



ESCUELA DE INGENIERÍA DE FUENLABRADA

GRADO EN INGENIERÍA AEROESPACIAL EN AERONAVEGACIÓN

TRABAJO FIN DE GRADO

Diseño de redes de transporte: integrando la
competencia modal con técnicas de optimización
convexa

Autor: Fernando Real Rojas

Tutor: Prof. Dr. Antonio G. Marqués

Co-tutor: Prof. Dr. Luis Cadarso Morga

Curso académico 2023/2024

©2024 Fernando Real Rojas

Algunos derechos reservados

Este documento se distribuye bajo la licencia "Atribución 4.0 Internacional" de Creative Commons, disponible en:

<https://creativecommons.org/licenses/by/4.0/deed.es>

Resumen

Las redes de transporte (RdT en el resto del documento, tanto para singular como plural) son columnas vertebrales de la interacción humana y el desarrollo social. Desempeñan un papel clave en la generación de progreso económico a través del comercio y la movilidad laboral, siendo también un factor crítico para reducir las desigualdades económicas. Esto ha supuesto que el papel de las RdT en nuestra sociedad, en especial el del transporte público, haya crecido constantemente a lo largo de las últimas décadas. Más importante si cabe desde el punto de vista ingenieril, este crecimiento ha conllevado también un incremento del tamaño, la complejidad y la heterogeneidad de las RdT contemporáneas, haciendo que su planificación y operación sean más desafiantes.

El uso de herramientas de optimización para el diseño de RdT tiene una larga tradición en el ámbito de la investigación operativa, con técnicas de programación entera mixta (MIP, por sus siglas en inglés *Mixed Integer Programming*) jugando un papel preponderante. Problemas variados y relevantes, como la localización de estaciones o el diseño de rutas, han sido resueltos de manera exitosa mediante modelos MIP. No obstante, el creciente tamaño de las RdT y la presencia de incertidumbre y dinámicas no lineales suponen un desafío desde el punto de vista de la programación entera. En paralelo, el uso de técnicas de optimización no lineal y convexa ha ganado preponderancia en multitud de ámbitos de ciencia e ingeniería, revolucionando disciplinas como la ciencia de datos. Sin embargo, el uso de técnicas de optimización convexa para el control y el diseño de RdT es sorprendentemente escaso, debido en parte a que múltiples aspectos del diseño de RdT involucran variables de naturaleza discreta.

En este contexto, la primera mitad del presente Trabajo Fin de Grado (TFG) propone una nueva metodología de optimización convexa para el diseño de RdT, con el objetivo de: 1) decidir el conjunto de estaciones y enlaces que se construyen, así como sus capacidades, y 2) realizar dicho diseño teniendo en cuenta los costes de construcción de la infraestructura y los costes de operación, así como las decisiones de los usuarios, que utilizarán la nueva infraestructura en función de su precio y tiempo de viaje de acuerdo a un modelo *logit*. Dos son las innovaciones principales consideradas en el TFG. En primer lugar, el uso de regularizadores que promueven *sparsity*¹ en la infraestructura mediante el uso de la norma 1 reponderada. El aspecto clave en este caso consiste en modelar conjuntamente y de forma convexa tanto la decisión de construir un enlace (estación) como la capacidad del mismo. En segundo lugar, el uso de regularizadores basados en la función de entropía sobre los flujos de pasajeros para tener en cuenta los modelos *logit* de demanda. De esta manera, se consigue una formulación convexa que tiene en cuenta las interrelaciones entre los distintos elementos y actores de la red y que puede resolverse con una complejidad computacional polinómica. Los resultados teóricos y numéricos demuestran la validez de la propuesta, consiguiendo soluciones con pérdidas de optimalidad pequeñas y tiempos de ejecución sustancialmente menores que los de las alternativas MIP.

La segunda mitad de este TFG se centra en demostrar el valor práctico de la metodología propuesta aplicándola a dos casos de uso reales: una red de metro (de la ciudad de Sevilla) gestionada por un operador público y una red aérea (del mercado doméstico de EE.UU.) con topología de *hub & spoke* gestionada por un operador privado. En ambos casos resulta necesario realizar pequeñas modificaciones de la formulación original, con el objeto de adaptarse a la idiosincrasia y necesidades de cada una

¹De manera general, se intentará evitar el uso innecesario de extranjerismos en el documento. No obstante, en aquellos casos en los que el uso del término en inglés esté generalizado en el ámbito científico-tecnológico, se utilizará la versión inglesa, escribiéndola en cursiva. Así mismo, la memoria adoptará la convención inglesa de utilizar un punto en lugar de una coma para separar la parte entera de la parte decimal en una cifra numérica.

de estas redes. Más relevante desde un punto de vista tecnológico, para ambas redes los algoritmos propuestos son capaces de encontrar soluciones de calidad en tiempos de cómputo razonables. Esto abre la puerta a la realización de análisis de robustez o de incertidumbre con respecto a un conjunto amplio de variables, lo que permite a operadores y reguladores una mejor toma de decisiones.

El trabajo futuro incluye extensiones a otro tipo de RdT, incorporación de modelos probabilísticos o formulaciones de problemas de optimización multinivel, por nombrar algunos ejemplos relevantes.

Abstract

Transport Systems and Networks (TSNs) are the backbone of human interaction and social development. They play a key role in generating economic progress through trade and labor mobility and are also a critical factor in reducing economic inequalities. This has led to the role of TSNs in our society, particularly public transportation, growing steadily over the past few decades. More importantly, from an engineering perspective, this growth has also brought an increase in the size, complexity, and heterogeneity of contemporary TSNs, making their planning and operation more challenging.

The use of optimization tools for designing TSNs has a long tradition in the field of operations research, with Mixed-Integer Programming (MIP) techniques playing a predominant role. Various significant problems, such as station location or route design, have been successfully solved using MIP models. However, the increasing size of TSNs and the presence of uncertainty and non-linear dynamics pose a challenge for integer programming. Simultaneously, the use of non-linear and convex optimization techniques has gained prominence in many areas of science and engineering (revolutionizing disciplines such as data science). However, their application in TSN control and design is surprisingly limited, partly due to the fact that many aspects of TSN design involve discrete variables.

In this context, the first half of the present Bachelor's Thesis (BT) proposes a new convex optimization methodology for TSN design. The objective is: 1) to decide the set of stations and links to be constructed, along with their capacities, and 2) to perform the design considering infrastructure construction costs, operating costs, as well as the decisions of users, who will use the new infrastructure based on its price and travel time according to a logit model. There are two main innovations considered in the BT. Firstly, the use of regularizers that promote sparsity in the infrastructure through the use of reweighted norm-1. The key aspect in this case is to model both the decision to construct a link (or a station) and its capacity jointly and in a convex manner. Secondly, the use of entropy-based regularizers on passenger flow variables to account for logit demand models. This results in a convex formulation that takes into account the interrelations between the different elements and actors of the network, which can be solved with polynomial computational complexity. The theoretical and numerical results demonstrate the validity of the proposal, achieving solutions with small optimality losses and substantially shorter execution times compared to MIP alternatives.

The second half of this BT focuses on demonstrating the practical value of the proposed methodology by applying it to two real use cases: a metro network (in the city of Seville) managed by a public operator and an air network (in the U.S. domestic market) with a hub & spoke topology managed by a private operator. In both cases, it is necessary to make small modifications to the original formulation, adapting to the needs of each network. More relevant from a technological point of view, for both networks, the proposed algorithms are capable of finding high-quality solutions in reasonable computation times. This opens the door to robustness or uncertainty analysis regarding a broad set of variables, allowing operators to make better decisions.

Future work includes extensions to other types of transport networks, incorporation of probabilistic models, or formulations of multilevel optimization problems, to name a few relevant examples.

Agradecimientos

Debo comenzar mencionando a los responsables directos de este TFG: Antonio G. Marqués, Luis Cadarso y Víctor Manuel Tenorio. Sin su compromiso y atención desmesurada no habría sido posible realizar este trabajo que, no solo me ha hecho crecer como ingeniero, sino, cuanto menos al mismo nivel, como persona. He de darles también las gracias por brindarme la oportunidad de colaborar de forma activa en grupos de investigación de prestigio internacional. En este sentido, aprovecho estas líneas para agradecer a los profesores Geert Leus (TU Delft, Países Bajos), Tianyi Chen (Rensselaer Polytechnic Institute, EE.UU.) y Vikrant Vaze (Dartmouth College, EE.UU.) por guiarme en mis primeros pasos en el mundo de la investigación. Doy las gracias también a todos los profesores de la Escuela de Ingeniería de Fuenlabrada, quienes, a lo largo de mi formación académica, me han dotado de los conocimientos y herramientas necesarias para llevar a cabo este TFG.

Un agradecimiento especial merecen mis compañeros y amigos Guillermo y Bianca, que sin duda alguna son lo mejor que me llevo de esta etapa.

A mi familia, especialmente a mis padres, por su amor incondicional y su sacrificio constante. A los que están y los que ya no están. A mi hermano, por marcarme el camino desde el inicio. A Nuria, por ser mi apoyo diario y acompañarme en todo el proceso. A mis amigos, que son el motivo de tantos buenos momentos.

A todos los que, alguna vez en mi vida, me han ayudado, gracias.

Índice general

Resumen	I
Abstract	III
Agradecimientos	IV
Índice de figuras	VII
Índice de tablas	VIII
Glosario de acrónimos y abreviaturas	X
1. Introducción	1
1.1. Motivación y estado del arte	1
1.1.1. Planificación en TP	1
1.1.2. Modelos de demanda	2
1.1.3. Optimización en el diseño de RdT público	3
1.2. Contribuciones	5
1.3. Estructura del documento	6
1.4. Grupos de tareas y tiempo dedicado	7
1.5. Colaboraciones, publicaciones y repositorio de código	8
2. La optimización en RdT: formulaciones enteras y una propuesta convexa	10
2.1. Formalización del problema de diseño de RdT	10
2.1.1. Variables de estado	11
2.1.2. Variables de diseño	11
2.1.3. Restricciones	12
2.1.4. Función objetivo	13
2.1.5. Algoritmo de solución	13
2.2. Diseño de RdT vía MIP	14
2.2.1. Topología y variables de estado	14
2.2.2. Formulación basada en aproximación <i>all-or-nothing</i> para la demanda	16
2.2.3. Formulación basada en aproximación <i>logit</i> para la demanda	19
2.3. Algoritmos de <i>sparsity</i> aplicados a RdT	22
2.3.1. Introducción a <i>sparsity</i>	22
2.3.2. <i>Sparsity</i> aplicada a RdT	26
2.4. Funciones de entropía aplicadas a RdT: formulación convexa con reparto <i>logit</i>	26
2.4.1. Aplicación de regularizadores de entropía a un problema simplificado	27
2.4.2. Aplicación de regularizadores de entropía al problema original	28
3. Formulación general del problema de optimización convexa para RdT	32
3.1. Aproximación convexa a la congestión en la red	33
3.2. Variables y objetivos del modelo matemático	35

3.2.1.	Variables de estado	35
3.2.2.	Variables de diseño	36
3.2.3.	Descripción del objetivo	37
3.3.	Formulación convexa del problema	39
3.4.	Versión MIP del problema convexo con variables de construcción binarias	42
3.5.	Optimización multiobjetivo: búsqueda del punto de utopía y curva de Pareto	43
3.6.	Análisis cuantitativo de resultados	44
3.6.1.	Caso de estudio	45
3.6.2.	Valor de α en la función objetivo	46
3.6.3.	Comparación cuantitativa de la formulación convexa y la formulación MIP	46
4.	Optimización de una RdT de metro con operador público	49
4.1.	Estado del arte	49
4.2.	Datos básicos y descripción de la red	51
4.3.	Modificaciones en la formulación del problema de optimización	52
4.3.1.	Función objetivo	53
4.3.2.	Formulación del problema de optimización	53
4.4.	Presentación y análisis de los resultados numéricos	54
4.4.1.	Caso base y diseño óptimo	55
4.4.2.	Sobreponderación de usuarios provenientes de zonas deprimidas	61
4.4.3.	Crecimiento de la demanda y adaptación futura	64
5.	Optimización de una RdT aéreo con operador privado	67
5.1.	Estado del arte	67
5.2.	Datos básicos y descripción de la red	69
5.3.	Modificaciones en la formulación del problema	71
5.3.1.	Motivación de las modificaciones	71
5.3.2.	Variables de estado	72
5.3.3.	Parámetros	73
5.3.4.	Variables de diseño	74
5.3.5.	Formulación del modelo	75
5.4.	Presentación y análisis de los resultados numéricos	78
5.4.1.	Análisis del caso base	80
5.4.2.	Modificación de costes fijos y variables	82
5.4.3.	Restricción de la capacidad de <i>hubs</i>	83
5.4.4.	Restricción del tamaño de flota	86
5.4.5.	Consideración de distintos escenarios económicos	87
5.4.6.	Modificación de patrones de precios	90
6.	Conclusiones y líneas futuras	93
6.1.	Contribuciones, resultados e impacto	93
6.2.	Líneas de trabajo futuro	94
	Bibliografía	97
A.	Linealización por trozos de la función <i>logit</i>	101
B.	Análisis de sensibilidad de la formulación convexa	105

Índice de figuras

2.1.	Ejemplo de grafo que representa a una RdT de 9 nodos.	15
2.2.	Representación de la función $\alpha \log(\beta x + \varepsilon)$ para distintos valores de α y β	23
2.3.	Comparación de las normas $\ell_0(x)$, $\ell_1(x)$ y $\ell_{1_{rw}}(x)$	24
2.4.	Representación de $\hat{\ell}_{1_{rw}}^k(x x^{(k-1)})$, la linealización de la norma $\ell_{1_{rw}}(x)$	25
3.1.	Representación del tiempo de viaje en un arco t_{ij}^c en función de la congestión	34
3.2.	Comparativa entre los modelos convexo (3.26)-(3.23) y MIP (3.29)-(3.40)	48
4.1.	Distribución geográfica de los nodos de la red de Sevilla	52
4.2.	Red de Sevilla. Demanda atraída en función del presupuesto y γ	58
4.3.	Red de Sevilla. Tiempo de viaje medio por pasajero y ruta en función de γ	59
4.4.	Red de Sevilla. Topología de la red de metro para distintos γ	60
4.5.	Red de Sevilla. Topología de la red de metro para distintos escenarios de costes	61
4.6.	Red de Sevilla. Representación de los coeficientes asociados a cada nodo según el nivel de renta.	63
4.7.	Red de Sevilla. Topología de la red de metro priorizando la cobertura en barrios deprimidos para varios γ	64
4.8.	Red de Sevilla. Representación de los coeficientes asociados a cada nodo según el crecimiento demográfico esperado	65
4.9.	Red de Sevilla. Topología de la red de metro priorizando la cobertura en barrios con mayor crecimiento demográfico para varios γ	66
5.1.	Red aérea. Topología de la red de operación de la aerolínea para el caso base	81
5.2.	Red aérea. Porcentaje de vuelos en escala por aeropuerto	82
5.3.	Red aérea. Topología de la red para distintos escenarios de costes fijos y presupuesto.	83
5.4.	Red aérea. Topología de la red limitando la capacidad máxima de despliegue en los <i>hubs</i>	85
5.5.	Red aérea. Topología de la red limitando la capacidad máxima de despliegue en los <i>hubs</i> , manteniendo el presupuesto	86
5.6.	Red aérea. Topología de la solución propuesta para distintos escenarios de costes fijos y presupuesto, restringiendo el tamaño de flota	87
5.7.	Red aérea. Topología de la solución propuesta para distintos escenarios de costes fijos y presupuesto, considerando una previsión económica optimista	89
5.8.	Red aérea. Topología de la solución propuesta para distintos escenarios de costes fijos y presupuesto, considerando una previsión económica pesimista	90
5.9.	Red aérea. Topología de la solución propuesta para distintos escenarios de costes fijos y presupuesto, modificando el modelo de precios.	92
A.1.	Representación de la función logística y su linealización	102

Índice de tablas

3.1. Red de 9 nodos. Tamaño de los modelos convexo y MIP para el caso de referencia. . .	46
4.1. Ubicación de los nodos de la red de Sevilla.	52
4.3. Red de Sevilla. Tamaño del modelo para el caso de referencia.	56
4.2. Parámetros para el diseño de la red de metro de Sevilla.	56
4.4. Red de Sevilla. Métricas de las soluciones para el caso de referencia, haciendo un barrido en γ	57
4.5. Red de Sevilla. Métricas de las soluciones para distintos escenarios de costes fijos. . . .	62
4.6. Renta per cápita por distrito en Sevilla.	62
4.7. Red de Sevilla. Métricas de las soluciones para distintos escenarios priorizando cobertura en barrios deprimidos.	64
4.8. Promociones de viviendas de obra nueva por distrito en Sevilla.	65
4.9. Red de Sevilla. Métricas de las soluciones para distintos escenarios, considerando crecimiento demográfico.	66
5.1. Nodos considerados en el diseño de la red de una aerolínea	71
5.2. Parámetros para el diseño de la red para una aerolínea.	80
5.3. Red aérea. Tamaño del modelo para el caso de referencia.	80
5.4. Red aérea. Métricas de la solución para el caso base.	80
5.5. Red aérea. Métricas de las soluciones para distintos escenarios de costes fijos y presupuesto. .	82
5.6. Red aérea. Métricas de las soluciones para distintos escenarios de costes fijos y presupuesto, limitando la capacidad de los <i>hubs</i>	84
5.7. Red aérea. Métricas de las soluciones para distintos escenarios de costes fijos y presupuesto, limitando la capacidad de los <i>hubs</i> , manteniendo el presupuesto del caso base. .	84
5.8. Red aérea. Métricas de las soluciones para distintos escenarios de costes fijos y presupuesto, limitando el tamaño de flota.	86
5.9. Red aérea. Métricas de las soluciones para distintos escenarios de costes fijos y presupuesto, considerando una previsión económica optimista	91
5.10. Red aérea. Métricas de las soluciones para distintos escenarios de costes fijos y presupuesto, considerando una previsión económica pesimista	91
5.11. Red aérea. Métricas de las soluciones para distintos escenarios de costes fijos y presupuesto, modificando el modelo de precios	91
B.1. Análisis de sensibilidad para $\lambda = 5$ y $\beta = 1.25$, evaluando en el objetivo con los parámetros originales.	107
B.2. Análisis de sensibilidad para $\lambda = 5$ y $\beta = 1.25$, evaluando en el objetivo con los parámetros modificados con los que se calcula la solución.	107
B.3. Análisis de sensibilidad para $\lambda = 5$ y $\beta = 1.50$, evaluando en el objetivo con los parámetros originales.	108
B.4. Análisis de sensibilidad para $\lambda = 5$ y $\beta = 1.50$, evaluando en el objetivo con los parámetros modificados con los que se calcula la solución.	108

B.5. Análisis de sensibilidad para $\lambda = 6$ y $\beta = 1.25$, evaluando en el objetivo con los parámetros originales.	109
B.6. Análisis de sensibilidad para $\lambda = 6$ y $\beta = 1.25$, evaluando en el objetivo con los parámetros modificados con los que se calcula la solución.	109

Glosario de acrónimos y abreviaturas

AENA Aeropuertos Españoles y Navegación Aérea.

BT *Bachelor's Thesis.*

BTS *Bureau of Transportation Statistics.*

CASM *Cost per Available Seat and Mile.*

cf. *confer.*

e.g. *exempli gratia.*

EE.UU. Estados Unidos.

FMI Fondo Monetario Internacional.

GICDPS Grupo de Ciencia de Datos y Procesamiento de Señal.

GISAT Grupo de Investigación en Sistemas Aeroespaciales y Transporte.

HNDP *Hub & Spoke Network Design Problem.*

i.e. *id est.*

KKT Karush-Kuhn-Tucker.

MIP *Mixed Integer Programming.*

MM *Majorization-Minimization.*

NP *Nondeterministic Polynomial time.*

RASM *Revenue per Available Seat and Mile.*

RdT Red de Transporte.

TFG Trabajo de Fin de Grado.

TP Transporte Público.

TSN *Transport Systems and Networks.*

URJC Universidad Rey Juan Carlos.

Capítulo 1

Introducción

El Trabajo Fin de Grado (TFG) que se resume en esta memoria aborda el diseño de redes de transporte (RdT) mediante el uso de herramientas de optimización convexa, proponiendo formulaciones novedosas que permiten modelar e integrar de manera eficiente en la optimización tanto los costes construcción y operación de la infraestructura como las preferencias de los usuarios descritas a través de modelos de elección, los cuales explican y predicen las decisiones entre dos o más alternativas discretas, como elegir entre modos de transporte.

Tomando como punto de partida lo expuesto en el párrafo anterior, este capítulo introductorio se articula en torno a 5 secciones. La sección 1.1 proporciona una motivación general sobre el problema estudiado y sobre el papel de la optimización en el diseño de RdT, realizando una revisión histórica del estado del arte. La sección 1.2 detalla las contribuciones científicas y tecnológicas de este TFG. La sección 1.3 describe la estructura del resto de capítulos. La sección 1.4 describe de forma detallada las principales tareas llevadas a cabo y presenta una estimación del tiempo total dedicado a este TFG. Finalmente, la sección 1.5 proporciona detalles sobre las publicaciones científicas asociadas a las tareas realizadas y el código utilizado.

1.1. Motivación y estado del arte

En esta sección se motiva de forma general el problema estudiado y se proporciona una descripción del estado del arte de alto nivel en tres ámbitos: la planificación de los sistemas de transporte público (TP), los modelos de elección de la demanda y la optimización de los sistemas de TP.

1.1.1. Planificación en TP

El TP abarca los servicios de transporte compartido que utilizan infraestructuras como carreteras o vías férreas, junto con vehículos como autobuses o trenes [1]. El TP engloba tanto el transporte urbano como el interurbano, ambos caracterizados por tener rutas y horarios fijos disponibles para todos los usuarios que abonen la tarifa establecida [2]. Nótese que los servicios de TP los articula de forma general el estado, para ofrecer así a la sociedad una mejor calidad de vida y promover la igualdad de oportunidades. A menudo, estos servicios son financiados con recursos públicos, ya sea de manera directa o indirecta. Y aunque existan modos de transporte, como el aéreo, que no reciben financiación estatal (al menos en igualdad de condiciones que otros modos), se ajustan a la definición de servicios de TP proporcionada previamente. El TP ha ido adquiriendo importancia debido a que la sostenibilidad se ha convertido en uno de los principales objetivos de la sociedad. En comparación con otros modos de transporte motorizado, el TP muestra tasas de eficiencia más altas en términos de consumo de energía, emisiones de gases de efecto invernadero, contaminación acústica y uso del espacio público. Sin embargo, tanto el establecimiento como la operación de los sistemas de TP conllevan costes significativos. Además, el rendimiento de estos sistemas tanto desde la perspectiva de los operadores

como de los usuarios es fundamental para proporcionar un servicio exitoso, lo que implica una necesidad de metodologías de planificación efectivas que tengan en cuenta las sinergias entre los distintos actores.

La planificación de los sistemas de TP, que ofrece diversas oportunidades de optimización, suele dividirse en varias etapas. Dicha división da lugar a una secuencia de decisiones jerárquicas, que incluye tareas como el diseño de la red, la determinación de frecuencias y horarios o la programación de flotas y recursos humanos [3, 4, 5, 6, 7]. De acuerdo con este enfoque, el diseño de la red desempeña un papel crucial en el proceso de planificación en su conjunto, ya que impacta en cada etapa posterior y, por lo tanto, en cada componente de costes del sistema. En otras palabras, la forma en la que se diseña la infraestructura en una RdT constituye una parte fundamental en la planificación de los sistemas de transporte, estando factores como el rendimiento, la eficiencia y la robustez fuertemente influenciados por la selección de nodos (estaciones) y la forma en que se interconectan [8]. Por ejemplo, el objetivo principal de una red de tránsito rápido es mejorar la movilidad en áreas urbanas, de forma que se incentiven tiempos de viaje más cortos, menor contaminación y un consumo de energía reducido, promoviendo así una movilidad más sostenible. En este contexto, el concepto de red de TP adquiere significados específicos según el modo de transporte.

Por un lado, sistemas como los basados en autobuses que comparten la vía con vehículos privados (principalmente automóviles) se caracterizan por costes de construcción de infraestructura nulos o insignificantes. Por otro lado, los sistemas de transporte aéreo, ferrocarril y metro implican grandes inversiones debido a la construcción de infraestructuras exclusivas, tales como aeropuertos, estaciones, corredores dedicados, vías férreas y túneles. Estos elementos se conocen como “red física”. Además, en los sistemas de TP existe un segundo nivel de red, que se define por los servicios que operan sobre la red física, y que consiste en las rutas seguidas por los vehículos. Este nivel se conoce como “red de rutas”. La operación de estas rutas determina un componente importante del coste operativo del sistema, en términos de costes de vehículos y personal por distancia y unidad de tiempo, respectivamente. Así mismo, tanto la estructura topológica como la frecuencia (número de vehículos que pasan por unidad de tiempo) de las rutas determinan en gran medida el nivel de servicio ofrecido a los usuarios en términos de tiempo de viaje total, que incluye el tiempo dedicado a caminar hasta/desde las paradas o estaciones de origen y destino, el tiempo de espera y el tiempo a bordo del vehículo.

En un problema de diseño de redes, tanto los modelos de las redes físicas como los de las rutas representan las estaciones o aeropuertos como nodos de la red y las vías (ya sean férreas, aéreas, marítimas o públicas) como sus enlaces. En el caso más general, cada nodo tiene como atributo el coste de construcción y la capacidad, y cada enlace el coste de construcción, el tiempo de viaje y la capacidad. Dichos atributos pueden ser parámetros dados por la topología de la red subyacente (caracterizada por la conexión de todos los nodos potenciales entre sí), como el tiempo de viaje o el coste de construcción, o variables de decisión que el problema de optimización ha de obtener, como la capacidad de nodos y enlaces.

En lo que respecta a la demanda, se suele representar por grupos específicos de personas que viajan desde algún nodo de origen hasta otro nodo de destino. El objetivo es acercar la infraestructura a los potenciales usuarios, estando las decisiones relacionadas con la inclusión o exclusión de los nodos y de los enlaces en el diseño, así como con la capacidad que se le asigna a cada uno de ellos.

1.1.2. Modelos de demanda

Existen características distintivas de los sistemas de TP que presentan dificultades inherentes. Al modelar el diseño de las rutas, las decisiones no están relacionadas con enlaces individuales, sino que se refieren a una secuencia de enlaces. Además, el sistema está compuesto por varias rutas que pueden superponerse, es decir, pueden compartir enlaces comunes. Así mismo, la evaluación del rendimiento del sistema desde el punto de vista de la demanda implica modelar su comportamiento con respecto al conjunto de enlaces o rutas habilitadas. Esta es una característica muy particular de los modelos de diseño de redes de TP, ya que los usuarios se comportan de manera autónoma y sus intereses pueden entrar en conflicto con los del operador, lo que implica considerar características específicas en los

modelos correspondientes de diseño de redes.

Es común describir la demanda mediante pares de puntos origen-destino (usualmente denominados mercados), cada uno con un peso que representa el flujo de tráfico entre ellos. Esta demanda se representa típicamente mediante una matriz origen-destino. Al planificar una nueva red, es probable que ya exista una red en funcionamiento que sirva al mismo conjunto de pares origen-destino. A modo de ilustración, podría considerarse el diseño de un nuevo sistema de transporte rápido para mejorar la movilidad en una gran ciudad o área metropolitana donde ya exista otro sistema de transporte, además del transporte privado. Aunque el sistema de transporte actual pueda ser más denso que el planificado, es probable que sea más lento debido al hecho de compartir el mismo derecho de paso que el tráfico privado. Por consiguiente, ambos sistemas compiten entre sí y con el transporte privado, y en consecuencia, el flujo de tráfico entre un origen y un destino se distribuye entre los diferentes sistemas que ofrecen el servicio.

Principalmente, hay dos formas de modelar la distribución de la demanda. La primera es el enfoque binario de “todo o nada”, donde la demanda es atendida exclusivamente por una de las alternativas propuestas. De forma habitual, la demanda se considera satisfecha si recibe servicio dentro de un rango de calidad predefinido, como se indica en [9]. La segunda se basa en alguna función continua, utilizando, por ejemplo, una distribución de probabilidad *multi-logit* [10, 11]. Ambos métodos se fundamentan en la comparación de distancias, tiempos, costes o utilidades generalizadas.

1.1.3. Optimización en el diseño de RdT público

En el contexto de las RdT, la optimización desempeña un papel fundamental por varios motivos. En primer lugar, la optimización puede ayudar a minimizar costes y maximizar eficiencia. Factores como la distancia, la capacidad y la congestión pueden afectar significativamente a los costes operativos y la experiencia del usuario. Al optimizar la RdT, es posible encontrar configuraciones que minimicen los tiempos de viaje, reduzcan los costes operativos y de mantenimiento y maximicen la capacidad de la red para satisfacer el flujo del tráfico.

Así mismo, la optimización puede contribuir a mejorar la accesibilidad y la conectividad. Al diseñar una RdT, es importante garantizar que los puntos de origen y destino estén bien conectados y que haya rutas eficientes entre ellos. La optimización puede ayudar a identificar las localizaciones óptimas para aeropuertos, estaciones, paradas de autobús o intercambiadores, así como a planificar las rutas más eficientes entre ellas.

Sin embargo, a pesar de la importancia de la optimización en el diseño de las RdT, esta enfrenta una serie de desafíos significativos. Uno de los principales retos es la complejidad inherente de los sistemas de transporte, que pueden involucrar múltiples modos de transporte y una gran cantidad de variables interrelacionadas, como la demanda, las condiciones del tráfico y la infraestructura disponible. Otro desafío es la incertidumbre. Los patrones de viaje pueden variar de acuerdo con factores como el tiempo, eventos especiales o cambios en la demografía, lo que dificulta la predicción precisa de la demanda futura y, por lo tanto, la optimización de la RdT para satisfacer esa demanda de manera eficiente. A su vez, la optimización de las RdT a menudo implica tomar decisiones que afectan a múltiples partes interesadas y pueden tener consecuencias sociales, económicas y ambientales significativas. Por lo tanto, es crucial considerar una amplia gama de criterios y objetivos al diseñar y optimizar las RdT, lo que puede complicar aún más el proceso de optimización.

Además de todo lo anterior, múltiples elementos de la infraestructura de una RdT (o las decisiones asociadas al diseño, construcción y operación de la misma) son de naturaleza discreta e involucran costes fijos relacionados con la selección de nodos y enlaces, así como costes variables asociados con los flujos (de personas, bienes, información, etc.). Por este motivo, las metodologías clásicas de diseño de infraestructura en RdT se basan en formulaciones de programación entera mixta (MIP, por sus siglas en inglés *Mixed Integer Programming*). A pesar de sus ventajas, las formulaciones MIP presentan una serie de inconvenientes, que incluyen la dificultad de incorporar funciones o dinámicas no lineales, la

complejidad de manejar la incertidumbre o estocasticidad en las variables de estado o parámetros de la red y, sobre todo, el alto coste computacional asociado. Este último aspecto es especialmente crítico en RdT, donde, por ejemplo, el número potencial de arcos en una red escala cuadráticamente con el número de nodos y las variables de flujo (que dependen tanto de los enlaces como de los pares origen-destino) escalan como una potencia cuarta del número de nodos. Esto conlleva que, en la práctica, el diseño de RdT con más de unas pocas decenas de nodos continúe siendo hoy en día un reto significativo para la comunidad investigadora.

La optimización de problemas de diseño de RdT ha sido un campo muy productivo en la investigación operativa durante décadas recientes, abordando una variedad de problemas tanto teóricos como prácticos. En [12] se realiza una clasificación de los problemas fundamentales de diseño de redes, incluyendo problemas clásicos como el árbol de expansión mínima, el árbol de Steiner, el camino más corto, la localización de infraestructuras y los problemas del agente viajero. Estos se consideran casos específicos dentro de un modelo general de programación matemática.

A lo largo de la historia, la investigación en el diseño de RdT se ha enfocado en primer lugar en determinar una alineación única y la ubicación de estaciones a lo largo de ella. En [13] se proporciona una revisión que analiza los principales métodos disponibles para problemas de diseño de redes. En [14] se busca maximizar la cobertura de la demanda mediante la localización de una línea de tránsito rápido. Por otro lado, en [15] se incorporan datos de origen y destino de la demanda.

Posteriormente, en [16] se integran los problemas de localización de estaciones y enlaces en los modelos de diseño de RdT, proponiendo un conjunto de alternativas posibles para maximizar la cobertura de la demanda e incluyendo restricciones presupuestarias como limitación adicional. Nótese que, debido al elevado coste generalmente asociado con la conexión de todos los nodos potenciales, es crucial determinar una “subred” que pueda satisfacer de manera óptima la demanda. En [17] se presenta un modelo de diseño de red de tránsito rápido que tiene como objetivo minimizar el coste de viaje y maximizar la captura de tráfico, mientras que en [18] se explora una formulación robusta para el problema de diseño de redes de tránsito rápido, introduciendo restricciones para desarrollar una red con múltiples rutas alternativas para pares origen-destino específicos.

Debido a la naturaleza combinatoria y la complejidad de las funciones objetivo y restricciones de los problemas de diseño de RdT, estos son intrínsecamente difíciles (la mayoría de ellos son NP-difíciles). Por ello, los modelos y algoritmos para resolver problemas de diseño de redes requieren enfoques provenientes de diversas áreas de la optimización combinatoria y la programación matemática. Aunque los métodos exactos ofrecen una vía para resolver casos de estudio que han experimentado un crecimiento significativo en tamaño a lo largo del tiempo, es esencial reconocer que no constituyen la única forma de abordar estos problemas y que tampoco son capaces de resolver casos de estudio de gran escala, como aquellos comúnmente encontrados en el contexto de aplicaciones del mundo real. Resulta necesario, por lo tanto, explorar y desarrollar enfoques alternativos que puedan complementar y enriquecer la capacidad de resolver desafíos complejos a gran escala.

En las últimas décadas se han propuesto diversos enfoques (*e.g.*, heurísticas y metaheurísticas) para obtener soluciones aproximadas a varios tipos de problemas de diseño de redes. Estos enfoques se han utilizado principalmente para trabajar con instancias más grandes que no pueden ser tratadas con métodos exactos o en los que las restricciones de tiempo para obtener soluciones factibles de calidad para problemas complejos han sido significativas. Debido a que la mayoría de los problemas de diseño de redes son NP-difíciles (véase, por ejemplo, [19]), los esfuerzos de investigación más recientes se han centrado en la aplicación de algoritmos metaheurísticos para obtener soluciones en un tiempo computacional razonable. En el ámbito del diseño de RdT se han empleado algoritmos genéticos [20], procedimientos de búsqueda aleatoria adaptativa *greedy* [21], procedimientos de búsqueda en vecindario grande adaptativo [22] y matheurísticos [23], y se han aplicado con éxito en instancias de tamaño mediano. No obstante, las metodologías que se acaban de mencionar no proporcionan garantías de optimalidad y, además, son difíciles de caracterizar desde un punto de vista teórico.

