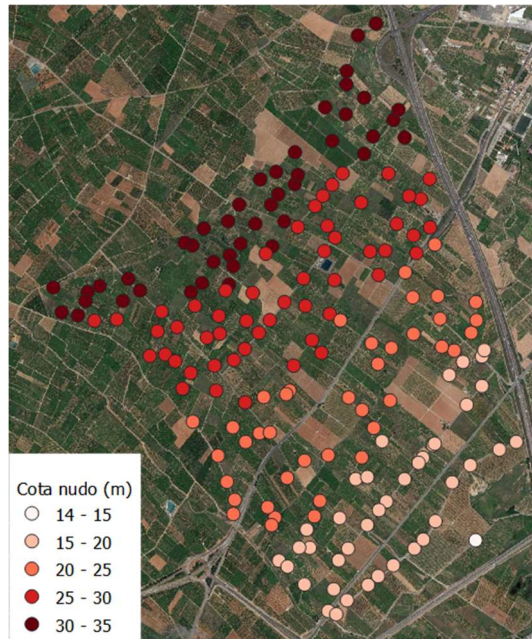


Figura 13. Distribución espacial de los nudos de consumo, clasificados de acuerdo con la cota.

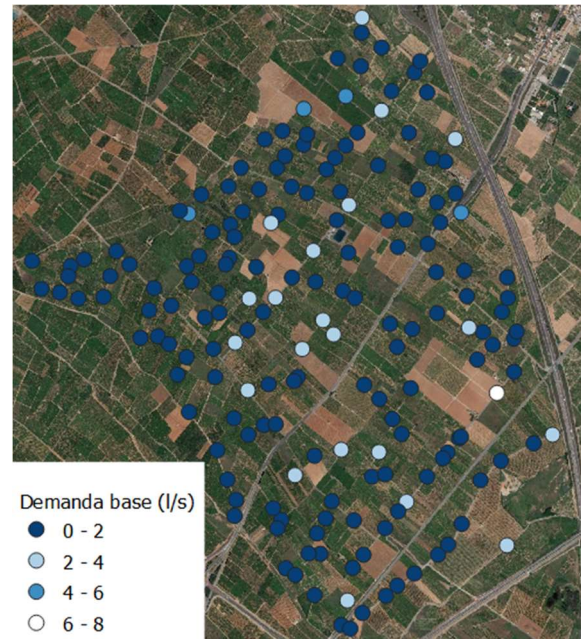


Figura 14. Distribución espacial de los nudos de consumo, clasificados de acuerdo con la demanda base

Adicionalmente, la metodología también requiere como entrada, la capa con la distribución espacial de los posibles ejes en los que sería viable la construcción de tuberías. Considerando que el caso de estudio se ubica en la comunidad de Valencia, para definir los posibles ejes de instalación se emplea la información proveniente del ICV (Institut Cartogràfic Valencià), la sede electrónica del catastro y del IGN (Instituto Geográfico Nacional)

Como insumos se emplean los siguientes archivos en formato shapefile del área de interés.

- Archivos shapefile catastrales de la Sede electrónica Catastral cuya fuente es el registro administrativo dependiente del Ministerio de Hacienda.
 - Ejes
 - Parcelas
 - Subparcelas
 - Construcciones.
- Archivos shapefile del instituto cartográfico valenciano dependiente de la Conselleria de Política Territorial, Obras Públicas y Movilidad de la Generalitat Valenciana.
 - Tramos viales.
 - Hidrografía
 - Edificaciones y construcciones
- Archivos shapefile del Centro de descargas del Instituto Geográfico Nacional dependiente del Ministerio de Transporte, Movilidad y Agenda Urbana.
 - Modelo Digital del Terreno - MDT02.
 - Redes de transporte

- Hidrografía
- Imagen satelital base de Google.

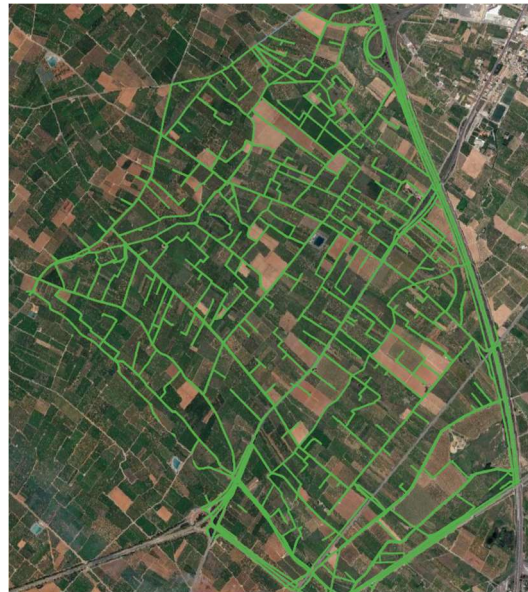


Figura 15. Ejes posibles para la instalación de tuberías en la localización de la red.

La Figura 15, presenta los posibles ejes en los que sería viable la instalación de la tubería, luego de un proceso de montaje y depuración además de los ejes viales y caminos se incluyen algunos ejes de acequias existentes como posibles ejes.

4.2. Etapa 1. Generación de clústeres o sectorización inicial.

Conocida la distribución espacial de los nudos de consumo y mediante la rutina de código implementada en la consola de Python en Qgis, se realiza el agrupamiento mediante clústeres de los nudos (datos/elementos) de consumo, creando una base de datos en la que además de las coordenadas se incluyen las demás variables (características/atributos) que los describen.

Para el caso de estudio solo las variables de demanda base y cota en la que se localizan los nudos son conocidas, debido que ambas condicionan tanto el dimensionado de la red como el consumo energético de la red, se deben analizar con el objetivo de identificar su comportamiento, patrones, e identificar posibles valores típicos (outliers).

La Figura 16 y Figura 17, presenta los diagramas de cajas y bigotes y el histograma para la variable elevación. De forma general la red deberá suministrar a nudos ubicados entre las cotas 14,39 msnm y 34,72 msnm, con un valor medio de 25,8m, siendo más frecuente la presencia de nudos sobre la cota de los 25m.

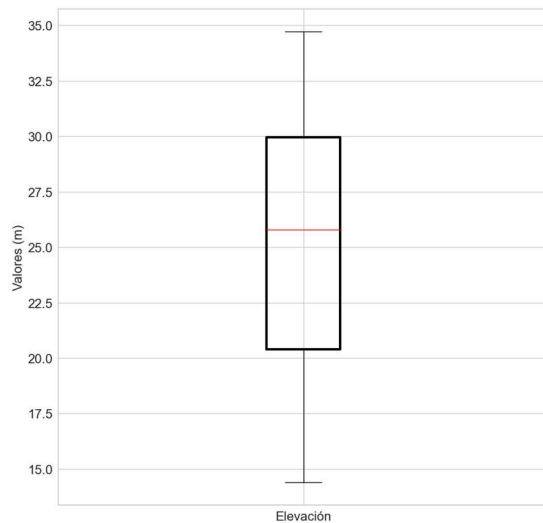


Figura 16. Diagrama de caja y bigotes para la característica elevación.

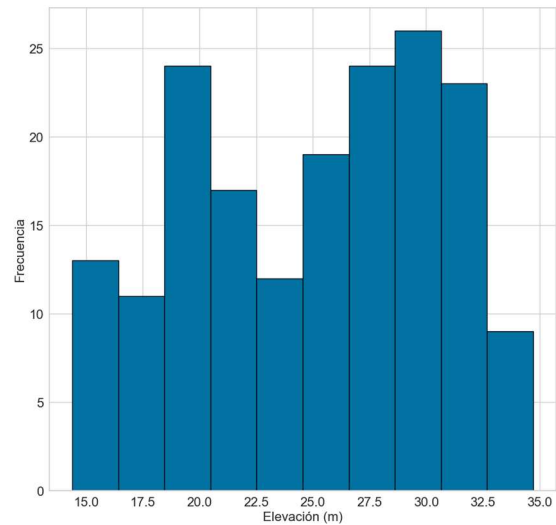


Figura 17. Histograma para la para la característica elevación.

Las Figura 18 y Figura 19, presenta los diagramas de cajas y bigotes y el histograma para la variable demanda base. En estos se observa la presencia de valores atípicos (Outlier), que no se descartaron ya que corresponden a demandas previstas en la red pero que son en magnitud altas respecto a las otras demandas. En general, la demanda se encuentra en el rango entre los 0.17 l/s y los 7.8 l/s, con un valor medio de 1.05 l/s. La mayor frecuencia en los valores de la demanda base se encuentra entre el valor mínimo y los 2 l/s.

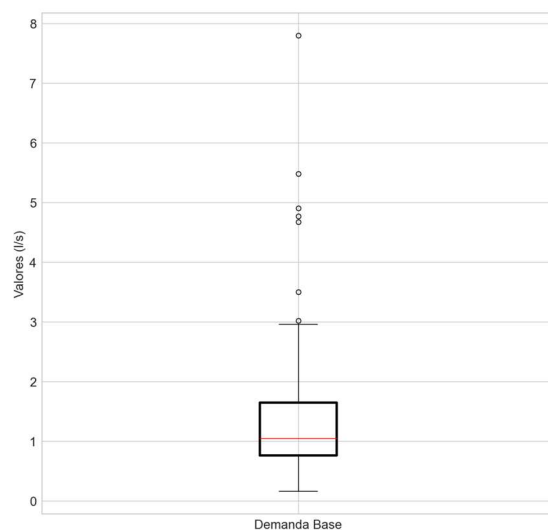


Figura 18. Diagrama de caja y bigotes para la característica elevación.

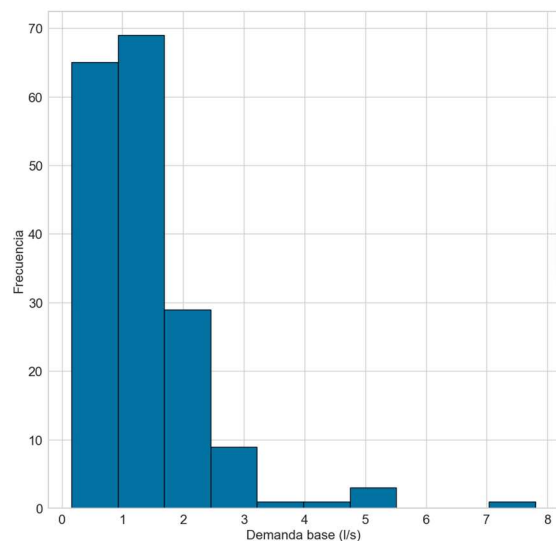


Figura 19. Histograma para la para la característica demanda base.

Estandarización de las variables.

Considerando que la localización espacial de los nudos, la cota a la que se encuentran y la demanda asignada a cada uno son fundamentales en la definición de los sectores

(DMA) a los cuales suministrará agua la RPDA, es necesario que previo al agrupamiento de los datos en clústeres las variables tengan una escala común.

Tabla 3. Rango de las variables que definen los nudos de consumo.

Variable	Valor mínimo	Valor máximo
Coordenada X	743256,41	745968,53
Coordenada Y	4416924,31	4420389,90
Elevación	14,39	34,72
Demanda base	0,16	7,80

Como se mencionó en la metodología, se realiza un proceso de estandarización Z, el cual resta la media de cada variable y luego divide la diferencia por la desviación estándar, obteniendo variables en una escala similar con una media de 0 y una desviación estándar de 1.

Análisis de componentes principales (PCA).

Debido a que se está trabajando con 4 variables (dimensiones) es deseable reducir su número mediante el análisis de componentes principales. Esto permite encontrar un conjunto de nuevas variables que representen la mayor parte de la varianza de los datos originales.