1.2. Contribuciones

De manera general, este TFG trata con el problema de cómo formular problemas de diseño de RdT que tengan en cuenta: 1) los intereses del operador, 2) las preferencias de los usuarios y 3) las condiciones de operación de la RdT en cuestión. El objetivo principal es el desarrollo de una metodología rigurosa capaz de abordar el diseño de RdT como un problema de optimización convexa. Esto abriría la puerta a que la optimización de la red pudiera llevarse a cabo con una complejidad computacional moderada (tiempo polinómico), permitiendo de esta manera su aplicación a RdT de gran tamaño, así como la integración eficiente de fuentes de incertidumbre y de métricas de diseño no lineales. Dentro de este contexto, pueden distinguirse dos categorías de contribuciones específicas asociadas a este TFG: unas relacionadas con aspectos metodológicos y teóricos y otras relacionadas con la aplicación. En lo que respecta a las contribuciones metodológicas fundamentales, deben destacarse dos:

- El uso de funciones logarítmicas seguidas de una aproximación de mayorización-minimización (MM) para tener en cuenta costes fijos y lineales y, simultáneamente, dar lugar a soluciones *sparse*¹. El uso de aproximaciones convexas o pseudoconvexas a la norma ℓ_0 para conseguir soluciones *sparse* se ha utilizado masivamente en aplicaciones de procesamiento de señal, regresión y aprendizaje automático (incluido *active learning*) [24]. Más recientemente, han sido utilizadas en aplicaciones como el aprendizaje de grafos, la selección de muestras, la elección de localizaciones para colocar sensores o la detección y seguimiento de fuentes. No obstante, el énfasis en todas esas aplicaciones está en conseguir soluciones frugales en las que solo unos pocos elementos se activen. En el caso de este TFG, las novedades en el uso de regularizadores que induzcan soluciones con *sparsity* son varias: i) no se persigue únicamente que el número enlaces sea *sparse*, sino que se implementa una estrategia conjunta en la topología de red obtenida para que sea *sparse* tanto en enlaces como en nodos, ii) se pretende que los aproximadores de la norma ℓ_0 no se centren únicamente en promover soluciones *sparse*, sino que sirvan para modelar, de forma simultánea, una componente de costes fijos y otra componente de costes variables y iii) se obtiene de manera aproximada la solución de un problema originalmente MIP con una fracción del coste computacional, permitiendo abordar problemas de mayor tamaño (en términos de número de nodos y de enlaces potenciales). En este sentido, el uso de técnicas de MM que aproximan, de forma sucesiva, el regularizador pseudoconvexo por un polinomio de orden uno (con una componente fija y otra lineal) resultan esenciales.
- El uso de funciones de entropía para tener en cuenta las preferencias de los usuarios. En lugar de imponer *a fortiori* un mecanismo de elección basado en modelos *logit*, que obliga a considerar restricciones no lineales y no convexas, utilizamos una estrategia de ingeniería inversa en la que se añade a la función objetivo una función de entropía cuyas condiciones KKT² dan lugar a un mecanismo de elección basado en modelos *logit*. Las funciones de entropía se utilizan ampliamente en el contexto de algoritmos de clasificación en ciencia de datos y aprendizaje automático. No obstante, el uso que se le da en este TFG supone una novedad en la formulación de problemas de optimización aplicados al diseño de RdT, ayudando a reducir de manera drástica el tiempo de cómputo de los modelos formulados.

En lo que respecta a las contribuciones relacionadas con la aplicación, el TFG presenta y analiza los resultados asociados al diseño óptimo de dos tipos de RdT:

- Diseño de RdT con operador público. Concretamente, este trabajo analiza en el primer caso de estudio el diseño de una red de metro en la ciudad de Sevilla. El diseño parte de un estudio de

¹*Sparse*: voz inglesa. Dicho del conjunto numérico o estructura matemática que posee *sparsity*, i.e., una matriz *sparse*. *Sparsity*: voz inglesa. Propiedad que poseen los conjuntos numéricos o estructuras matemáticas en las que la mayor parte de los elementos son nulos. En este TFG se aplica principalmente a matrices y vectores.

²En optimización matemática, las condiciones Karush-Kuhn-Tucker (KKT), son condiciones necesarias y suficientes de primer orden para que una solución en programación no lineal sea óptima, siempre y cuando se cumplan ciertas condiciones de regularidad [25].

atracción de demanda, que permite comprender los flujos de personas dentro de la ciudad y las áreas de mayor concentración de viajes. El estudio incluye factores como la densidad poblacional, los centros de actividad económica, las zonas residenciales y comerciales, entre otros. Además, en la estrategia de diseño propuesta en el TFG para este caso, se considera la situación de los barrios deprimidos o con dificultades de accesibilidad, ya que el diseño de la red de metro podría contribuir significativamente a mejorar su conectividad y calidad de vida. Esto puede implicar la inclusión de estaciones estratégicamente ubicadas para atender estas áreas y facilitar el acceso a servicios y oportunidades. Así mismo, se analizan los potenciales cambios en los patrones de movilidad de la población, especialmente aquellos impulsados por factores como el desarrollo urbano, cambios demográficos, tendencias de trabajo y estudios, entre otros. Estos cambios pueden influir en la demanda de TP y en la configuración óptima de la red de metro.

- Diseño de RdT aéreo con operador comercial. En el segundo caso de estudio se analiza el diseño de una RdT aéreo para el transporte de pasajeros en el mercado doméstico de EE.UU. Este problema presenta características distintivas en comparación con el transporte urbano mediante metro. En concreto, los costes fijos son menores, mientras que los costes variables asociados a la operación y mantenimiento de aeronaves son mucho más elevados. Para optimizar la eficiencia y la rentabilidad, es crucial considerar la necesidad de *hubs* o centros de concentración estratégicamente ubicados. Estos *hubs* permiten consolidar el tráfico de pasajeros y optimizar las rutas, reduciendo los costes operativos y maximizando la utilización de las aeronaves. En este sentido, este TFG representa una contribución significativa al campo de la ingeniería aeroespacial en el ámbito del transporte aéreo. Específicamente, se aporta en áreas clave como el funcionamiento y la gestión del sistema de transporte aéreo, así como en la distribución, la gestión y la economía del mismo. Esto se consigue tanto mediante la planificación de la ubicación de *hubs*, como mediante el modelado de rutas que facilitan el uso eficiente y la administración del tráfico aéreo.

Por último, debe señalarse que, debido a su carácter general y relativa versatilidad, la metodología de diseño presentada en este TFG puede: 1) aplicarse a un amplio abanico de RdT y 2) generalizarse desde distintas vías teóricas. Remitimos al lector al capítulo final (sección de líneas de trabajo futuro) para obtener más detalles con respecto a estos dos aspectos.

1.3. Estructura del documento

Este documento se articula en torno a cuatro capítulos principales. A continuación, se proporciona una breve descripción de cada uno de ellos.

- En el capítulo 2 se discute el rol de la optimización en las RdT, se analiza una primera formulación (simplificada) del problema y se presentan los dos pilares fundamentales en los que se sustenta la propuesta de diseño de RdT: los mecanismos de optimización *sparse* para las decisiones de despliegue de la infraestructura y el uso de funciones de entropía para tener en cuenta las preferencias de los usuarios.
- En el capítulo 3 se presenta, de forma rigurosa e integral, la metodología para el diseño de RdT. En concreto, se formula un problema de optimización cuyo objetivo está compuesto por dos términos principales (uno asociado a los costes del operador y otro a las utilidades de los usuarios) y en cuyas restricciones se codifican la conservación de flujo, las limitaciones de capacidad y la congestión. Además de analizar las condiciones de optimalidad del problema formulado, se realiza un análisis numérico para: i) encontrar el punto óptimo de equilibrio entre costes de operador y usuarios, ii) caracterizar la potencial pérdida de optimalidad con respecto a una formulación MIP y iii) realizar un análisis de sensibilidad de la solución con respecto a los parámetros más relevantes.

- En el capítulo 4 se aplica el marco genérico propuesto a una potencial red de metro en la ciudad de Sevilla operada por una sociedad concesionaria de la Junta de Andalucía, donde se exploran distintas configuraciones presupuestarias, llegando a un punto de equilibrio entre la inversión necesaria y el impacto social. Además, se tienen en cuenta diversos factores como pueden ser la renta media o el índice de crecimiento poblacional, desagregados por distritos, de cara a obtener formulaciones socialmente responsables y orientadas al futuro de la ciudad. Cabe destacar que la propuesta de diseño basada en optimización convexa consigue resolver el problema de diseño de infraestructuras en unos pocos minutos, mientras que los diseños existentes basados en MIP no son capaces de encontrar la solución después de un periodo de 48 horas.
- En el capítulo 5 el marco genérico de optimización convexa propuesto se aplica a una RdT aéreo, donde el operador es una aerolínea comercial que pretende comenzar a operar el mercado doméstico estadounidense. En este contexto, se diseñan las rutas a operar y se identifican los aeropuertos estratégicos para establecer bases operativas, conocidas como *hubs*. Se presenta la estructura inicial de datos, para posteriormente discutir las modificaciones necesarias en la formulación del modelo (particularmente relevantes dado que se trata de un medio de transporte distinto y la entidad operadora es privada). La parte final del capítulo se dedica a la exposición y el análisis de los resultados experimentales obtenidos.

Las conclusiones y las líneas futuras de trabajo, que se proporcionan en el capítulo 6, cierran este TFG.

1.4. Grupos de tareas y tiempo dedicado

El TFG que se resume en esta memoria se diseñó con unos objetivos relativamente ambiciosos, combinando una dimensión teórica (formulación de nuevos problemas de investigación y diseño de los algoritmos asociados a los mismos) con una dimensión práctica (diseño de dos RdT reales). Desde un punto de vista formativo, se considera que la realización de este TFG ha supuesto poner en práctica conocimientos, competencias y habilidades diversas, aprendidas a lo largo de toda la titulación (diseño y operación de redes aéreas, programación, optimización lineal...). También ha supuesto la adquisición de nuevos conocimientos y nuevas habilidades. Entre los nuevos conocimientos deben destacarse la optimización convexa y las técnicas de *sparsity* y, entre las nuevas habilidades, deben destacarse la lectura de artículos científicos, la redacción de documentos en inglés o la colaboración con grupos internacionales.

En lo que a la planificación y ejecución se refiere, el trabajo realizado puede articularse en torno a grupos de tareas, que se describen muy brevemente a continuación:

- Revisión del estado de arte (lectura de artículos de investigación en el ámbito de diseño de RdT, *sparsity*, redes *hub & spoke* y modelos de demanda). Esta tarea se llevó a cabo a través de la lectura de artículos científicos. Los artículos iniciales fueron propuestos por los directores del TFG y, según avanzó el trabajo, el estudiante encontró también referencias que resultaron relevantes para el diseño de la solución final propuesta.
- Aprendizaje de manejo de herramientas de optimización (CVX y MOSEK). Esta tarea se llevó a cabo principalmente mediante la lectura y visionado de tutoriales, la realización de pequeños ejemplos de juguete y la discusión con los directores del TFG.
- Formulación de los problemas de optimización y técnicas de resolución asociadas. El rol de los tutores fue crítico en los primeros pasos, ganando el estudiante autonomía en las fases de diseño final, donde se presentaban varias opciones a los tutores y se decidía de forma colegiada.

- Programación de algoritmos de optimización. La práctica totalidad de la programación se ha llevado a cabo en Matlab, comenzando por casos más sencillos y construyendo poco a poco versiones más complejas.
- Acceso y procesamiento de bases de datos públicas. Se utilizaron tanto bases de datos públicas como privadas (a disposición de los grupos de investigación dirigidos por los tutores de este TFG). El procesamiento se ha llevado a cabo fundamentalmente en Matlab, aunque para casos específicos se ha utilizado Excel y Python.
- Diseño de experimentos numéricos y ejecución de análisis de resultados. Las tareas en este caso van más allá de programar los experimentos, incluyendo también la selección de los casos de estudio que resultaban más significativos para mostrar los beneficios de la metodología propuesta, así como decidir la forma en la que se visualizarían y representarían los resultados.
- Redacción de documentos científico-técnicos, incluyendo la presente memoria, así como los artículos científicos derivados del trabajo de este TFG.

De forma general, las tareas previamente listadas se han ejecutado de una forma secuencial. No obstante, en ocasiones, los resultados intermedios en alguna de las tareas han revelado problemas o aspectos no previstos inicialmente, lo que ha exigido visitar las tareas precedentes. Entre los ejemplos en esta última categoría, pueden mencionarse la modificación de la formulación en función de los resultados obtenidos durante las simulaciones o la lectura de literatura adicional cuando, en la formulación, se ha decidido incorporar una nueva característica o particularidad de la RdT.

El trabajo aquí presentado se ha llevado a cabo durante 9 meses (agosto de 2023 a mayo de 2024) y ha supuesto la inversión de unas 900 horas de trabajo por parte del estudiante. De forma aproximada, las tareas de revisión del estado del arte y aprendizaje de herramientas consumieron un 10 % del tiempo. La formulación, diseño y programación de algoritmos consumió en torno al 40 % del tiempo. El procesamiento de datos, diseño y ejecución de simulaciones consumió un 25 % del tiempo y el análisis y visualización de resultados un 10 %. La redacción de los documentos consumió un 15 % del tiempo.

1.5. Colaboraciones, publicaciones y repositorio de código

El trabajo presentado en esta memoria se enmarca dentro de una colaboración entre dos grupos de investigación de alto rendimiento de la Universidad Rey Juan Carlos (URJC): el Grupo de Investigación en Sistemas Aeroespaciales y Transporte (GISAT-URJC) y el Grupo de Ciencia de Datos y Procesamiento de Señal (GICDPS-URJC). Más específicamente, las tareas de investigación y desarrollo realizadas en este TFG se asocian a paquetes de trabajo de los proyectos de investigación competitivos “Procesamiento de señal para datos definidos sobre grafos: aprovechando la estructura en dominios irregulares” (MCI/AEI/PID2019-105032GB-I00) e “Integración de las ciencias del transporte y de los datos para impulsar una movilidad en red socialmente responsable y un entorno sostenible” (MCI/AEI/TED2021-130347B-I00). Debo por tanto, agradecer la ayuda prestada por los directores de esos dos grupos (tutores de este TFG) como del resto de los miembros. También debo reconocer el apoyo financiero de la Agencia Estatal de Investigación, que me ha permitido llevar a cabo las tareas asociadas a este TFG, así como la asistencia a conferencias y las visitas a los grupos de investigación que colaboran en los proyectos anteriormente mencionados. En este sentido, quiero expresar mi gratitud a los profesores Geert Leus (Delft University of Technology, Países Bajos), Vikrant Vaze (Dartmouth College, EE.UU.) y Tianyi Chen (Rensselaer Polytechnic Institute, EE.UU.) por su ayuda y su colaboración en las tareas llevadas a cabo en este proyecto. Los consejos y contribuciones del Prof. Leus (métodos de optimización inductores de *sparsity*), el Prof. Vaze (modelado de preferencias de pasajeros en RdT) y el Prof. Chen (optimización binivel y multiobjetivo) en distintos ámbitos no solo me han ayudado a adquirir nuevos conocimientos teóricos y a manejar nuevas herramientas, sino que han sido extremadamente valiosos para llevar a cabo el trabajo que se detalla en los capítulos de esta memoria.

En lo que a publicaciones se refiere, en la actualidad se está trabajando en 3 artículos de revista (Q1) y 2 artículos de conferencia (uno de ellos ya aceptado). Los tres artículos de revista son:

- F. Real-Rojas, V. M. Tenorio, L. Cadarso, and A. G. Marques, “Rapid transit network design via convex optimization: The role of sparsity and entropy-based regularizers”, *Transportation Research Part B: Methodological* (to be submitted), 2024.
- P. Singh, F. Real-Rojas, V. M. Tenorio, L. Cadarso, A. G. Marques, and V. Vaze, “Optimal design of logit-based capacitated transportation networks: Convex vs. MIP approaches”, *Transportation Science* (to be submitted), 2024.
- F. Real-Rojas, V. M. Tenorio, L. Cadarso, and A. G. Marques, “A convex sparsity-promoting approach for designing hub-and-spoke air transportation networks”, *Transportation Research Part C: Emerging Technologies* (to be submitted), 2024.

Los dos primeros artículos están relacionados con los capítulos 3 y 4 de esta memoria. El primero de ellos se centra en la introducción de las funciones logarítmicas y las funciones de entropía para el diseño de RdT, así como su aplicación a redes de metro. El segundo de ellos aplica la metodología propuesta en este TFG a redes ferroviarias, estudiando también la modificación de las funciones de entropía, para poder cubrir un mayor espectro de modelos de preferencias por parte de los usuarios (esta aplicación no se ha incluido en el presente TFG por limitaciones de espacio). Por último, el tercer artículo se centra en el diseño de RdT aéreo del tipo *hub & spoke* (correspondiente al contenido del capítulo 5).

Finalmente, el código utilizado para ejecutar los experimentos numéricos presentados en los distintos capítulos se encuentra disponible en:

- <https://github.com/frealr/TFG-RdT>

Capítulo 2

La optimización en RdT: formulaciones enteras y una propuesta convexa

El crecimiento sin precedentes de los recursos computacionales a nuestro alcance, así como el acceso a una cantidad ingente de bases de datos, ha provocado que las herramientas de optimización se hayan convertido en un actor principal y transversal en el ámbito de la ingeniería y las ciencias aplicadas. Las RdT no son una excepción y, tal y como se mencionaba en el capítulo introductorio, los avances fundamentales en el diseño y control de RdT inteligentes exigen un planteamiento holístico riguroso en el que la optimización matemática desempeñe un papel protagonista. Desde este punto de vista, el presente capítulo discute el rol de la optimización en el diseño de RdT, describe las aproximaciones de diseño clásicas basadas en MIP y explica cómo el uso combinado de regularizadores que promueven *sparsity* en la infraestructura y de funciones de entropía aplicadas a flujos de pasajeros permiten atacar el diseño de RdT desde un punto de vista convexo.

Este capítulo se estructura como sigue. La sección 2.1 proporciona una breve introducción a la metodología del diseño de redes basada en optimización, identificando los distintos elementos clave a tener en cuenta. A continuación, la sección 2.2 presenta dos formulaciones MIP (simplificadas) que sirven para ilustrar tanto distintos modelos de demanda, como las limitaciones de este tipo de metodologías. La segunda mitad del capítulo (secciones 2.3 y 2.4) describe dos de las contribuciones fundamentales de este TFG: el uso de técnicas de *sparsity* para el modelado de los costes de construcción de la infraestructura de la red (estaciones y enlaces) y el uso de funciones de entropía para el modelado del comportamiento de los usuarios. Tal y como se detallará en el capítulo siguiente, la implementación simultánea de estas dos estrategias permite diseñar la RdT mediante técnicas de optimización convexa.

2.1. Formalización del problema de diseño de RdT

En este TFG los esquemas para el diseño de RdT se obtienen como soluciones de un problema de optimización meticulosamente formulado. En este sentido, la selección de una formulación de optimización adecuada (interpretable y que equilibre la precisión del modelo con la complejidad computacional asociada a su optimización) es fundamental para lograr un diseño satisfactorio. Esto no es trivial y, en general, requiere un diseño iterativo de la formulación del problema, replanteando o modificando formulaciones iniciales en función de los resultados intermedios que se van obteniendo. Para ser más precisos, hay 5 elementos clave a considerar cuando se diseña la formulación de un problema de optimización:

- E1. Variables de estado, que sirven de “entrada” para la optimización y dan cuenta de la naturaleza distribuida, dinámica (y potencialmente incierta) del estado de la red y de los actores que la utilizan.

- E2. Variables de diseño, que son la “salida” de la optimización y representan los recursos que pueden emplearse durante la monitorización y operación de la RdT -las variables de diseño dependen del segmento de la red, del horizonte temporal y del participante en el transporte-.
- E3. Restricciones, que modelan (hacen cumplir) las relaciones entre las variables de diseño y de estado y dan cuenta tanto de las condiciones de la red (tales como las restricciones de capacidad o de restricción de flujo) como de posibles requisitos de rentabilidad económica o de calidad de servicio.
- E4. La función objetivo a optimizar (beneficios, costes, satisfacción del usuario, etc.), que debe tener en cuenta los servicios prestados, los costes de explotación de la red y las necesidades de los distintos actores (operadores, usuarios, reguladores...), así como la variabilidad, la incertidumbre y los distintos horizontes temporales de las RdT.
- E5. La técnica de resolución a utilizar, que es particularmente importante en esta propuesta debido a la alta dimensionalidad y el fuerte acoplamiento entre las variables de una RdT.

Los elementos listados son críticos para la formulación de cualquier problema de ingeniería como un problema de optimización. A continuación, se proporcionan más detalles sobre cómo este TFG aborda cada uno de estos elementos dentro del contexto del diseño de una RdT.

2.1.1. Variables de estado

Una de las características de las RdT es que, en su diseño, intervienen tanto variables continuas como discretas. Entre las variables de estado discretas se incluyen: el listado de orígenes y destinos; el listado de estaciones y enlaces (tanto existentes como potenciales despliegues), la topología de la red competidora (en caso de que exista) o las rutas que siguen los usuarios en algunas de esas redes. Entre las variables de estado continuas más relevantes se encuentran: el tiempo de viaje entre estaciones, la demanda existente por parte de los pasajeros y las características o propiedades del servicio ofrecidas por la competencia (precio, tiempos de viaje...), sobre las cuales no tenemos control directo.

Otras variables de estado que impactan en el diseño incluyen los costes de construcción de la infraestructura, los costes de operación de la red, la capacidad de los vehículos de transporte (trenes, aviones...) o los coeficientes que modelan la sensibilidad de la utilidad percibida por los usuarios frente a distintos factores (precio, tiempo de viaje...). Es importante señalar que, mientras que en el capítulo 3 estas variables tomarán valores obtenidos del estado del arte [26, 27, 18, 28], en los capítulos 4 y 5 se estimarán a partir de bases de datos relacionadas con la aplicación de transporte estudiada.

Por último, es importante hacer notar que la (topología de la) red suele representarse a través de un grafo, donde los nodos se corresponden con las estaciones y los arcos indican enlaces directos entre estaciones. Así mismo, muchas de las variables de estado descritas anteriormente se representan comúnmente como vectores, cuya longitud coincide con el número de nodos del grafo (*e.g.*, el coste de construcción de cada estación), o como matrices cuadradas donde filas y columnas indexan nodos y cada entrada representa un par de nodos específico (*e.g.*, la matriz de tiempos de viaje a un solo salto o la matriz de demanda origen-destino).

2.1.2. Variables de diseño

En el contexto de este TFG, las variables asociadas al diseño de la RdT pueden agruparse en torno a tres categorías:

1. Variables asociadas con la construcción de la red. Podemos distinguir dos subgrupos: a) variables asociadas a la construcción de nodos en la red y b) variables asociadas a la construcción de enlaces entre nodos. Estas variables tienen un significado dicotómico, en el sentido de que, si valen cero,

indican que la estación (o el enlace) no se construye y, si toman un valor positivo, indican que la infraestructura se construye. En el estado del arte, este tipo de variables suele representarse con variables binarias [26, 27], aunque existen aproximaciones en las que son continuas [29, 30], correspondiéndose en este caso su valor con la capacidad dimensionada.

2. Variables asociadas con los flujos de pasajeros que utilizan la red para viajar de un determinado origen a un destino. Debe distinguirse entre variables de flujo agregado origen-destino (que indican el número de pasajeros asociados a un determinado mercado) y variables de flujo origen-destino por enlace (que indican el número de pasajeros de un determinado mercado que están utilizando un determinado enlace). Al igual que las variables de construcción, tienen un comportamiento dicotómico.
3. Variables secundarias. Este grupo engloba las variables auxiliares de los modelos, que son importantes para obtener modelos eficientes que se puedan resolver en tiempo polinómico. También incluye aquellas variables que pretenden capturar algunos de los efectos que se dan en RdT reales, como puede ser la congestión de los enlaces o la holgura en el diseño de la infraestructura. Por lo tanto, en función del tipo de variables, podrán estar asociadas, a un nodo, a un enlace, a un flujo origen-destino o a una combinación de varios de ellos.

2.1.3. Restricciones

De forma general, una restricción se puede formular como una condición que se impone al valor de una función que depende tanto de las variables de diseño como de las variables de estado. En su forma más sencilla, una restricción impone una condición al valor de la variable de diseño (*e.g.*, no ser negativa o ser binaria). Así mismo, las restricciones pueden modelar condiciones de operación de la red o requisitos de calidad de servicio para el operador o los usuarios. En el ámbito de las RdT, algunas de las restricciones que impactan de forma crítica en el diseño de la red son:

- Restricciones de conservación de flujo. Imponen que el flujo total que llega a un nodo (sumado a través de todos los enlaces entrantes) es igual al flujo total que sale del nodo (sumado a través de todos los enlaces salientes), exceptuando los nodos que ejercen de origen o de destino.
- Restricciones de capacidad para la infraestructura. Imponen que el uso de un determinado recurso no puede exceder la capacidad máxima del mismo. En el caso de los enlaces, establecen que el número de usuarios que utiliza un determinado enlace tiene que estar por debajo de la capacidad máxima de dicho enlace. En el caso de las estaciones, establecen que la estación debe tener una capacidad suficiente para gestionar todos los usuarios (o vehículos) que llegan a través de sus respectivos enlaces. Así mismo, dentro de este grupo de restricciones pueden incluirse otras como una restricción presupuestaria (*e.g.*, los costes de construcción no pueden superar un determinado nivel de inversión) o una restricción de flota (*e.g.*, la capacidad agregada de la red no puede superar el resultado de multiplicar el número de vehículos disponibles por la capacidad de estos vehículos).
- Restricciones de calidad de servicio. Destacan restricciones de satisfacción de la demanda, que garantizan que bien la totalidad o bien un porcentaje alto de la demanda de los usuarios queda satisfecha. Del mismo modo, pueden considerarse restricciones que estipulan que una RdT debe construirse para garantizar la atracción de un porcentaje mínimo de demanda o, en el caso de operadores privados, restricciones que garanticen que rutas con beneficios por debajo de un umbral mínimo no se operen. Así mismo, dentro de este grupo se consideran las restricciones de holgura, que imponen que la capacidad finalmente construida para un día tipo sea mayor que la inicialmente planificada. Se utilizan para mantener la congestión bajo control y ser capaces de hacer frente a cambios repentinos, fluctuaciones en la demanda o, incluso, disrupciones en la red.

Las restricciones anteriores son algunas de las más representativas en el diseño de RdT. A lo largo de esta memoria, se formularán otras que se irán introduciendo a medida que se vayan necesitando.

2.1.4. Función objetivo

La función objetivo es aquella que asigna un valor de mérito (o de coste) a cada una de las soluciones factibles y, por lo tanto, nos permite identificar la solución óptima. En el contexto de este TFG, debe tenerse en cuenta que las RdT son sistemas distribuidos con múltiples actores y horizontes temporales, lo que complica la definición de una única función objetivo que tenga en cuenta toda la complejidad de la red. Los objetivos a maximizar incluyen beneficios económicos y tasas de cobertura, entre otros, y los objetivos a minimizar incluyen costes fijos, costes de operación, congestión y retardos, entre otros. Los actores a considerar incluyen los operadores, los usuarios y, en ocasiones, el propio regulador. Una metodología razonable consiste en agrupar los distintos objetivos en función de los horizontes temporales o los actores involucrados, otorgando un peso distinto a cada uno de esos roles. En los problemas investigados en este TFG, se considerarán tres objetivos principales: i) costes de construcción de la red (véase, *e.g.*, [31]), ii) costes e ingresos de la operación de la red (véase, *e.g.*, por ejemplo [23]) y iii) satisfacción de los usuarios (véase, *e.g.*, [32]). En función del tipo de problema considerado, estos objetivos se podrán formalizar (modelar) de manera ligeramente distinta, siendo la función objetivo a optimizar una combinación lineal de cada uno de ellos, con los pesos asociados teniendo en cuenta tanto la importancia relativa del objetivo como el margen dinámico de la métrica que se está optimizando.

2.1.5. Algoritmo de solución

El último aspecto a considerar es la técnica utilizada para resolver el problema de optimización. El diseño de algoritmos de optimización es un área amplia y diversa, con resultados muy distintos en función de las propiedades del problema en cuestión.

Los problemas convexos con variables continuas y diferenciables se caracterizan por una estructura favorable que facilita la aplicación de algoritmos iterativos de descenso. Entre los más destacados se encuentran los algoritmos de gradiente, de Newton y los algoritmos acelerados, todos ellos reconocidos por su eficacia y versatilidad en la optimización de este tipo de problemas [33]. A pesar de no ser diferenciables, los problemas lineales con variables continuas y restricciones presentan una estructura que permite aplicar métodos eficientes. El método *simplex* [34], introducido hace más de 70 años, sigue siendo una opción viable y eficaz para resolver este tipo de problemas, demostrando consistentemente su eficacia a lo largo del tiempo. En el caso de problemas MIP, la elección del algoritmo adecuado puede depender de diversos factores, como la naturaleza específica del problema, el tamaño del modelo y la disponibilidad de recursos computacionales. Desde métodos exactos, como el *Branch-and-Bound* [35], hasta enfoques heurísticos, como la metaheurística de búsqueda local [36], existe una amplia gama de técnicas disponibles para abordar eficazmente los problemas MIP. Es crucial evaluar cuidadosamente las características del problema y las limitaciones computacionales para seleccionar la estrategia más adecuada que garantice resultados óptimos en un tiempo razonable. Un aspecto crítico en cualquiera de estos métodos consiste en cómo tratar con las restricciones. En formulaciones convexas es habitual, por ejemplo, utilizar métodos de punto interior basados en barreras logarítmicas. Otra aproximación consiste en recurrir a los métodos duales, los cuales son de aplicación general y definen un multiplicador de Lagrange asociado a cada una de las restricciones, para formar un objetivo aumentado sin restricciones y optimizar simultáneamente sobre variables primales y duales.

Cada uno de los algoritmos anteriores tiene una complejidad computacional y tasa de convergencia diferente. Para problemas en los que la complejidad es crítica (bien por el gran tamaño del modelo o bien porque resulta necesario resolverlo en tiempo real) los algoritmos pueden diseñarse de manera individualizada, teniendo en cuenta las particularidades del objetivo y restricciones del problema en cuestión. En la mayoría de las aplicaciones es más habitual utilizar software de optimización generalista (conocido en inglés como *solver*) en el que el problema de optimización se describe a alto

nivel (especificando las funciones consideradas en el objetivo y las restricciones) y el propio software se encarga de desarrollar y ejecutar los algoritmos asociados. Entre los *solvers* más populares pueden destacarse CPLEX, GUROBI o MOSEK (para problemas lineales o MIP) o SeDuMi, SCS, ECOS (para problemas convexos). Si bien dichos *solvers* permiten implementar directamente los problemas a resolver, es habitual utilizar herramientas de más alto nivel (como CVX), que se integran de manera natural en lenguajes de programación generalistas como Python o Matlab, y que llaman a los *solvers* anteriormente listados, ofreciendo una interfaz más amigable para el usuario.

En este sentido, es importante destacar que también pueden plantearse formulaciones híbridas que combinan ambas metodologías. Así, por ejemplo, las variables pueden agruparse en dos subconjuntos y optimizarse de forma iterativa, utilizando para un subconjunto un tipo de algoritmo (*e.g.*, un algoritmo de descenso de segundo orden) y utilizando un *solver* genérico para el segundo subconjunto.

Un aspecto fundamental cuando se elige cualquiera de estos algoritmos o *solvers* es la complejidad computacional asociada a los mismos. De manera general, la complejidad será mayor cuanto mayor sea el número de variables a optimizar y el número de restricciones a resolver. Si las variables son continuas y el problema de optimización es lineal o convexo, la complejidad computacional escala de forma polinómica, con un exponente de entre 3 y 4 [37]. El caso MIP es más complejo y depende enormemente del problema, siendo muchos de ellos NP-difíciles y, por lo tanto, teniendo una complejidad computacional exponencial. Gracias a las mejoras algorítmicas y el incremento de la capacidad de computación, en la actualidad es factible resolver una amplia gama de problemas MIP del mundo real que involucran millones de decisiones discretas y continuas, proporcionando soluciones óptimas o aproximadas según las necesidades específicas del problema.

La enorme diferencia en términos de complejidad computacional entre problemas convexos y problemas MIP es una de las principales motivaciones de este TFG. Puesto que el diseño de infraestructuras de RdT se ha venido modelando habitualmente como un problema MIP y una red involucra un gran número de variables y restricciones, obtener su solución en un tiempo reducido es prácticamente inviable. Como se demostrará en el resto de esta memoria, modelando el problema como una optimización convexa, se obtiene la solución óptima de manera mucho más rápida, pudiendo utilizar incluso ordenadores personales para la ejecución de los algoritmos de diseño de RdT.

En lo que a los algoritmos propuestos en este TFG se refiere, se utilizará un método híbrido en el que, por un lado, se emplearan algoritmos de MM iterativos para un subconjunto de variables [38] y el *solver* genérico MOSEK-CVX para el resto de variables.

2.2. Diseño de RdT vía MIP

En esta sección se presentan dos variantes de formulación matemática para abordar el problema del diseño de una RdT. En primer lugar, se introduce al lector en la estructura de las RdT y en la notación utilizada en la definición del modelo matemático para su optimización. Posteriormente, se describe un modelo matemático que se fundamenta en el reparto binario de la demanda, también conocido como “todo o nada” (*all-or-nothing*, en inglés). A continuación, se procede con una formulación que se basa en la distribución de la demanda utilizando modelos del tipo *logit*, típicamente empleados para modelar la elección de rutas por parte de los usuarios de una red. Este tipo de modelo se basa en la función logística, que relaciona la probabilidad de que un individuo (pasajero) elija una opción particular con una combinación lineal de variables explicativas.

2.2.1. Topología y variables de estado

Antes de comenzar con la formulación es importante familiarizarse con la estructura de una RdT. En la figura 2.1 se puede ver una red de ejemplo. Se plantea la construcción de una RdT que pretende comunicar los nodos recogidos en el conjunto $\mathcal{S}$ mediante los enlaces pertinentes recogidos en el conjunto $\mathcal{A}$. El número de nodos es, por lo tanto, $|\mathcal{S}|$ y el número de enlaces es $|\mathcal{A}|$. La red construida se dedicará

a transportar flujos de pasajeros asociados a distintos mercados. Los mercados existentes se recogen en el conjunto $\mathcal{K}$. Puesto que cada mercado se corresponde con un par origen-destino, el número máximo de mercados es $|\mathcal{K}| = |\mathcal{S}|^2 - |\mathcal{S}|$. A continuación se definen formalmente los conjuntos y parámetros del modelo. Para denotar los conjuntos, se utilizarán letras caligráficas, mientras que los parámetros se denotan con letras minúsculas del alfabeto latino y griego.

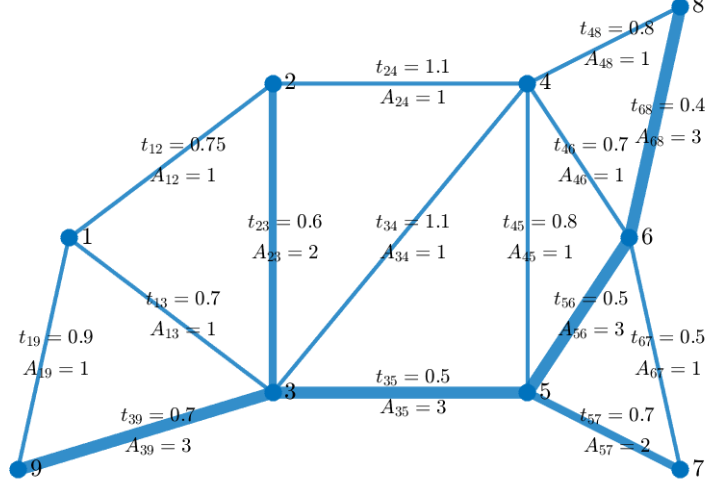


Figura 2.1: Ejemplo simplificado de grafo que representa a una RdT, donde $\mathcal{S}$ es el conjunto de nodos y $\mathcal{A} \subseteq \mathcal{S} \times \mathcal{S}$ es el conjunto de enlaces. En el grafo representado el conjunto de nodos es $\mathcal{S} = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ y, por lo tanto, el número de nodos es $|\mathcal{S}| = 9$. El número de enlaces es $|\mathcal{A}| = 30$ (15 segmentos con dos potenciales sentidos cada uno), frente a un máximo posible de 72. Así, por ejemplo, $(1, 2) \in \mathcal{A}$ pero $(1, 4) \notin \mathcal{A}$. Para cada enlace se muestra tanto la capacidad del mismo como el tiempo de viaje entre los nodos incidentes al enlace. Pese a que de forma general la matriz de capacidad de enlaces $\mathbf{A}$ no será simétrica, para facilitar su representación, en este ejemplo se asume que $A_{ij} = A_{ji} \forall (i, j) \in \mathcal{A}$.

Conjuntos

- $\mathcal{S}$ es el conjunto de estaciones candidatas para ser construidas. Las estaciones se indexan normalmente utilizando el índice $i = 1, \dots, |\mathcal{S}|$ y, ocasionalmente, utilizando el índice j .
- $\mathcal{A} \subseteq \mathcal{S} \times \mathcal{S}$ es el conjunto de enlaces candidatos para construir. Los elementos de $\mathcal{A}$ se denotan típicamente como (i, j) y suponemos que (i, j) es diferente de (j, i) .
- $\mathcal{N}_{out}(i)$ y $\mathcal{N}_{in}(i)$ son los subconjuntos de estaciones candidatas a las que se puede construir un enlace desde o hacia i , respectivamente. Es decir, $\mathcal{N}_{out}(i) = \{j \mid (i, j) \in \mathcal{A}\}$ y $\mathcal{N}_{in}(i) = \{j \mid (j, i) \in \mathcal{A}\}$.
- $\mathcal{K}$ es el conjunto de mercados formados por pares origen-destino entre los que existe demanda. Los pares origen-destino son referidos con el par (o, d) , que indexa el nodo origen y el nodo destino de dicha demanda, respectivamente.
- $\mathcal{F}$ es el conjunto de flujos de pasajeros que viajan a través de los enlaces. Dado que se consideran múltiples mercados, tenemos que $\mathcal{F} = \mathcal{A} \times \mathcal{K}$. Los elementos de $\mathcal{F}$ se indexan mediante dos pares $((i, j), (o, d))$.
- $\mathcal{Q}$ es el conjunto de alternativas de viaje entre las que escogen los pasajeros.

Parámetros

- c_{s_i} es el coste fijo de construcción de la estación $i \in \mathcal{S}$.
- $\bar{c}_i$ es el coste variable (por unidad de capacidad) de construcción de la estación $i \in \mathcal{S}$.

- $c_{A_{ij}}$ es el coste fijo de construcción del enlace $(i, j) \in \mathcal{A}$.
- $\bar{c}_{A_{ij}}$ es el coste variable (por unidad de capacidad) de construcción del enlace $(i, j) \in \mathcal{A}$.
- $\sigma_i \in (0, 1]$ es el factor de carga asociado a la estación $i \in \mathcal{S}$. Este parámetro tiene en cuenta el hecho de que no todos los pasajeros (vehículos) que pasan por la estación i durante el horizonte de planificación lo hacen simultáneamente.
- a_{nom} es la capacidad nominal del vehículo, que es igual al número de pasajeros que el vehículo puede transportar de forma simultánea.
- a_{max} es la capacidad máxima de la red. Representa una cota superior al número de vehículos a operar, aproximado como el valor máximo que puede tomar la suma de las capacidades en los enlaces $(i, j) \in \mathcal{A}$ de la red. Si tal límite no existe, entonces su valor se fija como $a_{max} = \infty$.
- t_{ij} es el tiempo de viaje asociado al enlace $(i, j) \in \mathcal{A}$ en la nueva red.
- t_{ext}^{od} es el coste de viaje entre o y d para los pasajeros cuando utilizan una RdT alternativa, es decir, cuando viajan con la competencia.
- ω_t es el coeficiente de sensibilidad de los pasajeros al tiempo de viaje en la función de utilidad.
- w^{od} es la demanda total del mercado $(o, d) \in \mathcal{K}$, es decir, el número de pasajeros que desean viajar del nodo o al nodo d .
- b es el presupuesto disponible para la construcción de la nueva RdT.
- Ξ es un parámetro auxiliar al que se le da un valor alto. Se utiliza para relacionar variables binarias y continuas, normalmente para restricciones en las que para activar la variable continua es necesario activar también la binaria.

Una vez que se han presentado los conjuntos y parámetros más relevantes para nuestro problema, se procede a presentar la formulación del modelo matemático (formulación) considerando dos aproximaciones diferentes para la demanda. Los modelos de demanda de transporte se utilizan para obtener pronósticos sobre el uso de la infraestructura que, para cada uno de los mercados, los pasajeros llevarán a cabo. Habitualmente estos pronósticos se basan en modelos no lineales que tienen en cuenta atributos tales como las tarifas promedio, las frecuencias de servicio, las características demográficas o las condiciones económicas [39]. Una vez estimada la demanda total, se utilizan modelos de elección para estimar, para cada operador competidor y cada mercado, la proporción de demanda que ese operador captura en ese mercado. Para ello, se consideran las características específicas del mercado, incluyendo las preferencias de modo de transporte y de operador de los pasajeros, las tarifas, las frecuencias de servicio y otros potenciales. Las secciones 2.2.2 y 2.2.3 presentan la formulación MIP para dos modelos (aproximaciones) diferentes en lo que a la elección de mercado por parte de los pasajeros se refiere. En la primera de ellas, se presenta el modelo con una aproximación de elección de mercado “todo o nada” (*all-or-nothing*, en inglés). En la segunda, se presenta el modelo para una aproximación tipo *logit*.

2.2.2. Formulación basada en aproximación *all-or-nothing* para la demanda

En esta sección se presenta un modelo matemático para abordar el problema de diseño de redes, donde el objetivo es maximizar la demanda atraída o, equivalentemente, minimizar el número de usuarios que optan por redes competidoras para sus viajes.

Para modelar la elección de los pasajeros, se emplea un enfoque de elección discreta conocido como *all-or-nothing*. Los modelos de elección discreta han desempeñado un papel importante en la modelización del transporte en las últimas décadas. Estos modelos consideran que la demanda es el resultado de las decisiones de cada individuo en la población bajo consideración (*i.e.*, de cada potencial

pasajero de la RdT). Estas decisiones suelen consistir en una elección entre un conjunto finito de alternativas.

En primer lugar, es necesario definir las características de los individuos, las alternativas como las opciones posibles del individuo, los atributos de cada alternativa potencial que el individuo está considerando y las reglas de decisión que describen el comportamiento del individuo.

Los individuos con características similares se agregan en grupos. Es posible distinguir diferentes grupos de pasajeros, por ejemplo, según su forma de pago de transporte (*e.g.*, el grupo de personas jubiladas, el grupo de trabajadores activos, el grupo de turistas y el grupo de estudiantes).

El modelo *all-or-nothing* parte del supuesto de que los pasajeros fundamentan su elección en la comparación de un coste generalizado o función de utilidad (u_q^{od}) de la alternativa q , la cual es una función de los atributos de la misma y del individuo en el mercado (o, d) . En este caso, los pasajeros siempre optan por la alternativa que presenta el menor coste o la mayor utilidad. Bajo la hipótesis de que los pasajeros se agregan en grupos homogéneos, este modelo supone que la demanda para cada mercado no se divide entre las opciones disponibles, sino que el grupo elige una única alternativa para viajar. Este enfoque simplificado permite una representación clara y directa de las decisiones de los usuarios, dando lugar, además, a problemas de optimización más sencillos.

Usando t_q^{od} para denotar el tiempo de viaje de la alternativa q para el mercado (o, d) y asumiendo que la función de utilidad de una alternativa es función únicamente del tiempo de viaje necesario para trasladarse desde el origen hasta el destino ($u_q^{od} = f(t_q^{od})$), un pasajero en el mercado (o, d) optará por una alternativa $q \in \mathcal{Q}$ si y solo si se satisface que $u_q^{od} \geq u_{q'}^{od}$ para toda alternativa q' compatible con q .

Una vez definido el modelo de elección de los usuarios, se formula a continuación el problema de diseño óptimo de la RdT. Para ello, se comienza declarando las variables sobre las que se va a optimizar (distinguiendo entre las del operador y las de los usuarios) y, a continuación, se detalla el modelo de optimización asociado.

Variables

Variables de construcción

- $S_i^b \in \{0, 1\}$ toma el valor de 1 si se construye una estación en el nodo $i \in \mathcal{S}$. Toma el valor 0 en cualquier otro caso.
- $S_i \in \mathbb{R}^{\geq 0}$ es la capacidad que se construye en la estación del nodo $i \in \mathcal{S}$.
- $A_{ij}^b \in \{0, 1\}$ toma el valor de 1 si se construye el enlace entre los nodos $(i, j) \in \mathcal{A}$. Vale 0 en cualquier otro caso.
- $A_{ij} \in \mathbb{R}^{\geq 0}$ es la capacidad que se construye para el enlace $(i, j) \in \mathcal{A}$.

Variables de flujo

- $F_{ext}^{od} \in \{0, 1\}$ toma el valor de 1 si los pasajeros del mercado $(o, d) \in \mathcal{K}$ viajan por la RdT alternativa para ir a su destino. Vale 0 en cualquier otro caso.
- $F^{od} \in \{0, 1\}$ toma el valor de 1 si los pasajeros del mercado $(o, d) \in \mathcal{K}$ viajan por la red construida para ir a su destino. Vale 0 en cualquier otro caso.
- $F_{ij}^{od} \in \{0, 1\}$ toma el valor de 1 si los pasajeros del mercado $(o, d) \in \mathcal{K}$ recorren el enlace $(i, j) \in \mathcal{S}$ de la red construida en la ruta hacia su destino. Vale 0 en cualquier otro caso.

Con el objetivo de simplificar la notación, se introducen los conjuntos de variables $\mathcal{X}_{cons}$ y $\mathcal{X}_{flow}$, definidos en (2.1) y (2.2), respectivamente. Esta definición permite una representación más compacta y clara de los distintos problemas estudiados en este TFG.

$$\mathcal{X}_{cons} = \{S_i\}_{\forall i \in \mathcal{S}} \cup \{S_i^b\}_{\forall i \in \mathcal{S}} \cup \{A_{ij}\}_{\forall (i,j) \in \mathcal{A}} \cup \{A_{ij}^b\}_{\forall (i,j) \in \mathcal{A}} \quad (2.1)$$

$$\mathcal{X}_{flow} = \{F_{ext}^{od}\}_{\forall (o,d) \in \mathcal{K}} \cup \{F^{od}\}_{\forall (o,d) \in \mathcal{K}} \cup \{F_{ij}^{od}\}_{\forall ((i,j),(o,d)) \in \mathcal{F}} \quad (2.2)$$

Introducidas las variables de optimización y las variables de estado (véase la sección 2.2.1), el siguiente paso consiste en formular el problema de optimización, lo que requiere especificar la función objetivo y las restricciones. La función objetivo se define en (2.3) y las restricciones se definen en (2.4)-(2.12). Tanto en este como en el resto de problemas de optimización de este documento, “mín” indica “mínimo” y “s. a” indica “sujeto a”.

$$\min_{\mathcal{X}_{cons}, \mathcal{X}_{flow}} \sum_{\forall (o,d) \in \mathcal{K}} w^{od} F_{ext}^{od} \quad (2.3)$$

s. a:

$$\frac{\sum_{\forall (o,d) \in \mathcal{K}} F_{ij}^{od} w^{od}}{a_{nom}} \leq A_{ij} \quad \forall (i,j) \in \mathcal{A} \quad (2.4)$$

$$\sigma_i \sum_{\forall j \in \mathcal{N}_{in}(i)} A_{ji} \leq S_i \quad \forall i \in \mathcal{S} \quad (2.5)$$

$$\sigma_i \sum_{\forall j \in \mathcal{N}_{out}(i)} A_{ij} \leq S_i \quad \forall i \in \mathcal{S} \quad (2.6)$$

$$\sum_{\forall (i,j) \in \mathcal{A}} A_{ij} \leq a_{max} \quad (2.7)$$

$$A_{ij} \leq \Xi A_{ij}^b \quad \forall (i,j) \in \mathcal{A} \quad (2.8)$$

$$S_i \leq \Xi S_i^b \quad \forall i \in \mathcal{S} \quad (2.9)$$

$$\sum_{\forall i \in \mathcal{S}} (c_{s_i} S_i^b + \bar{c}_i S_i) + \sum_{\forall (i,j) \in \mathcal{A}} (c_{A_{ij}} A_{ij}^b + \bar{c}_{A_{ij}} A_{ij}) \leq b \quad (2.10)$$

$$F^{od} + F_{ext}^{od} = 1 \quad \forall (o,d) \in \mathcal{K} \quad (2.11)$$

$$\sum_{\forall (i,j) \in \mathcal{A}} F_{ij}^{od} t_{ij} \leq (1 - F_{ext}^{od}) t_{ext}^{od} \quad \forall (o,d) \in \mathcal{K} \quad (2.12)$$

$$\sum_{\forall j \in \mathcal{N}_{out}(i)} F_{ij}^{od} - \sum_{\forall j \in \mathcal{N}_{in}(i)} F_{ji}^{od} = \begin{cases} F^{od} & i = o, \\ 0 & i \neq o \neq d, \\ -F^{od} & i = d \end{cases} \quad \forall i, (o,d) \in \mathcal{S} \times \mathcal{K} \quad (2.13)$$

La función objetivo (2.3) minimiza el número de pasajeros que utilizan la red alternativa de transporte (o, de forma equivalente, maximiza el número de pasajeros que utiliza la red a construir). Las restricciones de capacidad (2.4)-(2.6) limitan el número de pasajeros que pueden utilizar un enlace y el número de vehículos que pueden procesar las estaciones, desagregando en tráfico de llegada y salida. Si resultase necesario, estas restricciones se podrían reescribir para que se aplicase al volumen total de tráfico en la estación. La restricción (2.7) limita la capacidad máxima que se puede asignar a los enlaces, lo que equivale a limitar el número máximo de vehículos que pueden desplegarse en la red. Las restricciones (2.8) y (2.9) aseguran que se activa la variable binaria de construcción de enlaces y estaciones cuando se asigna capacidad a los mismos. La restricción (2.10) impone el límite presupuestario en la construcción de la infraestructura. El grupo de restricciones (2.11) obliga a todos los pasajeros a escoger entre la red nueva y la alternativa, puesto que la suma de los flujos para la red nueva y la alternativa ha de ser igual a uno para todos los mercados (o,d) , mientras que las restricciones (2.12)

establecen que los pasajeros eligen la ruta con la mayor utilidad (*i.e.*, la de menor tiempo de viaje). Por último, las restricciones de conservación de flujo (2.13) para cada estación y mercado garantizan que, si la estación no es origen ni destino del mercado, se conserva el flujo (los pasajeros que llegan a la estación procedentes del mercado (o, d) son iguales a los que salen), mientras que las estaciones origen y destino del mercado (o, d) serán, respectivamente, fuente y sumidero para pasajeros de dicho mercado.

Tal y como se ha mencionado al inicio de la sección, el modelo (2.3)-(2.13) genera soluciones donde todos los pasajeros de un mismo mercado (o, d) eligen la misma ruta, ya sea por la red alternativa o por la construida. Sin embargo, es importante tener en cuenta que este modelo no refleja completamente la complejidad del comportamiento humano en la realidad, donde las decisiones están influenciadas por una variedad de factores que hacen que no todos los pasajeros en un mercado sean puramente racionales y elijan la misma alternativa para viajar. Por este motivo, en la sección 2.2.3 se describe un modelo de demanda más sofisticado, bajo el cual los pasajeros eligen una alternativa u otra de forma probabilística, estando las probabilidades determinadas por el nivel de utilidad que los pasajeros asocian (perciben) a cada una de las alternativas. Como se describirá a continuación, esta distribución de la demanda se realiza a través de una función *logit* o *softmax*.

2.2.3. Formulación basada en aproximación *logit* para la demanda

Esta sección introduce un modelo *logit* multinomial, ampliamente aceptado en la práctica, para representar el comportamiento de los pasajeros [10, 40, 41].

El modelo *all-or-nothing* presentado en la sección anterior, además de solo considerar un proceso de elección determinista, no considera que puedan existir efectos no observables por los atributos considerados. El modelo *logit* multinomial que se presenta a continuación salva estas dos limitaciones. Por un lado, es un modelo de elección probabilístico y, por otro lado, considera que la función de utilidad se compone de un término determinista y de un término de error. La parte determinista de la utilidad que un individuo en un mercado (o, d) percibe de una alternativa $q \in \mathcal{Q}$ es v_q^{od} , mientras que ϵ_q^{od} es una desviación aleatoria que contiene todos los factores no observados de la utilidad.

La parte determinista de la función de utilidad para cada alternativa se calcula como la suma de diferentes términos, dependientes de métricas como la distancia del viaje, la tarifa del viaje y el tiempo de viaje. Estos valores pueden, a su vez, depender de la ruta en cuestión. Por ejemplo, el tiempo de viaje se corresponde con la suma del tiempo de recorrido de cada enlace de la ruta, el tiempo de transbordo y el tiempo de espera. Independientemente de este último aspecto, la función de utilidad se calcula como una suma ponderada de los distintos parámetros a los que los usuarios son sensibles. Matemáticamente, el nivel de utilidad “determinista” de la alternativa de viaje q en el mercado (o, d) se expresa como

$$v_q^{od} = \omega_d d_q^{od} + \omega_p p_q^{od} + \omega_t t_q^{od} \quad \forall (o, d) \in \mathcal{K}, q \in \mathcal{Q}^{od}, \quad (2.14)$$

donde d_q^{od} es la distancia de viaje, p_q^{od} es la tarifa del viaje y t_q^{od} es el tiempo de viaje, todos ellos para la alternativa q en el mercado (o, d) . Los coeficientes $\{\omega_d, \omega_p, \omega_t\}$ representan la sensibilidad de los usuarios a la distancia, tarifa y tiempo de viaje, respectivamente. Nótese que, puesto que los usuarios reaccionan negativamente ante un incremento en los tres parámetros y el término (2.14) representa una utilidad (métrica positiva), los coeficientes $\{\omega_d, \omega_p, \omega_t\}$ serán negativos. El subconjunto $\mathcal{Q}^{od}$ contiene las alternativas compatibles que sirven el mercado (o, d) . Finalmente, el nivel de utilidad “real” de una alternativa q para un individuo en el mercado (o, d) queda como:

$$u_q^{od} = v_q^{od} + \epsilon_q^{od} \quad \forall (o, d) \in \mathcal{K}, q \in \mathcal{Q}^{od}, \quad (2.15)$$

donde es ϵ_q^{od} un término aleatorio de error. Suponiendo que los errores de la función de utilidad ϵ_q^{od} son i) distribuidos de manera independiente, ii) distribuidos de manera idéntica y iii) distribuidos Gumbel,

la *probabilidad* de elegir la alternativa $q \in \mathcal{Q}^{od}$ para en el mercado (o, d) por parte de los pasajeros es [10, 40]:

$$P(q|od) = \frac{e^{(\alpha_q^{od} + v_q^{od})}}{\sum_{q' \in \mathcal{Q}^{od}} e^{(\alpha_{q'}^{od} + v_{q'}^{od})}} \quad \forall (o, d) \in \mathcal{K}, q \in \mathcal{Q}^{od}, \quad (2.16)$$

donde $P(q|od)$ es la probabilidad de que un pasajero del mercado (o, d) elija la alternativa q . Aquí, α_q^{od} es un parámetro que representa la atracción relativa de una alternativa q para la demanda (o, d) debido a factores no incluidos en la función de utilidad, como pueden ser la seguridad, la comodidad, la emisión de boletos, etc. Para introducir el modelo *logit* multinomial de captura de la demanda descrito en (2.16) dentro del problema de optimización formulado en (2.3)-(2.13), se realizan las siguientes suposiciones (hipótesis):

- Todos los pasajeros en un mismo mercado $(o, d) \in \mathcal{K}$ forman parte de un único grupo de iguales características.
- Se considera como único atributo en la función de utilidad el tiempo de viaje t_q^{od} , por lo que $v_q^{od} = \omega_t t_q^{od}$. Así mismo, los pasajeros no tienen en cuenta otros factores no recogidos en la función de utilidad, implicando que $\alpha_q^{od} = 0, \forall q \in \mathcal{Q}, \forall (o, d) \in \mathcal{K}$.
- Se considera una única red competidora frente a la nueva red a diseñar, cuya función de utilidad para el mercado (o, d) vendrá dada por $v_{ext}^{od} = \omega_t t_{ext}^{od}$.

Es importante destacar que las hipótesis enunciadas no constituyen una limitación para la aproximación presentada y no implican una pérdida de generalidad en el análisis. En el caso de contar con los datos pertinentes, estas hipótesis pueden relajarse sin introducir una mayor complejidad tanto en los modelos matemáticos como en las aproximaciones de solución expuestas. La flexibilidad para ajustar las hipótesis según la disponibilidad de datos contribuye a mejorar la robustez y la utilidad de la metodología introducida en la resolución de problemas del mundo real.

La sustitución de la aproximación *all-or-nothing* en el modelo matemático dado por (2.3)-(2.13) por el modelo *logit* multinomial presentado en esta sección, requiere la introducción de nueva notación y restricciones, como se expone a continuación.

Variables de flujo

- $F_{ext}^{od} \in [0, 1]$ indica la proporción de pasajeros del mercado $(o, d) \in \mathcal{K}$ que viajan por la red alternativa para ir a su destino.
- $F^{od} \in [0, 1]$ indica la proporción de pasajeros del mercado $(o, d) \in \mathcal{K}$ que viajan por la red construida para ir a su destino.
- $F_{ij}^{od} \in \{0, 1\}$ es una variable binaria que es igual a 1 en caso que pasajeros del mercado (o, d) viajen por el enlace $(i, j) \in \mathcal{A}$, siendo 0 en caso de que no usen dicho enlace para su viaje.
- $Z^{od} \in \{0, 1\}$ es una variable binaria que es igual a 1 si al menos una fracción de pasajeros del mercado (o, d) viaja por la nueva red, siendo 0 en caso de que no usen la nueva red.

Restricciones

La función objetivo permanece coincide con la formulada en (2.3) para en el modelo *all-or-nothing*, al igual que las restricciones (2.4)-(2.11). Por el contrario, las restricciones (2.12)-(2.13) se sustituyen por las restricciones (2.17)-(2.19). Las restricciones (2.17) son equivalentes a (2.13), pero utilizando la variable binaria Z^{od} en lugar de la variable F^{od} que, en este modelo, es continua. Para cada mercado,

en las estaciones que no son ni origen ni destino de dicho mercado se conserva el flujo. Además, las estaciones origen y destino del mercado actúan como fuente y sumidero del flujo, respectivamente. Las restricciones (2.18) introducen el modelo *logit* multinomial para la elección de la demanda. Por último, las restricciones (2.19) relacionan las variables de flujo para cada mercado, F^{od} , con su proyección binaria, Z^{od} , multiplicada por un número de gran escala, Ξ , de tal forma que F^{od} sea mayor que cero únicamente si Z^{od} está activa.

$$\sum_{\forall j \in \mathcal{N}_{out}(i)} F_{ij}^{od} - \sum_{\forall j \in \mathcal{N}_{in}(i)} F_{ji}^{od} = \begin{cases} Z^{od} & i = o, \\ 0 & i \neq o \neq d, \\ -Z^{od} & i = d \end{cases} \quad \forall i, (o, d) \in \mathcal{S} \times \mathcal{K} \quad (2.17)$$

$$F^{od} \leq \frac{e^{\sum_{\forall (i,j) \in \mathcal{A}} F_{ij}^{od} \omega_{t_{ij}}}}{e^{\omega_{t_{ext}}^{od}} + e^{\sum_{\forall (i,j) \in \mathcal{A}} F_{ij}^{od} \omega_{t_{ij}}}} \quad \forall (o, d) \in \mathcal{K} \quad (2.18)$$

$$F^{od} \leq \Xi Z^{od} \quad \forall (o, d) \in \mathcal{K} \quad (2.19)$$

Resolviendo el modelo basado en la distribución *logit* encontramos soluciones en las que los pasajeros eligen la nueva red o la alternativa con una probabilidad que viene dada por la función logística de la restricción (2.18), es decir, la probabilidad de elegir una red depende de la utilidad de la misma y de la competencia. Tal y como se formula el objetivo en (2.3), siempre que haya capacidad suficiente por la nueva red, la restricción (2.18) se satisfará con igualdad. En cambio, en caso de no haber capacidad suficiente en la red para todos los mercados, la red absorberá una demanda menor a la que le vendría dada por el reparto *logit* en algunos casos.

Estrictamente, la *no linealidad* de la restricción (2.18) impide, en principio, la resolución del problema de optimización mediante un *solver* MIP. Una de las metodologías más aceptadas para sortear este obstáculo consiste en la linealización por trozos de la función no lineal, lo que, en este caso, supone linealizar la función logística que aparece a la derecha de la restricción (2.18). En el apéndice A se explica detalladamente el procedimiento a seguir para linealizar por trozos la función logística, proporcionando un modelo equivalente a (2.3)-(2.11), (2.17)-(2.19) que puede ser abordado por un *solver* MIP. La linealización a trozos obliga a dividir el dominio de entrada en un conjunto $\mathcal{R}$ de regiones no solapantes, una por trozo. Cuantas más regiones o trozos se consideren, mayor será la precisión de la linealización, pero mayor coste computacional implicará a la hora de resolver el problema, dado que es necesario introducir 2 variables, una de ellas binaria, para cada región. Así, para $|\mathcal{R}|$ regiones y $|\mathcal{K}|$ pares origen-destino, deben introducirse $2|\mathcal{R}||\mathcal{K}|$ nuevas variables. Por lo tanto, la linealización de la función logística supone una sobrecarga sustancial a la hora de resolver el problema MIP, incrementando el tiempo de cómputo respecto al modelo *all-or-nothing*. A modo de ejemplo, para un problema relativamente pequeño en el que se intenta optimizar la RdT mostrada en la figura 2.1 considerando 72 posibles mercados (tantos como pares origen-destino), el tiempo de solución del MIP utilizando el *solver* comercial CPLEX en un portátil promedio pasa de ser de 5 minutos para el modelo *all-or-nothing* a no ser capaz de obtener la solución del problema transcurridas 48 horas para el modelo que linealiza la restricción (2.18) con $|\mathcal{R}| = 6$ regiones.

Complejidad computacional de la formulación MIP con decisiones *logit*

Tal y como se ha visto en esta sección, el número de variables binarias necesario para resolver la optimización de redes de $|\mathcal{S}|$ nodos escala de manera polinómica con $|\mathcal{S}|^4$ debido a las variables de flujo. Así mismo, la linealización de las restricciones de preferencias de usuarios suponen la inclusión de $2|\mathcal{R}|(|\mathcal{S}|^2 - |\mathcal{S}|)$ nuevas variables, la mitad de ellas binarias. Desafortunadamente, los problemas de optimización que involucran variables binarias o enteras carecen de convexidad, con lo que la complejidad computacional de los mismos crece exponencialmente a medida que aumenta el tamaño de la

red. Por este motivo, no hay *solvers* actuales que puedan resolver problemas de este tipo para redes de unas dimensiones significativas en un tiempo moderado.

Como se ha mencionado en el capítulo introductorio, en este TFG se busca abordar la resolución de este tipo de problemas desde la formulación de modelos continuos y convexos. Para ello, las principales contribuciones para resolver esta problemática consisten en:

- Formulación no lineal, utilizando funciones convexas o convexificables para modelar comportamientos más realistas.
- Uso de logaritmos seguidos de un algoritmo de MM para considerar las variables de decisión discretas, así como los costes fijos y lineales asociados a la construcción de nodos y arcos.
- Uso de funciones de entropía para obtener un modelo convexo de absorción de demanda basado en una distribución *logit*.

En las dos secciones siguientes, se describen en mayor detalle el segundo y tercer punto.

2.3. Algoritmos de *sparsity* aplicados a RdT

Una de las contribuciones más importantes de este TFG al estado del arte es la optimización de RdT con variables de decisión binarias mediante la formulación y resolución de problemas continuos relacionados. Para ello, se han aplicado técnicas de *sparsity*, ampliamente extendidas en otros campos como la ciencia de datos, pero que hasta ahora apenas se habían utilizado para problemas de RdT.

2.3.1. Introducción a *sparsity*

La posibilidad más simple para recuperar la continuidad en problemas de optimización con variables binarias es relajarlas al intervalo $[0, 1]$ y, luego, proyectar la solución relajada al conjunto discreto $\{0, 1\}$. Un problema asociado a este método es que, al trabajar con cada variable de forma individual, la solución proyectada no tiene por qué encontrarse necesariamente cerca de la solución óptima al problema binario. Un segundo problema de mayor gravedad es que la solución proyectada no suele ser factible, por lo que se hace necesario recurrir a técnicas heurísticas como por ejemplo *feasibility pump* [42]. En lugar de simplemente proyectar la solución relajada sobre $\{0, 1\}$, los algoritmos basados en *sparsity* aumentan el objetivo con un regularizador que fomenta que las variables relajadas sean cero. El beneficio de potenciar soluciones con muchos ceros es triple:

1. Las soluciones obtenidas tienden a estar más cerca de las del problema MIP.
2. Incluso cuando las variables no son binarias, tener soluciones *sparse* tiene un valor práctico en muchas configuraciones del mundo real (*e.g.*, si no se desea construir muchos enlaces o estaciones simultáneamente).
3. En casos en los que existe una variable de decisión (*e.g.*, construir o no el enlace) conjuntamente con una variable de dimensionamiento (*e.g.*, capacidad del enlace en caso de que se construya), el uso de técnicas *sparse* permite combinar ambas variables en una sola. De esta forma si la variable de capacidad es cero, significa que no se construye. Por otro lado, si es positiva, indica que se construye, así como su capacidad.

De forma rigurosa, para fomentar la *sparsity* de un conjunto de variables, deberá aumentarse la función objetivo a minimizar con la norma ℓ_0 del conjunto de variables que se desea sea *sparse*. La (pseudo-)norma ℓ_0 de un vector $\mathbf{x} \in \mathbb{R}^I$ se define como el número de entradas del vector que son

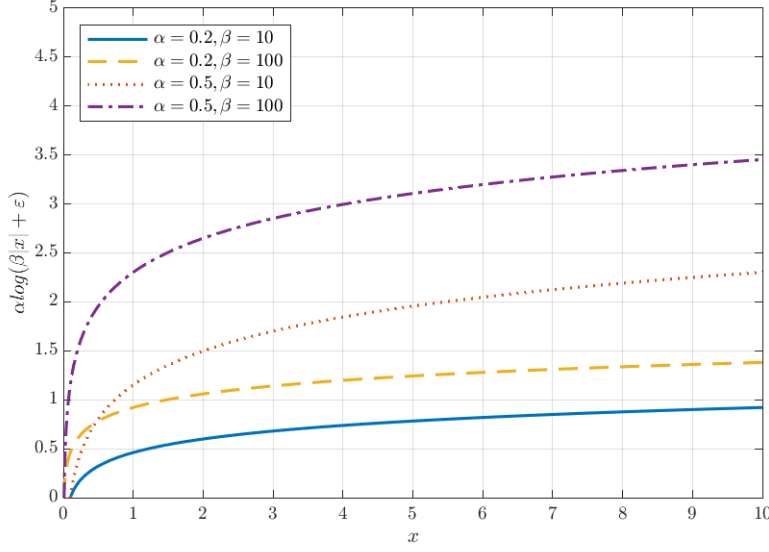


Figura 2.2: Ejemplo de $\alpha \log(\beta|x| + \varepsilon)$ (una aproximación cóncava y *quasi* convexa de la norma ℓ_0) para una variable escalar, para varios valores de α y β , siendo ε un número pequeño.

distintas de cero, es decir,

$$\ell_0(\mathbf{x}) = \|\mathbf{x}\|_0 = \sum_{i=1}^I \mathbb{I}_{\{x_i \neq 0\}}, \quad (2.20)$$

donde $\mathbb{I}_{\{\cdot\}} \in \{0, 1\}$ es la función indicador. Como la norma ℓ_0 es discontinua y, por lo tanto, no convexa, es conveniente encontrar un enfoque manejable (si es posible convexo) para inducir soluciones *sparse*. Es bien sabido que la norma ℓ_1 , definida como

$$\ell_1(\mathbf{x}) = \|\mathbf{x}\|_1 = \sum_{i=1}^I |x_i| = \sum_{i=1}^I |x_i|, \quad (2.21)$$

es el sustituto convexo más cercano a la norma ℓ_0 . Desde un punto de vista de optimización, el gradiente de la norma ℓ_1 es discontinuo en torno al origen, lo que hace que la minimización de la norma ℓ_1 promueva soluciones que igualen varias de las variables a cero [43]. Sin embargo, en escenarios complejos con dependencias cruzadas entre variables, la minimización de ℓ_1 puede ser menos efectiva para generar soluciones *sparse*. Una solución eficaz para mitigar este efecto, que es la utilizada en este TFG, consiste en utilizar una norma ℓ_1 reponderada, que se describe a continuación.

En el enfoque reponderado de ℓ_1 , conocido como *reweighted* ℓ_1 , de ahora en adelante $\ell_{1_{rw}}$, la norma ℓ_0 no se sustituye por la norma ℓ_1 , sino por la función no convexa

$$\ell_{1_{rw}}(\mathbf{x}) = \sum_i \alpha \log(\beta|x_i| + \varepsilon), \quad (2.22)$$

siendo ε un número pequeño y (α, β) dos parámetros positivos utilizados para determinar el ajuste de la función logarítmica a la norma ℓ_0 . En la figura 2.2 se representa la función logarítmica para varios valores de estos parámetros, mientras que en la figura 2.3 es posible comparar la norma ℓ_0 con las normas ℓ_1 y $\ell_{1_{rw}}$. El logaritmo no es una función convexa (de hecho, es cóncava). No obstante, sí es una función continua y monótona, lo que supone que su *level-set*¹ sea convexo. Por ello, tal y como demuestra [38], el algoritmo para minimizar la norma $\ell_{1_{rw}}$ siempre aporta una solución al problema

¹*level-set*: Para una función $f(x)$, se define como el conjunto de valores de x tal que $f(x)$ es menor a una cota superior L . Matemáticamente, se denota como $\{x|f(x) \leq L\}$.