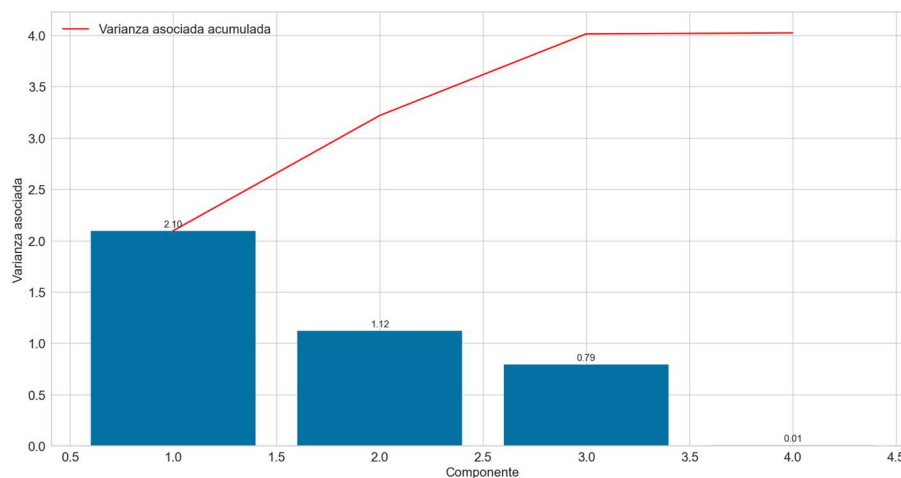


Figura 20. Diagrama de Pareto para los componentes principales

La varianza asociada se refiere a la fracción de variabilidad presente en los datos originales que es explicada por cada componente principal derivado del análisis de componentes principales (PCA). Con este análisis buscamos transformar los datos originales en un nuevo conjunto de variables llamadas componentes principales, que son combinaciones lineales de las variables originales.

En la Figura 20 se presenta el diagrama de Pareto obtenido, compuesto por un gráfico de barras y una línea de acumulación que permite visualizar la varianza asociada de cada componente principal. En este caso se observa que los nuevos componentes 1,2 y 3 capturan la mayor parte de la varianza de los datos, mientras

que la variable 4 no contribuye de forma significativa a la varianza total de los datos por lo que no se considera.

Tabla 4. Varianza por componente principal.

Componente	Varianza	(%)
1	2,1	52,24
2	1,12	27,86
3	0,79	19,65
4	0,01	0,25
Total	4,02	100,00

La Tabla 4, presenta las varianzas correspondientes a cada variable, generadas por mediante el PCA. En la literatura, se sugiere mantener aquellos componentes principales que sean capaces de explicar al menos el 80% de la varianza total, por lo que se decide preservar los componentes 1 y 2, ya capturan un porcentaje mayor de la varianza total, asegurando que se conserve la mayor parte de la información importante de los datos.

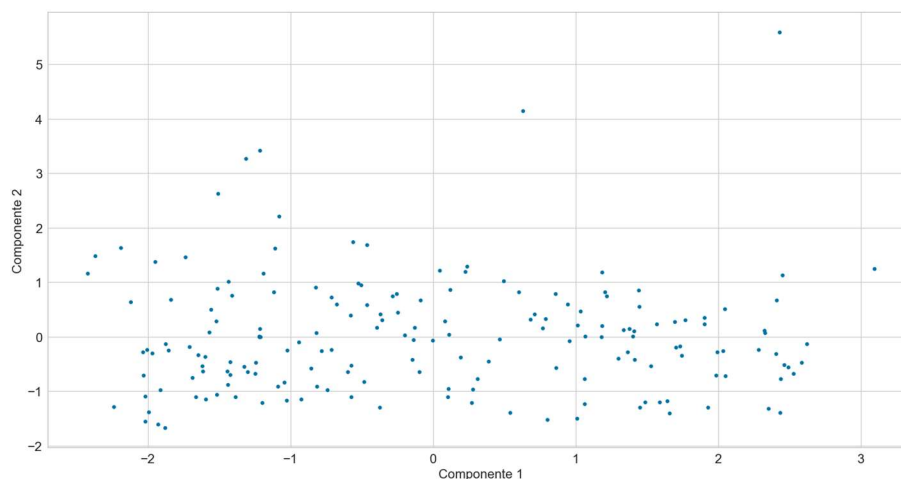


Figura 21. Gráfico de dispersión de los datos para los componentes principales 1 y 2.

La Figura 21 presenta el gráfico de puntos de dispersión de los datos de acuerdo con los dos componentes principales definidos en el PCA. La nueva distribución es la base para el proceso de clusterización, si se compara la Figura 13 y la Figura 14, se distingue que aún se conserva la forma de la distribución original.

Generación de los clústeres.

Tal como se mencionó en la metodología el agrupamiento en clústeres de los datos, se propone mediante la comparación de los resultados de dos algoritmos el jerárquico y K-means. La definición del número de clústeres en cada caso se realiza con el apoyo de varios métodos de acuerdo con la metodología.

Inicialmente y con base en los componentes principales previamente delineados, se aplica el algoritmo jerárquico aglomerativo con el fin de establecer agrupaciones. La Figura 22 presenta el dendrograma resultante, la cantidad de agrupaciones podría

fluctuar si se considera como criterio la distancia vertical entre las fusiones de clústeres o el número de enlaces presentes. A pesar de ello, la determinación del número de clústeres es subjetiva, dependiente de la naturaleza particular del problema y los datos, así como de los objetivos inherentes del análisis, lo hace que la definición de la cantidad de clústeres no sea exacta y única.

En este caso, el número de clústeres podría definirse en el rango entre 2 y 13, el cual permite tener una cantidad de posibles soluciones de agrupaciones. De forma que la solución no sea la considerada como “óptima” por representar el número mínimo de clústeres, sino que sea un número que sin que se pierda la representatividad del método se obtenga una mejor sectorización para la red.

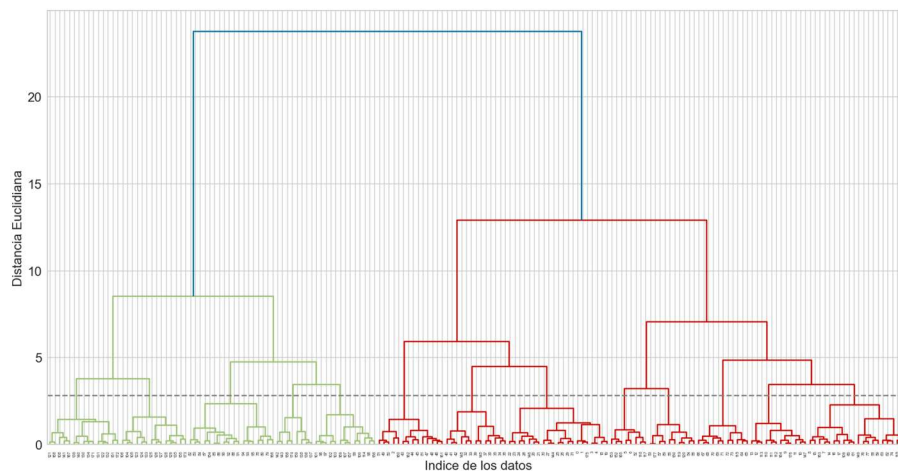


Figura 22. Dendrograma generado mediante el clusterización jerárquica.

A continuación, se aplica el algoritmo de K-meas de forma iterativa para obtener el gráfico del codo de Jambú que relaciona la inercia o suma de los cuadrados de las distancias dentro de cada clúster (WCSS), con la cantidad de clústeres de generados en cada iteración. La Figura 23, presenta los resultados para los datos del caso de estudio. Como es natural el valor de la inercia decrece a medida que aumenta el número de clústeres, lo que significa que grupos con menos elementos generarían mejores grupos.

El número de clústeres adoptado es 12, punto donde se observó la disminución menos significativa en la inercia y a partir del cual, dicho valor decreció de forma casi que constante y lo hizo con una tasa mínima.

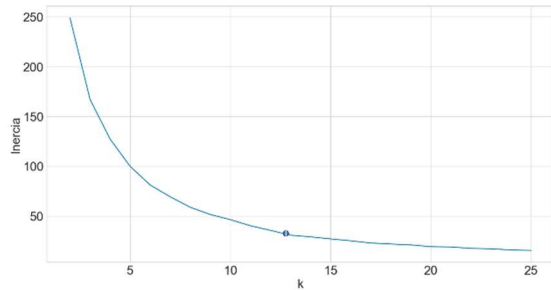


Figura 23. Codo de Jambú.

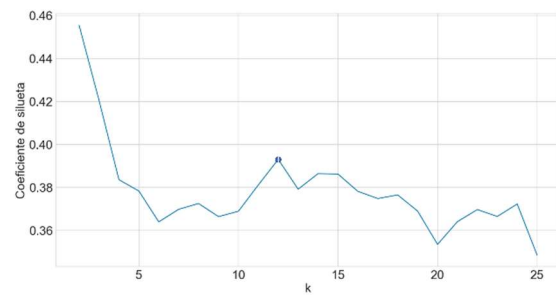


Figura 24. Gráfico del coeficiente de la silueta.

La Figura 24, presenta el gráfico del promedio de los coeficientes de silueta que se obtuvieron al aplicar el algoritmo K-meas de forma iterativa, los picos se interpretan como agrupaciones en los cuales los clústeres son más compactos y están mejor separados, en este caso nuevamente la solución con 12 clústeres resulta conveniente respecto a otras soluciones.

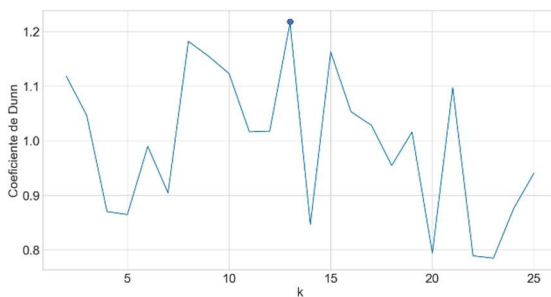


Figura 25. Gráfico del coeficiente de Dunn

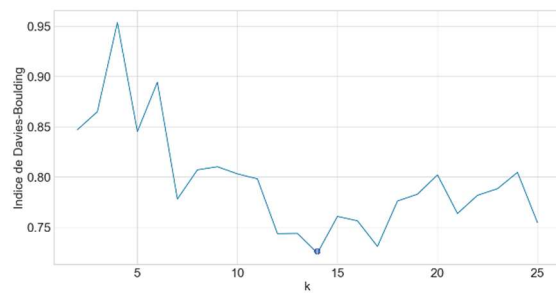


Figura 26. Gráfico del índice de Davies - Boulding.

La Figura 25 representa el gráfico para el coeficiente de Dunn y Figura 26 gráfico para índice Davies - Boulding, cuyo resultado se asemeja al obtenido mediante el gráfico del codo de Jambú y del coeficiente de la silueta, lo que valida la definición del número de clústeres.

Comparadas las dos metodologías se opta por el algoritmo de K-means para la generación de los clústeres, con un número de clústeres igual a 12, que a pesar de no corresponder con el valor que se consideraría "óptimo" si se aplicase el criterio tradicional del mínimo número de clústeres, si representa una buena solución en términos de calidad de los clústeres. El resultado del proceso de clusterización se muestra en la Figura 27, en la cual se presenta la distribución de los 12 posibles sectores que tendría la red, claramente definidos y conformados por entre 12 y 19 nudos.

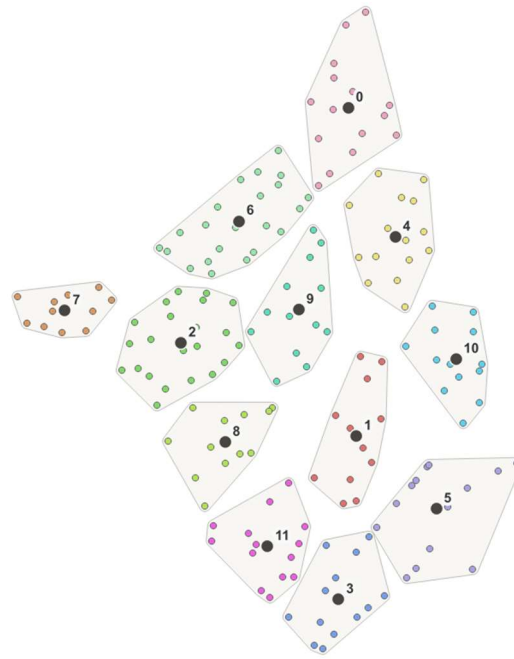


Figura 27. clústeres generados mediante el algoritmo de k-means.

Finalmente, la Figura 28 presenta la distribución espacial de los centroides que representan a los sectores, generados como parte de la aplicación del algoritmo de K-meas.



Figura 28. Centroides obtenidos mediante el algoritmo de K-means.

La Tabla 5 presenta las variables heredadas del grupo de nodos que conforman a cada sector, que para el caso es el acumulado de la demanda base de los nodos dentro de cada clúster y la cota que representa al sector. Los centroides y sus características definen los nodos de la RPDA.