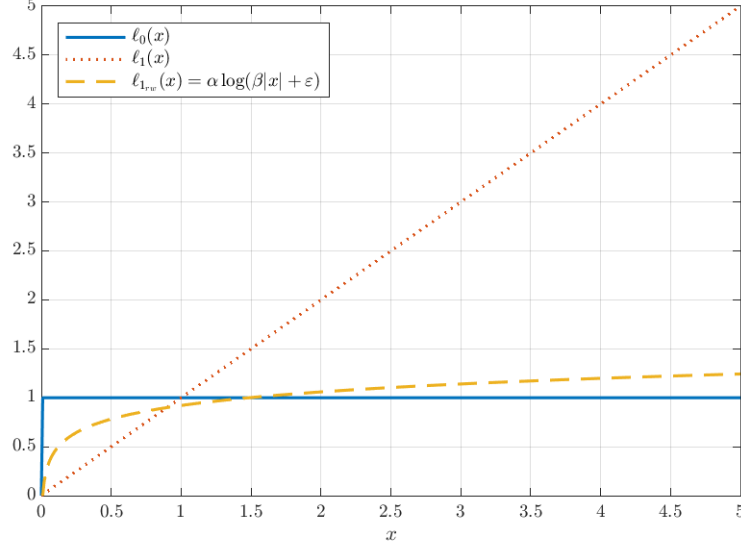


Figura 2.3: Ejemplo de la norma $\ell_0(x)$, $\ell_1(x)$ (su relajación convexa), y $\ell_{rw}(x) = \alpha \log(\beta|x| + \varepsilon)$ (una aproximación cóncava y quasi convexa) para una variable escalar, con $\alpha = 0.2$ y $\beta = 100$.

original igual o mejor que la que se consigue minimizando la norma ℓ_1 . Para hacer frente a la no convexidad del logaritmo, se implementa un enfoque de MM [38, 44], donde en cada iteración se utiliza una aproximación lineal del logaritmo (serie de Taylor de primer orden), como se puede ver en la figura 2.4.

Así, en lugar de minimizar la función no convexa $\ell_{1rw}(\mathbf{x})$, se implementa un algoritmo recursivo donde, en la iteración k , el objetivo es minimizar

$$\hat{\ell}_{1rw}^k(\mathbf{x}|\mathbf{x}^{(k-1)}) = \sum_{i=1}^I \left(|x_i^{(k-1)}| + \frac{1}{\varepsilon + |x_i^{(k-1)}|} |x_i| \right). \quad (2.23)$$

Dado que los términos constantes en el objetivo son irrelevantes para la optimización, se eliminan, por lo que el algoritmo puede centrarse en minimizar

$$\hat{\ell}_{1rw}^k(\mathbf{x}|\mathbf{x}^{(k-1)}) = \sum_{i=1}^I \frac{1}{\varepsilon + |x_i^{(k-1)}|} |x_i|, \quad (2.24)$$

que puede entenderse como una minimización ponderada ℓ_1 , donde los pesos de cada término están relacionados con la solución en la iteración anterior. Claramente, los valores bajos de $|x_i^{(k-1)}|$ aumentarán mucho el coste de la variable i para la siguiente iteración, y por lo tanto, el problema de optimización establecerá un valor para la variable i muy cercano a cero, logrando una solución *sparse*.

No obstante, es conocido que los algoritmos que utilizan la norma ℓ_1 (o sus versiones reponderadas) para inducir *sparsity* suelen inducir un sesgo en la solución. Como ilustración, considérese un escenario donde se quiere obtener una solución *sparse* para un problema diagonal (separable) de ajuste por mínimos cuadrados por medio de la penalización de la norma ℓ_1 . La función objetivo a minimizar en este caso viene dada por

$$\min_{\mathbf{x}} \|\mathbf{y} - \text{diag}(\mathbf{a})\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1 = \min_{\mathbf{x}} \sum_{i=1}^I (y_i - a_i x_i)^2 + \lambda |x_i|, \quad (2.25)$$

donde λ es el valor del regularizador de la *sparsity*. La solución óptima para este problema se muestra en (2.26) y, tal y como es apreciable, está sesgada con respecto de la solución natural al problema de

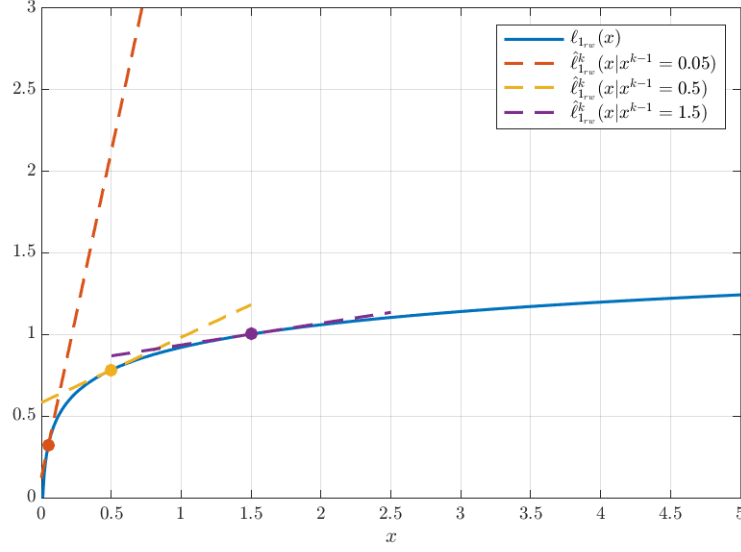


Figura 2.4: Linealización $\hat{\ell}_{1_{rw}}^k(x|x^{(k-1)})$ de $\ell_{1_{rw}}(x)$ alrededor de distintos puntos para una variable escalar.

ajuste de mínimos cuadrados por un factor $\frac{\lambda}{2}$:

$$x_i^* = \text{sign}(y_i/a_i) \max \left\{ 0, |y_i/a_i| - \frac{\lambda}{2a_i^2} \right\}, \quad \forall i. \quad (2.26)$$

A partir de la solución en (2.26) se pueden obtener los conjuntos de índices $\mathcal{I}_0 = \{i \mid x_i^* = 0\}$ (que identifica las variables que serán cero) e $\mathcal{I}_0^c = \{i \mid x_i^* \neq 0\}$. El procedimiento habitual para eliminar el sesgo en las variables indexadas por $\mathcal{I}_0^c$, consiste en resolver una última optimización en la que: i) el objetivo se corresponde con el objetivo original sin regularizador -es decir, el primer término en (2.25)- y ii) se fuerzan a cero las entradas x_i tales que $i \in \mathcal{I}_0$. Expuesto de manera rigurosa, la solución refinada insesgada $\tilde{\mathbf{x}}^*$ se obtiene como $\tilde{x}_i^* = 0$ para todo $i \in \mathcal{I}_0$ y como $\{\tilde{x}_i^*\}_{i \in \mathcal{I}_0^c} = \arg \min_{\{x_i\}_{i \in \mathcal{I}_0^c}} \sum_{i \in \mathcal{I}_0^c} (y_i - a_i x_i)^2$ para el resto de índices. Puesto que el problema vuelve a ser separable, la expresión de la solución óptima para $i \in \mathcal{I}_0^c$ es simplemente $\tilde{x}_i^* = y_i/a_i$, lo que ilustra que, efectivamente, se ha eliminado el sesgo presente en (2.26).

Algoritmo 1: Algoritmo MM para minimizar la norma $\ell_{1_{rw}}$

Entradas: $\lambda, \mathcal{X}$
1 Inicialización: $x_i^{(0)} = 10^6 \quad \forall i \in \mathcal{X}$
2 para $k = 1, 2, \dots, K-1$ **hacer**
3 Obtener $x_i^{*(k)} \quad \forall i \in \mathcal{X}$ resolviendo

$$\tilde{\mathbf{x}}^{*(k)} = \arg \min_{\mathcal{X}} f(\mathbf{x}) + \lambda \sum_{i=1}^{|\mathcal{X}|} \frac{1}{\varepsilon + |x_i^{(k-1)}|} |x_i|,$$

 utilizando $x_i^{(k-1)} \quad \forall i \in \mathcal{X}$.
4 Establecer $x_i^{(k)} = x_i^{*(k)} \quad \forall i \in \mathcal{X}$.
5 fin
6 Construir el conjunto $\mathcal{X}_0 := \{i \mid x_i^{(K-1)} \leq \varepsilon_i\}$
7 Obtener la solución al problema resolviendo

$$\tilde{\mathbf{x}}^{*(K)} = \arg \min_{\mathcal{X}} f(\mathbf{x})$$

s. a: $x_i = 0 \quad \forall i \in \mathcal{X}_0$.

Salidas: $\tilde{\mathbf{x}}^{*(K)}$

Los pasos para llevar a cabo la optimización se explican detalladamente en el algoritmo 1, donde

se pretende minimizar la función genérica $f(\mathbf{x})$ y, a la vez, se buscan soluciones *sparse*. Para ello, se aumenta el objetivo con la norma $\ell_{1_{rw}}$ y se implementa un esquema de MM. El refinamiento de la solución es un aspecto importante en la mayoría de algoritmos que promueven *sparsity* y, tal y como se describe en las secciones subsiguientes, se utiliza también en los algoritmos de diseño de RdT desarrollados en este TFG.

2.3.2. *Sparsity* aplicada a RdT

Uno de los principales obstáculos al que se enfrentan los problemas de optimización de diseño de RdT es la necesidad de emplear variables binarias para modelar la construcción de estaciones y enlaces en la red, dado que se pierde la continuidad en el modelo y, por lo tanto, la dificultad computacional del problema crece sustancialmente. Estas variables de construcción aparecen asociadas a unos costes fijos, cuyo modelado requiere la definición de una variable binaria (o, alternativamente, el uso de la función indicador).

La idea de aplicar *sparsity* a estos problemas tiene como objetivo evitar la utilización de variables binarias en la optimización de RdT, recuperando la continuidad del problema. Concretamente, se propone: i) aumentar la función objetivo con la norma $\ell_{1_{rw}}$ aplicada a las variables asociadas a la construcción de estaciones S_i y enlaces A_{ij} y ii) adoptar un algoritmo iterativo basado en MM. De esta manera, se obtienen expresiones como [cf. (2.24)]

$$\sum_{i \in \mathcal{S}} \frac{c_i}{\varepsilon + S_i^{(k-1)}} S_i, \quad \sum_{ij \in \mathcal{A}} \frac{c_{ij}}{\varepsilon + A_{ij}^{(k-1)}} A_{ij}, \quad (2.27)$$

donde c_i representa el coste fijo de activar S_i y c_{ij} representa el coste fijo de activar A_{ij} . Para entender este último comentario, nótese que si la solución converge y se tiene que $S_i^* \approx S_i^{(k-1)}$, el primer término de (2.27) puede aproximarse como $\sum_{i \in \mathcal{S}} c_i \frac{S_i^*}{\varepsilon + S_i^{(k-1)}} \approx \sum_{i \in \mathcal{S}} c_i \mathbb{I}_{\{S_i^* > 0\}}$, que refleja el coste fijo de construcción. Análogamente, c_{ij} es el coste fijo de construir el arco entre los nodos del par (i, j) . La justificación de la adopción de la norma $\ell_{1_{rw}}$ es, por lo tanto, doble: i) por un lado, da lugar a soluciones más *sparse* y ii) por otro lado, para el problema en cuestión, permite modelar de manera más fidedigna la presencia de unos costes fijos y unos costes variables de construcción. Esto contrasta con la norma ℓ_1 , que únicamente modela de forma fidedigna los costes variables. La utilización de una variante del algoritmo 1 adaptada a nuestro problema permite reducir el tiempo de cómputo de la optimización de forma exponencial, mejorando la escalabilidad del problema de una forma notoria, lo que abre la puerta a la optimización de RdT de mayor complejidad y dimensiones. Además, como se verá en el capítulo 3, tras comparar las soluciones que generan la formulación MIP y la formulación con *sparsity* para una red de ejemplo, los resultados son también muy prometedores en términos de optimalidad (valor de la función objetivo).

2.4. Funciones de entropía aplicadas a RdT: formulación convexa con reparto *logit*

Otra contribución importante de este TFG al estado del arte consiste en la formulación de problemas de optimización computacionalmente eficientes que permiten la inclusión de modelos *logit* de reparto de demanda en el diseño de RdT. En concreto, nuestra propuesta consiste en formular un problema de optimización que, en vez de imponer el reparto *logit* a través de una restricción no convexa como en (2.18), regulariza el objetivo con funciones convexas de entropía para los flujos de pasajeros. Esta propuesta, inspirada por el trabajo seminal de [45], implementa una metodología de “ingeniería inversa”, en la que, en lugar de forzar unas restricciones de forma explícita, se introduce un objetivo cuyas condiciones de optimalidad (basadas en los gradientes de las funciones de entropía) dan lugar a la forma explícita deseada.

Una ventaja adicional de la introducción de las funciones de entropía en el objetivo es que la restricción (2.18), que incluía variables binarias y que, por lo tanto, era no convexa, se elimina del modelo. Combinando la formulación con funciones de entropía en el objetivo junto con algoritmos de *sparsity* para el modelado de costes fijos, es posible emular un problema MIP de una forma continua y convexa, reduciendo el tiempo de cómputo en varios órdenes de magnitud.

2.4.1. Aplicación de regularizadores de entropía a un problema simplificado

Para explicar esta nueva formulación de forma más didáctica, consideremos en primer lugar una versión simplificada del problema de diseño de RdT en el que se incluye un solo mercado (es decir, un único par origen-destino) y en el que sólo se mantienen en la formulación el objetivo (2.28) y la restricción (2.29). Las únicas variables de optimización necesarias para este problema simplificado son:

- $F_{ext} \in [0, 1]$, que es una variable que indica la proporción de la demanda que elige la red alternativa para realizar su viaje.
- $F \in [0, 1]$, que es una variable que indica la proporción de la demanda que elige la nueva red para realizar su viaje.

Así mismo, se emplea ν para denotar el multiplicador de Lagrange asociado a (2.29) y v para la parte determinista de la utilidad percibida por los pasajeros en la nueva red, correspondiéndose v_{ext} con la utilidad percibida en la red alternativa. El problema de optimización que, tras su resolución, resulta en un reparto *logit* entre ambas redes en función de la utilidad en cada una de ellas es

$$\min_{F, F_{ext}} \quad F(\log(F) - 1) + F_{ext}(\log(F_{ext}) - 1) - Fv - F_{ext}v_{ext} \quad (2.28)$$

$$\text{s. a:} \quad F + F_{ext} = 1 \quad (\nu) \quad (2.29)$$

Este problema utiliza funciones de entropía (cóncavas) invertidas (signo opuesto, y por lo tanto convexas) en el objetivo y funciones lineales en el objetivo y la restricción, por lo que se trata de un problema estrictamente convexo en todo su dominio. Esto supone que el óptimo global del problema puede encontrarse como la solución única de las condiciones KKT. Dichas condiciones se corresponden con un punto estacionario del lagrangiano asociado al problema de optimización (2.28)-(2.29). Para obtener el lagrangiano, denotado como $\mathcal{L}(F, F_{ext}, \nu)$, se dualizan las restricciones del problema y se “suben” a la función objetivo [45]. Puesto que el problema (2.28)-(2.29) tiene una sola restricción, el lagrangiano $\mathcal{L}$ asociado es:

$$\mathcal{L}(F, F_{ext}, \nu) = F(\log(F) - 1) + F_{ext}(\log(F_{ext}) - 1) - Fv - F_{ext}v_{ext} + \nu(F + F_{ext} - 1). \quad (2.30)$$

Las condiciones de optimalidad KKT del modelo son:

1. **Condiciones de factibilidad primal.** Debido a que el modelo cuenta con una única restricción y a que esta debe estar obligatoriamente activa, la solución óptima del problema satisface

$$F^* + F_{ext}^* = 1. \quad (2.31)$$

2. **Condiciones de factibilidad dual.** Puesto que no hay restricciones de desigualdad en esta formulación, no hay condiciones de este tipo.
3. **Condiciones de holgura complementaria.** Puesto que no hay restricciones de desigualdad en esta formulación, no hay condiciones de este tipo.
4. **Condiciones de gradiente nulo.** El gradiente del lagrangiano ha de ser el vector nulo, puesto que la solución óptima del problema de optimización se encuentra en un punto de silla del

lagrangiano $\mathcal{L}$ (mínimo en las variables primales F y F_{ext} , máximo en el multiplicador de Lagrange ν). Esto implica que

$$\frac{\partial \mathcal{L}}{\partial F}(F^*, F_{ext}^*, \nu^*) = \log(F^*) - \nu + \nu^* = 0, \quad (2.32)$$

$$\frac{\partial \mathcal{L}}{\partial F_{ext}}(F^*, F_{ext}^*, \nu^*) = \log(F_{ext}^*) - \nu_{ext} + \nu^* = 0. \quad (2.33)$$

Despejando F^* y F_{ext}^* de las ecuaciones (2.32) y (2.33), se llega a las expresiones:

$$F^* = \frac{e^v}{e^{\nu^*}}, \quad F_{ext}^* = \frac{e^{v_{ext}}}{e^{\nu^*}}. \quad (2.34)$$

Sustituyendo estas expresiones en (2.31), se obtiene:

$$F^* + F_{ext}^* = \frac{e^v}{e^{\nu^*}} + \frac{e^{v_{ext}}}{e^{\nu^*}} = 1 \Rightarrow e^{\nu^*} = e^v + e^{v_{ext}}. \quad (2.35)$$

Introduciendo esta última expresión en (2.34) se llega al reparto *logit* de la demanda:

$$F^* = \frac{e^v}{e^v + e^{v_{ext}}}, \quad F_{ext}^* = \frac{e^{v_{ext}}}{e^v + e^{v_{ext}}}, \quad (2.36)$$

demostrando la validez de la metodología propuesta. A modo de recapitulación, en lugar de forzar la expresión en (2.36) *para todo valor de flujos* mediante una restricción, se sugiere la introducción de unas funciones de regularización que garantizan que la *solución óptima* satisface dicha restricción.

2.4.2. Aplicación de regularizadores de entropía al problema original

Una vez explicado el efecto de introducir los términos de entropía en la formulación de un problema sencillo, la parte final de este capítulo ilustra cómo aplicar esta metodología al problema formulado en la sección 2.2.3 sin tener aún en cuenta los costes fijos. Para ello, se plantean y resuelven las nuevas condiciones KKT considerando todos los flujos y restricciones. Cabe recordar que, aunque el número de variables primales y duales sea relativamente elevado, como el problema regularizado con entropías es continuo y convexo, la optimización podrá ejecutarse con una complejidad computacional moderada. El primer paso para reformular el problema de optimización de la sección 2.2.3 con entropías consiste en identificar las variables de estado y optimización. Los parámetros y conjuntos coinciden con los descritos en la sección 2.2.1 y las variables de optimización se agrupan, una vez más, en los conjuntos

$$\mathcal{X}_{cons} = \{S_i\}_{\forall i \in \mathcal{S}} \cup \{A_{ij}\}_{\forall (i,j) \in \mathcal{A}}, \quad (2.37)$$

$$\mathcal{X}_{flow} = \{F_{ext}^{od}\}_{\forall (o,d) \in \mathcal{K}} \cup \{F^{od}\}_{\forall (o,d) \in \mathcal{K}} \cup \{F_{ij}^{od}\}_{\forall ((i,j),(o,d)) \in \mathcal{F}}, \quad (2.38)$$

donde no figura ninguna variable binaria asociada a la construcción de arcos y estaciones. Como se ha mencionado, aún no se tienen en cuenta los costes fijos de construcción, puesto que el propósito de esta sección es demostrar la validez de la metodología propuesta para imponer un reparto *logit* de la demanda. El objetivo a formular se define en (2.39), mientras que las restricciones se especifican en (2.40)-(2.46), donde, junto a cada restricción, se denota entre paréntesis el multiplicador de Lagrange asociado.

$$\begin{aligned} \min_{\mathcal{X}_{cons}, \mathcal{X}_{flow}} & \sum_{\forall (o,d) \in \mathcal{K}} w^{od} F_{ext}^{od} (\log(F_{ext}^{od}) - 1) + \sum_{\forall (o,d) \in \mathcal{K}} w^{od} F^{od} (\log(F^{od}) - 1) \\ & - \sum_{\forall (o,d) \in \mathcal{K}} w^{od} \omega_t t_{ext}^{od} F_{ext}^{od} - \sum_{\forall (i,j),(o,d) \in \mathcal{F}} w^{od} \omega_t t_{ij} F_{ij}^{od} \end{aligned} \quad (2.39)$$

s. a:

$$\frac{\sum_{\forall(o,d) \in \mathcal{K}} F_{ij}^{od} w^{od}}{a_{nom}} \leq A_{ij} \quad (\delta_{A_{ij}}) \quad \forall(i,j) \in \mathcal{A} \quad (2.40)$$

$$\sigma_i \sum_{\forall j \in \mathcal{N}_{in}(i)} A_{ji} \leq S_i \quad (\delta_{in_{S_i}}) \quad \forall i \in \mathcal{S} \quad (2.41)$$

$$\sigma_i \sum_{\forall j \in \mathcal{N}_{out}(i)} A_{ij} \leq S_i \quad (\delta_{out_{S_i}}) \quad \forall i \in \mathcal{S} \quad (2.42)$$

$$\sum_{\forall(i,j) \in \mathcal{A}} A_{ij} \leq a_{max} \quad (\kappa) \quad (2.43)$$

$$\sum_{\forall i \in \mathcal{S}} \bar{c}_i S_i + \sum_{\forall(i,j) \in \mathcal{A}} \bar{c}_{A_{ij}} A_{ij} \leq b \quad (\beta) \quad (2.44)$$

$$F^{od} + F_{ext}^{od} = 1 \quad (\nu^{od}) \quad \forall(o,d) \in \mathcal{K} \quad (2.45)$$

$$\sum_{\forall j \in \mathcal{N}_{out}(i)} F_{ij}^{od} - \sum_{\forall j \in \mathcal{N}_{in}(i)} F_{ji}^{od} = \begin{cases} F^{od} & i = o, \\ 0 & i \neq o \neq d, \\ -F^{od} & i = d \end{cases} \quad (\mu_i^{od}) \quad \forall i, (o,d) \in \mathcal{S} \times \mathcal{K}. \quad (2.46)$$

Cabe mencionar que las restricciones (2.46) utilizan únicamente variables continuas, puesto que, gracias a la introducción de las funciones de entropía en el objetivo, se consigue modelar el reparto *logit* de la demanda desapareciendo tanto las restricciones (2.18)-(2.19) como las variables binarias asociadas (incluyendo la variable Z^{od} utilizada en (2.18)). En cuanto a la restricción de presupuesto (2.44), su principal diferencia respecto a (2.10) es la ausencia de costes fijos, dado que en este caso de ejemplo no se tienen en cuenta. Estos costes se incluirán en las formulaciones propuestas en el capítulo 3.

Al igual que en el problema simplificado en (2.28)-(2.29), dualizamos las restricciones para obtener el lagrangiano $\mathcal{L}(\mathcal{X}_{cons}, \mathcal{X}_{flow}, \mathbf{\Lambda})$, donde $\mathbf{\Lambda}$ denota un conjunto que contiene todos los multiplicadores de Lagrange asociados a (2.40)-(2.46).

$$\begin{aligned} \mathcal{L}(\mathcal{X}_{cons}, \mathcal{X}_{flow}, \mathbf{\Lambda}) = & \sum_{\forall(o,d) \in \mathcal{K}} w^{od} F_{ext}^{od} (\log(F_{ext}^{od}) - 1) + \sum_{\forall(o,d) \in \mathcal{K}} w^{od} F^{od} (\log(F^{od}) - 1) \\ & - \sum_{\forall(o,d) \in \mathcal{K}} w^{od} \omega_t t_{ext}^{od} F_{ext}^{od} - \sum_{\forall(i,j),(o,d) \in \mathcal{F}} w^{od} \omega_t t_{ij} F_{ij}^{od} + \sum_{(i,j) \in \mathcal{A}} \delta_{A_{ij}} \left(\frac{\sum_{\forall(o,d) \in \mathcal{K}} F_{ij}^{od} w^{od}}{a_{nom}} - A_{ij} \right) \\ & + \sum_{i \in \mathcal{S}} \delta_{in_{S_i}} \left(\sigma_i \sum_{\forall j \in \mathcal{N}_{in}(i)} A_{ji} - S_i \right) + \sum_{i \in \mathcal{S}} \delta_{out_{S_i}} \left(\sigma_i \sum_{\forall j \in \mathcal{N}_{out}(i)} A_{ij} - S_i \right) + \kappa \left(\sum_{(i,j) \in \mathcal{A}} A_{ij} - a_{max} \right) \\ & + \beta \left(\sum_{\forall i \in \mathcal{S}} \bar{c}_i S_i + \sum_{\forall(i,j) \in \mathcal{A}} \bar{c}_{A_{ij}} A_{ij} - b \right) + \sum_{(o,d) \in \mathcal{K}} \nu^{od} (F^{od} + F_{ext}^{od} - 1) \\ & + \sum_{(o,d) \in \mathcal{K}} (\mu_o^{od} \left(\sum_{\forall j \in \mathcal{N}_{out}(o)} F_{oj}^{od} - \sum_{\forall j \in \mathcal{N}_{in}(o)} F_{jo}^{od} - 1 + F_{ext}^{od} \right) + \mu_d^{od} \left(\sum_{\forall j \in \mathcal{N}_{out}(d)} F_{dj}^{od} - \sum_{\forall j \in \mathcal{N}_{in}(d)} F_{jd}^{od} + 1 - F_{ext}^{od} \right) \\ & + \sum_{i \neq o \neq d \in \mathcal{S}} \mu_i^{od} \left(\sum_{\forall j \in \mathcal{N}_{out}(i)} F_{ij}^{od} - \sum_{\forall j \in \mathcal{N}_{in}(i)} F_{ji}^{od} \right)). \end{aligned} \quad (2.47)$$

Describiendo las condiciones KKT del problema², se obtiene:

1. **Condiciones de factibilidad primal.** Imponen que todas las restricciones del problema, planteadas en (2.40)-(2.46), han de satisfacerse.

²Para aligerar la notación, en el resto de la sección se obviará la dependencia de $\mathcal{L}$ de $(\mathcal{X}_{cons}, \mathcal{X}_{flow}, \mathbf{\Lambda})$ y se considerará que todos las variables primales y duales son óptimas, prescindiendo así del uso de la notación *.

2. **Condiciones de factibilidad dual.** Todos los multiplicadores de Lagrange asociados a restricciones de desigualdad de la forma menor o igual ($\leq$) han de ser positivos. Matemáticamente se traduce en

$$\delta_{A_{ij}} \geq 0 \quad \forall (i, j) \in \mathcal{A}; \quad \delta_{in_{S_i}} \geq 0 \quad \forall i \in \mathcal{S}; \quad \delta_{out_{S_i}} \geq 0 \quad \forall i \in \mathcal{S}; \quad \kappa \geq 0 \quad \beta \geq 0. \quad (2.48)$$

3. **Condiciones de holgura complementaria.** El producto del multiplicador de Lagrange asociado a una restricción de desigualdad por la diferencia entre los términos de la desigualdad ha de ser siempre nulo. Es decir, el multiplicador de Lagrange de dicha restricción será distinto de cero únicamente cuando la restricción se satisfaga con igualdad. Para el problema en cuestión, esto supone que

$$\delta_{A_{ij}} \left(\frac{\sum_{\forall (o,d) \in \mathcal{K}} F_{ij}^{od} w^{od}}{a_{nom}} - A_{ij} \right) = 0 \quad \forall (i, j) \in \mathcal{A} \quad (2.49)$$

$$\delta_{in_{S_i}} \left(\sigma_i \sum_{\forall j \in \mathcal{N}_{in}(i)} A_{ji} - S_i \right) = 0 \quad \forall i \in \mathcal{S} \quad (2.50)$$

$$\delta_{out_{S_i}} \left(\sigma_i \sum_{\forall j \in \mathcal{N}_{out}(i)} A_{ij} - S_i \right) = 0 \quad \forall i \in \mathcal{S} \quad (2.51)$$

$$\kappa \left(\sum_{(i,j) \in \mathcal{A}} A_{ij} - a_{max} \right) = 0 \quad (2.52)$$

$$\beta \left(\sum_{\forall i \in \mathcal{S}} \bar{c}_i S_i + \sum_{\forall (i,j) \in \mathcal{A}} \bar{c}_{A_{ij}} A_{ij} - b \right) = 0. \quad (2.53)$$

4. **Condiciones de gradiente nulo.** Tomando el gradiente de $\mathcal{L}$ con respecto a las variables primales e igualando a cero, se obtienen las condiciones estacionarias del problema:

$$\frac{\partial \mathcal{L}}{\partial F_{ext}^{od}} = w^{od} \log(F_{ext}^{od}) - w^{od} \omega_t t_{ext}^{od} + \mu_o^{od} - \mu_d^{od} + \nu^{od} = 0 \quad \forall (o, d) \in \mathcal{K} \quad (2.54)$$

$$\frac{\partial \mathcal{L}}{\partial F^{od}} = w^{od} \log(F^{od}) + \nu^{od} = 0 \quad \forall (o, d) \in \mathcal{K} \quad (2.55)$$

$$\frac{\partial \mathcal{L}}{\partial F_{oi}^{od}} = -w^{od} \omega_t t_{oi} + \mu_o^{od} - \mu_i^{od} + \delta_{A_{oi}} a_{nom}^{-1} w^{od} = 0 \quad \forall (o, i), (o, d) \in \mathcal{A} \times \mathcal{K} \quad (2.56)$$

$$\frac{\partial \mathcal{L}}{\partial F_{ij}^{od}} = -w^{od} \omega_t t_{ij} + \mu_i^{od} - \mu_j^{od} + \delta_{A_{ij}} a_{nom}^{-1} w^{od} = 0 \quad \forall (i, j), (o, d) \in \mathcal{A} \times \mathcal{K} \quad (2.57)$$

$$\frac{\partial \mathcal{L}}{\partial F_{jd}^{od}} = -w^{od} \omega_t t_{jd} + \mu_j^{od} - \mu_d^{od} + \delta_{A_{jd}} = 0 \quad \forall (j, d), (o, d) \in \mathcal{A} \times \mathcal{K}. \quad (2.58)$$

Despejando μ_i^{od} de (2.56), se obtiene:

$$\mu_i^{od} = -w^{od} \omega_t t_{oi} + \mu_o^{od} + \delta_{A_{oi}} a_{nom}^{-1} w^{od}. \quad (2.59)$$

Si se introduce (2.59) en (2.57), se llega a que

$$\mu_j^{od} = -w^{od} \omega_t t_{oi} - w^{od} \omega_t t_{ij} + \delta_{A_{oi}} a_{nom}^{-1} w^{od} + \delta_{A_{ij}} a_{nom}^{-1} w^{od} + \mu_o^{od}. \quad (2.60)$$

Por otro lado, despejando μ_j^{od} de (2.58), se obtiene:

$$\mu_j^{od} = \mu_d^{od} + w^{od} \omega_t t_{jd} - \delta_{A_{jd}}. \quad (2.61)$$

Igualando (2.60) y (2.61) y despejando $\mu_o^{od} - \mu_d^{od}$, se deduce que

$$\mu_o^{od} - \mu_d^{od} = w^{od} \omega_t(t_{oi} + t_{ij} + t_{jd}) - a_{nom}^{-1} w^{od} (\delta_{A_{oi}} + \delta_{A_{ij}} + \delta_{A_{jd}}). \quad (2.62)$$

Una vez que se tiene la expresión de $\mu_o^{od} - \mu_d^{od}$, se introduce en (2.54) y se despeja F_{ext}^{od} :

$$F_{ext}^{od} = \frac{e^{\omega_t t_{ext}^{od}}}{e^{\frac{\mu_o^{od} - \mu_d^{od}}{w^{od}}}} e^{\frac{\nu}{w^{od}}} = \frac{e^{\omega_t t_{ext}^{od}}}{e^{\omega_t(t_{oi} + t_{ij} + t_{jd}) - a_{nom}^{-1}(\delta_{A_{oi}} + \delta_{A_{ij}} + \delta_{A_{jd}})}} e^{-\frac{\nu}{w^{od}}}. \quad (2.63)$$

Despejando F^{od} de (2.55), se obtiene:

$$F^{od} = e^{-\frac{\nu^{od}}{w^{od}}}. \quad (2.64)$$

Por último, agrupando (2.63) y (2.64), teniendo en cuenta (2.45), se concluye que

$$e^{-\nu^{od}} = \frac{e^{\omega_t(t_{oi} + t_{ij} + t_{jd}) - a_{nom}^{-1}(\delta_{A_{oi}} + \delta_{A_{ij}} + \delta_{A_{jd}})}}{e^{\omega_t(t_{oi} + t_{ij} + t_{jd}) - a_{nom}^{-1}(\delta_{A_{oi}} + \delta_{A_{ij}} + \delta_{A_{jd}})} + e^{\omega_t t_{ext}^{od}}}, \quad (2.65)$$

y, por lo tanto, se llega a la expresión del reparto *logit* de la demanda:

$$F^{od} = \frac{e^{\omega_t(t_{oi} + t_{ij} + t_{jd}) - a_{nom}^{-1}(\delta_{A_{oi}} + \delta_{A_{ij}} + \delta_{A_{jd}})}}{e^{\omega_t(t_{oi} + t_{ij} + t_{jd}) - a_{nom}^{-1}(\delta_{A_{oi}} + \delta_{A_{ij}} + \delta_{A_{jd}})} + e^{\omega_t t_{ext}^{od}}}, \quad (2.66)$$

$$F_{ext}^{od} = \frac{e^{\omega_t t_{ext}^{od}}}{e^{\omega_t(t_{oi} + t_{ij} + t_{jd}) - a_{nom}^{-1}(\delta_{A_{oi}} + \delta_{A_{ij}} + \delta_{A_{jd}})} + e^{\omega_t t_{ext}^{od}}}. \quad (2.67)$$

Hay que tener en cuenta que el multiplicador de Lagrange $\delta_{A_{ij}}$ hace referencia a la restricción de capacidad en el enlace entre los nodos $(i, j) \in \mathcal{A}$. De esta forma, en caso de no haber capacidad suficiente para acomodar a todos los pasajeros, el multiplicador $\delta_{A_{ij}}$ se activará, incrementando el coste percibido por los pasajeros para viajar por la nueva red. En caso de haber capacidad suficiente, el valor de dicho multiplicador será 0, de forma que el reparto de demanda se basará únicamente en el tiempo de viaje. Este comportamiento puede relacionarse con la restricción de reparto de demanda (2.18) presente en la formulación MIP. La activación del multiplicador $\delta_{A_{ij}}$ supondría no saturar la restricción (2.18) debido a que los pasajeros perciben un coste adicional asociado a la saturación de capacidad (congestión) de algunos de los enlaces de la ruta. Esta congestión también estaba contemplada en el modelo MIP, a través de la restricción (2.4). Por el contrario, cuando $\delta_{A_{ij}} = 0$ y el reparto de demanda se basa únicamente en el tiempo de viaje, la restricción (2.18) se cumplirá con igualdad.

Nótese que el resultado obtenido en (2.66) y (2.67) depende solamente de los multiplicadores $\delta_{A_{ij}}$ y de los tiempos de viaje t_{ij} , puesto que, en aras de simplificar el desarrollo, el tiempo de viaje es el único atributo que se ha considerado en la función de utilidad. Sin embargo, en escenarios realistas, la función de utilidad de los pasajeros depende de más atributos, como, por ejemplo, el precio o la distancia. Así, denotando el precio de un par origen destino como p^{od} , con coeficiente en la función de utilidad asociado ω_p , e introduciéndolo en el objetivo de forma análoga a como se ha realizado con el tiempo en (2.39), se obtiene el siguiente reparto:

$$F^{od} = \frac{e^{\omega_t(t_{oi} + t_{ij} + t_{jd}) + \omega_p p^{od} - a_{nom}^{-1}(\delta_{A_{oi}} + \delta_{A_{ij}} + \delta_{A_{jd}})}}{e^{\omega_t(t_{oi} + t_{ij} + t_{jd}) + \omega_p p^{od} - a_{nom}^{-1}(\delta_{A_{oi}} + \delta_{A_{ij}} + \delta_{A_{jd}})} + e^{u^{od}}}, \quad (2.68)$$

$$F_{ext}^{od} = \frac{e^{u^{od}}}{e^{\omega_t(t_{oi} + t_{ij} + t_{jd}) + \omega_p p^{od} - a_{nom}^{-1}(\delta_{A_{oi}} + \delta_{A_{ij}} + \delta_{A_{jd}})} + e^{u^{od}}}. \quad (2.69)$$

Por lo tanto, una vez conocido el resultado de introducir los términos de entropía en el objetivo, queda demostrada la posibilidad de plantear un modelo alternativo para la optimización de RdT con reparto de demanda basado en modelos tipo *logit*. Esta aproximación, en combinación con técnicas de *sparsity*, permite la posibilidad de obtener una formulación convexa y continua para la optimización de RdT, computacionalmente mucho más eficiente que el modelo clásico MIP. Este es, precisamente, el objetivo del siguiente capítulo de esta memoria.

Capítulo 3

Formulación general del problema de optimización convexa para RdT

En este capítulo se formaliza la metodología de diseño de RdT con técnicas de optimización convexa, empleando regularizadores logarítmicos para modelar los costes de construcción de estaciones y enlaces y el uso de funciones de entropía para modelar las preferencias de los usuarios.

A diferencia del capítulo anterior, en este caso se consideran todas las variables de diseño y estado, discutiendo de forma detallada la formulación de todas las restricciones y términos de la función objetivo, dotando al modelo de más realismo e introduciendo nuevos tipos de variables relacionadas con la congestión de la red. El modelo de la congestión se introduce en la sección 3.1, lo que cierra el marco teórico de la formulación propuesta. A continuación, la sección 3.2 describe las variables y objetivos planteados. La sección 3.3 presenta de forma detallada la formulación convexa propuesta, mientras que la sección 3.4 introduce el modelo MIP equivalente, enfatizando las principales diferencias con la formulación convexa planteada en este TFG. Dado que el problema formulado combina objetivos de operadores y pasajeros, la sección 3.5 describe la metodología empleada para tratar con más de una función objetivo. Por último, la sección 3.6 presenta un análisis cuantitativo de la metodología de diseño que se propone en una red sintética de 9 nodos. El motivo fundamental para elegir una red de este tamaño es poder estudiar las diferencias entre el planteamiento convexo y un diseño basado en programación entera.

La evaluación y el estudio numérico presentados en este capítulo incluyen:

- La búsqueda del punto de utopía, determinado por el valor óptimo de los distintos objetivos (operador y pasajeros) de forma independiente, y la determinación de la curva de Pareto con el fin de ajustar la ponderación de cada objetivo.
- El análisis de la topología de red construida para distintos valores de costes fijos y presupuesto disponible.
- La comparación con la solución proporcionada por la formulación MIP, tanto en términos de valor de la función objetivo, como de tiempo de ejecución.
- La realización de análisis de sensibilidad para estudiar las variaciones del valor óptimo de la función objetivo frente a cambios en los precios, la demanda, costes de operación y construcción. Por motivos de espacio, los resultados numéricos de este análisis se detallan en el apéndice B. Se insta al lector a revisarlo una vez concluido el presente capítulo.

3.1. Aproximación convexa a la congestión en la red

En las RdT, la congestión es un fenómeno con impacto significativo en el tiempo de viaje y en el grado de satisfacción de los pasajeros. Cuando el volumen de tráfico operado por la red está cerca o incluso supera la capacidad de diseño, los niveles de congestión y, consecuentemente, los tiempos de viaje aumentan considerablemente. Numerosos trabajos en el estado del arte se han dedicado a modelar los efectos de la congestión, desarrollando modelos matemáticos basados en datos empíricos para estimar el impacto en los pasajeros. En [46] se realiza una revisión sistemática de la literatura, analizando diversos modelos de congestión como, por ejemplo, la linealización por tramos de una función de coste de congestión tipo *Kleinrock* [47], la integración de decisiones de capacidad con efectos en el coste de congestión [48], o la consideración de teoría de colas para extender la función de utilidad con atributos adicionales como tiempos de espera y probabilidades de no poder acceder al servicio [49], dando lugar en ambos casos a funciones no lineales de las variables de tráfico.

Los modelos más simplificados expresan el nivel de congestión como una función no lineal dependiente del volumen de tráfico. En [50] se agrega un término de congestión explícitamente en la función objetivo que crece exponencialmente a medida que aumenta el tráfico. Sin embargo, la congestión está más relacionada con la diferencia relativa entre el flujo y la capacidad que con el flujo únicamente [47]. Cuanto más cerca esté el flujo total de la capacidad, más severa será la congestión. Una alternativa consiste en formular la tasa de congestión como el cociente entre el flujo y la diferencia entre la capacidad y el flujo. Este modelo hiperbólico se puede obtener de forma independiente cuando se considera un servidor para el que la tasa de llegada del tráfico sigue una distribución de Poisson y los tiempos de procesamiento siguen una distribución exponencial, pudiéndose entonces modelar la dinámica como una red de colas M/M/1 [47]. En el estacionario, siendo el tráfico inferior a la capacidad, el tiempo de espera promedio se calcula como la inversa de la diferencia entre la capacidad y el flujo, mientras que el tiempo de espera promedio total se calcula como el cociente entre el flujo y la diferencia entre la capacidad y el flujo.

De forma similar, la propuesta en este TFG para modelar la congestión consiste en un modelo hiperbólico. Bajo este prisma, el nivel de congestión en una determinada estación o enlace tiende a valores muy altos cuando se opera una capacidad cercana a la diseñada, como se muestra en la figura 3.1. Recordando que F_{ij}^{od} y A_{ij} denotan el volumen de tráfico del par (o, d) en el enlace (i, j) y la capacidad diseñada para el enlace (i, j) , respectivamente, el nivel de congestión total en dicho enlace se aproxima como

$$\frac{\sum_{\forall(o,d) \in \mathcal{K}} F_{ij}^{od}}{(A_{ij} - \sum_{\forall(o,d) \in \mathcal{K}} F_{ij}^{od})h_A + \epsilon_H}, \quad (3.1)$$

donde h_A y ϵ_H son constantes. La primera cuantifica el impacto que la congestión tiene sobre el tiempo de viaje, mientras que la segunda (que se fija a un número pequeño) establece una cota superior al nivel de congestión (véase la figura 3.1). El nivel de congestión total se modela por tanto como un cociente entre el volumen de tráfico ($\sum_{\forall(o,d) \in \mathcal{K}} F_{ij}^{od}$) y la holgura ($A_{ij} - \sum_{\forall(o,d) \in \mathcal{K}} F_{ij}^{od}$) en el enlace (i, j) . Y consecuentemente, el tiempo de viaje total t_{ij}^c por pasajero viene dado por

$$t_{ij}^c = t_{ij} + \frac{1}{(A_{ij} - \sum_{\forall(o,d) \in \mathcal{K}} F_{ij}^{od})h_A + \epsilon_H}, \quad (3.2)$$

donde t_{ij} es el tiempo de viaje en ausencia de congestión.

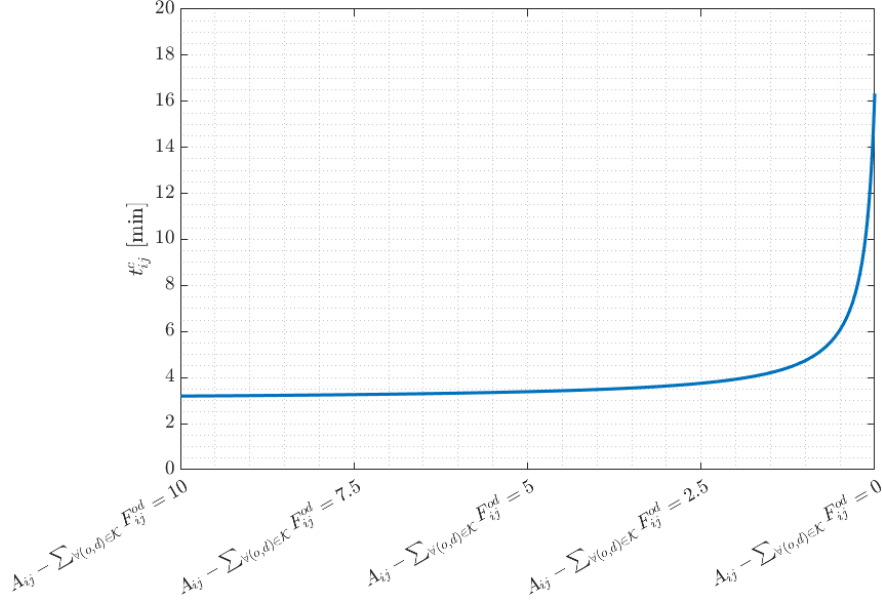


Figura 3.1: Modelo hiperbólico de congestión. Se representa el tiempo de viaje t_{ij}^c por pasajero para el arco (i, j) en función de la diferencia entre capacidad diseñada A_{ij} y capacidad operada $\sum_{(o,d) \in \mathcal{K}} F_{ij}^{od}$ para dicho arco. Se ha fijado $A_{ij} = 10$, $t_{ij} = 3$, $h_A = 0.5$ y $\epsilon_H = 0.15$.

Tal y como se ha introducido en esta sección, existen distintas alternativas con niveles de resolución diferentes para tener en cuenta la congestión. Por ejemplo, la modificación de los repartos *logit* teniendo en cuenta el nuevo tiempo de viaje. O la penalización en la función objetivo con los niveles (o tiempos) de congestión, de forma que el operador haga un esfuerzo explícito por reducirlos. Para implementar esta última opción (que es la estrategia seguida en este TFG), conviene tener en cuenta que el modelo de congestión en (3.1) es convexo en F_{ij}^{od} , pudiéndose incorporar en problemas en los que la capacidad esté fijada o pertenezca a un conjunto finito [47]. Sin embargo, el modelo en (3.1) no es convexo conjuntamente en F_{ij}^{od} y A_{ij} . Por ello, en este TFG se considera el incremento en el tiempo de viaje para un único pasajero en un enlace debido a la congestión, el cual viene determinado por la expresión (*cf.* término derecho en (3.2))

$$\frac{1}{(A_{ij} - \sum_{(o,d) \in \mathcal{K}} F_{ij}^{od})h_A + \epsilon_H}, \quad (3.3)$$

que únicamente depende de la holgura entre capacidad diseñada y volumen de tráfico soportado en el enlace $(i, j) \in \mathcal{A}$ (*i.e.*, $A_{ij} - \sum_{(o,d) \in \mathcal{K}} F_{ij}^{od}$). La expresión en (3.3) es conjuntamente convexa en A_{ij} y F_{ij}^{od} , por lo puede incorporarse fácilmente en el problema de diseño de RdT tratado en este TFG. Concretamente, se propone aumentar la función objetivo con un término de la forma de (3.3) para cada uno de los enlaces (i, j) y estaciones i .

Definiendo ΔA_{ij} como la diferencia entre la capacidad dimensionada y la utilizable en cada enlace $(i, j) \in \mathcal{A}$, y ΔS_i como la diferencia entre la capacidad dimensionada y la utilizable en la estación $i \in \mathcal{S}$, el modelo de congestión en la ecuación (3.3) se sustituye por el modelo

$$\sum_{(i,j) \in \mathcal{A}} \frac{1}{\Delta A_{ij} h_{A_{ij}} + \epsilon_H} + \sum_{i \in \mathcal{S}} \frac{1}{\Delta S_i h_{S_i} + \epsilon_H}. \quad (3.4)$$

En este nuevo modelo, se detalla de forma explícita un primer término para los enlaces y un segundo término para las estaciones. Dado que los sistemas de transporte habitualmente tienen como objetivo maximizar la cobertura, la capacidad utilizable es una buena aproximación al tráfico en el enlace o

estación, y la diferencia entre la capacidad dimensionada y la capacidad utilizable es una cota inferior razonablemente ajustada a la holgura entre la capacidad diseñada y el volumen de tráfico soportado. Por lo tanto, el modelo en (3.4) proporciona una cota superior a la congestión que experimenta el usuario, proporcionando en cualquier caso soluciones en las que el nivel de servicio es igual o superior al del proporcionado por el modelo en (3.3). Además, la dualidad en la capacidad (dimensionada-utilizable) en el contexto de los sistemas de transporte tiene un propósito adicional al enunciado anteriormente. En la planificación de este tipo de sistemas, el *día tipo* es un día representativo que se utiliza como modelo para analizar y planificar operaciones y servicios. Este día se selecciona porque sus patrones de demanda y operación son típicos o promedio para el sistema de transporte en cuestión. Por lo tanto, desde la concepción de estos tipos de sistemas se admite que habrá determinados periodos en los que operarán por encima de la capacidad nominal.

Es en este contexto, congestión percibida por el usuario y operación del sistema, en el que se incorporan al modelo los conceptos de capacidades utilizable y dimensionada (véase apartado de variables de construcción y auxiliares de la sección 3.2.2 para una descripción completa).

3.2. Variables y objetivos del modelo matemático

A lo largo de esta sección se describen de forma detallada las variables y objetivos del problema introducido en el capítulo 2. Para facilitar la lectura, se listan y describen brevemente todas las variables participantes en la formulación (incluidas las ya definidas en capítulos anteriores). En el listado se diferencia entre variables de estado (las cuales vendrán dadas *a priori*) y variables de diseño (sobre las que se optimizará). Estas últimas se dividen, según su naturaleza, en variables de construcción, variables de flujo y variables secundarias o auxiliares. Concluido el listado, se explica en cierto nivel de detalle cada uno de los términos de la función objetivo, incidiendo en los aspectos más relevantes para entender las propiedades de la solución óptima del modelo matemático que se presenta en la sección 3.3.

3.2.1. Variables de estado

Las variables de estado y parámetros que intervienen en el modelo se listan a continuación:

- w^{od} es la demanda del mercado $(o, d) \in \mathcal{K}$.
- t_{ext}^{od} es el tiempo de viaje para el par $(o, d) \in \mathcal{K}$ por la red alternativa.
- p_{ext}^{od} es el precio que pagan los pasajeros del par $(o, d) \in \mathcal{K}$ por la red alternativa.
- t_{ij} es el tiempo de viaje por el arco entre los nodos $(i, j) \in \mathcal{A}$ por la nueva red.
- ω_t es el coeficiente asociado al tiempo de viaje en la función de utilidad de los pasajeros.
- p_{ij} es el precio que pagan los pasajeros por recorrer el enlace $(i, j) \in \mathcal{A}$ en la nueva red.
- ω_p es el coeficiente asociado al precio en la función de utilidad de los pasajeros.
- $h_{A_{ij}}$ es el coeficiente asociado a la congestión del enlace entre los nodos $(i, j) \in \mathcal{A}$.
- h_{S_i} es el coeficiente asociado a la congestión de la estación en el nodo $i \in \mathcal{S}$.
- ϵ_H es un parámetro que impone una cota superior a la congestión.
- $c_{o_{ij}}$ es el coste de operación por unidad de capacidad del enlace $(i, j) \in \mathcal{A}$.
- $\bar{c}_{A_{ij}}$ es el coste por unidad de capacidad construida asociado al enlace $(i, j) \in \mathcal{A}$.

- $\bar{c}_{S_i}$ es el coste por unidad de capacidad construida asociado a la estación en el nodo $i \in \mathcal{S}$.
- $c_{A_{ij}}$ es el coste fijo de construcción del enlace $(i, j) \in \mathcal{A}$.
- c_{S_i} es el coste fijo de construcción de la estación en el nodo $i \in \mathcal{S}$.
- b es el presupuesto del que se dispone para la construcción de la nueva red.
- τ_{ij} es el factor de carga u ocupación promedio en el enlace $(i, j) \in \mathcal{A}$, es decir, la relación entre la carga de pasajeros y la capacidad total disponible. Se trata de un parámetro que depende del modo de transporte. A modo de ejemplo, en transporte aéreo, el factor de carga promedio suele estar en torno al 80-85 %, mientras que en transporte terrestre varía más ampliamente.
- σ_i es el porcentaje de vehículos sobre el total que llegan a la estación $i \in \mathcal{S}$ en hora punta. Sirve para relacionar la capacidad dimensionada en enlaces y estaciones.
- a_{nom} es igual a la capacidad nominal de los vehículos, es decir, el número de pasajeros que un vehículo puede transportar simultáneamente.
- a_{max} es la capacidad total máxima que se puede diseñar para los enlaces. Representa el número máximo de vehículos o frecuencias que la red puede operar.

3.2.2. Variables de diseño

Es posible distinguir entre tres tipos de variables de diseño, según su naturaleza: variables flujo, variables de construcción y variables auxiliares. Estas últimas están relacionadas con las variables de construcción, por lo que se detallan de forma conjunta.

Variables de flujo

Las variables de flujo modelan el movimiento de los pasajeros por la red, además de indicar el reparto de los mismos entre la red a construir y la alternativa. Dichas variables son:

- $F_{ext}^{od} \in [0, 1]$ es la proporción de pasajeros de un mercado $(o, d) \in \mathcal{K}$ que viajan por la red alternativa.
- $F^{od} \in [0, 1]$ es la proporción de pasajeros de un mercado $(o, d) \in \mathcal{K}$ que viajan por la nueva red.
- $F_{ij}^{od} \in [0, 1]$ es la proporción de pasajeros de un mercado $(o, d) \in \mathcal{K}$ que viajan por el enlace $(i, j) \in \mathcal{A}$ en su ruta.

Variables de construcción y auxiliares

Las variables de construcción modelan la capacidad constituida en estaciones y enlaces. Aunque se han introducido en capítulos anteriores, en la formulación definitiva resulta necesario redefinir ligeramente las variables, de forma que se facilite la incorporación de los términos de congestión. Concretamente, en la formulación definitiva, se consideran las siguientes variables de construcción:

- $A'_{ij} \in \mathbb{R}^{\geq 0}$ es la capacidad construida en el enlace $(i, j) \in \mathcal{A}$.
- $S'_i \in \mathbb{R}^{\geq 0}$ es la capacidad construida en la estación $i \in \mathcal{S}$.

Asociadas a las anteriores, se consideran así mismo las siguientes variables auxiliares:

- $A_{ij} \in \mathbb{R}^{\geq 0}$ es la capacidad utilizable por los usuarios en el enlace $(i, j) \in \mathcal{A}$. Esta variable representa, por lo tanto, el flujo máximo permitido en el enlace (i, j)

- $\Delta A_{ij} \in \mathbb{R}^{\geq 0}$ es igual a la diferencia entre la capacidad dimensionada y la utilizable en el enlace $(i, j) \in \mathcal{A}$. Puesto que A_{ij} representa el flujo máximo, la variable ΔA_{ij} representa una cota inferior a la holgura en el enlace (i, j) [c.f. (3.3)]. Por consiguiente, $\frac{1}{\Delta A_{ij} h_{A_{ij}} + \epsilon_H}$ representa una cota superior a la congestión que existe en el enlace (i, j) .
- $S_i \in \mathbb{R}^{\geq 0}$ es la capacidad utilizable por los usuarios en la estación $i \in \mathcal{S}$.
- $\Delta S_i \in \mathbb{R}^{\geq 0}$ es igual a la diferencia entre la capacidad dimensionada y la utilizable en la estación $i \in \mathcal{S}$. Al igual que en el caso de ΔA_{ij} , esta variable representa una cota inferior a la holgura en la estación, sirviendo así para limitar la congestión sufrida por los pasajeros.

De forma similar a como se ha procedido anteriormente, se definen los siguientes conjuntos compactos de variables:

$$\mathcal{X}_{flow} = \{F_{ext}^{od}\}_{\forall (o,d) \in \mathcal{K}} \cup \{F^{od}\}_{\forall (o,d) \in \mathcal{K}} \cup \{F_{ij}^{od}\}_{\forall ((i,j),(o,d)) \in \mathcal{F}}, \quad (3.5)$$

$$\mathcal{X}_{cons} = \{S'_i\}_{\forall i \in \mathcal{S}} \cup \{A'_{ij}\}_{\forall (i,j) \in \mathcal{A}}, \quad (3.6)$$

$$\mathcal{X}_{aux} = \{S_i\}_{\forall i \in \mathcal{S}} \cup \{\Delta S_i\}_{\forall i \in \mathcal{S}} \cup \{A_{ij}\}_{\forall (i,j) \in \mathcal{A}} \cup \{\Delta A_{ij}\}_{\forall (i,j) \in \mathcal{A}}. \quad (3.7)$$

3.2.3. Descripción del objetivo

A continuación se proporciona una explicación detallada de todos los términos que forman parte de la función objetivo, distinguiendo los intereses de los pasajeros y del operador. Además, se incluyen los términos necesarios para obtener el modelo de elección de la demanda basado en la distribución *logit*, así como para modelar la congestión en la red.

Se considera un caso de un operador público, cuyo objetivo es maximizar la demanda capturada o servida (bienestar social) con el mínimo coste de operación posible (intereses del operador). Por esta razón, se busca una solución de compromiso entre los costes de operación de la RdT y el número de pasajeros que viajan por la nueva red.

Objetivo del operador

El objetivo del operador es minimizar los costes de operación, que van asociados a la capacidad dimensionada en los enlaces de la nueva red. Por ello, en la función objetivo del problema aparece el término OP , definido como

$$OP(\mathcal{X}_{cons}) = \sum_{\forall (i,j) \in \mathcal{A}} c_{o_{ij}} A'_{ij}. \quad (3.8)$$

En lo referente a los costes de construcción, tal y como se discute en detalle en la siguiente sección, se formulan o bien como una restricción o bien como parte del objetivo. Por último, cabe hacer notar que, desde el punto de vista de nuestra formulación, podríamos haber considerado cualquier función de coste convexa. No obstante, siguiendo la tradición de formulaciones MIP, se ha adoptado una función lineal.

Objetivo de los pasajeros

Al considerar un operador público, el diseño de la RdT debe perseguir también prestar un servicio a la sociedad. Esto supone que hay que tener en cuenta la preferencia de los usuarios por rutas con bajo coste en términos de tiempo de viaje y de precio del billete. Tal y como se demostró en la sección 2.4, para obtener un modelo de elección de la demanda basado en una distribución tipo *logit* mediante una formulación convexa, hay que añadir a la función objetivo: i) términos de entropía relacionados con los flujos y ii) la función de utilidad de los pasajeros para cada una de las redes competidoras.

Adicionalmente, para considerar los efectos de la congestión del sistema, es necesario incorporar a la función objetivo el modelo introducido en la sección 3.1.

Por tanto, los costes asociados a los pasajeros que se tendrán en cuenta en la función objetivo a minimizar se pueden clasificar en cuatro grupos:

- Función de utilidad de los pasajeros en la nueva red. Uno de los objetivos del problema de optimización es maximizar la utilidad percibida por los pasajeros en la nueva red. Puesto que el problema de optimización se formula como una minimización, la utilidad de los pasajeros se multiplica por -1 y, tras esa multiplicación, se añade al objetivo de la minimización. De esta forma, el término de *coste* asociado a los pasajeros considerado en la función a *minimizar* es

$$- \sum_{((i,j),(o,d)) \in \mathcal{F}} w^{od}(\omega_t t_{ij} + \omega_p p_{ij}) F_{ij}^{od}. \quad (3.9)$$

- Función de utilidad de los pasajeros en la red competidora. Puesto que la proporción de demanda absorbida por cada red en un mercado $(o, d) \in \mathcal{K}$ viene dada por la utilidad de las distintas redes para ese mercado, es necesario incluir la utilidad de cada una de las redes competidoras en la función objetivo. Esto supone que debe incluirse

$$- \sum_{\forall (o,d) \in \mathcal{K}} w^{od}(\omega_t t_{ext}^{od} + \omega_p p_{ext}^{od}) F_{ext}^{od}. \quad (3.10)$$

En este punto es importante señalar que en la aproximación que se presenta en este TFG, el diseño de la red alternativa es conocido y fijo, es decir, no es susceptible de ser optimizado.

- Entropías de los flujos en las redes nueva y competidora. Tal y como se demostró en la sección 2.4, para obtener un modelo de elección de la demanda basado en una distribución tipo *logit* entre distintas redes competidoras, es necesario incluir en la función objetivo términos de entropía de las variables que modelan el flujo por las distintas redes. En este caso, esto supone tener en cuenta los siguientes términos:

$$\sum_{\forall (o,d) \in \mathcal{K}} w^{od} F^{od} (\log(F^{od}) - 1) + \sum_{\forall (o,d) \in \mathcal{K}} w^{od} F_{ext}^{od} (\log(F_{ext}^{od}) - 1). \quad (3.11)$$

- Congestión en los arcos de la nueva red. En este TFG se utiliza el modelo de congestión basado en una función hiperbólica descrito en la sección 3.1, que proporciona una cota superior a la congestión del sistema. De esta forma, se añade a la función de coste los siguientes términos de congestión:

$$\sum_{\forall (i,j) \in \mathcal{A}} \frac{1}{\Delta A_{ij} h_{A_{ij}} + \epsilon_H} + \sum_{i \in \mathcal{S}} \frac{1}{\Delta S_i h_{S_i} + \epsilon_H}. \quad (3.12)$$

Una vez definidos todos los términos que modelan el interés de los pasajeros, se define de forma completa el *coste* de los pasajeros, función *PAX*, como:

$$\begin{aligned} PAX(\mathcal{X}_{flow}, \mathcal{X}_{aux}) = & \sum_{\forall (o,d) \in \mathcal{K}} w^{od} F^{od} (\log(F^{od}) - 1) - \sum_{((i,j),(o,d)) \in \mathcal{F}} w^{od}(\omega_t t_{ij} + \omega_p p_{ij}) F_{ij}^{od} \\ & + \sum_{\forall (o,d) \in \mathcal{K}} w^{od} F_{ext}^{od} (\log(F_{ext}^{od}) - 1) - \sum_{\forall (o,d) \in \mathcal{K}} w^{od}(\omega_t t_{ext}^{od} + \omega_p p_{ext}^{od}) F_{ext}^{od} \\ & + \sum_{\forall (i,j) \in \mathcal{A}} \frac{1}{\Delta A_{ij} h_{A_{ij}} + \epsilon_H} + \sum_{i \in \mathcal{S}} \frac{1}{\Delta S_i h_{S_i} + \epsilon_H}. \end{aligned} \quad (3.13)$$

Objetivo combinado

El último paso consiste en definir una función objetivo que balancee adecuadamente el objetivo de los pasajeros y el coste del operador. La manera clásica de abordar el diseño consiste en combinar los dos términos de una forma lineal, asignando un coeficiente positivo a cada uno de ellos.

En problemas en los que, como en el que aquí se presenta, se tiene únicamente dos objetivos, el procedimiento habitual consiste en: i) definir un parámetro escalar $\alpha \in (0, 1)$ y ii) fijar que los coeficientes de cada uno de los dos términos en la función objetivo sean α y $(1 - \alpha)$. Esto supone que el objetivo a minimizar es

$$\alpha PAX(\mathcal{X}_{flow}, \mathcal{X}_{aux}) + (1 - \alpha)OP(\mathcal{X}_{cons}). \quad (3.14)$$

Esta aproximación plantea un interrogante adicional: la selección del valor de α . En determinados escenarios, su valor viene dado en función de algún criterio preestablecido, como, por ejemplo, la importancia estratégica de los objetivos o las preferencias del usuario, ya sea gestor u operador. Sin embargo, es habitual realizar un análisis multiobjetivo de suma ponderada [51] que, haciendo un barrido en valores de α , permite obtener el conjunto de soluciones Pareto dominantes. A partir de ellas, el usuario puede escoger cuál de esas soluciones se ajusta más a la realidad de su escenario de aplicación. Para facilitar la exposición, la metodología seguida en este TFG para fijar el valor de α (que involucra la obtención de la curva de Pareto y del punto de utopía) no se detallará hasta la sección 3.5, una vez que ya se haya introducido el problema de optimización completo.

3.3. Formulación convexa del problema

Habiendo introducido la notación, las variables de estado y diseño y la función objetivo, es momento de describir el modelo de optimización completo. En esta sección se detallan las distintas restricciones y se proporcionan los detalles necesarios para la implementación del algoritmo de MM para promover soluciones *sparse* para la infraestructura.

Teniendo en cuenta lo expuesto en las secciones anteriores, el modelo matemático que describe el problema de optimización a resolver es el siguiente:

$$\min_{\mathcal{X}_{cons}, \mathcal{X}_{flow}, \mathcal{X}_{aux}} \quad \alpha PAX(\mathcal{X}_{flow}, \mathcal{X}_{aux}) + (1 - \alpha)OP(\mathcal{X}_{cons}) \quad (3.15)$$

s. a:

$$\frac{\sum_{\forall(o,d) \in \mathcal{K}} F_{ij}^{od} w^{od}}{a_{nom}} \leq \tau_{ij} A_{ij} \quad \forall(i,j) \in \mathcal{A} \quad (\delta_{ij}) \quad (3.16)$$

$$\sigma_i \sum_{\forall j \in \mathcal{N}_{in}(i)} A'_{ji} \leq S_i \quad \forall i \in \mathcal{S} \quad (\zeta_{in_i}) \quad (3.17)$$

$$\sigma_i \sum_{\forall j \in \mathcal{N}_{out}(i)} A'_{ij} \leq S_i \quad \forall i \in \mathcal{S} \quad (\zeta_{out_i}) \quad (3.18)$$

$$A'_{ij} = A_{ij} + \Delta A_{ij} \quad \forall(i,j) \in \mathcal{A} \quad (\chi_{ij}) \quad (3.19)$$

$$S'_i = S_i + \Delta S_i \quad \forall i \in \mathcal{S} \quad (\gamma_i) \quad (3.20)$$

$$\sum_{\forall(i,j) \in \mathcal{A}} A'_{ij} \leq a_{max} \quad (\xi) \quad (3.21)$$

$$\sum_{\forall j \in \mathcal{N}_{out}(i)} F_{ij}^{od} - \sum_{\forall j \in \mathcal{N}_{in}(i)} F_{ji}^{od} = \begin{cases} F^{od} & i = o, \\ 0 & i \neq o \neq d, \\ -F^{od} & i = d \end{cases} \quad \forall i, (o,d) \in \mathcal{S} \times \mathcal{K} \quad (\mu_i^{od}) \quad (3.22)$$

$$F^{od} + F_{ext}^{od} = 1 \quad \forall(o,d) \in \mathcal{K} \quad (\nu^{od}) \quad (3.23)$$

$$\sum_{\forall i \in \mathcal{S}} \lambda c_{S_i} \mathbb{I}_{\{S'_i > 0\}} + \bar{c}_{S_i} S'_i + \sum_{\forall(i,j) \in \mathcal{A}} \lambda c_{A_{ij}} \mathbb{I}_{\{A'_{ij} > 0\}} + \bar{c}_{A_{ij}} A'_{ij} \leq b \quad (\beta). \quad (3.24)$$

La función objetivo en (3.15) minimiza la suma ponderada de los términos *PAX* y *OP*. Las restricciones (3.16)-(3.18) limitan la cantidad de pasajeros en los enlaces y estaciones, respectivamente. El grupo de restricciones (3.16) establece que el número de pasajeros usando un enlace (transformado a unidades de capacidad) ha de ser menor o igual a la capacidad operada multiplicada por el factor de carga de los vehículos τ_{ij} . Las restricciones (3.17) y (3.18) fuerzan que la capacidad de operación de una estación sea mayor o igual al flujo de vehículos que recibe en hora punta, para operaciones de llegada y salida, respectivamente. Las restricciones (3.19) y (3.20) relacionan las variables de capacidad en enlaces y estaciones con las variables auxiliares utilizadas para modelar la congestión, respectivamente. En concreto, fuerzan a que la capacidad dimensionada en enlaces y estaciones sea igual a la suma de la capacidad operable y el término de holgura, limitando la congestión máxima de la infraestructura. La restricción (3.21) limita la capacidad máxima agregada de la red (*e.g.*, debido a la existencia de un número máximo de vehículos disponible). Las restricciones (3.22) garantizan la conservación de flujo en todos los nodos de la red para todos los mercados. Nótese que el flujo generado en origen ha de ser igual a la proporción de pasajeros que viajan por la red que se construye y que lo mismo debe ocurrir con el tráfico absorbido en destino. Por el contrario, en nodos intermedios no puede ni aparecer ni desaparecer tráfico. Las restricciones (3.23) obligan a clasificar a todos los pasajeros de un mercado $(o,d) \in \mathcal{K}$ de acuerdo con su elección de viaje, ya sea por la nueva red o su alternativa. Por último, (3.24) es la restricción de presupuesto, que limita la inversión máxima en la construcción de la red. El parámetro λ indica la importancia relativa del coste fijo con respecto a los costes lineales de construcción. Estos costes son, habitualmente, muy difíciles de cuantificar y existe un nivel de incertidumbre sobre los mismos, por lo que, durante las simulaciones, se considerarán distintos valores de λ . Tal y como se ha explicado en capítulos anteriores, la estrategia adoptada para lidiar con la restricción de presupuesto consiste en mover los términos de la restricción de presupuesto a la función objetivo (siendo β un multiplicador de Lagrange / hiperparámetro que refleja el peso relativo que debe darse a la restricción) y, además, utilizar un algoritmo de MM que permita reemplazar la norma ℓ_0 con una norma ℓ_1 reponderada (véase la sección 2.3 para obtener detalles adicionales con respecto a este

Algoritmo 2: Algoritmo MM para el diseño de una RdT

Entradas: $\alpha, \beta, K, \mathcal{S}, \mathcal{A}$, variables de estado y parámetros.

1 **Inicialización:** $A'_{ij}{}^{(0)} = 10^6 \forall (i, j) \in \mathcal{A}$ y $S'_i{}^{(0)} = 10^6 \forall i \in \mathcal{S}$

2 **para** $k = 1, 2, \dots, K - 1$ **hacer**

3 Obtener $A'^{* (k)}_{ij}$ y $S'^{* (k)}_i$ resolviendo

$$(\check{\mathcal{X}}^{* (k)}_{cons}, \check{\mathcal{X}}^{* (k)}_{flow}, \check{\mathcal{X}}^{* (k)}_{aux}) = \arg \min_{\mathcal{X}_{cons}, \mathcal{X}_{flow}, \mathcal{X}_{aux}} \quad (3.25) \quad \text{s. a: (3.16) - (3.23),}$$

 utilizando $A'^{(k-1)}_{ij}, S'^{(k-1)}_i$.