Tabla 5. Variables asociadas a cada sector

Id clúster	Elevación	Demanda base
0	31,89	17,41
1	21,23	20,26
2	28,34	23,28
3	16,77	22,85
4	26,01	24,62
5	17,06	12,47
6	31,25	23,19
7	30,70	17,62
8	24,27	22,39
9	26,70	20,05
10	20,51	17,94
11	20,28	19,27
<i>Total</i>		241,33

Evaluación de la calidad de los clústeres.

La evaluación de la calidad de los clústeres generados se realizó por medio del gráfico de silueta, el cual proporciona información sobre la separación relativa entre los clústeres y cómo los puntos individuales se ajustan a sus clústeres respectivos.

En el gráfico se representan los valores del coeficiente de silueta para cada punto, ordenados de acuerdo con su magnitud de manera descendente y de acuerdo con el clúster al que pertenecen. Además, se traza una línea vertical para representar el valor promedio del coeficiente de silueta de todos los puntos en la agrupación.

Las siguientes figuras representan los gráficos de la silueta generados por los algoritmos de K-means y jerárquico respectivamente. En la Figura 29, los clústeres se ven mejor formados que los clústeres representados en la Figura 30, con valores de silueta más cercanos a 1, lo que se interpreta como datos mejor asignados a sus clústeres y con una buena separación a otros clústeres.

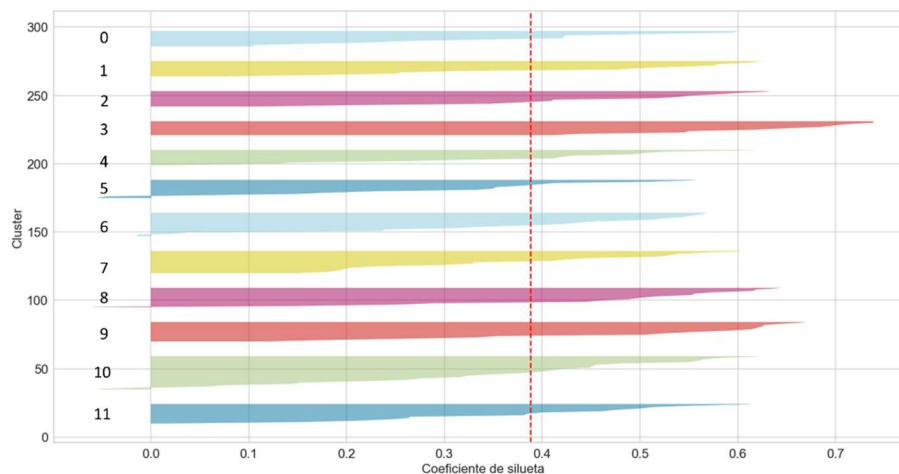


Figura 29. Gráfico de silueta para los clústeres generados mediante el algoritmo de K-means.

Otro aspecto para considerar son las siluetas con valores menores a 0, cuya interpretación está asociada a la posibilidad de algunos nodos que podrían hacer parte de otros clústeres, y que para el método jerárquico presentan una proporción de los datos más significativa, comparado con los datos con iguales características de la primera metodología.

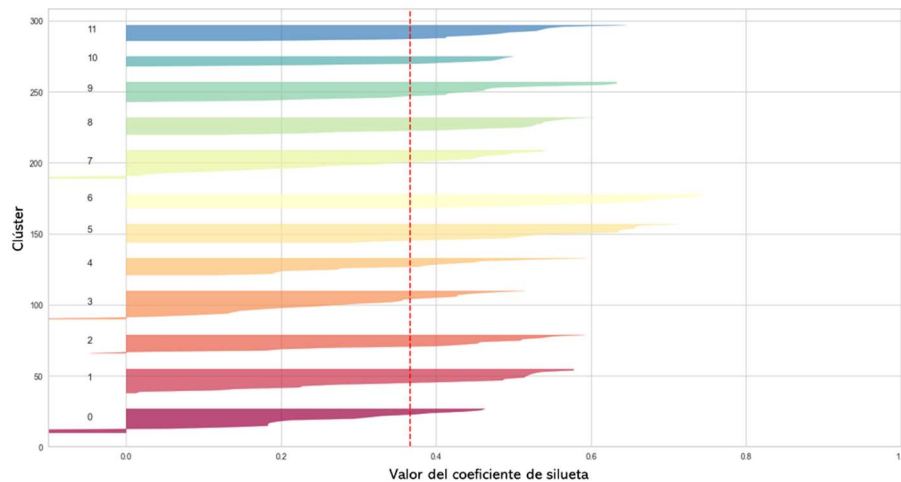


Figura 30. Gráfico de silueta para los clústeres generados mediante el algoritmo jerárquico.

Fundamentado en lo previamente expuesto, se concluye que a pesar de que la agrupación de clústeres generada mediante el método K-means presenta un coeficiente de silueta promedio que no es considerablemente alto, la cantidad de grupos es suficiente para definir clústeres más cohesivos y de mayor calidad en comparación con los grupos generados por el método jerárquico. Además, el número de clústeres elegido presenta el valor del coeficiente de silueta promedio más alto, valor asociado a la pertinencia de los datos a sus clústeres, con agrupaciones que maximizan tanto la calidad intrínseca de cada clúster como la distinción interclústeres.

La Tabla 6, presenta los resultados de diferentes evaluaciones empleadas para la evaluación de la calidad del agrupamiento mediante clústeres y que permiten concluir que la mejor alternativa es la obtenida por K-means a pesar de que la alternativa jerárquica presenta valores cercanos.

El coeficiente silueta será mejor entre mayor sea el valor, para el índice de Davies Bouldin este será mejor entre más cercano a cero se este y para el índice de Calinski Harabasz será mejor entre mayor sea el valor, en los tres casos la clusterización generada mediante K-means resulta mejor.

Tabla 6. resultados de la evaluación de la agrupación mediante clústeres.

Método	Coefficiente de silueta	Índice de Davies Bouldin	Índice de Calinski_Harabasz
K-means	0,40	0,75	215,60
Jerárquico	0,37	0,85	194,62

4.3. Etapa 2. Generación de la topología base.

A partir de la capa con los ejes en los que sería viable la instalación de tubería y la sectorización inicial de la RDA con los centroides que representan a cada uno de los sectores, se realizó una primera selección de los ejes con una mejor posibilidad de conectar a los sectores y que por su localización e interconexión con otros ejes resultaban convenientes.

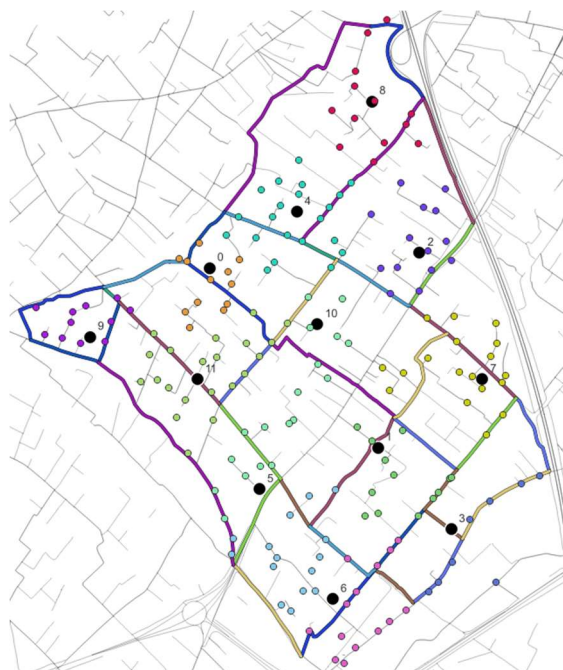


Figura 31. Primera selección de los ejes principales para la topología base.

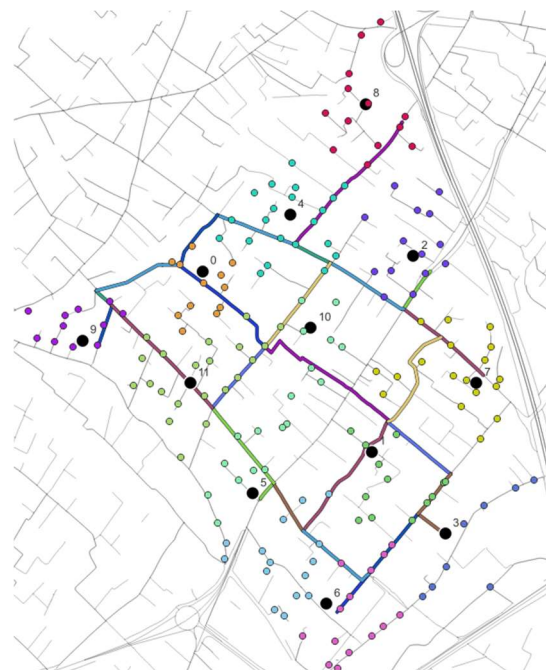


Figura 32. Definición de los ejes principales para la topología base.

La Figura 31 presenta la primera selección de ejes para la RDPA a partir de la cual se realiza un nuevo proceso de selección. En este segundo proceso se consideró la menor longitud del eje, la cercanía a los nudos que representan a los sectores y cuidando que se consiga una configuración de una red mixta (malla con brazos). La Figura 32, presenta los ejes seleccionados para el análisis del trazado y que podrían definirse como principales.

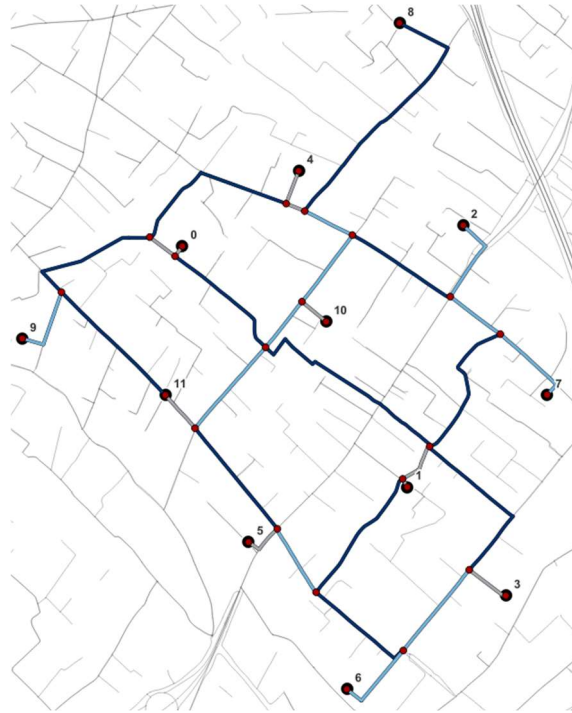


Figura 33. topología base.

A partir de los ejes principales, se generó la topología base completa con las aristas y vértices que definen el grafo, el resultado se presenta en la Figura 33. Finalmente, se crea el archivo de texto en el que se incluyen las características de los elementos que posteriormente representarían la red, como la longitud de las líneas, la cota y demanda de los nudos.

4.4. Etapa 3. Generación de los posibles trazados.

La generación de los trazados se realiza mediante el código desarrollado para tal fin en VBA de Excel, cuya entrada es el archivo de texto con la topología base. Debido a que el código utiliza la API de Epanet, el archivo de entrada se configura como un archivo “inp” de Epanet el cual representa a la topología base como una red del tipo mixta.

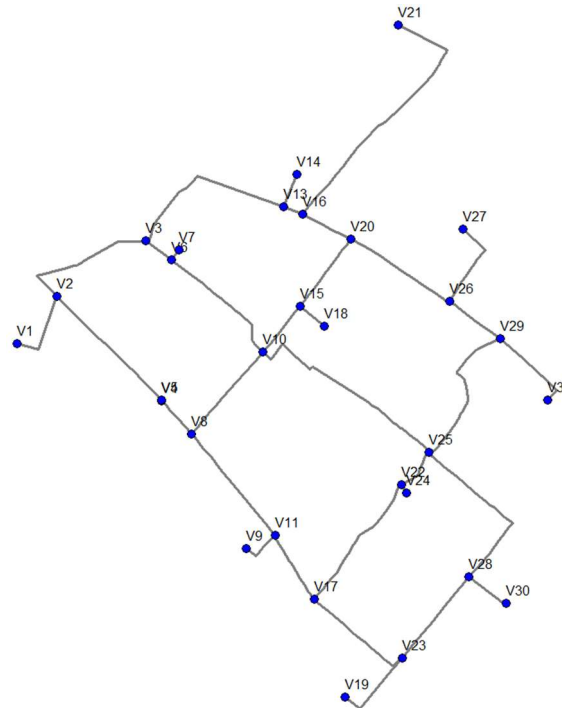


Figura 34. topología base en Epanet.