4 Establecer $A'_{ij}{}^{(k)} = A'^{* (k)}_{ij} \forall (i, j) \in \mathcal{A}$ y $S'_i{}^{(k)} = S'^{* (k)}_i \forall i \in \mathcal{S}$.

5 **fin**

6 Construir los conjuntos $\mathcal{S}_0 := \{i \mid S'^{(K-1)}_i \leq \varepsilon_S\}$ y $\mathcal{A}_0 := \{(i, j) \mid A'^{(K-1)}_{ij} \leq \varepsilon_A\}$.

7 Fijar $c_{S_i} = 0 \forall i \in \mathcal{S}$, $c_{A_{ij}} = 0 \forall (i, j) \in \mathcal{A}$.

8 Obtener la solución al problema resolviendo

$$(\check{\mathcal{X}}^{* (K)}_{cons}, \check{\mathcal{X}}^{* (K)}_{flow}, \check{\mathcal{X}}^{* (K)}_{aux}) = \arg \min_{\mathcal{X}_{cons}, \mathcal{X}_{flow}, \mathcal{X}_{aux}} \quad (3.25)$$

s. a:

$$(3.16) - (3.23)$$

$$A'_{ij} = 0 \quad \forall (i, j) \in \mathcal{A}_0$$

$$S'_i = 0 \quad \forall i \in \mathcal{S}_0.$$

Salidas: $(\check{\mathcal{X}}^{* (K)}_{cons}, \check{\mathcal{X}}^{* (K)}_{flow}, \check{\mathcal{X}}^{* (K)}_{aux})$

último asunto). Teniendo en cuenta esta estrategia, el objetivo penalizado del problema pasa a ser

$$\alpha PAX(\mathcal{X}_{flow}, \mathcal{X}_{aux}) + (1 - \alpha) OP(\mathcal{X}_{cons}) + \beta \left(\sum_{\forall i \in \mathcal{S}} \lambda \frac{c_{S_i}}{S'^{(k-1)}_i + \varepsilon_S} S'_i + \bar{c}_{S_i} S'_i + \sum_{\forall (i, j) \in \mathcal{A}} \lambda \frac{c_{A_{ij}}}{A'^{(k-1)}_{ij} + \varepsilon_A} A'_{ij} + \bar{c}_{A_{ij}} A'_{ij} \right). \quad (3.25)$$

donde k denota el índice de iteración asociado al algoritmo MM y el valor de β se escogerá para que se satisfaga la restricción de presupuesto.

De esta manera, el problema a resolver queda como se muestra a continuación:

$$\min_{\mathcal{X}_{cons}, \mathcal{X}_{flow}, \mathcal{X}_{aux}} \quad (3.25) \quad \text{s. a: (3.16) - (3.23).} \quad (3.26)$$

Este problema es convexo y, por lo tanto, puede resolverse de forma óptima de múltiples maneras (incluyendo la utilización de un *solver* genérico de programación convexa) en un tiempo polinómico. El algoritmo final utilizado en este TFG para el diseño de una RdT se resume en el algoritmo 2.

El algoritmo, cuyo cuerpo principal se corresponde a un esquema de MM en el que en cada iteración se resuelve una instancia de la minimización de (3.26), resuelve el problema de optimización de RdT planteado en este capítulo. Es importante tener en cuenta dos aspectos. El primero es que, finalizadas las K iteraciones del bucle, en el paso 6 se identifican los conjuntos $\mathcal{S}_0$ y $\mathcal{A}_0$, que representan las estaciones y enlaces que el método de MM ha decidido no construir (con una pequeña tolerancia representada por ε_S y ε_A). Una vez identificados las estaciones y enlaces en $\mathcal{S}_0$ y $\mathcal{A}_0$, el algoritmo resuelve una última instancia del problema (pasos 7 y 8) en la que se fijan a cero las capacidades de las estaciones y enlaces que no se van a construir y se eliminan todas las penalizaciones asociadas a la *sparsity* de la infraestructura (se remite al lector a la discusión de la eliminación del sesgo en la parte final de la sección 2.3.1). El segundo aspecto es que el algoritmo 2 resuelve un problema para un valor de β determinado. Si se desea alcanzar un nivel de presupuesto específico, deberá realizarse una búsqueda sobre β (e.g., con un método dual o de bisección) para encontrar la solución asociada a un presupuesto determinado.

3.4. Versión MIP del problema convexo con variables de construcción binarias

En esta sección se presenta un modelo de optimización híbrido que: i) considera las funciones no lineales de entropía para llevar a cabo el reparto *logit* y ii) no incorpora *sparsity*, sino que considera variables *binarias* para decidir si se construye o no la infraestructura. El modelo en esta sección puede verse como una versión MIP del modelo convexo propuesto en la sección 3.3. La principal motivación detrás de este modelo es poder comparar su solución con la generada por el algoritmo 2, para poder así validar la aproximación presentada.

Para presentar el modelo MIP, es necesario en primer lugar redefinir alguna de las variables de trabajo. En concreto, se introducen nuevas variables de diseño (de naturaleza binaria) para modelar la construcción de estaciones y enlaces. Las variables de estado, por otro lado, son iguales a las descritas en la sección 3.2. Una vez explicadas las nuevas variables de diseño, se procede a presentar el nuevo modelo de optimización.

En lo que a las variables de diseño se refiere, las variables de flujo (agrupadas en $\mathcal{X}_{flow}$) y las variables auxiliares (agrupadas en $\mathcal{X}_{aux}$ y orientadas a modelar la congestión), no sufren cambios. Las variables de construcción de enlaces y estaciones (agrupadas en $\mathcal{X}_{cons}$), sí sufren cambios. Concretamente, en esta sección se consideran las siguientes variables:

- $A'_{ij} \in \mathbb{R}^{\geq 0}$, que hace referencia a la capacidad construida en el enlace entre los nodos $(i, j) \in \mathcal{A}$.
- $S'_i \in \mathbb{R}^{\geq 0}$, que es igual a la capacidad construida en la estación en el nodo $i \in \mathcal{S}$.
- $A^b_{ij} \in \{0, 1\}$, que es una variable binaria que vale 1 si se construye el enlace entre los nodos $(i, j) \in \mathcal{A}$.
- $S^b_i \in \{0, 1\}$, que es una variable binaria que vale 1 cuando se construye la estación en el nodo $i \in \mathcal{S}$.

Para distinguir estas variables de las definidas en $\mathcal{X}_{cons}$ en (3.6), el nuevo conjunto de variables de construcción se define como:

$$\mathcal{X}_{cons}^b = \{S'_i\}_{\forall i \in \mathcal{S}} \cup \{S^b_i\}_{\forall i \in \mathcal{S}} \cup \{A'_{ij}\}_{\forall (i,j) \in \mathcal{A}} \cup \{A^b_{ij}\}_{\forall (i,j) \in \mathcal{A}} \quad (3.27)$$

El siguiente paso consiste en introducir el modelo de optimización. Como ya se ha visto en la sección 3.2, la función objetivo se puede agrupar en dos términos: objetivo del operador, OP , que permanece definido de la misma forma que en (3.8), y objetivo de los pasajeros, ahora PAX^b , debido a que el término asociado a la congestión cambia ligeramente, tal y como se muestra a continuación:

$$\begin{aligned} PAX^b(\mathcal{X}_{flow}, \mathcal{X}_{aux}) = & \sum_{\forall (o,d) \in \mathcal{K}} w^{od} F^{od} (\log(F^{od}) - 1) - \sum_{((i,j),(o,d)) \in \mathcal{F}} w^{od} (\omega_t t_{ij} + \omega_p p_{ij}) F^{od}_{ij} \\ & + \sum_{\forall (o,d) \in \mathcal{K}} w^{od} F^{od}_{ext} (\log(F^{od}_{ext}) - 1) - \sum_{\forall (o,d) \in \mathcal{K}} w^{od} (\omega_t t^{od}_{ext} + \omega_p p^{od}_{ext}) F^{od}_{ext} \\ & + \sum_{\forall (i,j) \in \mathcal{A}} \frac{1}{\Xi(1 - A^b_{ij}) + h_{A_{ij}} \Delta A_{ij} + \epsilon_H} + \sum_{i \in \mathcal{S}} \frac{1}{\Xi(1 - S^b_i) + h_{S_i} \Delta S_i + \epsilon_H}. \end{aligned} \quad (3.28)$$

En los dos últimos términos de la expresión (3.28), la constante Ξ es un número de gran magnitud. El rol de esta constante es doble: i) que el nivel de congestión en enlaces y estaciones que no se construyen es, en la práctica, nulo y ii) que si la estación o enlace se construye, el término asociado a la congestión sea igual que el considerado en la sección anterior. Teniendo en cuenta esta ligera

redefinición, el problema a resolver es

$$\min_{\mathcal{X}_{cons}^b, \mathcal{X}_{flow}, \mathcal{X}_{aux}} \alpha PAX^b(\mathcal{X}_{flow}, \mathcal{X}_{aux}) + (1 - \alpha)OP(\mathcal{X}_{cons}^b) \quad (3.29)$$

s. a:

$$\frac{\sum_{\forall (o,d) \in \mathcal{K}} F_{ij}^{od} w^{od}}{a_{nom}} \leq \tau_{ij} A_{ij} \quad \forall (i,j) \in \mathcal{A} \quad (3.30)$$

$$\sigma_i \sum_{\forall j \in \mathcal{N}_{in}(i)} A'_{ji} \leq S_i \quad \forall i \in \mathcal{S} \quad (3.31)$$

$$\sigma_i \sum_{\forall j \in \mathcal{N}_{out}(i)} A'_{ij} \leq S_i \quad \forall i \in \mathcal{S} \quad (3.32)$$

$$A'_{ij} \leq M A_{ij}^b \quad \forall (i,j) \in \mathcal{A} \quad (3.33)$$

$$S'_i \leq M S_i^b \quad \forall i \in \mathcal{S} \quad (3.34)$$

$$A'_{ij} = A_{ij} + \Delta A_{ij} \quad \forall (i,j) \in \mathcal{A} \quad (3.35)$$

$$S'_i = S_i + \Delta S_i \quad \forall i \in \mathcal{S} \quad (3.36)$$

$$\sum_{\forall (i,j) \in \mathcal{A}} A'_{ij} \leq a_{max} \quad (3.37)$$

$$\sum_{\forall j \in \mathcal{N}_{out}(i)} F_{ij}^{od} - \sum_{\forall j \in \mathcal{N}_{in}(i)} F_{ji}^{od} = \begin{cases} F^{od} & i = o, \\ 0 & i \neq o \neq d, \\ -F^{od} & i = d \end{cases} \quad \forall i, (o,d) \in \mathcal{S} \times \mathcal{K} \quad (3.38)$$

$$F^{od} + F_{ext}^{od} = 1 \quad \forall (o,d) \in \mathcal{K} \quad (3.39)$$

$$\sum_{\forall i \in \mathcal{S}} \lambda c_{S_i} S_i^b + \bar{c}_{S_i} S'_i + \sum_{\forall (i,j) \in \mathcal{A}} \lambda c_{A_{ij}} A_{ij}^b + \bar{c}_{A_{ij}} A'_{ij} \leq b. \quad (3.40)$$

Las restricciones (3.30)-(3.32), (3.35)-(3.36) y (3.37)-(3.39) son iguales a las que se impusieron en la sección anterior, por lo que no hace falta comentarlas en más detalle aquí. Las restricciones (3.33)-(3.34), no presentes en la sección anterior, garantizan que si la variable binaria de construcción es cero, la capacidad asociada también es cero. Por último, la restricción (3.40) limita el presupuesto máximo a emplear y, tal y como se puede apreciar, su forma es ligeramente distinta de la restricción (3.24) de la sección anterior, puesto que en (3.40) se hace uso de las variables binarias.

La solución de (3.29)-(3.40) deberá atacarse con un *solver* que sea capaz de manejar simultáneamente variables enteras, variables lineales y funciones convexas. La complejidad del problema escalará *exponencialmente* con el número de variables enteras, que, en este caso, se corresponde con $|\mathcal{S}| + |\mathcal{A}|$.

3.5. Optimización multiobjetivo: búsqueda del punto de utopía y curva de Pareto

Antes de presentar los resultados numéricos que analizan el comportamiento de los modelos propuestos en este capítulo, en esta sección se explica la propuesta adoptada para fijar el valor del coeficiente α que determina el balance entre los costes del operador y de los pasajeros en la función objetivo.

El primer paso para determinar el valor de α consiste en normalizar los dos términos de la función objetivo de forma que los rangos dinámicos sean similares. Para esto, se calculará el punto de utopía [52]. En un problema de optimización cuya función objetivo es la combinación lineal de L objetivos, el punto de utopía se define como el punto perteneciente a un espacio de L dimensiones dado por el valor óptimo de cada una de las funciones objetivo optimizadas por separado.

En un problema de minimización de una única variable x con L funciones objetivo, denotadas como $f_1, \dots, f_L$, el procedimiento para calcular el punto de utopía y normalizar las L funciones objetivo sería el siguiente:

1. Resolver el problema optimizando $f_1, \dots, f_L$ por separado, obteniendo la solución óptima $\{x_{f_i}^*\}_{i=1}^L$ para cada uno de ellos, donde $x_{f_i}^* = \arg \min_x f_i(x)$. Obtener el punto de utopía, $\mathbf{f}^{UP} \in \mathbb{R}^L$, como:

$$\mathbf{f}^{UP} = (f_1(x_{f_1}^*), \dots, f_L(x_{f_L}^*))^T. \quad (3.41)$$

2. Calcular el margen dinámico de cada función objetivo, $\Delta \mathbf{f}$, que se define como la diferencia entre el valor óptimo y el peor valor (máximo en este caso) que toma la función objetivo al ser evaluada en las soluciones óptimas del resto de funciones objetivo:

$$\Delta \mathbf{f} = \left(\max\{f_1(x_{f_i}^*)\}_{i=1}^L - [\mathbf{f}^{UP}]_1, \dots, \max\{f_L(x_{f_i}^*)\}_{i=1}^L - [\mathbf{f}^{UP}]_L \right)^T. \quad (3.42)$$

Finalmente, los márgenes dinámicos se utilizan para normalizar las funciones objetivo, de forma que se minimicen las funciones $f_1/[\Delta \mathbf{f}]_1, \dots, f_L/[\Delta \mathbf{f}]_L$. Es decir, ahora se considera la minimización de la combinación lineal de las funciones como sigue:

$$\min_x \sum_{i=1}^L \alpha_i \frac{f_i(x)}{[\Delta \mathbf{f}]_i}, \quad (3.43)$$

donde la suma de todos los coeficientes $\alpha_i \geq 0$ debe valer uno. Resolviendo el modelo multiobjetivo para diferentes combinaciones de α_i , se obtiene la frontera Pareto, la cual contiene el conjunto de soluciones no dominadas.

Nótese que cuando, como en el caso que se presenta en este TFG, se trata con un problema con $L = 2$ funciones objetivo, la frontera Pareto colapsa en una curva en el plano. Así pues, para calcular la curva (frontera) de Pareto hay que tener en cuenta que $L = 2$, que $f_1(\cdot) = PAX^b(\cdot)$, que $f_2(\cdot) = OP(\cdot)$ y que las variables de optimización son las incluidas en los conjuntos $(\mathcal{X}_{cons}^b, \mathcal{X}_{flow}, \mathcal{X}_{aux})$, definidos en (3.27), (3.5) y (3.7). Teniendo esto en cuenta, en primer lugar es necesario resolver dos problemas de optimización para el mismo escenario: uno minimiza el objetivo $PAX^b(\cdot)$ y otro el objetivo $OP(\cdot)$. Cada problema aporta una solución diferente, de donde se obtiene el punto de utopía (PAX^{b*}, OP^*) y los márgenes dinámicos ΔPAX^b y ΔOP . Con ellos, se resuelve el problema de optimización cuya función objetivo es

$$\min_{\mathcal{X}_{cons}^b, \mathcal{X}_{flow}, \mathcal{X}_{aux}} \alpha \frac{PAX^b(\mathcal{X}_{flow}, \mathcal{X}_{aux})}{\Delta PAX^b} + (1 - \alpha) \frac{OP(\mathcal{X}_{cons}^b)}{\Delta OP}, \quad (3.44)$$

sujeto a las restricciones (3.30)-(3.40). Como se ha indicado anteriormente, este problema se resuelve para distintos valores del parámetro α . De este barrido se obtiene una serie de soluciones al problema, que se representan en el mismo plano que el punto de utopía. Esta representación se conoce como diagrama de Pareto.

Una vez que se ha obtenido el diagrama de Pareto, es habitual elegir como valor de α aquel que se corresponde con el punto de la frontera que se encuentra a una menor distancia Euclídea del punto de utopía [52]. Dicho valor será el que se utilice en las simulaciones que se presentan en la siguiente sección.

3.6. Análisis cuantitativo de resultados

A lo largo de esta sección se realiza una serie de simulaciones con los modelos propuestos en este capítulo, de cara a validar la formulación convexa propuesta en la sección 3.3. Para ello, en primer lugar

se describe detalladamente el escenario sobre el que se han realizado las simulaciones. En segundo lugar, se realiza un análisis multiobjetivo para determinar el valor numérico del parámetro que pondera los términos de la función objetivo. Para finalizar, se realiza un análisis comparativo entre las soluciones propuestas por la formulación convexa en combinación con el algoritmo 2 y por la versión MIP en (3.29)-(3.40) para distintas configuraciones de las variables de estado. Además, en el apéndice B se proporciona un análisis de sensibilidad, en el que se estudia la variación de la solución óptima dada por el algoritmo 2 frente a perturbaciones en las variables de estado, describiendo así el comportamiento del modelo frente a la incertidumbre en la estimación de los parámetros que intervienen en el mismo.

El objetivo principal de esta sección consiste en validar la metodología de diseño de RdT propuesta en este TFG. Dicha validación justifica la aplicación de nuestra metodología a los casos de estudio realistas que se presentarán en los capítulos 4 y 5.

3.6.1. Caso de estudio

Los experimentos realizados en este capítulo se apoyan en la red presentada en la figura 2.1, cuyo estudio se plantea en [53]. La red máxima se compone de un conjunto $|\mathcal{S}|$ de 9 nodos y un conjunto $|\mathcal{A}|$ de 30 enlaces. Se consideran $|\mathcal{K}| = 72$ mercados origen-destino cuya demanda es conocida. Cada estación $i \in \mathcal{S}$ y enlace $(i, j) \in \mathcal{A}$ tienen asociado un coste fijo de construcción c_{S_i} y $c_{A_{ij}}$, respectivamente. Puesto que la modelización con la que se trabaja en este capítulo contempla el diseño de una RdT capacitada, es necesario introducir también costes de construcción dependientes de la capacidad diseñada para las estaciones $i \in \mathcal{S}$ y enlaces $(i, j) \in \mathcal{A}$. En este caso, se considera que los costes son lineales, por lo que su valor por unidad de capacidad viene dado por $\bar{c}_{S_i}$ y $\bar{c}_{A_{ij}}$, y se obtiene como un porcentaje de los costes fijos. Además, también se introducen costes de operación sobre los enlaces $(i, j) \in \mathcal{A}$ de la red, de nuevo linealmente dependientes de la capacidad. Su coste unitario está representado por $c_{o_{ij}}$, y se establece en base a los costes operativos de los vehículos de RENFE Cercanías. Al igual que ocurrirá en el capítulo siguiente, los valores de referencia se han obtenido a través de colaboraciones previas entre RENFE y el grupo de investigación GISAT-URJC. La distancia entre las estaciones de la red también viene dada en [53], por lo que asumiendo una velocidad media de los vehículos de 30 km/h, es posible determinar el tiempo t_{ij} de viaje para cualquier enlace $(i, j) \in \mathcal{A}$. El precio asociado a cada enlace $(i, j) \in \mathcal{A}$, p_{ij} , se calcula mediante una función cóncava monótonamente creciente dependiente del tiempo de viaje asociado al mismo. Por su parte, la distancia de viaje alternativa para cada mercado $(o, d) \in \mathcal{K}$ se proporciona en [53]. Asumiendo de nuevo una velocidad media de 30 km/h, se obtiene el tiempo de viaje t_{ext}^{od} . De igual forma, el precio del viaje para el mercado $(o, d) \in \mathcal{K}$ asociado a la red alternativa, p_{ext}^{od} , se determina mediante una función cóncava monótonamente creciente sobre el tiempo de viaje t_{ext}^{od} . De acuerdo con los datos proporcionados por RENFE Cercanías, se estima que la ocupación media de los trenes, τ , es del 57 %, y que en la hora punta del día se operan el 25 % de los viajes, tomando σ este valor. La capacidad nominal a_{nom} de los vehículos se extrae también de los datos de RENFE Cercanías, donde cada unidad de capacidad (cada vehículo) puede acomodar hasta 588 pasajeros. El valor de todos los parámetros y variables de estado que forman parte del modelo se puede consultar en el repositorio de código de este TFG (<https://github.com/frealr/TFG-RdT>), por lo que los experimentos expuestos a continuación son reproducibles.

Finalmente, en lo que a los *solvers* se refiere, tanto el modelo convexo como la versión MIP se han resuelto utilizando MOSEK (versión 10.1.13 para Matlab), al que se llama dentro del entorno CVX (versión 2.2) utilizando Matlab (versión 2023b). Los algoritmos han sido ejecutados en un ordenador con sistema operativo MacOS con 1 procesador de 2 núcleos y frecuencia de 2.8 GHz, y 8GB de memoria RAM. Parámetros como el tiempo de ejecución dependerán, naturalmente, del tipo de ordenador y el tipo y versión de *solvers* utilizados. En las simulaciones realizadas, se ha fijado el tiempo máximo de cómputo a 1000 segundos.

Modelo	Núm. V.C.	Núm. V.B.	Núm. Restr.	Núm. NZE
Convexo	6789	0	17092	4606
MIP	14663	81	10313	6873

Tabla 3.1: Comparación entre el modelo convexo y MIP para el caso de referencia, con $\lambda = 5$ y $\beta = 1.25$. Se reporta el número de: variables continuas (“Núm. V.C.”), variables binarias (“Núm. V.B.”), restricciones (“Núm. Restr.”) y *non-zero elements* (“Núm. NZE”), entradas distintas de cero en el modelo matemático de optimización.

3.6.2. Valor de α en la función objetivo

Tal y como se describió en la sección 3.5, se ha llevado a cabo un análisis multiobjetivo para la selección del valor de α . Puesto que se pretende validar el algoritmo de optimización convexo tomando como referencia el modelo MIP de la sección 3.4, dicho análisis se ha realizado utilizando las funciones objetivo y restricciones del modelo MIP. Para la red descrita en la sección 3.6.1, el punto de utopía se corresponde con el origen (0,0) y los valores de la curva de Pareto más cercanos al mismo se obtienen en el rango $\alpha \in [0.4, 0.6]$, con el punto de mínima distancia al origen estando muy cercano a $\alpha = 0.5$. Tras un análisis preliminar de resultados, en los experimentos presentados en el resto del capítulo, se utilizará un valor de $\alpha = 0.6$, que prioriza sensiblemente a los pasajeros con respecto al punto $\alpha = 0.5$. La principal motivación tras esta decisión es que la solución óptima del diseño de la RdT es más sensible cuando el valor de α se fija a 0.6 (en parte porque para este valor la nueva red capta una mayor proporción de la demanda). Esta mayor sensibilidad hace que el análisis numérico al cambiar los valores de los distintos parámetros y la configuración de los experimentos sea más rico, lo que permite entender mejor las ventajas y limitaciones del diseño propuesto.

3.6.3. Comparación cuantitativa de la formulación convexa y la formulación MIP

Para comparar la formulación convexa frente a la formulación MIP, se han simulado una serie de escenarios para la red de 9 nodos descrita en la sección 3.6.1, variando: i) el presupuesto disponible b (o, de forma equivalente, el valor del parámetro β) y ii) la importancia de los costes fijos de construcción en el presupuesto, cuantificada en el parámetro λ . En concreto, λ es un coeficiente que ajusta los costes fijos de construcción de estaciones y enlaces, permitiendo explorar cómo varían estos costes en diferentes escenarios. Al modificar λ , se ponderan los valores base de los costes fijos. Antes de realizar un barrido de parámetros, en la tabla 3.1 se muestran los tamaños de las formulaciones convexa y MIP.

Con el fin de entender mejor el comportamiento de ambos modelos, en la figura 3.2 (compuesta por 4 gráficas) se proporcionan los valores de 4 métricas distintas (una por gráfica) para los modelos convexo y MIP y para distintas combinaciones de λ y β . Más concretamente, se consideran dos valores de λ y se hace un barrido sobre β . De esta manera, el eje vertical de cada una de las gráficas se corresponde con la métrica analizada (valor objetivo, arcos o enlaces construidos, demanda absorbida y tiempo de cómputo), el eje horizontal se corresponde con el barrido en β y las distintas líneas se corresponden con diferentes combinaciones de modelo y valor de λ . Más concretamente, se utilizan líneas de color azul para $\lambda = 5$, líneas de color rojo para $\lambda = 6$, líneas continuas para los resultados del modelo convexo y líneas discontinuas para los resultados del modelo MIP. A continuación se comentan los resultados clave observados en cada una de las gráficas de la figura 3.2 :

- En la gráfica 1 se compara el número de arcos o enlaces construidos en ambas soluciones. Comparando las soluciones de los dos métodos, se comprueba que esta métrica es muy similar para las soluciones de ambas formulaciones. Observando las variaciones con cambios en λ y β , se aprecia que menores valores de β (mayores presupuestos) dan lugar a más arcos construidos y que mayores valores de λ (costes fijos más altos) dan lugares a redes más *sparse*.
- En la gráfica 2 de la figura 3.2 se representa la demanda atraída, donde se observa que: a) ambos modelos generan resultados similares para todos los presupuestos y b) la demanda atraída decrece cuando el presupuesto es más limitado.

- La gráfica 3 de la figura 3.2 representa el valor objetivo obtenido. Cabe destacar que ambas soluciones se evalúan en la función objetivo de la formulación MIP, para poder realizar así una comparación homogénea. Pese a que la formulación MIP siempre obtiene un mejor valor objetivo (como es de esperar, porque se evalúa en su objetivo y por tanto representa un límite inferior), la diferencia es muy poco significativa, alcanzando el 0.6 % en el peor de los casos. Así mismo, se aprecia que, cuanto menor es el presupuesto, mayor es el valor del coste en la función objetivo (es decir, peor es la solución obtenida).
- Por último, en la gráfica 4 de la figura 3.2 se representa el tiempo computacional necesario para obtener cada una de las soluciones. En este aspecto, se observa cómo el tiempo de cómputo de la formulación convexa es básicamente plano, mientras que el tiempo de cómputo de la formulación MIP crece de manera exponencial (recuérdese que el tiempo máximo de cómputo se ha fijado a 10^3 segundos). Es interesante observar que, cuando el límite presupuesto es alto (valores bajos de β), la restricción (3.40) del problema MIP (que es la que involucra combinaciones lineales de variables binarias) es bastante laxa, lo que reduce el tiempo de cómputo necesario para encontrar la solución. Sin embargo, al aumentar el valor de β (restringir el presupuesto en el problema MIP), la resolución del problema se vuelve computacionalmente más compleja, puesto que la restricción (3.40) toma gran relevancia.

Las cuatro conclusiones fundamentales de los resultados presentados son: i) el diseño convexo es efectivamente capaz de generar soluciones *sparse*, ii) el comportamiento de distintas métricas con respecto a la variación de parámetros del presupuesto sigue lo esperado de una forma intuitiva, iii) las soluciones y el valor objetivo del modelo convexo y el modelo MIP son muy similares y iv) la complejidad computacional del modelo convexo es significativamente menor. Esto no solo valida la metodología de diseño propuesta en este TFG, sino que permite poder atacar problemas de mayor escala que traten con RdT reales.

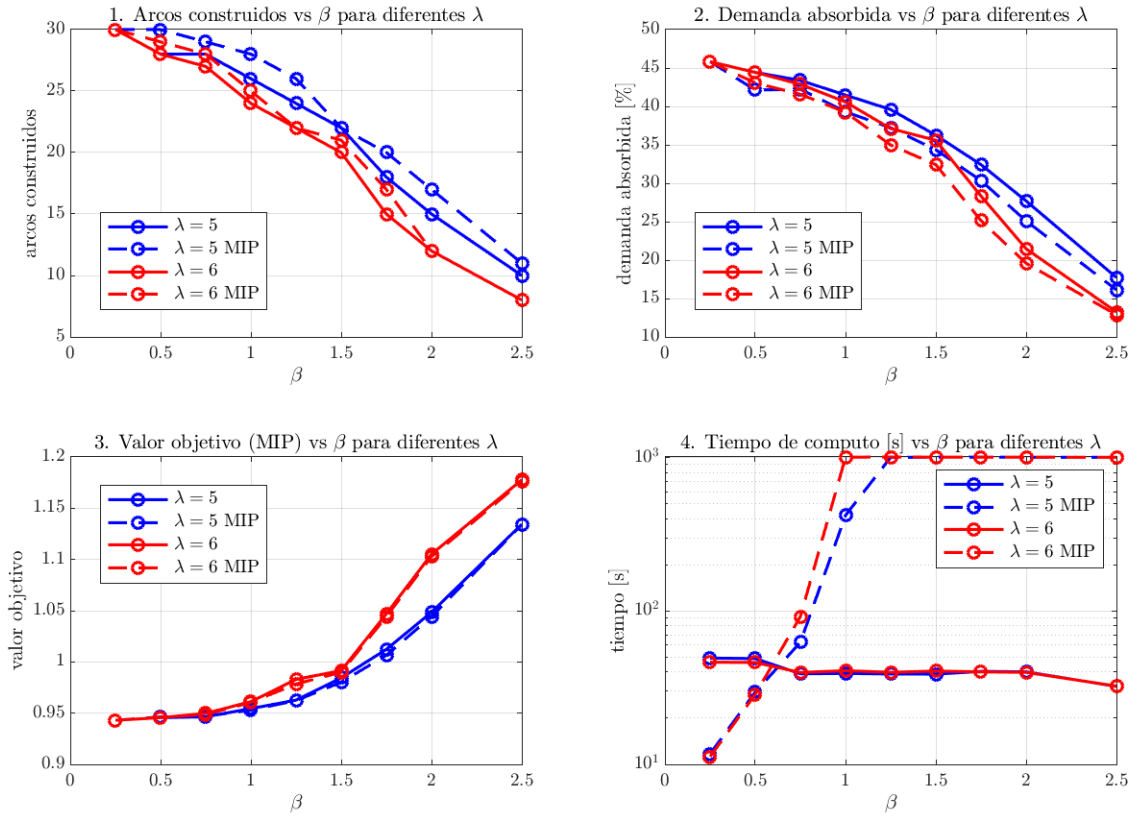


Figura 3.2: Comparativa de distintas métricas entre la propuesta convexa definida en la sección 3.3 y su versión MIP definida en la sección 3.4, para una red de 9 nodos. Las métricas estudiadas son: 1. Número de arcos o enlaces construidos. 2. Porcentaje de demanda atendida. 3. Valor objetivo evaluado en la formulación MIP. 4. Tiempo de cómputo.

Capítulo 4

Optimización de una RdT de metro con operador público

En este capítulo, se aplica la metodología propuesta de diseño convexo de RdT a un escenario real. El caso de estudio que se va a analizar se corresponde con el diseño de la red de metro de la ciudad de Sevilla, explotada por una sociedad concesionaria de la Junta de Andalucía. Concretamente, se considera la construcción de hasta 24 estaciones localizadas en lugares estratégicos de la ciudad, así como la potencial interconexión de estas. Los datos de demanda se corresponden con valores estimados por el propio municipio y, tal y como se hizo en el capítulo anterior, los valores de construcción y capacidad se fijarán a partir de valores realistas utilizados por otros operadores de TP.

La sección 4.1 proporciona una breve introducción al problema y revisa el estado del arte específicamente relacionado con el diseño de RdT de tránsito rápido. La sección 4.2 describe en mayor detalle la RdT considerada. La sección 4.3 se dedica a explicar las modificaciones (relativamente menores) que deben realizarse a la formulación presentada en el capítulo anterior para adaptarse a las necesidades y particularidades del escenario considerado en este capítulo. Tal y como se expondrá en detalle, la diferencia principal consiste en aplicar una metodología de diseño multiobjetivo en la que uno de los objetivos se tiene en cuenta por medio de una restricción de desigualdad, de forma que se establezcan límites en su rango dinámico. Así, se llevará a cabo un diseño de la red que permite estudiar cómo evolucionan los costes de construcción en función de la demanda atraída. Este análisis permite al operador público elegir el punto de funcionamiento de la red, de forma que pueda valorar adecuadamente el esfuerzo económico a realizar frente a los potenciales beneficios en términos de movilidad y cohesión social. La sección 4.4 se dedica a implementar los algoritmos de optimización de la RdT considerando distintos escenarios de estudio. Entre los aspectos que se analizarán se encuentran el impacto de los costes fijos, el impacto del potencial crecimiento de la demanda en los barrios de nueva construcción o cómo debería modificarse el diseño de la red si el regulador decide primar a los pasajeros de las zonas más deprimidas o desfavorecidas de la ciudad, de acuerdo con el objetivo 11 de los objetivos de desarrollo sostenible (“lograr que las ciudades sean más inclusivas, seguras, resilientes y sostenibles”).

4.1. Estado del arte

Esta sección presenta una revisión del estado del arte en el problema de diseño de RdT de tránsito rápido. Este tipo de RdT está habitualmente gestionado por operadores públicos cuyo propósito es impulsar la conectividad socio-económica de una región específica, donde, a pesar de que la minimización del coste de inversión es importante, se busca garantizar un mínimo de demanda atendida por la nueva RdT.

Debido al aumento de la población y la expansión de las zonas urbanas, muchas ciudades y áreas metropolitanas alrededor del mundo están planificando, construyendo o ampliando sus sistemas de TP.

Entre estos, los sistemas de metro son de los más eficientes porque consumen menos energía y pueden transportar una gran cantidad de pasajeros por unidad de superficie. Los sistemas de metro ayudan a disminuir el tráfico de automóviles privados, reduciendo así la congestión y la contaminación. A veces se utiliza el término metro como sinónimo de transporte/tránsito rápido, pero este último tiene una aceptación más amplia. El transporte rápido generalmente abarca no solo el metro, sino también trenes de cercanías, metro ligero, tranvía, monorraíl y otros sistemas urbanos de TP rápido masivo. La planificación del transporte rápido ferroviario se encuentra dentro del campo más amplio de la planificación de redes. El diseño de la RdT rápido está compuesto por dos problemas interrelacionados: la localización de las estaciones y la determinación de los enlaces entre las mismas.