La Tabla 7 presenta la cantidad de elementos que hacen parte del conjunto V de vértices y E de aristas del grafo base G . En esta configuración no se contabiliza el nodo que representan a la fuente para el suministro de agua y su conexión con la red.

Tabla 7. Conjunto de elementos que componen el grafo base G .

Número total de vértices $ V $	30
Número de vértices extremos	12
Número de aristas $ E $	33

Una vez se ejecuta el código se obtiene la cantidad de árboles de expansión posibles, que para este caso fue de 4484 subgrafos. Además, se efectúa el proceso de simplificación del grafo de forma automática, cambiando las matrices y listas que definen al grafo G , por las nuevas representaciones del grafo G' para la elaboración de las combinaciones de aristas. El tiempo de cálculo necesario para generar la salida e impresión de los resultados en los archivos de salida fue de 15 minutos aproximadamente.

Tabla 8. Fragmento de los resultados que definen los árboles de expansión válidos.

Conjunto E' que definen un árbol de expansión T valido.	Conjunto de aristas $E \setminus E'$ que no conforman T
3;5;21;4;27;2;25;31;1;35;17;34;6;11;22;15;19	16;23;24;28;30
3;5;21;4;27;2;25;31;1;35;17;34;6;11;22;15;23	16;19;24;28;30
3;5;21;4;27;2;25;31;1;35;28;11;6;17;22;15;19	16;23;24;30;34
3;5;21;4;27;2;25;31;1;35;28;11;6;17;22;15;23	16;19;24;30;34

3;5;21;4;27;2;25;31;1;35;28;34;6;11;22;15;19	16;17;23;24;30
3;5;21;4;27;2;25;31;1;35;28;34;6;11;22;15;23	16;17;19;24;30
3;5;21;4;27;2;25;31;1;35;28;34;6;17;22;15;19	11;16;23;24;30
3;5;21;4;27;2;25;31;1;35;28;34;6;17;22;15;23	11;16;19;24;30
3;5;21;4;27;2;25;31;16;35;17;34;6;11;22;15;19	1;23;24;28;30
3;5;21;4;27;2;25;31;16;35;17;34;6;11;22;15;23	1;19;24;28;30

La Tabla 8, presenta un fragmento del conjunto de resultados que se obtuvieron en la ejecución de la rutina para esta etapa de la metodología. La salida del código es un par de listas, la primera que contiene cada uno de los conjuntos de aristas $E'(T) \subseteq E(G)$, que definen un árbol de expansión válido. La segunda lista está conformada por los conjuntos de aristas $E \setminus E'$, tal que hacen parte del conjunto de aristas del grafo base pero que no conforman el árbol de expansión.

4.5. Etapa 4. Cálculo de redes.

Previo al dimensionado y cálculo de las posibles redes, el aplicativo desarrollado realiza la conformación de cada una de las redes a partir de los trazados definidos por los árboles de expansión y el “inp” con la topología base, generando un nuevo archivo con el modelo hidráulico base de la red a analizar. No obstante, el número de redes a calcular no es equivalente al número de árboles de expansión posibles, considerando que se analiza cada nudo como nudo de entrada a la red lo que conforma tantas posibilidades como nudos exista por cada alternativa de trazado. En este caso el número de redes se incrementa a 146.520 al tratarse de una red conformada por 30 nudos.

Tabla 9. Elementos que conforman la RPDA.

Nudos de consumo	12
Nº total nudos de caudal	30
Nudos de presión	1
Nº total nudos	31
Tuberías	30

La Tabla 9, presenta los elementos que conforman a la RPDA. La Figura 35 y la Figura 36, presentan dos ejemplos de configuración para la RPDA generadas siguiendo la metodología, previo a su dimensionado. Cada una de las alternativas que se generan presentan particularidades en su configuración, que influyen en el consumo energético de la red y los excesos de energía y por consiguiente en el coste de la red.

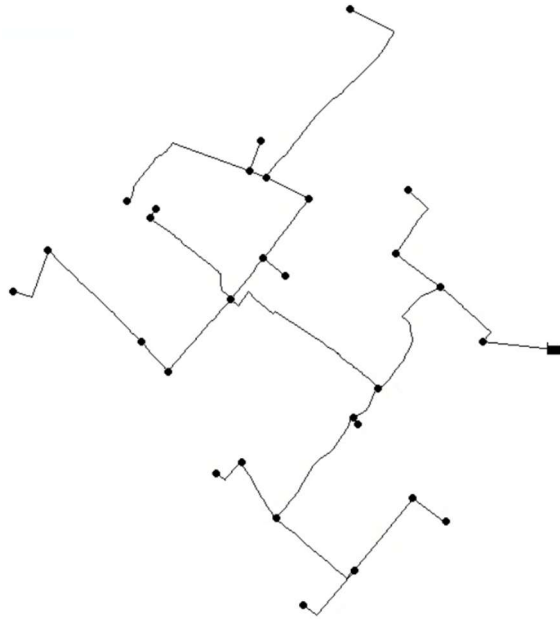


Figura 35. Ejemplo de trazado generado a partir de un árbol de expansión del grafo G.

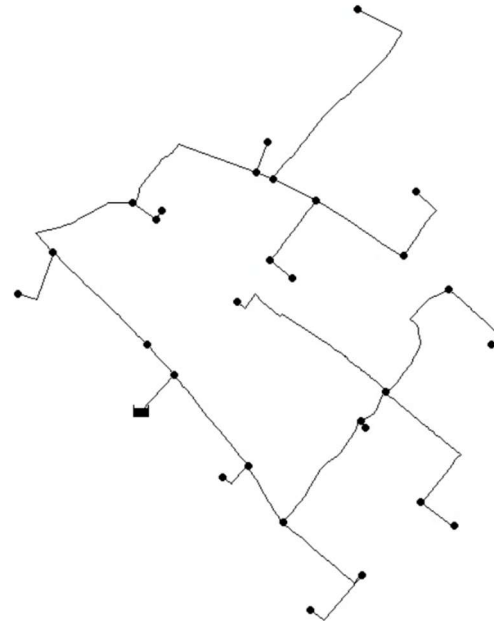


Figura 36. Ejemplo de trazado generado a partir de un árbol de expansión del grafo G.

El dimensionado y cálculo de la red, se realiza de forma automática a partir de la pendiente de diseño. Para este caso se probó un rango de 10 pendientes entre 1 y 10 m/km, a fin de tener valores para análisis futuros. No obstante, el rango en consideración puede ser menor si se considera que la pendiente óptima de diseño es de aproximadamente 1,1 m/km. La Tabla 10. presenta los valores necesarios empleados en el cálculo de la pendiente óptima para el caso de estudio, para el cual se estima una vida útil de 30 años y un coste promedio de la energía de 0,223 €/kW·h incluido el valor de la potencia contrada.

Tabla 10. Pendiente óptima para la instalación.

Q [l/s]	242,33
f	0,015
m	1,25
η	0,8
γ (N/m³)	9800,00
K_{mat}	540,20
T (h)	24,00
CkW.h (€/kW · h)	0,19
CkW (€/kW)	12,04
r	1%
i	5%
N (años)	30,00
J_{opt} (m/Km)	1,12

En total se evaluaron 1.465.200 redes, teniendo en cuenta que para cada combinación de red y nudo de cabecera se prueban 10 pendientes de diseño para la

red. El tiempo de cálculo empleado por el algoritmo fue de aproximadamente 11 horas, generando como salida un archivo con los resultados para cada una de las redes numerados de acuerdo con el índice de su red, el nudo de cabecera y la pendiente de diseño, con 61 variables y más de 80 millones de datos.

Tabla 11. Parámetros de diseño de la RPDA

Diámetro mínimo (mm)	50
Material de la tubería	Hierro fundido
Rugosidad del material	0,1
Presión de servicio (m.c.a)	20

La Tabla 11, muestra los valores de los parámetros de diseño empleados para el dimensionado y cálculo de la red.

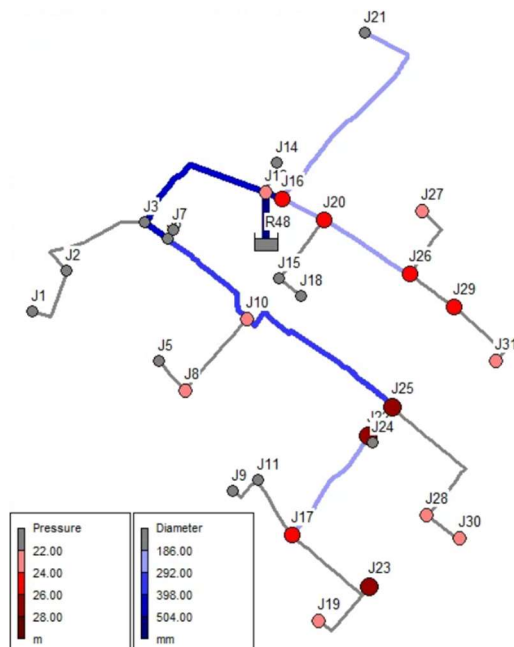


Figura 37. Ejemplo RPDA dimensionada y calculada con cabecera en el nudo 13.

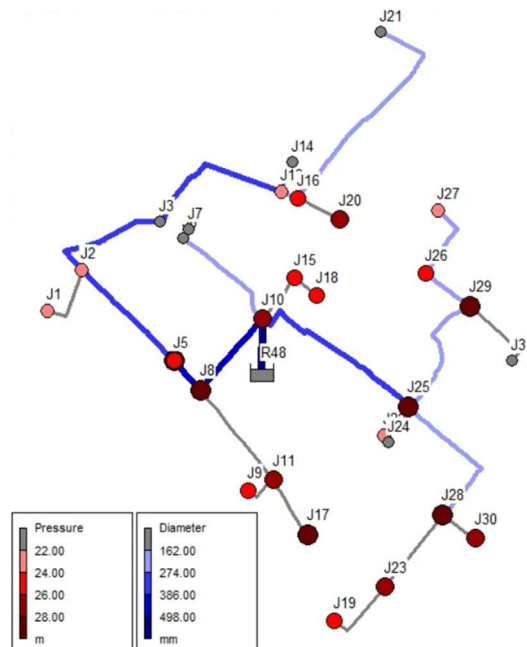


Figura 38. Ejemplo RPDA dimensionada y calculada con cabecera en el nudo 10.

La Figura 37 y Figura 38 corresponden a dos propuestas de RPDA donde se presentan las presiones y diámetros que tendría la red para la condición de diseño. A mayor intensidad de rojo los nudos tienen mayor presión y por ende mayor exceso de energía, por su parte tuberías con un azul más intenso corresponden a diámetros más grandes en la red.

4.6. Etapa 5. Análisis de resultados y selección de alternativas de trazados.

Con base en el archivo generado en la etapa anterior se crea una base de datos (DataBese) en Rstudio, para realizar el análisis de los resultados. Primero se revisa el rango en el que oscilan las energías requeridas en cabecera (E_{rc}) obtenidas para

todas las configuraciones de redes (ver Tabla 12) y su variación de acuerdo con la pendiente de diseño.

Tabla 12. Rango en que oscilan las E_{rc} para las RPDA.

E_{rc}	[kWh]
Mínima	125.838
Media	155.887
Máxima	264.564

La Figura 39 y la Figura 40, presenta el histograma de frecuencias y el diagrama de cajas y bigotes para la E_{rc} respectivamente, la forma de los gráficos muestra una distribución sesgada hacia la izquierda o negativa, donde los valores tienden a estar por debajo de los 156 Kwh. No obstante, también se identifica que existe un grupo amplio de soluciones con valores de energía mínimos o muy cercanos.

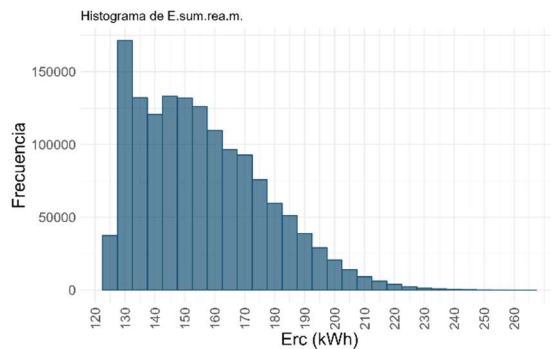


Figura 39. Histograma de frecuencias E_{rc} .