El objetivo principal de un sistema de transporte colectivo es mejorar la movilidad de la población. Otra característica importante es su influencia en la estructuración de las ciudades, ya que proporcionan la columna vertebral para el desarrollo de áreas residenciales, comerciales y empresariales. Los sistemas de transporte rápido requieren inversiones de alto nivel, tanto para su construcción como para su mantenimiento, generalmente relacionadas con la construcción de túneles, vías elevadas o a nivel del suelo, sistemas de comunicación y la compra de material rodante. Los costes operativos incluyen costes fijos y variables en base diaria.

Sin pérdida de generalidad, los proyectos de construcción de RdT rápido pueden clasificarse en tres grupos: aquellos en los que se planifica una sola línea desde cero, aquellos en los que se planifican simultáneamente varias líneas desde cero y aquellos en los que se planea extender una red existente, lo que corresponde a un problema de expansión de capacidad.

Aquí se considera el problema de diseñar una RdT rápido desde cero. El primer intento de modelar y resolver el problema general de diseño de RdT rápido se presenta en [16], que proporciona un enfoque computacionalmente viable dividido en tres etapas. La primera es la selección de estaciones clave, que son los principales puntos de atracción (estaciones de tren o autobús, aeropuertos, hospitales, campus universitarios, centros comerciales, áreas densamente pobladas lejos del área central de la ciudad, etc.). La segunda etapa consiste en conectar las estaciones clave para formar una red central. Finalmente, las estaciones intermedias se localizan en la alineación resultante de la segunda etapa. En el mismo trabajo, se utiliza un modelo lineal de programación entera con el objetivo de maximizar la cobertura de viajes para resolver el problema de diseño de la red central en presencia de un modo de transporte alternativo. Con el objetivo de modelar el comportamiento del usuario, [54] introduce una función *logit* para distribuir los viajeros entre la RdT rápido y los modos privados. Para mantener el carácter lineal del modelo, considera una interpolación lineal por trozos de la función *logit*. Por otro lado, [17] resuelve el problema de diseño de RdT rápido teniendo en cuenta que la red está compuesta por arcos que luego se ensamblan en líneas. Los autores aplican dos criterios: minimizar el coste de construcción y maximizar la captura de tráfico origen-destino, obteniendo soluciones óptimas de Pareto. En [55] se estudia un problema de expansión de capacidad multi-periodo, de forma que la infraestructura que se abrirá en cada periodo se determina teniendo en cuenta una función objetivo que combina términos relativos a la comunidad, a los pasajeros y al operador. Dado que el problema general no puede resolverse de manera exacta, los autores diseñan un procedimiento heurístico. Otras aproximaciones para resolver el modelo de programación matemática para problemas de diseño de RdT rápido se basan en la descomposición de Benders [28], en algoritmos genéticos [56] o en la búsqueda adaptativa de vecindarios grandes [23].

Otra línea de investigación trata aspectos de robustez de la RdT. Se han estudiado varias formas de tratar la robustez: mediante la aplicación de teoría de juegos [57], proporcionando rutas alternativas para ser utilizadas en caso de interrupción [18], a través del concepto de robustez recuperable [27] o mediante la aplicación de un *greedy randomized adaptive search procedure* al problema de diseño de la red ferroviaria [21].

Todas las referencias anteriores contemplan el coste del viaje como el tiempo empleado en viajar (o la distancia recorrida), sin considerar ningún coste de transbordo. Algunos autores estudian el problema de diseño de red teniendo en cuenta los transbordos entre diferentes líneas. En este contexto, [58], [26] y [59] abordan el problema haciendo algunas suposiciones sobre la velocidad de los trenes y la frecuencia

de los servicios, aunque no logran resolver problemas de diseño de redes a escala real.

4.2. Datos básicos y descripción de la red

En este capítulo se diseña la construcción de una potencial red de metro para la ciudad de Sevilla. Tal y como se menciona en la sección introductoria de este capítulo, al tratarse del diseño de una red de tránsito rápido por parte de un operador público, el beneficio social, asociado al número de pasajeros atendido, tendrá una importancia mayor que los costes de inversión (que, aun así, serán tenidos en cuenta).

Los nodos de la red son las hipotéticas ubicaciones donde se emplazarán las estaciones a considerar, y se corresponden con los centroides de demanda considerados. Los enlaces indican las estaciones a unir de manera directa, dimensionando a su vez la frecuencia de viaje. El problema de optimización a resolver se basa en los datos de demanda para un *día tipo*, por lo que hay que adaptar los costes de construcción y calcular el coste repercutido diario. Para dimensionar la capacidad necesaria en las estaciones, se utiliza el máximo volumen de tráfico simultáneo diario. A continuación, se detallan las fuentes de información consultadas para determinar el valor de las variables de estado a la hora de resolver el problema de optimización planteado.

- El conjunto de potenciales estaciones $\mathcal{S}$ y los candidatos a enlaces directos $\mathcal{A}$, así como la matriz de demanda W entre centroides, proceden del estudio realizado en [53]. La ubicación de los nodos se detalla en la tabla 4.1.
- Los valores de distancia para los enlaces entre estaciones se han establecido mediante la distancia de *Haversine* para coordenadas sobre la superficie terrestre. Para obtener el tiempo de viaje $t_{ij} \forall (i, j) \in \mathcal{A}$ se ha considerado una velocidad media de 30 km/h.
- El precio de viaje para cada enlace, p_{ij} , se calcula mediante una función cóncava monótonamente creciente dependiente únicamente de la distancia de dicho enlace.
- Los costes de operación, $c_{o_{ij}}$, se establecen de acuerdo al coste operativo en la red de RENFE Cercanías, incluyendo gasto de combustible y coste de personal, entre otros. Estos datos se obtienen de colaboraciones previas entre RENFE y el grupo de investigación GISAT-URJC.
- Los costes fijos de construcción de estaciones $c_{S_i} \forall i \in \mathcal{S}$ y enlaces $c_{A_{ij}} \forall (i, j) \in \mathcal{A}$ se extraen del estudio realizado en [27]. Por su parte, los costes por unidad de capacidad tanto en estaciones $\bar{c}_{S_i} \forall i \in \mathcal{S}$ como en enlaces $\bar{c}_{A_{ij}} \forall (i, j) \in \mathcal{A}$ se establecen como un porcentaje de los costes fijos.
- El máximo volumen de tráfico simultáneo soportado en cada estación, σ , se estima en un 25 % sobre el tráfico total diario en cada estación, según datos de la red de RENFE Cercanías obtenidos del GISAT-URJC.
- El factor de ocupación medio de cada vehículo, τ , toma un valor del 57 % sobre la capacidad nominal del vehículo, según datos procedentes de la red de RENFE Cercanías obtenidos del GISAT-URJC.
- La capacidad nominal asociada a cada vehículo, a_{nom} , es de 588 pasajeros. Esta es la capacidad máxima promedio para un tren de la red de RENFE Cercanías obtenida del GISAT-URJC.

1	Bellavista	7	Triana - Cartuja	13	Camas	19	Triana - Cartuja
2	Bellavista	8	Remedios	14	Macarena	20	San Jerónimo - Norte
3	Centro	9	Centro	15	Cerro Amate	21	Centro
4	Sur	10	Triana - Cartuja	16	Santa Justa	22	Sevilla Este
5	Camas	11	Sevilla Este	17	Triana - Cartuja	23	Sur
6	Sevilla Este	12	San Jerónimo - Norte	18	Nervión	24	Tomares

Tabla 4.1: Ubicación de los nodos de la red de Sevilla.

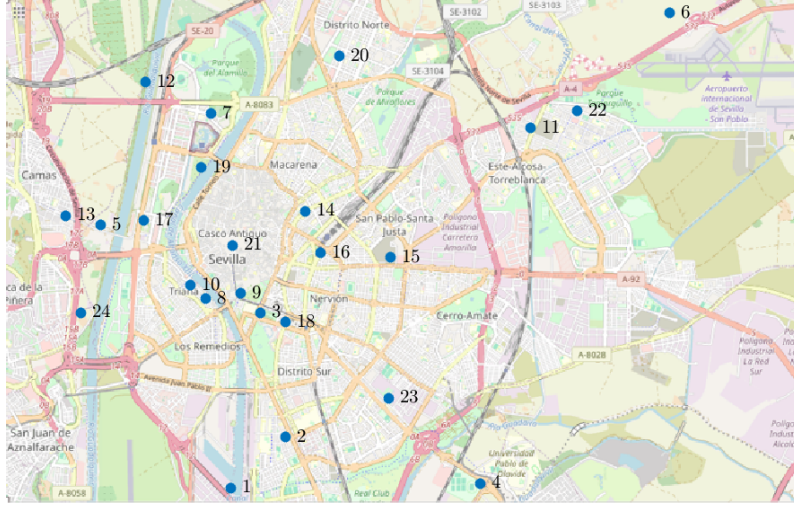


Figura 4.1: Distribución geográfica de los nodos (estaciones candidatas) para el diseño de una red de metro. La imagen del mapa de Sevilla ha sido extraída de Google Maps (<https://www.google.es/maps/@37.3833902,-6.0051874,12z/data=!5m1!1e2?entry=ttu>).

Explicados los valores de los parámetros, en las siguientes secciones: i) se detalla la metodología para la optimización de la red, ii) se presenta la formulación completa del problema de optimización y iii) se presentan y analizan los resultados de los experimentos numéricos llevados a cabo.

4.3. Modificaciones en la formulación del problema de optimización

Tal y como se ha mencionado, la filosofía (finalidad) de la optimización en este capítulo es ligeramente diferente de la de la sección 3.3. En este caso, se busca minimizar el presupuesto utilizado, *garantizando* un mínimo de demanda atendida por la nueva infraestructura. Esta garantía se modela habitualmente mediante la restricción

$$\sum_{(o,d) \in \mathcal{K}} w^{od} F^{od} \geq \eta \sum_{(o,d) \in \mathcal{K}} w^{od}, \quad (4.1)$$

donde η establece el porcentaje mínimo de la demanda total que debe ser atendido por la nueva red. De esta manera, la atracción de demanda se garantiza y, solo entonces, se permite al modelo reducir los costes. Esto contrasta con los capítulos anteriores, donde el presupuesto disponible se imponía mediante una restricción. No obstante, aunque la restricción (4.1) es tratable y preserva la convexidad, introducirla de forma directa como restricción provoca que su multiplicador de Lagrange asociado aparezca en el reparto logístico de la demanda. Para evitar esto, se propone eliminar esta restricción y asegurar la demanda atendida mediante una búsqueda exhaustiva sobre un nuevo parámetro γ , que represente el *trade-off* entre el objetivo de los pasajeros y del operador. Esto permitirá estudiar distintos niveles de despliegue (dimensión de la red) en función de la demanda mínima que se quiera

atender (parámetro η). En lo que resta de sección se formula el modelo de optimización, centrándose principalmente en las diferencias con respecto a capítulos anteriores.

4.3.1. Función objetivo

Las variables de estado, los parámetros y las variables de diseño (flujos, construcción y auxiliares) son las mismas que las del capítulo anterior (véanse las ecuaciones (3.5)-(3.7)), por lo que se remite al lector a dicho capítulo para la descripción de las mismas y se centra la discusión en la formulación de la función objetivo.

Tal y como se menciona en la sección introductoria de este capítulo, el objetivo del problema de optimización viene dado únicamente por los intereses del operador público: reducir los costes. Sin embargo, también es necesario considerar el objetivo de los pasajeros para asegurar el mínimo de demanda atendida (véase la discusión de la sección anterior) y el reparto logístico de la demanda.

Objetivo del operador

El objetivo del operador se define como

$$OP(\mathcal{X}_{cons}) = \sum_{\forall(i,j) \in \mathcal{A}} c_{oij} A'_{ij} + \sum_{\forall i \in \mathcal{S}} \lambda \frac{c_{S_i}}{S_i'^{(k-1)} + \epsilon_S} S'_i + \bar{c}_{S_i} S'_i + \sum_{\forall(i,j) \in \mathcal{A}} \lambda \frac{c_{A_{ij}}}{A_{ij}'^{(k-1)} + \epsilon_A} A'_{ij} + \bar{c}_{A_{ij}} A'_{ij}. \quad (4.2)$$

En esta ocasión se agregan tanto los costes de construcción como los de operación, puesto que el objetivo del operador es minimizar el conjunto de los costes, incluyendo la inversión inicial.

Objetivo de los pasajeros

En este caso, el objetivo de los pasajeros se mantiene igual que en la sección 3.2.3, definido como en la ecuación (3.13) y reproducido a continuación por comodidad de lectura.

$$\begin{aligned} PAX(\mathcal{X}_{flow}, \mathcal{X}_{aux}) = & \sum_{\forall(o,d) \in \mathcal{K}} w^{od} F^{od} (\log(F^{od}) - 1) - \sum_{((i,j),(o,d)) \in \mathcal{F}} w^{od} (\omega_t t_{ij} + \omega_p p_{ij}) F_{ij}^{od} \\ & + \sum_{\forall(o,d) \in \mathcal{K}} w^{od} F_{ext}^{od} (\log(F_{ext}^{od}) - 1) - \sum_{\forall(o,d) \in \mathcal{K}} w^{od} (\omega_t t_{ext}^{od} + \omega_p p_{ext}^{od}) F_{ext}^{od} \\ & + \sum_{\forall(i,j) \in \mathcal{A}} \frac{1}{\Delta A_{ij} h_{A_{ij}} + \epsilon_H} + \sum_{i \in \mathcal{S}} \frac{1}{\Delta S_i h_{S_i} + \epsilon_H}. \end{aligned} \quad (4.3)$$

4.3.2. Formulación del problema de optimización

Una vez que las variables involucradas en el modelo se han presentado y el objetivo está definido, es momento de introducir la formulación del problema de optimización a resolver, dado por (4.4)-(4.13) donde, nuevamente, se utilizan los conjuntos de variables definidos en (3.6)-(3.7). Así, el problema viene definido por

$$\min_{\mathcal{X}_{cons}, \mathcal{X}_{flow}, \mathcal{X}_{aux}} \gamma PAX(\mathcal{X}_{flow}, \mathcal{X}_{aux}) + OP(\mathcal{X}_{cons}) \quad (4.4)$$

s. a:

$$\frac{\sum_{\forall (o,d) \in \mathcal{K}} F_{ij}^{od} w^{od}}{a_{nom}} \leq \tau_{ij} A_{ij} \quad \forall (i,j) \in \mathcal{A} \quad (4.5)$$

$$\sigma_i \sum_{\forall j \in \mathcal{N}_{in}(i)} A'_{ji} \leq S_i \quad \forall i \in \mathcal{S} \quad (4.6)$$

$$\sigma_i \sum_{\forall j \in \mathcal{N}_{out}(i)} A'_{ij} \leq S_i \quad \forall i \in \mathcal{S} \quad (4.7)$$

$$A'_{ij} = A_{ij} + \Delta A_{ij} \quad \forall (i,j) \in \mathcal{A} \quad (4.8)$$

$$S'_i = S_i + \Delta S_i \quad \forall i \in \mathcal{S} \quad (4.9)$$

$$\sum_{\forall (i,j) \in \mathcal{A}} A'_{ij} \leq a_{max} \quad (4.10)$$

$$A'_{ij} = A'_{ji} \quad \forall (i,j) \in \mathcal{A} \quad (4.11)$$

$$\sum_{\forall j \in \mathcal{N}_{out}(i)} F_{ij}^{od} - \sum_{\forall j \in \mathcal{N}_{in}(i)} F_{ji}^{od} = \begin{cases} F^{od} & i = o, \\ 0 & i \neq o \neq d, \\ -F^{od} & i = d \end{cases} \quad \forall i, (o,d) \in \mathcal{S} \times \mathcal{K} \quad (4.12)$$

$$F^{od} + F_{ext}^{od} = 1 \quad \forall (o,d) \in \mathcal{K}, \quad (4.13)$$

donde las únicas novedades en la formulación respecto al modelo dado por (3.26) son la función objetivo (4.4) y la restricción (4.11). Como ya se ha mencionado, con el objetivo (4.4) se pretende minimizar los gastos del operador $OP(\mathcal{X}_{cons})$ garantizando un mínimo de demanda atendida. Para mantener la demanda atendida sobre el umbral deseado, se realizará una búsqueda exhaustiva sobre el valor de γ , el cual indica la importancia del objetivo de los pasajeros en relación con el objetivo del operador. Al aumentar el valor de dicho parámetro, se atraerá a más demanda, construyendo una mayor infraestructura que dé lugar a menores tiempos de viaje y por tanto mayores utilidades. Ajustando dicho coeficiente, se consigue el efecto deseado: minimizar costes asegurando un valor umbral de demanda atendida. Por otro lado, la restricción (4.11) impone que la red ha de ser simétrica, es decir, que las capacidades dimensionadas para los enlaces han de ser idénticas en ambos sentidos, lo cual es habitual en el diseño de redes urbanas de metro.

4.4. Presentación y análisis de los resultados numéricos

Los resultados numéricos y las simulaciones en esta sección se estructuran en torno a tres subsecciones:

1. *Caso base y diseño óptimo.* En primer lugar, se describe el caso base y se resuelve el modelo para obtener la red óptima para distintos niveles de despliegue. Esto corresponde a un barrido sobre el parámetro γ relativo a la importancia dada al beneficio social y la atracción de demanda. Dicho barrido permite explorar la relación entre el beneficio social y el presupuesto invertido, revelando los límites de operación y permitiendo al responsable del despliegue elegir el nivel de servicio en el que desea operar la red, buscando un punto de equilibrio entre ambos objetivos.
2. *Sobreponderación de usuarios provenientes de zonas deprimidas.* La segunda parte de la sección analiza cómo cambia el diseño al sobreponderar el beneficio social por la atracción de la demanda procedente de los barrios más deprimidos de la ciudad.
3. *Crecimiento de la demanda y adaptación futura.* Por último, se analiza cómo cambia el diseño de la red cuando se tienen en cuenta las previsiones de crecimiento demográfico en cada uno de los distritos de la ciudad. Los dos pasos clave para llevar a cabo este análisis son: i) para cada nodo,

se define un coeficiente de crecimiento de demanda asociado al plan de desarrollo urbanístico y al plan de nueva vivienda en el distrito donde se ubica el nodo y ii) se modifica el patrón de demanda, multiplicando la demanda actual por el coeficiente calculado en el paso anterior. Esto permite entender qué cambios serían necesarios en el diseño de la RdT si se deseara hacer frente a la demanda futura de la ciudad.

Debido a las limitaciones en el número de páginas de este documento, la presentación y el análisis de los resultados se circunscriben a los aspectos más importantes. Se remite al lector al repositorio de código y a los artículos relacionados para consultar detalles adicionales. Para el análisis, se reportarán métricas que aporten intuición sobre la evolución de las soluciones a medida que se modifica el parámetro γ . Además, también se presentará la topología de la red construida para varias simulaciones, comparándolas entre sí y visualizando la evolución de la infraestructura a construir a medida que crece el presupuesto disponible. A continuación, se listan las métricas (figuras de mérito) que se analizarán en el resto del capítulo:

- Demanda atraída (“D.A.”) por la nueva red: mide el impacto que la nueva red tiene sobre la población de la ciudad. También sirve para medir la magnitud (dimensión) de la infraestructura construida.
- Presupuesto utilizado (“Pres.”) para la construcción de la nueva red: reporta el gasto monetario en el que se incurre al construir una infraestructura determinada. Junto con la demanda atraída, sirve para estimar la magnitud de la red.
- Número de estaciones construidas (Nodos) para la nueva red: ayuda a comprender la topología de la solución, así como la magnitud de la red construida.
- Número de enlaces construidos (Arcos) en la nueva red: sirve para medir el grado de interconexión de las estaciones, así como la posible robustez de la red ante disrupciones.
- Capacidad operada (“Cap.”) por la nueva red: se trata de una medida importante, ya que hace alusión a los recursos humanos y físicos que serán necesarios para operar la red diseñada.
- Tiempo medio por pasajero en la nueva red y la red competidora ($\bar{t}_{PAX}$ y $\bar{t}_{PAX_{ext}}$ [min], respectivamente): resulta interesante analizar la evolución de estas métricas para comprender el crecimiento y la interconexión de la nueva red a medida que aumenta el valor de γ , así como el impacto que tiene dicho crecimiento sobre las decisiones de los pasajeros y, por lo tanto, para medir el impacto en la red competidora.
- Tiempo medio por ruta ($\bar{t}_{ruta}$ [min]) en la nueva red: es una métrica relevante para evaluar la magnitud y el desempeño de la red construida.
- Tiempo de cómputo (t_{comp}): resulta crítico obtener un tiempo de cómputo reducido cuando el nivel de incertidumbre es significativo, siendo en estos casos conveniente el estudio de numerosas configuraciones y escenarios. En este aspecto, la formulación propuesta en este TFG aporta resultados muy prometedores. Las simulaciones presentadas en este capítulo se han ejecutado en una máquina con 64 GB de RAM y un procesador de 8 núcleos a 3 GHz de frecuencia de reloj. El software utilizado ha sido el entorno CVX (versión 2.2) ejecutado desde Matlab (versión 2023b) y seleccionando MOSEK (versión 10.1.13) como *solver*.

4.4.1. Caso base y diseño óptimo

El caso base de estas simulaciones consiste en la resolución del problema de optimización planteado en este capítulo para distintas configuraciones y escenarios. Estas configuraciones hacen alusión al grado de importancia otorgado al objetivo de los usuarios en la función objetivo (medido por el parámetro γ),

Núm. V.C.	Núm. V.B.	Núm. Restr.	Núm. NZE
333282	0	726275	304563

Tabla 4.3: Métricas del problema de optimización para la red de Sevilla, para el caso concreto de $\lambda = 10$ y $\gamma = 6.7$. Se reporta el número de variables continuas (“Núm. V.C.”), número de variables binarias (“Núm. V.B.”), número de restricciones (“Núm. Restr.”) y número de *non-zero elements* (“Núm. NZE”), entradas distintas de cero en el modelo matemático de optimización.

de donde el responsable de la construcción podrá elegir la magnitud de la red. Además, se consideran varios escenarios donde se varía el valor de λ , la constante de regularización relacionada con los costes fijos, debido a que no se conoce con exactitud la relación entre dichos costes fijos y los lineales. Así pues, se hará un barrido del parámetro γ en el problema declarado en (4.4)-(4.13), para distintos valores del regularizador λ . Antes de presentar los primeros resultados, resulta pertinente establecer el valor de los parámetros que permanecen estáticos durante las simulaciones. Dichos valores están recogidos en la tabla 4.2.

Parámetro	ω_t	ω_p	$\varepsilon_S = \varepsilon_A = \varepsilon_H$	h_{S_i}	$h_{A_{ij}}$
Valor	-0.1	-1	1×10^{-4}	0.1	0.1

Tabla 4.2: Parámetros para el diseño de la red de metro de Sevilla.

Una vez fijados los valores de los parámetros (tabla 4.2) y de las variables de estado (estudios referenciados en la sección 4.2), es momento de comenzar a explicar en detalle las simulaciones ejecutadas y los resultados numéricos obtenidos. En la tabla 4.3 se muestran varias métricas del modelo para el caso concreto de $\lambda = 10$ y $\gamma = 6.7$, relacionadas con las dimensiones del modelo, como el número de variables, de restricciones y de *non-zero elements* (*i.e.*, el número de entradas activas en el modelo de optimización [60]).

Barrido de configuraciones de red: determinación del punto de operación

El primer estudio consiste en explorar distintas configuraciones de la red según el nivel de servicio, medido como demanda atendida, que se pretenda obtener y el presupuesto invertido. Este tipo de estudio es habitual en el estado del arte cuando se busca optimizar un problema multiobjetivo. Sin embargo, dado el tiempo que requieren los modelos MIP para ser resueltos en casos de estudio realistas, hay que elegir un número limitado de escenarios. Una de las ventajas de la formulación propuesta en este TFG es la reducción sustancial del tiempo de cómputo, beneficio amplificado por la necesidad de resolver el problema en múltiples ocasiones, por ejemplo, debido a la existencia de incertidumbre en parámetros o, como es el caso, cuando se quieren obtener distintas configuraciones de la red (una por cada valor de γ) para compararlas.

Así pues, en este estudio se realiza un barrido sobre parámetro γ y, para cada valor, se resuelve una versión del algoritmo 2 (véase sección 3.3) adaptada a las particularidades del problema de optimización en (4.4)-(4.13). En lo que respecta a la importancia de los costes fijos, el valor del parámetro λ se fija a 10. Los resultados mostrados en esta sección se agrupan en tres categorías: a) resumen de métricas obtenidas (tabla 4.4), b) gráficas detalladas de demanda frente a presupuesto invertido y γ , y tiempos frente a γ (figuras 4.2 y 4.3, respectivamente) y c) gráficas con las topologías de la red construida (figura 4.4).

En primer lugar, la tabla 4.4 resume las principales métricas para distintos valores de γ . Como es de esperar, cuanto mayor es el valor de γ mayor es la demanda atraída y el presupuesto necesario. También es interesante analizar el número de arcos y nodos construidos. Se observa un salto notable de demanda atraída cuando la red alcanza el 50 % de los nodos (12 estaciones), así como cuando alcanza el 100 % de los mismos (24 nodos para $\gamma = 7.1$). Por otro lado, se observa que, una vez que

γ	D.A. [%]	Pres. [€]	Arcos	Nodos	Cap.	$\bar{t}_{PAX}$	$\bar{t}_{PAX_{ext}}$	$\bar{t}_{ruta}$	t_{comp} [s]
5.0	0.00	0.00×10^0	0	0	0	-	18.86	-	2474
5.9	1.99	1.07×10^5	12	6	33	5.62	19.12	6.04	2484
6.3	10.93	2.47×10^5	40	12	212	9.66	20.02	10.50	3098
6.7	28.54	4.52×10^5	60	19	680	15.90	20.02	16.72	2354
7.1	43.40	6.25×10^5	80	24	1098	18.84	18.92	20.68	3045
10	45.84	6.78×10^5	96	24	1126	18.58	18.88	20.30	2290
100	52.21	1.32×10^6	230	24	1456	17.98	18.32	18.64	2272

Tabla 4.4: Métricas extraídas de las soluciones para la red de metro de Sevilla con un valor de $\lambda = 10$, haciendo un barrido en γ .

se alcanza el número máximo de nodos, pasar de 80 a 240 arcos, supone un incremento modesto de la demanda atraída, mientras que requiere duplicar el presupuesto. Por último, cabe destacar que el tiempo de ejecución en todos los casos está por debajo de los 3600 segundos. Esto quiere decir que, con la metodología propuesta en este TFG, el diseño de una red real puede realizarse en menos de una hora en un ordenador de prestaciones intermedias.

Con el fin de ganar más intuición sobre la evolución y sensibilidad de algunas de las métricas, en las figuras 4.2 y 4.3 se representan una serie de gráficas que resultan particularmente interesantes. En concreto, se representa y analiza:

- La demanda atraída frente a presupuesto. Es una gráfica que resulta muy útil para medir el esfuerzo necesario para atraer a cierta demanda. La inversa de su derivada indicará el coste marginal necesario para captar más pasajeros. Los resultados se muestran en la gráfica superior de la figura 4.2. La gráfica se analiza de izquierda a derecha y se distinguen tres regímenes diferentes. En primer lugar, se observa que, al principio (valores bajos de presupuesto), la demanda crece muy lentamente con la inversión. De hecho, los resultados parecen indicar que es necesario que la inversión supere un umbral mínimo para comenzar a atraer un valor perceptible de demanda. Esto es razonable, puesto que, para poder ofertar múltiples mercados origen-destino y aprovechar las economías de red, es necesario construir varias estaciones. La parte central de la gráfica muestra que, una vez que dicha infraestructura mínima existe, el coste marginal decrece, aumentando la pendiente de la función representada. Esto se debe que las nuevas estaciones y arcos que se incorporan a la red obtienen una conexión inmediata a múltiples destinos. Finalmente, la parte final de la gráfica (valores altos de presupuesto) revela que, cuando la red comienza a tener un alto grado de conexión, la adición de nuevos enlaces tiene un impacto pequeño en la distancia de viaje y, por lo tanto, la cantidad de nuevos pasajeros atraídos es menor. Esto supone que para atraer más demanda sea necesario aumentar considerablemente la inversión.
- La demanda atraída frente a γ . Sirve para ver la evolución de la demanda atraída de acuerdo a la evolución del coeficiente asociado al beneficio social en el objetivo. Está representada en la gráfica inferior de la figura 4.2. Como es de esperar, se obtiene una función monótonamente creciente donde, a medida que aumenta la importancia del beneficio social en (4.4), crece la red construida y, por lo tanto, la demanda atendida. La gráfica también muestra que la pendiente en el intervalo $\gamma \in [5, 10]$ es muy alta, lo que indica que la infraestructura construida es altamente sensible a la variación del parámetro γ . Esta observación pone en valor los beneficios asociados a poder ejecutar múltiples simulaciones con una complejidad computacional moderada, pudiendo realizar un barrido más fino y garantizando así que el valor de γ se escoge para trabajar en la zona deseada.
- El tiempo de viaje medio por pasajero $\bar{t}_{PAX}$ frente a γ . Se representa tanto para la red de metro construida como para las alternativas competidoras. Sirve para visualizar el tipo de rutas operadas en cada configuración y las sinergias obtenidas cuando se construye una red considerablemente

grande. Los resultados se muestran en las dos gráficas de la columna izquierda de la figura 4.3. En la gráfica superior izquierda de dicha figura se observa cómo, para valores altos de γ , a medida que aumenta el valor de γ y por tanto aumenta la infraestructura construida, los pasajeros ven disminuido su tiempo de viaje por ambas redes. Para la red de metro, esto se debe a que los pasajeros obtienen continuamente mejores rutas para realizar sus viajes. Para la competencia, el tiempo medio por pasajero disminuye debido a que, a medida que la red de metro crece, los pasajeros solo utilizan las rutas más atractivas de la alternativa competidora. Cuando el valor de γ es bajo, se observa cómo, al principio, la red construida se basa en conectar centroides de demanda cercanos entre sí, expandiéndose hacia centroides más lejanos a medida que crece la red. Llama la atención la fase de transición que ocurre en la región $\gamma \in [4, 8.5]$, donde el tiempo de viaje por ambas redes se iguala y se observa una pérdida de monotonidad en el tiempo medio por pasajero para la red competidora. Para analizar mejor este efecto, la gráfica inferior izquierda de la figura 4.3 proporciona una visualización centrada únicamente en el intervalo $\gamma \in [4, 8.5]$. Esta región se corresponde con un escenario en el que la topología de la nueva red se focaliza en conectar las estaciones del centro de la ciudad, no proporcionando servicio a las zonas periféricas. Esto supone que la nueva red “roba” a la red competidora pasajeros que realizan trayectos cortos mientras que la red competidora tiene que seguir dando servicio a los pasajeros de la periferia que, en general, recorren rutas más largas.

- El tiempo medio por ruta $\bar{t}_{ruta}$ frente a γ . Ayuda a entender el tipo de rutas que se operan en las distintas configuraciones obtenidas según el nivel de despliegue de la red. Los resultados se muestran en las dos gráficas de la columna derecha de la figura 4.3. En la gráfica superior derecha de dicha figura se validan las conclusiones discutidas en el punto anterior. En concreto, cuando la red de metro construida es pequeña (poca inversión y pocas estaciones), los destinos que se unen están más cercanos entre sí. Por otro lado, a medida que aumenta la red, comienzan a operarse rutas más largas. Esto es aún más visible en la gráfica inferior derecha de la figura 4.3, que se centra en los resultados obtenidos en la región $\gamma \in [4, 8.5]$. Por último, una vez que se supera el valor de $\gamma = 7.1$, el tiempo medio por ruta operada disminuye, debido a que se ofrecen mejoras para dichas rutas mediante la construcción de nuevos enlaces.

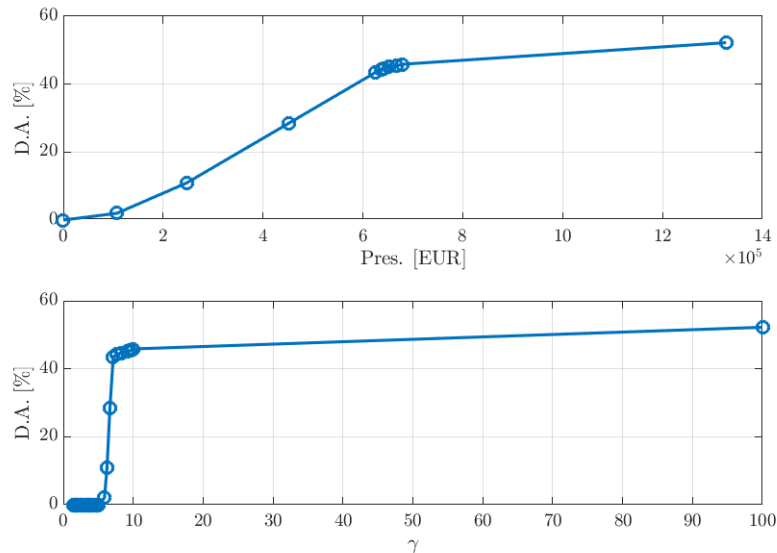


Figura 4.2: Representación de la demanda en función del presupuesto gastado (gráfica superior) y del parámetro γ (gráfica inferior) para el caso base $\lambda = 10$.

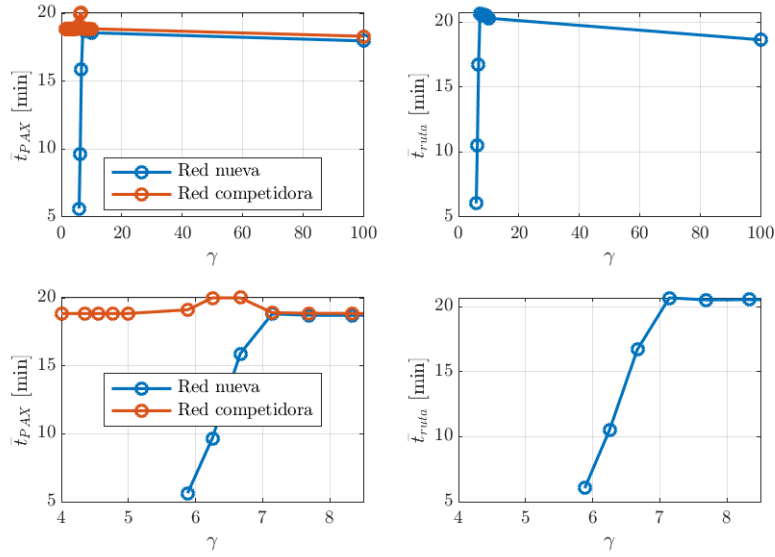


Figura 4.3: Representación del tiempo medio percibido por pasajero (gráfica superior izquierda) y el tiempo medio por ruta operada (gráfica superior derecha) en función del parámetro γ para el caso $\lambda = 10$. De color azul se reportan las métricas para la red de metro, mientras que en naranja aparecen las métricas para las redes competidoras alternativas al metro. Las gráficas de la parte inferior de la figura amplían la información mostrada en las gráficas de la parte superior, focalizándose en la región $\gamma \in [4, 8.5]$.