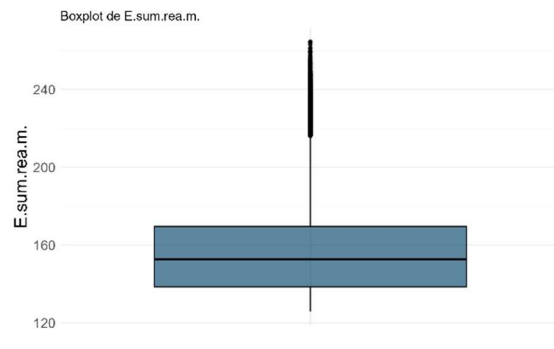


Figura 40. Diagrama de cajas y bigotes para la E_{rc} .

Del total de resultados se seleccionaron las dos configuraciones de RPDA (Figura 41 y Figura 42) que presenta el mínimo valor de E_{rc} . Sin embargo, las características de partida del problema (un terreno con poco desnivel y una distribución relativamente homogénea de los nudos) generan un número significativo de soluciones con valores de E_{rc} muy semejantes al mínimo, por lo que es conveniente incluir otras variables como el exceso de energía en los nudos $\Sigma\Delta H_i$ y el coste total de la red C_{Total} , para poder acotar el número de soluciones y definir las más convenientes energéticamente.

La Figura 41 y Figura 42, presentan configuraciones semejantes, con un tramo principal del cual se derivan subredes hacia cada uno de los nudos y con apenas un cambio en el tramo que abastece al nudo J24. Es evidente que cualquier otra configuración con características similares, nudo crítico en J21, cabecera en J13 y un tramo principal desde el nudo J13 hasta el nudo J14, requerirá una E_{rc} similar.

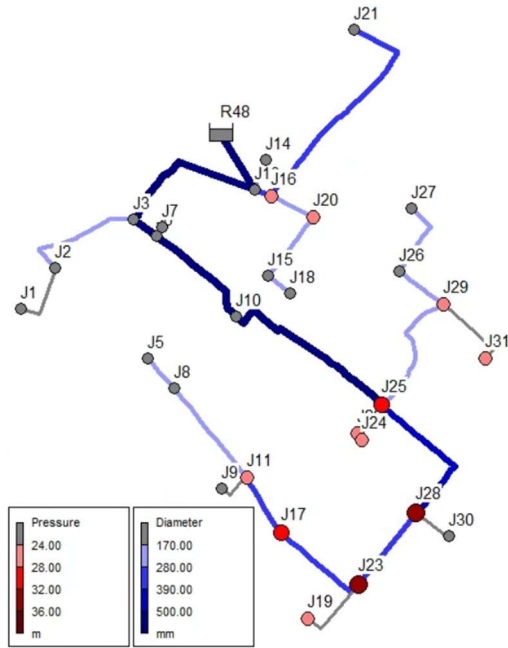


Figura 41. Configuración 4453 de la RPDA con mínimo valor de E_{RC} .

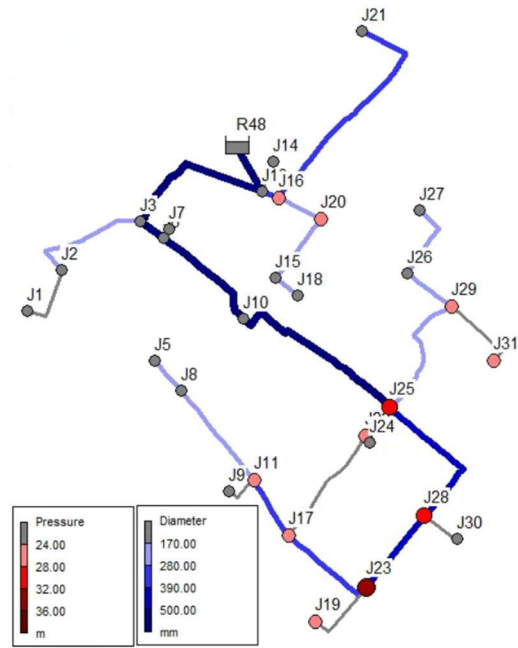


Figura 42. Configuración 4471 de la RPDA con mínimo valor de E_{RC} .

La Tabla 13 presenta el resumen de los resultados para las dos configuraciones con mínimo valor de ERC obtenidas al aplicar la metodología, ambas con $J_{diseño}$ de 2 mca/km, nudo de entrada en J13 y crítico en J3. De esta se es posible concluir que los costes de energía para el periodo de evaluación son iguales al tener un valor de E_{rc} semejante, no obstante, la Configuración 4453-C13-J2 podría ser más deseable en términos de coste de capital y reducción del exceso de carga en los nodos.

Tabla 13. Resumen resultados para la configuración 1 y 2 con menor E_{RC}

Configuración		4453-C13-J2	4453-C13-J2
Energía requerida en cabecera (kWh)	E_{RC}	125.838	125.838
Coste energía (euros)	$C_{energía}$	7389542	7389542
Coste capital (euros)	$C_{capital}$	910283.4	944219.1
Coste total (euros)	C_{Total}	8299826	8333761
Σ exceso carga en los nudos (m)	$\Sigma \Delta H_i$	121.92	114.87

Considerando que el número de soluciones con un valor de E_{rc} similar al mínimo es del orden de 40.000, se hace evidente que es necesario ampliar el análisis a fin de identificar el grupo de soluciones que sean más convenientes.

La Figura 43 representa el gráfico de barras apiladas que relaciona la frecuencia con la que se presenta un valor en una categoría de energía requerida y el valor de la pendiente de diseño. Como es natural es más frecuente que soluciones con valores de energía más bajos se generen a partir de pendientes de diseño menores. La Figura 44 presentan la dispersión de los valores de E_{rc} de acuerdo con la pendiente de diseño, en el que valores bajos de pendiente conducen a soluciones con un menor consumo energético. No obstante, existen soluciones con bajos valores E_{rc} para pendientes de diseño más altas que la óptima que resultan interesantes.

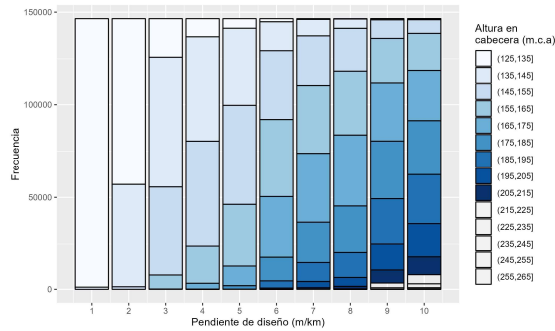


Figura 43. Variación de la Erc de acuerdo con la pendiente de diseño.

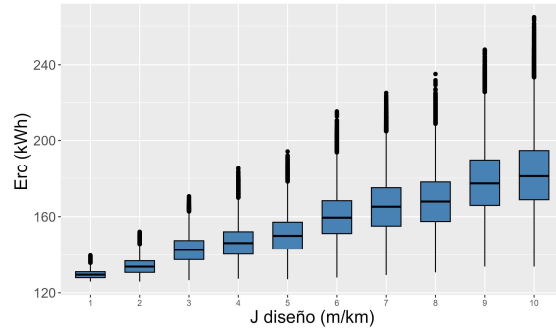


Figura 44. Diagrama de cajas y bigotes para la pendiente y la Erc.

La figura 47, corresponde a la topología base de la red. En escala de azul se presentan los nudos clasificados de acuerdo con su cota, donde el tono azul más intenso representa a los nudos con mayor cota y el azul claro a los nudos con menor cota. Las tuberías se clasifican de acuerdo con su longitud, en donde el gris más intenso representa a los tramos de mayor longitud y tono claro a los de menor longitud.

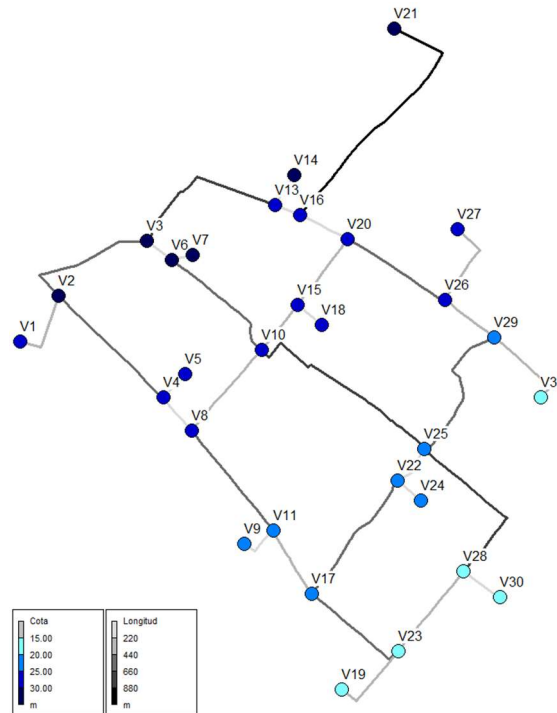


Figura 45. Topología base con distribución de cotas en los nudos y longitud en los arcos.

La Figura 46 presenta el histograma de frecuencias para el nudo crítico y todas las posibles RPDA. Notablemente el nudo J21 es el nudo que para la mayoría de los casos resulta ser el más restrictivo. Sin embargo, la Figura 47 muestra como soluciones con nudo crítico en J10, J2, J20, J3, J4 y J7 pueden también conducir a soluciones con un valor bajo de ERC.

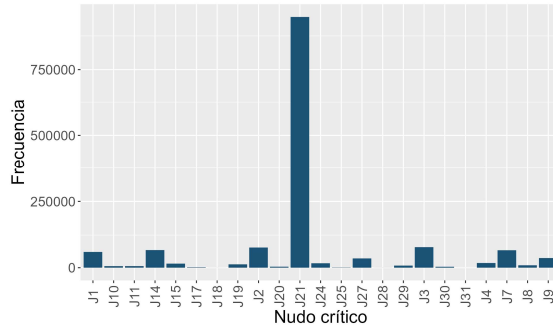


Figura 46. Histograma para el nudo crítico.

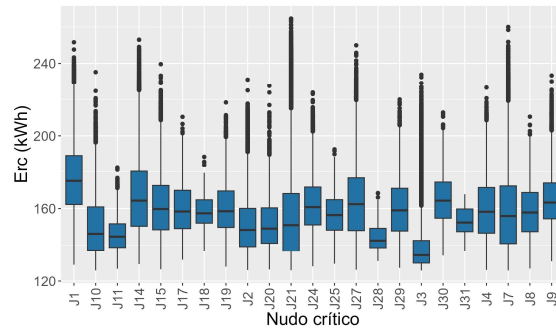


Figura 47. Diagrama de cajas y bigotes para el nudo crítico y la energía mínima requerida en cabecera.

Comparando este resultado con lo mostrado en la Figura 45, se constata que los nudos con mayor cota tienden a ser nudos críticos en un mayor número de configuraciones de red. En cuanto al nudo de entrada a la red y de acuerdo con la Figura 48 y Figura 49, las soluciones con nudo de cabecera en J13, J14, J16 y J20 presentan una mayor tendencia a generar trazados con un menor requisito de energía en cabecera.

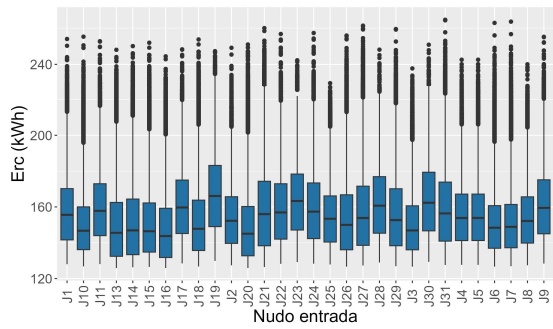


Figura 48. Diagrama de cajas y bigotes para el nudo de entrada y la energía mínima requerida en cabecera.

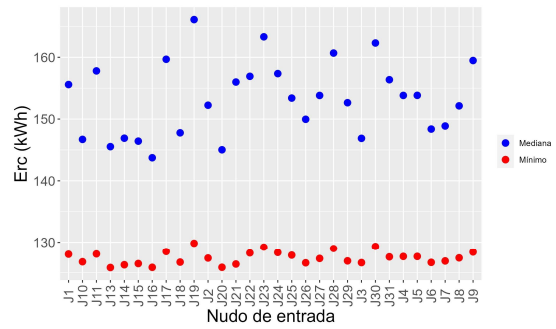


Figura 49. Valores mínimos y mediana de ERC para los nudos de entrada a la red.

De nuevo, comparando con la Figura 45 se refleja una tendencia a que el nudo de cabecera de la red se ubique entre los dos nudos más altos, el nudo J21 (32,12 m.s.n.m) y el nudo J3 (32,15 m.s.n.m), intentando ser equidistante a ambos.

De la base de datos de resultados, se seleccionaron las soluciones que se consideran mejores asociadas a la variable E_{RC} , comprendidas en el rango entre 125 y 127 kWh, equivalente a el mínimo valor de E_{RC} obtenido y un delta de energía de 2kWh. El número de soluciones que cumplen con este criterio se reduce a 21909, equivalente al 1.5% del total de soluciones.

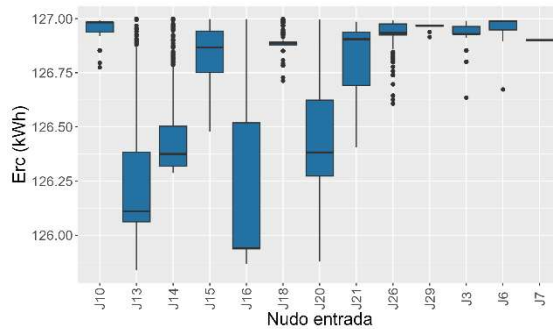


Figura 50. Diagrama de cajas y bigotes para nudo de entrada a la red y E_{RC} , para las mejores soluciones.

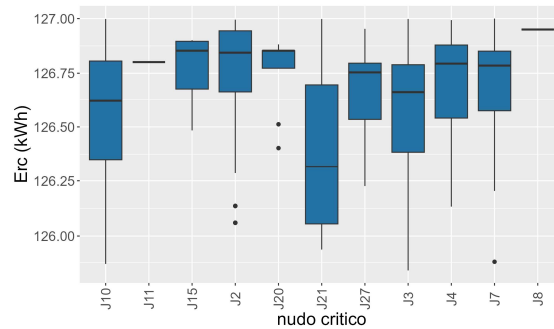


Figura 51. Diagrama de cajas y bigotes para nudo crítico de la red y E_{RC} , para las mejores soluciones.

A partir del grupo de mejores soluciones, se obtuvieron los diagramas de cajas y bigotes para el valor de la E_{RC} con relación con el nudo de entrada y el nudo crítico, presentados en la Figura 50 y la Figura 51. De estos se infiere que existe una tendencia a que las soluciones con nudo de cabecera en J16 presenten un menor valor de E_{RC} , al igual que soluciones con nudo crítico en J21, por lo que soluciones con estas condiciones son candidatas a ser convenientes energéticamente.

Como ya se comentó, el nudo J21 es el segundo nudo más alto, si se compara la longitud del tramo que existe entre el nudo del nudo J16 (nudo de cabecera) y el nudo J21 y J3, se encuentra que el nudo J21 resulta más crítico al estar a 1097,02m del nudo j16, respecto a 784.58m del tramo entre J16 y J21.

Como se señaló previamente, para seleccionar la o las alternativas de trazado que resulten energéticamente convenientes, no solo se consideraron las soluciones con menor valor de E_{RC} , además se incluyeron dentro el análisis el valor de la energía de exceso ($\Sigma\Delta H_i$), para valorar la forma en que la red energéticamente se acopla mejor a las condiciones del terreno, y el coste total (C_{total}) de la instalación para el periodo de vida útil de la instalación que incluye el coste de construcción y el coste de la energía consumida.

La Figura 52 y la Figura 53 presentan los diagramas de cajas y bigotes para los criterios adicionales de selección la $\Sigma\Delta H_i$ y C_{total} , comparados respecto al nudo de entrada a la red. Para el primer criterio adicional, las soluciones con nudo de cabecera en J13, J16 y J21 presentan una significativa tendencia a generar soluciones con un menor valor de exceso de energía en los nudos. Esto se relaciona con soluciones que requieren menos energía para funcionar y por tanto tienen menores excesos de energía en los nudos, siendo más eficientes.

Para el segundo criterio adicional, las soluciones con nudo de entrada en J21, J26, J29, J3, J6 y J7 resultan tener un mayor coste total. En términos generales al tener un valor E_{rc} semejante, el valor determinante es el $C_{capital}$, y no el coste de la energía; por consiguiente, al ser los caudales invariantes en el tiempo, el dimensionado de la red dependerá en gran medida de la longitud de la ruta crítica, y esta a su vez al ser esta la ruta que concentra los caudales será la que tenga mayores diámetros, lo que a su

vez se relaciona con mayores costes por metro lineal de tubería. Rutas críticas más cortas corresponderán con redes económicas en $C_{capital}$ y un menor C_{total} .

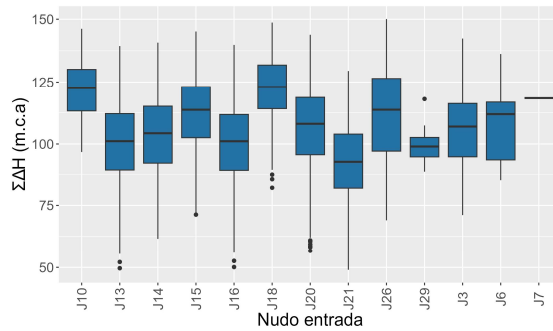


Figura 52. Diagrama de cajas y bigotes para nudo de entrada a la red y $\Sigma\Delta H_i$, para el grupo de las mejores soluciones.

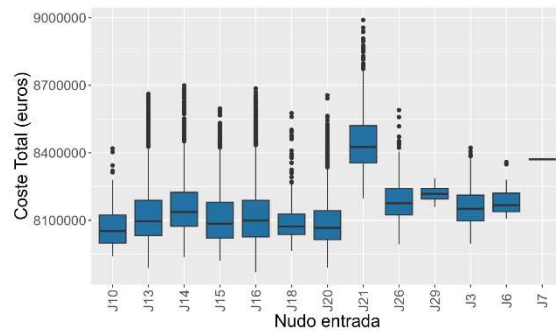


Figura 53. Diagrama de cajas y bigotes para nudo de entrada a la red y C_{Total} , para el grupo de las mejores soluciones.

Para identificar las soluciones más convenientes, se analizó la relación existente entre $\Sigma\Delta H_i$ y C_{Total} para el grupo de mejores soluciones. La Figura 54 presenta la nube de puntos generada para las variables consideradas. Debido al número de puntos y a que el objetivo es identificar las soluciones óptimas por tener valores mínimos de $\Sigma\Delta H_i$ y C_{Total} se filtraron los resultados generando un frente óptimo de Pareto (ver Figura 55).

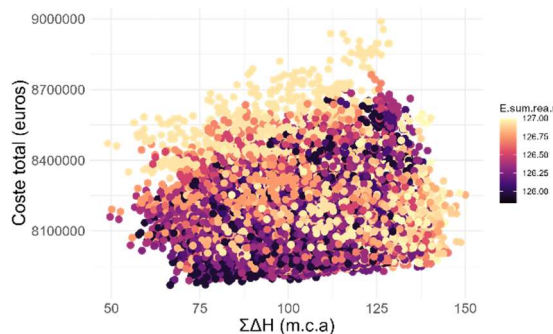


Figura 54. Diagrama de nube de puntos para $\Sigma\Delta H_i$ y C_{Total} , para el grupo de soluciones con menor E_{RC} .

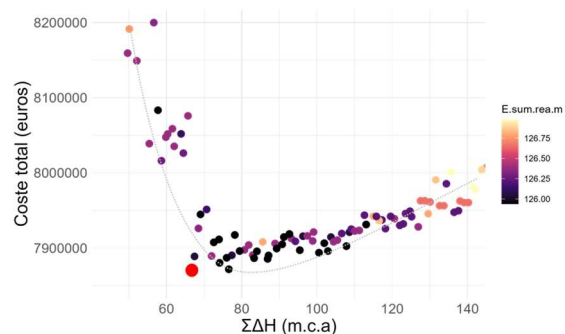


Figura 55. Frente Óptimo de Pareto para $\Sigma\Delta H_i$ y C_{Total} , para el grupo de soluciones con menor E_{RC} .

A partir del frente generado con las soluciones que se considerarían como óptimas y no dominadas, se identifican aquellas que se encuentren más cercanas al origen de los ejes del gráfico (por tener valores mínimos para las dos variables), por lo que podrían considerarse como óptimas.

Como resultado de la aplicación, se seleccionaron las cuatro mejores alternativas de trazado para la RPDA a partir de la Figura 55. Estas corresponden a los trazados más convenientes energéticamente, al tener un valor mínimo de E_{RC} , presentar menores excesos de energía en los nodos $\Sigma\Delta H_i$ y requerir un menor $C_{capital}$ y por consiguiente un C_{total} bajo. Esto se traduce RDA más eficientes y sostenibles en el tiempo.

Las 4 alternativas tienen configuraciones muy semejantes, el nudo de entrada es J16 y nudo crítico es el J21, el valor de la pendiente de diseño $J_{diseño}$ está en el rango entre 1 y 2 m.c.a./Km. La Tabla 14, presenta el resumen de resultados para los trazados de

la RPDA energéticamente convenientes y las Figura 56, Figura 57, Figura 58 y Figura 59 presentan la topología de las redes.

Tabla 14. Resumen de resultados para los trazados de la RPDA energéticamente convenientes.

Id. Configuración.	H (m.c.a)	ERC (kWh)	Cenergía (euros)	Ccapital (euros)	Ctotal (euros)	$\Sigma\Delta H_i$ (m.c.a)
2633-C16-1	53,218	125,939	7.395.473,08	474.619,60	7.870.092,67	66,63
2609-C16-1	53,218	125,939	7.395.473,08	476.124,21	7.871.597,29	76,56
2645-C16-1	53,218	125,939	7.395.473,08	484.721,49	7.880.194,57	74,1
2615-C16-1	53,218	125,939	7.395.473,08	489.983,97	7.885.457,05	86,88

A pesar de que la elección entre estas configuraciones dependerá de las prioridades específicas del proyecto y los objetivos del diseño de la red de abastecimiento de agua, del grupo de soluciones resalta la configuración 2633-C16-1, tener el coste total mínimo y menores excesos de energía. Por consiguiente, se podría considerar como el mejor trazado para la RPDA y a base para realizar el diseño definitivo de la RDA completa, incluyendo los elementos adicionales que sean necesarios y realizar la conexión hacia todos los nudos de consumo en los diferentes sectores.

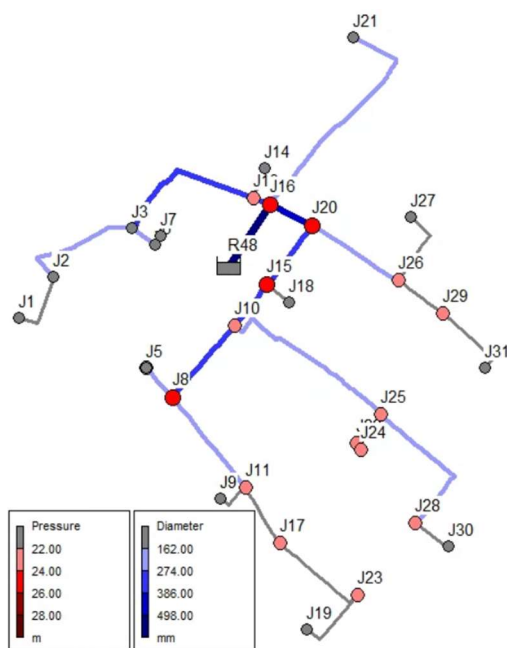


Figura 56. Trazado energéticamente conveniente 2633-C16-1

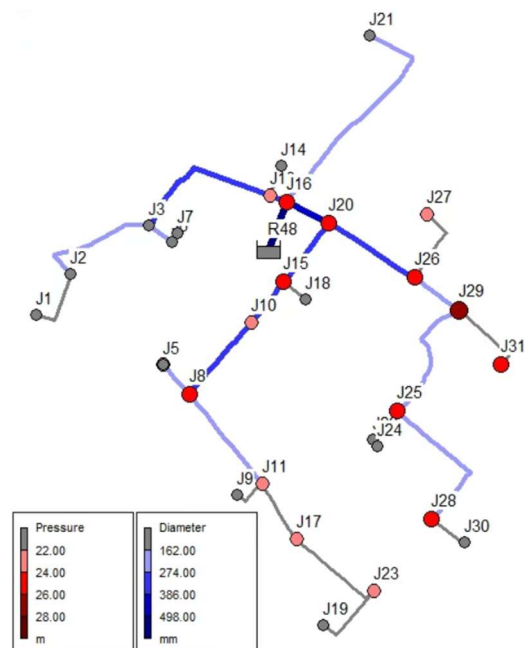


Figura 57. Trazado energéticamente conveniente 2609-C16-1

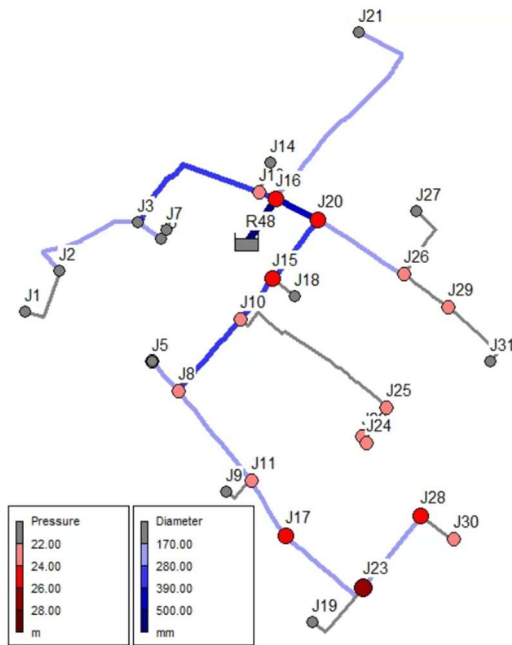


Figura 58. Trazado energéticamente conveniente 2645-C16-1

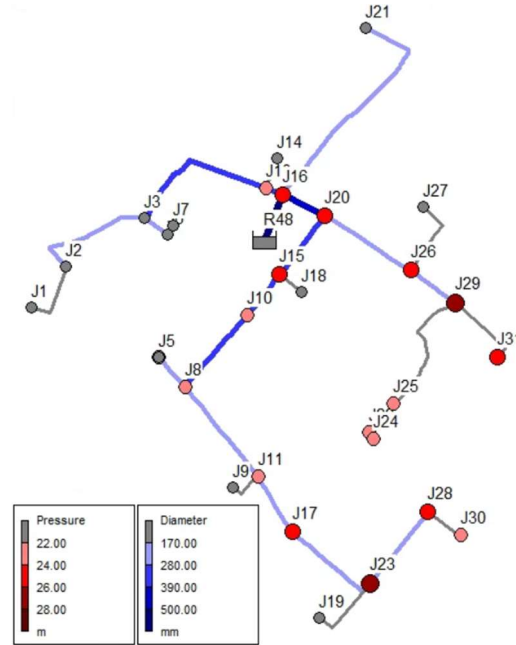


Figura 59. Trazado energéticamente conveniente 2615-C16-1

4.7. Discusión y análisis de resultados.

En este trabajo se ha desarrollado una metodología para la generación de trazados energéticamente convenientes en redes principales de distribución de agua, lo que minimiza los requerimientos energéticos, abarata los costes a lo largo de la vida útil del proyecto y permite conseguir sistemas más eficientes y sostenibles.

La aplicación en un caso de estudio ha conducido a obtener un grupo de soluciones que podrían considerarse como óptimas para el objetivo y las condiciones del problema. No obstante, como se demostró utilizar la energía requerida en cabecera (ERC) como criterio único para la definición del trazado, es insuficiente al conducir a un número considerable de soluciones con valores muy semejantes; por lo que es necesario que se analicen otras variables relacionadas con la eficiencia energética de la red para reducir el número de soluciones, en este caso la sumatoria de los excesos de energía ($\Sigma\Delta H_i$) y el coste total del proyecto (C_{Total}).

La metodología propone el análisis de dos variables que resultan definitorias para el consumo energético de la red, el punto de entrada a esta o cabecera y la pendiente de diseño. La definición del punto de entrada a la red se realiza a partir de los nodos que la componen, lo que trae como consecuencia que se incremente el número de soluciones a analizar. En este sentido, resulta conveniente reducir el número de posibilidades adicionales, a partir del descarte de nodos de acuerdo con sus características.

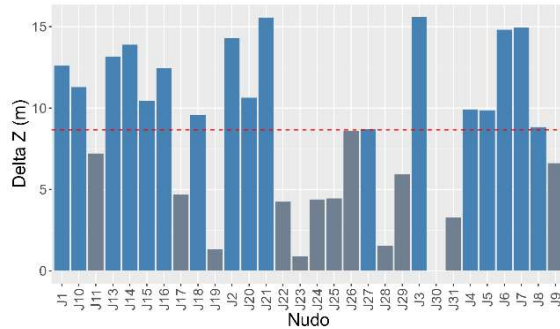


Figura 60. Delta Z ($Z_i - Z_{\min}$) de los nudos de la RPDA.

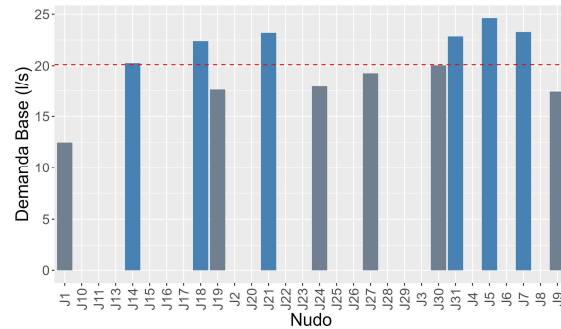


Figura 61. Demanda base en los nudos de consumo.

La Figura 60, presenta un gráfico de barras con la diferencia entre la cota de cada nudo y la cota del nudo más baja, y el valor medio de la cota de todos los nudos. Por su parte la Figura 61, presenta la demanda base en los nudos de consumo de la red y su media. Si se superponen los nudos con cota por encima de la media, y los nudos de entrada a la red con mayor frecuencia de generar redes con un valor bajo de ERC (Ver Figura 62) se observa que existe relación.

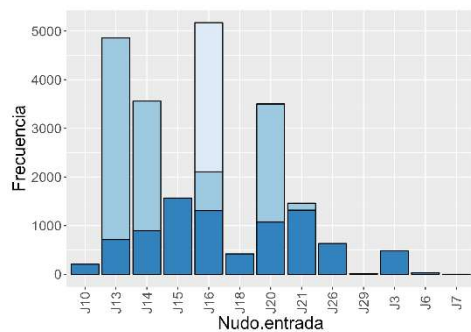


Figura 62. Gráfico de frecuencia de barras apiladas para valores de ERC y el nudo de entrada.

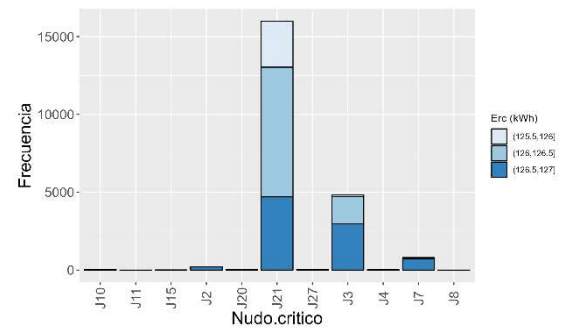


Figura 63. Gráfico de frecuencia de barras apiladas para valores de ERC y el nudo de entrada.

Según lo expuesto previamente, el número de soluciones a evaluar se reduciría significativamente, si se evalúan únicamente como nudo de entrada los 17 nudos más altos de red, pasando a 830.280 soluciones lo que reduce en un 43% el total de soluciones. En cuanto a la definición del nudo crítico (ver Figura 51), es claro que existe una relación con los nudos con cota más desfavorable, en este caso los nudos J21, J3 y J7, respecto a la demanda base los nudos J21 y J7 tienen valores asociados por encima del valor medio.

Respecto a la pendiente de diseño y como ya se mencionaba, si bien la tendencia es que valores de pendiente mayores conduzcan a soluciones con un mayor valor de ERC, existen configuraciones generadas a partir de pendientes de diseño superiores a la óptima con valores bajos de ERC que pueden resultar interesantes, si el objetivo es encontrar trazados con un coste total bajo y de capital mínimo.

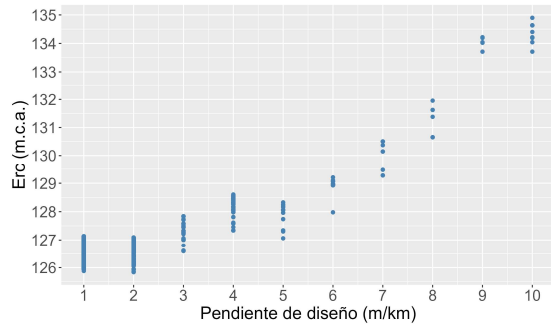


Figura 64. Conjunto de mejores con mínimo valor de E_{RC} para cada $J_{diseño}$.

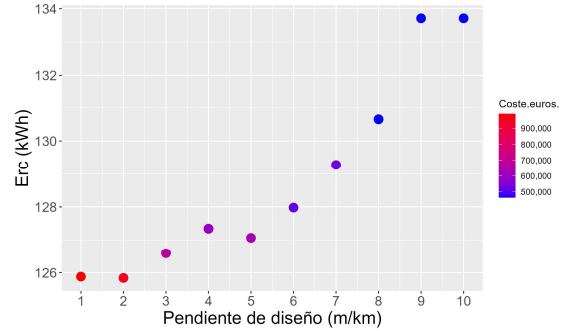


Figura 65. Soluciones con mínimo valor de E_{RC} para cada $J_{diseño}$.

La Figura 64 presenta el conjunto de soluciones con los mejores valores de E_{RC} que para cada $J_{diseño}$ analizada, y a partir de las cuales se buscaría la configuración con los costes mínimos para cada valor de pendiente. La Figura 65 extrae las soluciones óptimas de cada uno de los grupos anteriores y las categoriza en términos $C_{capital}$, en este caso la mejor solución sería la generada a partir de la pendiente 9 m.c.a./km. Sin embargo, si la comparamos con la obtenida por la metodología propuesta, se hace evidente que no corresponde a óptimo global en términos de coste Total (ver Tabla 15), esto debido a que pueden existir otras configuraciones más económicas para otros valores de pendiente o aun para el mismo valor de pendiente.

Tabla 15. Resumen de resultados para la solución óptima con $J_{Diseño}$ 9 m.c.a./km.

Id.	H	E_{RC}	$C_{energía}$	$C_{capital}$	C_{total}	$\Sigma\Delta H_i$
Configuración.	(m.c.a)	(kWh)	(euros)	(euros)	(euros)	(m.c.a)
2761-C13-9	56,479	133,712	7.851.924	465.050,1	8.316.974	113,97

Dadas las múltiples variables interdependientes y su alta dimensionalidad, obtener un único trazado que se considere como óptimo y conecte a todos los nudos de la red puede ser utópico, especialmente en redes complejas y con una magnitud considerable, donde la cantidad de posibles combinaciones que desborda la capacidad computacional actual o representa largos periodos de cálculo. Por lo tanto, es conveniente optar por análisis que permitan seleccionar configuraciones para la RPDA que muestren ventajas energéticas evidentes y a partir de esta base, extender el trazado hasta cada nodo de la red.

No obstante, la aplicabilidad de esta metodología se encuentra condicionada por diversos factores como la suposición de que es factible entrar por cualquier punto a la red, la existencia de una única fuente de suministro y el análisis en condiciones estacionarias son restricciones significativas. La inclusión de estas variables podría sobrepasar la capacidad de análisis dentro de los límites actuales de potencia de cómputo y tiempo de cálculo. A pesar de esto, iniciar el proceso con un trazado "optimizado" con relación a un conjunto de variables específicas puede influir positivamente en la obtención de soluciones más robustas cuando sea necesario considerar estas condiciones restrictivas.



En adición, la metodología no aborda el envejecimiento de la red ni se consideran pérdidas operativas como las fugas, factores que podrían impactar el balance energético de la red. Estos aspectos merecen atención en futuras investigaciones para lograr una comprensión más completa de la eficiencia energética en redes de agua.

Capítulo 5.

Conclusiones y desarrollos futuros.

5.1. Conclusiones.

A partir del estudio realizado y la aplicación de la metodología en el caso de estudio se obtuvieron las siguientes conclusiones:

1. La metodología propuesta representa una novedosa alternativa para el diseño de redes principales de distribución de agua (RPDA), que sean energéticamente convenientes. Mediante un análisis detallado de los trazados posibles, esta metodología busca optimizar la energía topográfica (pérdidas estructurales), al considerar factores como el punto de entrada a la red y la pendiente de diseño. Los resultados obtenidos demuestran su capacidad para generar configuraciones de RPDA, que no solo se ajustan mejor al entorno de implantación, sino que también exhiban valores de energía topográfica inevitable que podrían considerarse mínimos.
2. La metodología emplea técnicas provenientes del machine learning, específicamente la clusterización para la generación inicial de los sectores que conformaran la RPDA. Además, se realiza la configuración de los trazados a partir de algoritmos desarrollados a partir de la teoría de grafos, esquematizados a través de flujo gramas para la aplicación en casos de estudio. El uso de estos algoritmos resulta en la reducción del tiempo computacional requerido para llegar a una respuesta óptima.
3. Dada la versatilidad de la técnica de clusterización, es posible que, a partir de bases de datos disponibles en fuentes oficiales e insumos básicos, analizar conjuntos de datos y agruparlos de acuerdo con sus características, lo que en RDA resulta conveniente para la conformación de sectores homogéneos y espacialmente distintos, que a través de la estandarización y el estudio de los componentes principales capturan la información más relevante.
4. En un entorno ideal, la energía inyectada en la red se destinará completamente a satisfacer las necesidades energéticas mínimas de los nudos. No obstante, en condiciones reales, las pérdidas energéticas operaciones y estructuras no pueden ser eliminadas por completo. En consecuencia, es indispensable que incluso en etapas de proyección de la red

se realice un análisis que contemple estos factores, para identificar los trazados más eficientes energéticamente para la RPDA, resultando en sistemas más sostenibles.

5. La definición del nudo de entrada a la red es un aspecto que pocas veces se considera, su definición podría realizarse a partir del grupo de nudos con mayor cota, pero como se demostró es condicionante en la generación de trazados energéticamente más eficientes y económicos. A pesar de que este punto no concuerda con la posición de la fuente de suministro, es conveniente que se estudie a fin de reducir los costes.
6. El valor de la pendiente de diseño que define a una red energéticamente eficiente varía según el trazado que se realice, con tendencia a obtener trazados con un menor requerimiento energético a partir de valores de pendiente mínimos. Sin embargo, cada valor de pendiente tiene asociado un trazado óptimo específico, por lo que pueden resultar interesantes analizar pendientes con valores mayores, priorizando aspectos como el coste de capital.
7. Un aporte significativo es la definición de red principal para la RDA, a partir de la sectorización inicial realizada. Siendo esta la infraestructura que, partiendo desde la cabecera, sirve para conducción del agua de forma exclusiva a los sectores de la red, sin incluir otras derivaciones. Esta definición resulta útil a la hora de realizar la categorización del sistema y gestionar el servicio de agua eficientemente.
8. La sostenibilidad de un servicio de agua está condicionando por el coste energético durante su operación, por consiguiente, se debe considerar junto con el coste de capital a la hora de evaluar un proyecto. Incluir este aspecto en a la definición de la red energéticamente conveniente, no solo garantiza que tenga un consumo bajo de energía, sino que también, sea construible y sostenible a lo largo del periodo de operación.
9. Encontrar una solución óptima única, podría no ser factible dadas las limitaciones prácticas del problema. Maximizar la eficiencia energética en las redes de agua, donde se considere además de los aspectos mencionados en la metodología otros aspectos, como múltiples fuentes de abastecimiento, almacenamientos, etc. Representa un desafío significativo, aún más si se intenta definir la totalidad del trazado, debido a la multidimensionalidad del problema, la interrelación de las variables y al elevado coste computacional.
10. El análisis detallado de la topología y el estudio energético en las redes de distribución de agua es una herramienta indispensable para alcanzar los objetivos de eficiencia y sostenibilidad en los servicios de agua y en la gestión de los recursos hídricos y energéticos.

5.1. Desarrollos futuros

Este trabajo presenta un enfoque para la determinación del trazado de la red principal en redes de distribución de agua que integra el análisis energético. Este enfoque ha demostrado ser eficaz para generar trazados energéticamente eficientes, que pueden contribuir a la reducción de los costes operativos y a la creación de sistemas más sostenibles. Sin embargo, el desarrollo deja aspectos que se pueden analizar a futuro, algunos de estos son:

El análisis de más casos de estudio tanto en proyección de redes como en operación a fin de comprobar los resultados y analizar la influencia de variables nuevas, como el estado de las tuberías.

Incorporar al análisis del trazado el estudio de alternativas de recuperación energética y otras alternativas para mejorar su eficiencia.

Considerar la resiliencia de la red en el análisis de la definición del trazado. Esto podría ayudar a garantizar que la red siga proporcionando un servicio de agua fiable, incluso en condiciones adversas.

Utilizar modelos no estacionarios para reflejar cambios en la demanda de los nodos a lo largo del tiempo. Esto podría ayudar a mejorar la precisión de la planificación de la red.

Incorporar la eficiencia hídrica en el sistema operativo de la red. Esto podría proporcionar una comprensión más completa de la sostenibilidad del sistema en su conjunto.

Capítulo 6.

Referencias

- [1] UN, CEPAL, «Perspectivas de la población mundial 2019: metodología de las Naciones Unidas para las estimaciones y proyecciones de población», 04 2020, [En línea]. Disponible en: <https://hdl.handle.net/11362/45989>
- [2] UNESCO y WWAP, «Informe mundial de las Naciones Unidas sobre el desarrollo de los recursos hídricos 2018: soluciones basadas en la naturaleza para la gestión del agua», 2018, [En línea]. Disponible en: <https://unesdoc.unesco.org/ark:/48223/pf0000261494>
- [3] Programa Mundial de la UNESCO de Evaluación de los Recursos Hídricos, *Informe mundial de las Naciones Unidas sobre el desarrollo de los recursos hídricos 2019: no dejar a nadie atrás*. Paris, francia: UNESCO, 2019, 2019. [En línea]. Disponible en: <https://unesdoc.unesco.org/ark:/48223/pf0000367304>
- [4] «World Energy Outlook 2019 – Analysis», IEA. <https://www.iea.org/reports/world-energy-outlook-2019> (accedido 25 de agosto de 2023).
- [5] T. de O. Bredariol, «Energy and water – Topics», IEA, 2023. <https://www.iea.org/topics/energy-and-water> (accedido 25 de agosto de 2023).
- [6] IEA, *World Energy Outlook 2020*. Paris, francia: IEA. [En línea]. Disponible en: <https://www.iea.org/reports/world-energy-outlook-2020>
- [7] «Electricity consumption in the water sector by process, 2014-2040 – Charts – Data & Statistics», IEA. <https://www.iea.org/data-and-statistics/charts/electricity-consumption-in-the-water-sector-by-process-2014-2040> (accedido 25 de agosto de 2023).
- [8] UNESCO, WWAP y Azoulay A, «Informe mundial de las Naciones Unidas sobre el desarrollo de los recursos hídricos 2019: no dejar a nadie atrás», 2019.
- [9] D. F. Birol, «World Energy Outlook 2022».
- [10] G. Klein, M. Krebs, V. Hall, T. O'Brien, y B. B. Blevins, «CALIFORNIA ENERGY COMMISSION».
- [11] «Curso de ingeniería hidráulica aplicada a los sistemas de distribución de agua. Tomo I», *Curso de ingeniería hidráulica aplicada a los sistemas de distribución de agua. Tomo I*. Universidad Politécnica de Valencia, Valencia, 1992.
- [12] A. Di Nardo, A. Di Nardo, M. Di Natale, G. F. Santonastaso, V. Tzatchkov, y V. H. A. Yamanaka, «Divide and Conquer Partitioning Techniques for Smart Water Networks», *Procedia Engineering*, vol. 89, pp. 1176-1183, ene. 2014, doi: 10.1016/j.proeng.2014.11.247.
- [13] A. Di Nardo, M. Di Natale, y G. F. Santonastaso, «A comparison between different techniques for water network sectorization», *Water Science & Technology: Water Supply*, vol. 14, n.º 6, pp. 961-970, dic. 2014, doi: 10.2166/ws.2014.046.

- [14] Cabrera E., Gómez E., Soriano J., y Del Teso R., «Eco-Layouts in Water Distribution Systems», *Journal of Water Resources Planning and Management*, vol. 145, n.º 1, p. 04018088, ene. 2019, doi: 10.1061/(ASCE)WR.1943-5452.0001024.
- [15] R. Del Teso March, «Ecodiseño y ecogestión de redes de distribución de agua a presión», Universitat Politècnica de València, Valencia (Spain), 2020. doi: 10.4995/Thesis/10251/153135.
- [16] J. Delua, «Supervised vs. Unsupervised Learning: What's the Difference?», *IBM Blog*, 12 de marzo de 2021. <https://www.ibm.com/blog/supervised-vs-unsupervised-learning/> (accedido 2 de agosto de 2023).
- [17] J. H. Ward, «Hierarchical Grouping to Optimize an Objective Function», *Journal of the American Statistical Association*, vol. 58, n.º 301, pp. 236-244, mar. 1963, doi: 10.1080/01621459.1963.10500845.
- [18] J. B. MacQueen, «Some methods for classification and analysis of multivariate observations», vol. 1, pp. 281-297, ene. 1967.
- [19] *Clustering con K-Means. Explicación Matemática y Mucho más...*, (1 de mayo de 2020). Accedido: 4 de agosto de 2023. [En línea Video]. Disponible en: <https://www.youtube.com/watch?v=n98fnSEoRiM>
- [20] C. Jordán Lluch, *Teoría de grafos y modelización: problemas resueltos*. San Fernando de Henares, Madrid: Paraninfo, 2022.
- [21] D. Ivanov, «Kirchhoff's theorem. Finding the number of spanning trees - Technical Info». <https://d-kivanov.github.io/docs/Algorithms/Graph%20Algorithms/Kirchhoff%20Theorem/> (accedido 29 de agosto de 2023).
- [22] E. Cabrera, E. Gómez, E. Cabrera, J. Soriano, y V. Espert, «Energy Assessment of Pressurized Water Systems», *Journal of Water Resources Planning and Management*, vol. 141, n.º 8, p. 04014095, 2015, doi: 10.1061/(ASCE)WR.1943-5452.0000494.
- [23] R. del Teso, E. Gómez, E. Estruch-Juan, y E. Cabrera, «Topographic Energy Management in Water Distribution Systems», *Water Resour Manage*, vol. 33, n.º 12, pp. 4385-4400, sep. 2019, doi: 10.1007/s11269-019-02375-9.
- [24] A. Di Nardo, M. Di Natale, C. Giudicianni, D. Musmarra, y G. F. Santonastaso, «An innovative approach based on recursive clustering to design optimal districts in water distribution networks», ene. 2018.
- [25] L. Perelman y A. Ostfeld, «Short communication: Topological clustering for water distribution systems analysis», *Environmental Modelling and Software*, vol. 26, n.º 7, pp. 969-972, jul. 2011, doi: 10.1016/j.envsoft.2011.01.006.
- [26] X. Khoa Bui, M. S. Marlim, y D. Kang, «Water Network Partitioning into District Metered Areas: A State-Of-The-Art Review», *Water*, vol. 12, n.º 4, 2020, doi: 10.3390/w12041002.
- [27] Dinesh Kumar Gautam, Prakash Kotecha, y S. Subbiah, «Efficient k-means clustering and greedy selection-based reduction of nodal search space for optimization of sensor placement in the water distribution networks.», *Water Research*, pp. 118666-118666, may 2022, doi: 10.1016/j.watres.2022.118666.
- [28] Federico Spizzo, Giovanni Venaruzzo, M. Nicolini, y D. Goi, «Water Distribution Network Partitioning Based on Complex Network Theory: The Udine Case Study», *Water*, 2023, doi: 10.3390/w15081621.



- [29] D. Arsene, A. Predescu, C.-O. Truică, E.-S. Apostol, M. Mocanu, y C. Chiru, «Profiling consumers in a water distribution network using K-Means clustering and multiple pre-processing methods», en *2021 13th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, 2021, pp. 1-6. doi: 10.1109/ECAI52376.2021.9515193.
- [30] Athanasios V. Serafeim, George Kokosalakis, R. Deidda, Nikolaos Th. Fourniotis, y A. Langousis, «Combining Statistical Clustering with Hydraulic Modeling for Resilient Reduction of Water Losses in Water Distribution Networks: Large Scale Application Study in the City of Patras in Western Greece», *Water*, vol. 14, n.º 21, pp. 3493-3493, nov. 2022, doi: 10.3390/w14213493.