Por último, la figura 4.4 proporciona representaciones de la topología de la RdT obtenida para distintos valores de γ , permitiendo visualizar de una forma alternativa la evolución de la construcción de la red y proporcionando evidencias adicionales para algunos de los comportamientos que se han discutido en los párrafos precedentes. En concreto, se representa la red obtenida para tres valores de γ , todos en la región crítica en torno a $\gamma = 7$. En la gráfica superior de la figura 4.4 (asociada a $\gamma = 6.3$, el más pequeño de los tres valores escogidos) se verifica que, inicialmente, los nodos conectados por la red de metro son los más cercanos entre sí, ubicados además en las inmediaciones del centro de la ciudad. En la gráfica central de la figura 4.4 (asociada a $\gamma = 6.7$, correspondiente a un valor intermedio) se observa la expansión de la red de metro hacia los barrios situados en la periferia, como es el caso de los nodos 11 y 22, ubicados en Sevilla Este. Finalmente, en la gráfica inferior de la figura 4.4 (asociada a $\gamma = 7.1$, el mayor de los tres valores escogidos) se observa cómo no solo se construyen los nodos 6 (Aeropuerto) y 20 (Pino Montano), sino que, además, se crean nuevos enlaces entre nodos ya existentes, incluyendo los de los nodos 7 (Isla Mágica) o 14 (Macarena). Esta última observación, confirma que la mejora de la conexión de varios nodos con el resto de la red conlleva el descenso de tiempo de viaje percibido por los pasajeros que se observó en la figura 4.3.

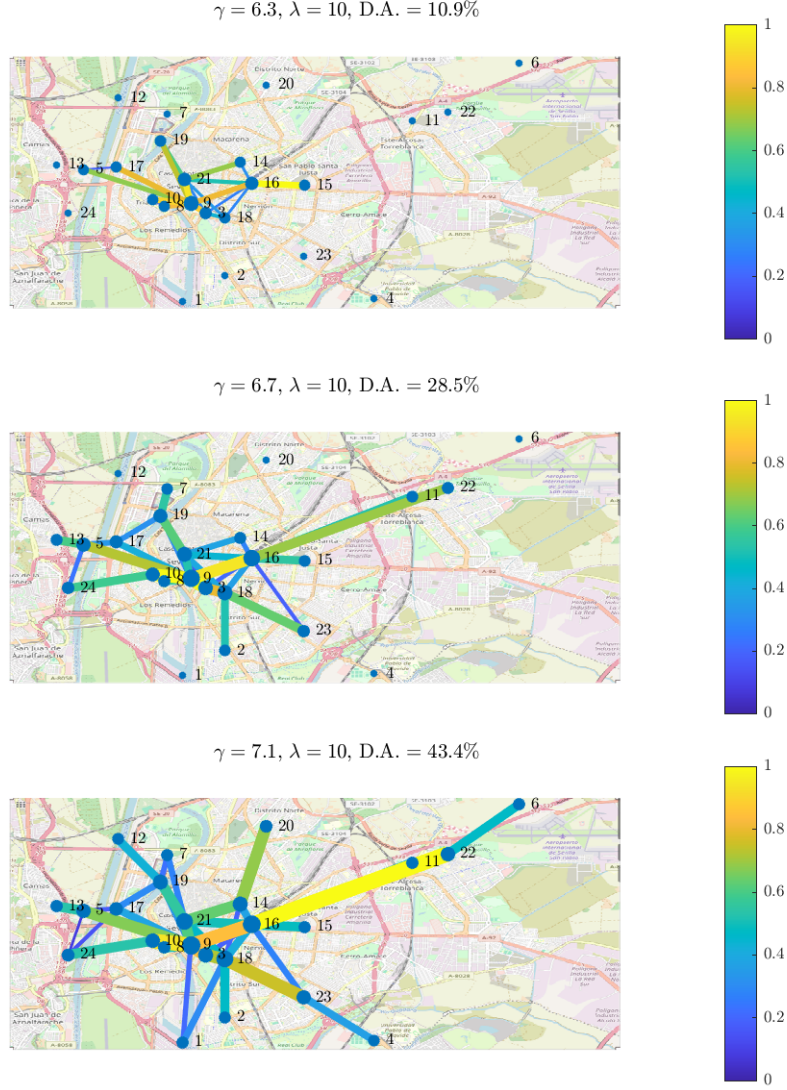


Figura 4.4: Topología de la red de metro para $\gamma = [6.3, 6.7, 7.1]$ y $\lambda = 10$. El tamaño de los nodos hace referencia a la capacidad construida en cada estación y el grosor y color de los enlaces se representa en función de la capacidad diseñada para los mismos. Los colores se han normalizado para que el amarillo se corresponda con el enlace de mayor capacidad (valor normalizado unitario).

Resultados para distintos escenarios de costes fijos de construcción

Para completar el caso de estudio base, se presentan los resultados de una serie de simulaciones en las que se varía el valor del parámetro λ que pondera los costes fijos de construcción de enlaces y estaciones. El valor de γ se adapta de tal forma que el porcentaje de la demanda atraída no varíe de forma significativa (con una tolerancia máxima del 5%) con respecto al caso base, donde $\lambda = 10$ y $\gamma = 6.7$. De esta manera, se facilita el análisis y comparación de los resultados. Los valores de λ elegidos han sido $\lambda \in \{5, 10, 12\}$. Para hacer más sencilla la interpretación de los resultados, se emplea una visualización similar a la de la figura 4.4, donde se representa la topología de la red obtenida. Las topologías asociadas a los tres valores de λ considerados se muestran en la figura 4.5, mientras que las métricas más importantes se reportan en la tabla 4.5. La principal diferencia con respecto al caso de estudio anterior se observa en el número de arcos construidos, que decrece con el valor de λ . Valores de λ mayores suponen mayores costes fijos y, por lo tanto, aumentando el presupuesto para mantener nivel de demanda atraída, la red tiene: i) menor cantidad de arcos construidos y ii) arcos con mayor capacidad. Es posible apreciar cómo, en la gráfica superior de la figura 4.5, los nodos 2 (Sur)

y 24 (Tomares) tienen más conexiones directas que en las soluciones obtenidas en la gráfica central e inferior de la figura 4.5. Por otro lado, la mayor diferencia entre las redes para $\lambda = 10$ y $\lambda = 12$ es que, para la red de $\lambda = 12$, no se construye estación en el nodo 7 (Isla Mágica) y se construyen menos enlaces en torno al nodo 23 (Bellavista).

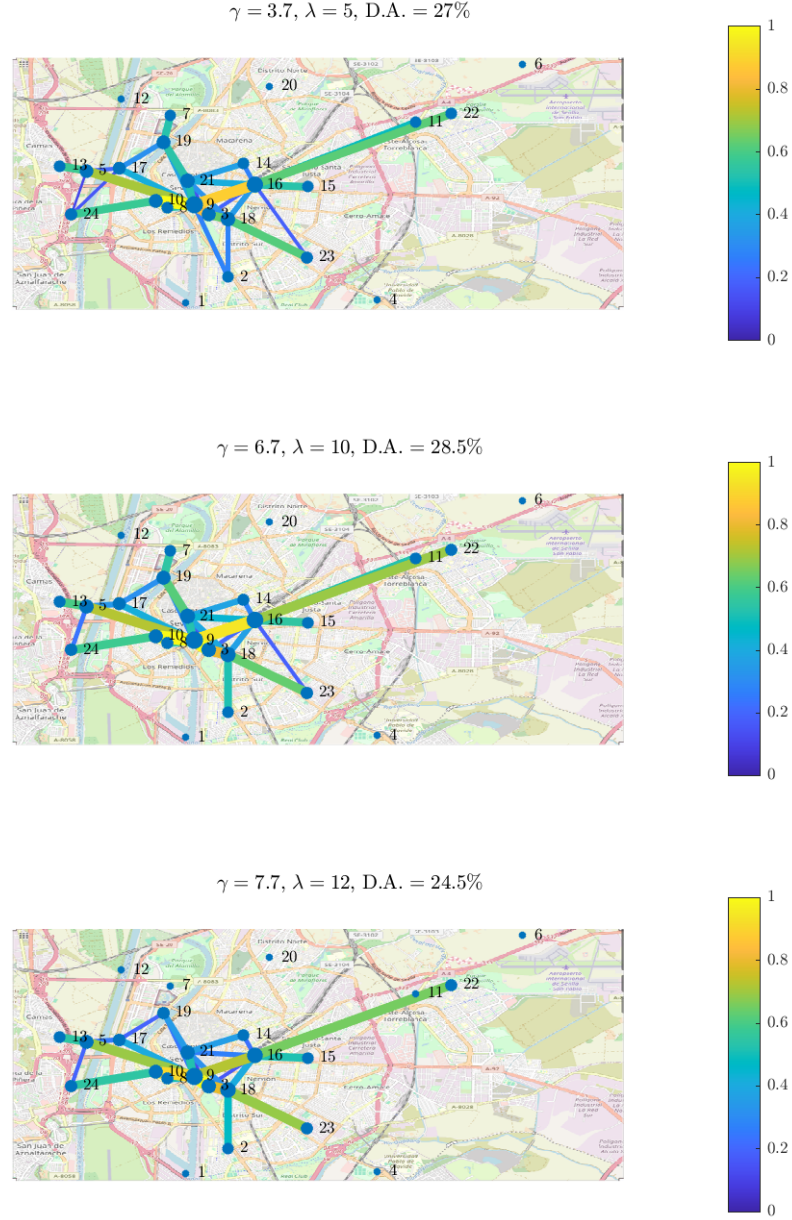


Figura 4.5: Topología de la red de metro construida para $\lambda = 5$ y $\gamma = 3.7$ (gráfica superior), $\lambda = 10$ y $\gamma = 6.7$ (gráfica intermedia) y $\lambda = 12$ y $\gamma = 7.7$ (gráfica inferior). El valor de γ en cada escenario se han ajustado para dar lugar a una atracción de demanda en el entorno del 25 %. El tamaño de los nodos hace referencia a la capacidad construida en cada estación y el grosor y color de los enlaces se representa en función de la capacidad diseñada para los mismos. Los colores se han normalizado para que el amarillo se corresponda con el enlace de mayor capacidad (valor normalizado unitario).

4.4.2. Sobreponderación de usuarios provenientes de zonas deprimidas

La tabla 4.6, obtenida de la base de datos del INE para 2021 [61], pone de manifiesto que la renta *per cápita* de los distritos de la ciudad de Sevilla es muy heterogénea. Esto supone que haya una gran disparidad económica entre los usuarios de cada una de las estaciones de la red de metro. La figura 4.6

λ	γ	D.A. [%]	Pres. [€]	Arcos	Nodos	Cap.	$\bar{t}_{PAX}$	$\bar{t}_{PAX_{ext}}$	$\bar{t}_{ruta}$	t_{comp} [s]
5	3.7	26.88	2.64×10^5	64	19	596	15.74	20.32	16.60	2510
10	6.7	28.54	4.52×10^5	60	19	680	15.90	20.02	16.72	2354
12	7.7	24.20	4.63×10^5	54	17	569	15.12	20.40	15.26	2233

Tabla 4.5: Métricas reportadas para simulaciones considerando distintos escenarios de coste fijo y ajustando el valor de γ para obtener una cobertura de demanda en el entorno del 25 %.

Distrito	Renta per cápita [€]	Distrito	Renta per cápita [€]
Centro	17577	Macarena	11124
Nervión	18554	Cerro Amate	8809
Sur	13270	Triana - Cartuja	15123
San Jerónimo - Norte	10612	Santa Justa	14035
Sevilla Este	11388	Bellavista	14309
Remedios	18682	Camas	9419
Tomares	14906		

Tabla 4.6: Renta per cápita por distrito en Sevilla.

realiza una representación de esta disparidad, asumiendo que la renta de los pasajeros que usan una estación se corresponde con la renta media del distrito en el que la estación se encuentra.

En este caso de estudio, se pretende analizar cómo cambia la topología de la red si en el diseño de la misma se priman a zonas económicamente más deprimidas. La metodología propuesta para llevar a cabo este estudio supone: 1) postular la existencia de un coeficiente asociado a cada mercado $(o, d) \in \mathcal{K}$, 2) establecer un mecanismo que haga que los pasajeros de zonas deprimidas tengan un coeficiente más alto y 3) formular un nuevo problema de optimización en el que los términos de la función objetivo asociados a la captación de la demanda (3.9)-(3.11) de cada mercado $(o, d) \in \mathcal{K}$ estén multiplicados por los coeficientes obtenidos en el paso anterior. Para llevar a cabo el diseño, se establece lo siguiente:

- Cada distrito tiene asociado un coeficiente entre 1 y 3, estando el valor máximo asociado al distrito más pobre (Camas) y el valor mínimo asociado al más rico (Remedios). Para el resto de distritos se asignan valores entre 1 y 3 de forma proporcional a la diferencia de renta con los distritos extremos.
- Cada estación i tiene como coeficiente multiplicador de demanda (denotado como η_i) el coeficiente del distrito en el que esté ubicada.
- Cada mercado $(o, d) \in \mathcal{K}$ tiene como coeficiente (denotado como η_{od}), que se define a partir de los coeficientes de estación mediante la regla $\eta_{od} = \max\{\eta_o, \eta_d\}$.

Debe hacerse notar que la decisión de diseño $\eta_{od} = \max\{\eta_o, \eta_d\}$ promueve soluciones donde la red de metro se expande hacia los barrios con menor nivel económico (sin importar si dicho barrio es origen o destino del flujo), facilitando la comunicación de estos distritos con el resto de la ciudad, para así ayudar a un mayor desarrollo social en estas áreas. Por último, antes de pasar a la presentación y análisis de resultados, conviene señalar que el mecanismo de sobreponderación de demanda propuesto es una solución particular para dar mayor prioridad a la construcción de infraestructura en zonas deprimidas. No obstante, otras muchas aproximaciones son posibles y, en el futuro, se estudiarán otras alternativas.

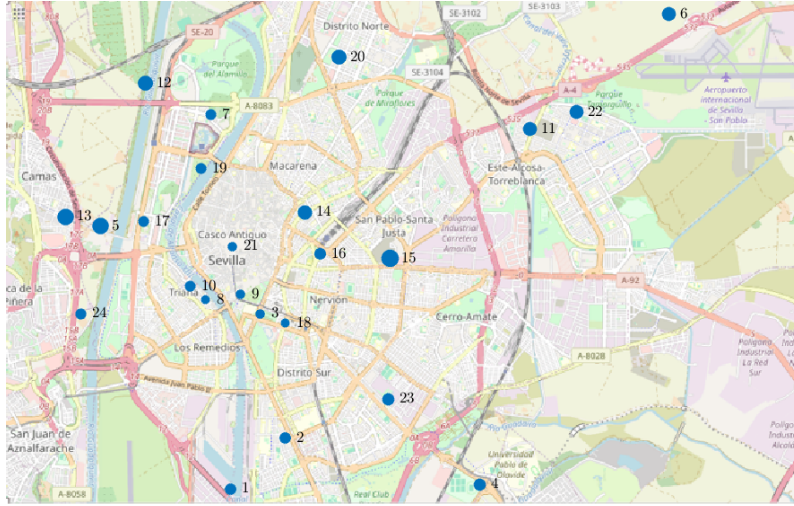


Figura 4.6: Representación geográfica de los nodos potenciales de la red de metro. El tamaño de los nodos hace referencia al coeficiente asociado a los mismos según la renta per cápita media del distrito al que pertenecen (a mayor tamaño, menor renta).

La metodología de sobreponderación de zonas deprimidas se aplica a dos escenarios: uno donde $\gamma = 3.1$ (menor nivel de despliegue) y otro donde $\gamma = 3.3$ (mayor nivel de despliegue). En ambos escenarios se fija $\lambda = 10$, siendo también iguales los valores para el resto de parámetros y variables de estado. Los resultados obtenidos para estas dos configuraciones se reportan en la tabla 4.7 y la figura 4.7. Los resultados confirman que para menores valores de γ el despliegue de la red (capacidad, presupuesto y demanda atraída) es mayor y que el tiempo de cómputo es similar. Resulta especialmente interesante comparar los resultados con los obtenidos para el caso base original (sección 4.4.1), donde no se primaban los pasajeros de las zonas deprimidas. Estas diferencias se aprecian mejor en la figura 4.7. Concretamente, en la gráfica superior de dicha figura se observa que la red tiende a conectar los nodos 22 (Sevilla Este), 13 y 5 (ambos pertenecientes a Camas) con el centro de la ciudad, siendo estos dos de los distritos con menor nivel de renta. Comparando la gráfica superior de la figura 4.7 con la gráfica superior de la figura 4.4 y revisando el presupuesto utilizado en las tablas 4.4 y 4.7, se comprueba como, para valores similares de presupuesto, la red en este caso de estudio es más descentralizada, buscando atender zonas deprimidas. Esto contrasta con el comportamiento del caso base, donde para valores intermedios de presupuesto el foco se centraba en absorber la demanda de la mayoría de nodos situados en el centro de la ciudad. En la gráfica inferior de la figura 4.7 se representa la red obtenida para un mayor nivel de despliegue (un mayor γ , lo que se traduce en una mayor capacidad, demanda y presupuesto). En este caso, cabe destacar el mayor desarrollo de la red en Sevilla Este, conectando el nodo 11 además del 6 (Aeropuerto - Sevilla Este), así como la inclusión del nodo 20 (Pino Montano - Norte), otro de los barrios más deprimidos de la ciudad. Comparando la gráfica inferior de la figura 4.7 con la gráfica central de la figura 4.4, se puede observar cómo, al tener en cuenta la renta de los distritos, se incluye antes al nodo 20 (Pino Montano - Norte) que a nodos procedentes de distritos con mayor nivel de renta, como pueden ser 24 (Tomares), 2 (Sur) o 7 (Cartuja - Triana).

γ	D.A. [%]	Pres. [€]	Arcos	Nodos	Cap.	$\bar{t}_{PAX}$	$\bar{t}_{PAX_{ext}}$	$\bar{t}_{ruta}$	t_{comp} [s]
3.1	14.43	2.89×10^5	36	14	298.73	14.78	19.96	14.94	3240
3.3	28.09	4.56×10^5	56	19	666.39	17.94	19.52	19.24	2189

Tabla 4.7: Métricas reportadas para las simulaciones priorizando demanda según nivel de renta para un valor de $\lambda = 10$.

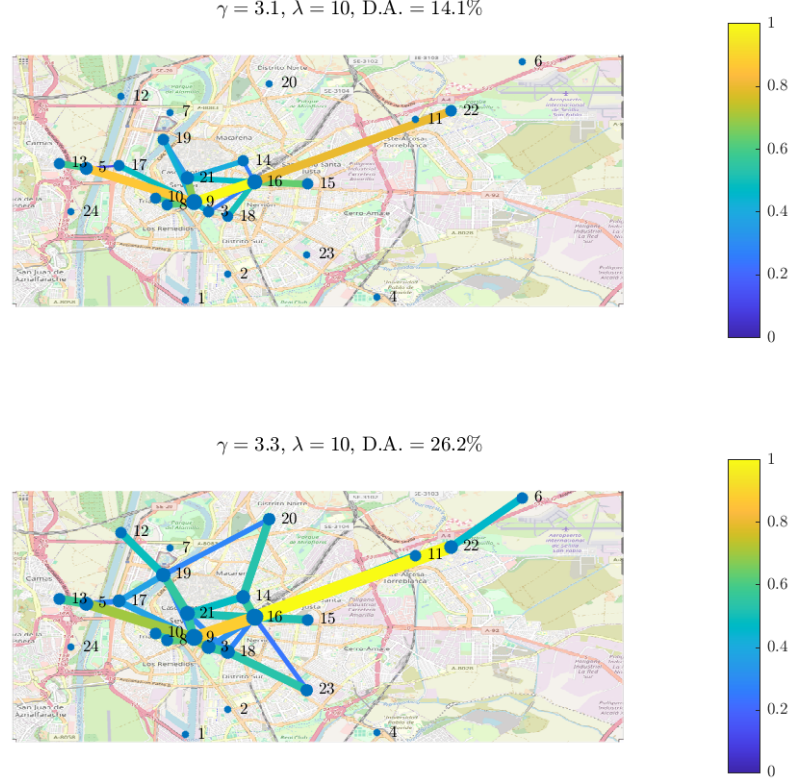


Figura 4.7: Topología de la red de metro para $\lambda = 10$ y $\gamma = \{3.1, 3.3\}$, priorizando la demanda procedente de barrios con menor nivel de renta. El tamaño de los nodos hace referencia a la capacidad construida en cada estación y el grosor y color de los enlaces se representa en función de la capacidad diseñada para los mismos. Los colores se han normalizado para que el amarillo se corresponda con el enlace de mayor capacidad (valor normalizado unitario).

4.4.3. Crecimiento de la demanda y adaptación futura

En esta subsección, en lugar de trabajar con los datos de demanda obtenidos de [27], se ha decidido hacer una predicción demográfica a medio plazo, observando la tasa de crecimiento para cada distrito y las promociones de viviendas de obra nueva existentes. De esta forma, se puede adaptar la construcción de la red de metro a la demanda esperada en el futuro. Los resultados de dicha predicción permiten calcular los coeficientes de crecimiento de cada distrito en función de las promociones de obra nueva en el mismo (véase tabla 4.8). Estos coeficientes toman un valor entre 0.9 y 1.6. Con esta predicción, la demanda futura para un par origen-destino (de forma genérica, el par $(o, d) \in \mathcal{K}$) se obtiene mediante un sencillo modelo multiplicativo, donde la nueva demanda es igual a la demanda original multiplicada por el producto entre el coeficiente de crecimiento asociado al distrito donde se encuentre el origen o y el coeficiente asociado al distrito donde se encuentre el destino d . Al igual que se comentó en el caso de estudio anterior, este mecanismo de asignación de coeficientes es una aproximación particular, cuyo fin último es ilustrar los beneficios de la metodología de diseño de RdT presentada, siendo importante tener en cuenta que existen muchas otras formas alternativas para predecir el crecimiento de la demanda y evaluar su impacto en el diseño de la red.

Distrito	Promociones de obra nueva	Distrito	Promociones de obra nueva
Centro	10	Macarena	0
Nervión	16	Cerro Amate	11
Sur	0	Triana	6
San Jerónimo - Norte	5	Santa Justa	0
Sevilla Este	5	Bellavista	21
Remedios	0	Camas	6
Tomares	4		

Tabla 4.8: Promociones de viviendas de obra nueva por distrito en Sevilla.

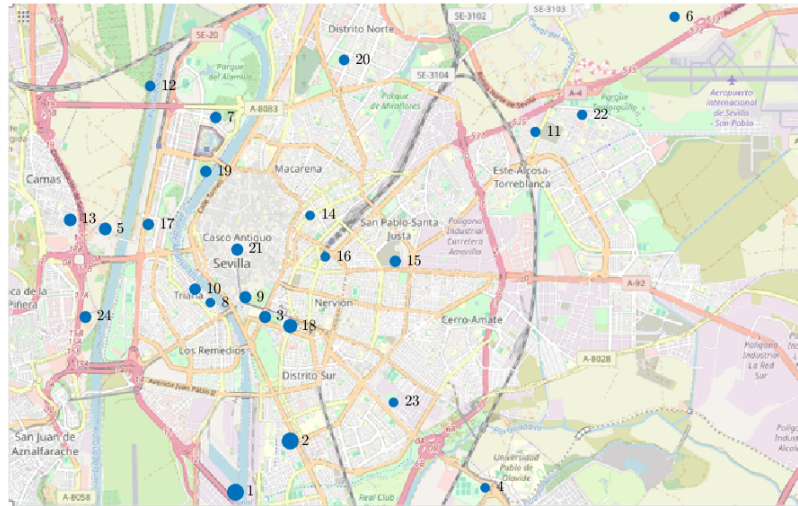


Figura 4.8: Representación geográfica de los nodos potenciales de la red de metro. El tamaño de los nodos hace referencia al coeficiente asociado a los mismos según el crecimiento demográfico esperado del distrito al que pertenecen.

Como en la subsección anterior, se simulan dos escenarios, cada uno con un despliegue de la red (valor del parámetro γ) distinto. En ambos escenarios se establece $\lambda = 10$, siendo también iguales los valores para el resto de parámetros y variables de estado. Los resultados obtenidos se resumen en la tabla 4.9 (métricas más relevantes) y en las dos gráficas de la figura 4.9. Las conclusiones más significativas se obtienen al comparar la topología de las redes construidas con la topología de la red para el caso base. Concretamente, comparando la gráfica superior de la figura 4.9 con la gráfica central de la figura 4.4, se observa cómo la red de metro en este caso se extiende hacia los barrios de Bellavista y Cerro Amate, puesto que estas son las zonas con una mayor proyección de crecimiento de población. Esto también es visible, aunque de manera más sutil, al comparar la gráfica inferior de la figura 4.9 con la gráfica inferior de la figura 4.4, dado que los nodos situados en la periferia, como pueden ser 20 (Pino Montano) o 2 (Bellavista) presentan un mayor nivel de conexión que en la solución presentada en el escenario base.

γ	D.A. [%]	Pres. [€]	Arcos	Nodos	Cap.	t_{PAX}	$t_{PAX_{ext}}$	t_{ruta}	t_{comp} [s]
5.9	29.99	4.77×10^5	62	19	812.21	15.54	20.34	16.04	2123
6.7	43.47	6.50×10^5	80	24	1274.13	18.70	18.96	20.84	2334

Tabla 4.9: Métricas reportadas para las simulaciones considerando crecimiento demográfico con un valor de $\lambda = 10$.

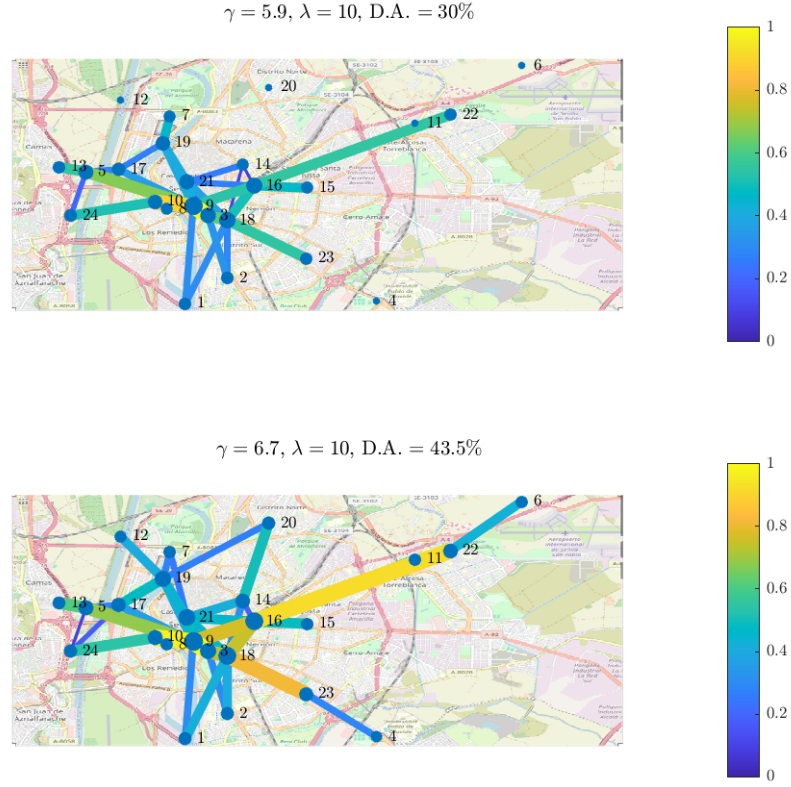


Figura 4.9: Topología de la red de metro para $\lambda = 10$ y $\gamma = [5.9, 6.7]$, modificando la demanda según previsiones demográficas. El tamaño de los nodos hace referencia a la capacidad construida en cada estación y el grosor y color de los enlaces se representa en función de la capacidad diseñada para los mismos. Los colores se han normalizado para que el amarillo se corresponda con el enlace de mayor capacidad (valor normalizado unitario).

Capítulo 5

Optimización de una RdT aéreo con operador privado

En este capítulo, la metodología de optimización convexa de RdT propuesta en este TFG se aplica al caso de un operador comercial de transporte aéreo de pasajeros. Para ello, se utilizan datos del mercado doméstico de EE.UU., incluyendo un listado de 25 potenciales aeropuertos y la demanda de tráfico entre ellos. El diseño en este caso consiste en decidir en qué aeropuertos se ofertará servicio, los aeropuertos que ejercerán de *hub*, así como las capacidades (frecuencias) de las conexiones entre los distintos aeropuertos.

Replicando la estructura del capítulo anterior, en la sección 5.1 se introduce el estado del arte en el diseño de RdT aéreas, dando una explicación detallada de las redes *hub & spoke*, y mencionando las metodologías que se han llevado a cabo en diversos estudios. La sección 5.2 se dedica a describir las características principales de la red de trabajo. La sección 5.3 detalla las modificaciones necesarias en la formulación de acuerdo con el estado del arte. Entre los aspectos más relevantes en este caso se incluyen la consideración de los ingresos del operador en la función objetivo (recuérdese que, en este caso, se trata de un operador privado), así como las modificaciones en las restricciones de rentabilidad y capacidad asociadas a la consideración de *hubs*. Finalmente, la sección 5.4 presenta y analiza los resultados. Al igual que en los capítulos anteriores, se consideran distintos escenarios. La primera parte del análisis se centra en el impacto que los costes fijos y las restricciones presupuestarias tienen en el dimensionamiento de la red. Un aspecto crucial en este capítulo es el número y localización de *hubs* establecidos, así como su relación con los modelos de demanda. Para ello, se considerará la presencia de competidores, así como potenciales limitaciones en la capacidad disponible para desplegar la red. En la segunda parte del análisis se estudia el impacto que distintas evoluciones económicas y modelos de precios (dependientes de la distancia y demanda) puedan tener en el diseño, dimensionamiento y viabilidad comercial de la red. Los resultados obtenidos revelan que, una vez más, los algoritmos propuestos en este TFG dan lugar a redes eficientes, tienen capacidad de adaptarse a cambios en los patrones de demanda o limitaciones de la red y son capaces de gestionar (optimizar) redes de tamaño realista con tiempos de optimización de varios órdenes de magnitud menores que sus equivalentes MIP.

5.1. Estado del arte

Los aspectos más innovadores abordados en este capítulo en comparación con los anteriores incluyen la inclusión de un operador privado y la exploración de una configuración en *hub & spoke*. En esta sección, se lleva a cabo una revisión histórica detallada de estos dos elementos, dedicando especial énfasis al último, el cual representa un desafío significativo en términos de optimización.

En las últimas décadas, el notable crecimiento de los sistemas de transporte y la intensa competencia en el sector han llevado a la exploración de diversas estrategias cooperativas, incluyendo alianzas y

coaliciones, sobre las cuales se han realizado numerosos estudios. En este contexto, las redes basadas en estructuras de concentrador y radio (*i.e.*, *hub & spoke*) han recibido especial atención, ofreciendo oportunidades para compartir capacidad de forma eficiente y gestionar flotas en diferentes tramos de las rutas de transporte [62].

Las redes *hub & spoke* se utilizan ampliamente en diversos sistemas de transporte para optimizar el envío de mercancías o pasajeros entre múltiples puntos de origen y destino, pero destaca su proliferación en el sector aéreo. Un elemento distintivo de estas redes es la presencia de nodos de transbordo o consolidación (*i.e.*, *hubs*), que permiten conectar numerosos pares origen/destino utilizando solo unos pocos enlaces. Este tipo de red hace más fácil agrupar mercancías o pasajeros que salen de orígenes similares pero que se dirigen a destinos diferentes. Esto ayuda a centralizar las operaciones de gestión y clasificación, reducir costes y aprovechar las ventajas de manejar grandes volúmenes de tráfico. El objetivo principal en una red de este tipo es minimizar los costes totales, maximizar la utilización de la capacidad y mejorar el nivel de servicio, siendo estas herramientas cruciales en la planificación por parte de las aerolíneas.

Desde una perspectiva estructural, las redes de *hub & spoke* pueden entenderse como redes jerárquicas con dos niveles fundamentales: una red de acceso, que conecta los nodos origen/destino con los *hubs*, y una subred de concentración, que enlaza los *hubs* entre sí. Una vez determinados los *hubs*, los nodos radio (*spoke*) se asignan a ellos para dirigir el flujo a través de la subred de concentración que incluye todos los nodos *hubs* y los enlaces que los conectan. El esquema de asignación puede ser único o múltiple, dependiendo de la aplicación específica. En un esquema de asignación única, cada nodo de radio se asigna a un solo *hub*, mientras que en un esquema de asignación múltiple esta restricción se relaja. Esta estructura de *hub & spoke* evita el viaje directo, lo que conduce a una infrautilización (o, al menos, una utilización muy ocasional) de la capacidad en algunos (o quizás muchos) de los enlaces directos, y favorece la concentración del flujo en los enlaces entre *hubs* para una mejor utilización de las infraestructuras. Como resultado, se pueden aprovechar las economías de escala mediante el uso más eficiente de los enlaces entre *hubs*. En aplicaciones como el transporte aéreo, la estructura de la subred de nivel de *hub* desempeña un papel fundamental en el nivel de servicio ofrecido a los usuarios y es determinante para la sostenibilidad del negocio en entornos muy competitivos [63]. Los problemas de diseño de redes *hub & spoke* (HNDPs, por sus siglas en inglés *Hub & Spoke Network Design Problems*) representan una categoría desafiante de problemas de optimización que involucran dos tipos de decisiones de diseño: la localización de los nodos *hub* y, la activación de diversos enlaces para conectar los puntos de origen/destino y los nodos *hub*. Dada la compleja naturaleza de la interacción entre estos dos tipos de decisiones, los HNDPs fueron inicialmente abordados desde una perspectiva centrada en la localización de los *hubs*. La mayoría de estos problemas se basan en un conjunto de suposiciones que simplifican las decisiones de diseño y enrutamiento, hasta el punto en que quedan completamente determinadas por las decisiones de asignación de nodos origen/destino a los *hubs*. Cuando estas suposiciones simplificadas se flexibilizan, los HNDPs se vinculan más estrechamente con los problemas de diseño de redes multiartículo. De hecho, los HNDPs constituyen una categoría específica de los problemas de diseño de redes multiartículo en la que se consideran las decisiones de localización de nodos. La amplia gama de aplicaciones existentes ha dado lugar a HNDPs con topologías específicas y a modelos más generales que involucran decisiones de diseño en ambos niveles, tanto de *hubs* como de *spokes*, así como decisiones adicionales de selección de nodos. En los HNDPs clásicos, se suelen considerar cuatro suposiciones principales [64]:

- La red de nivel de *hub* es completa.
- El coste por unidad de utilizar conexiones entre *hubs* es menor que el coste de utilizar conexiones de radio.
- No se permiten conexiones directas entre nodos de radio.
- Se cumple la desigualdad triangular en la estructura de costes, y los costes son proporcionales a la distancia.

En cuanto a las publicaciones del estado del arte, uno de los primeros trabajos en el campo es [65], donde se proponen modelos para el problema de encontrar P -hubs entre un conjunto de candidatos y asignar flujos de acuerdo a los hubs designados, minimizando el coste de transporte. Este es conocido como el problema de la P -mediana, que se define como un problema de optimización basado en distancias en el que P infraestructuras deben localizarse y asignarse a los puntos de demanda de manera que: i) cada punto de demanda se asigne a un único hub y ii) la suma de la distancia ponderada entre todos los puntos de demanda y los hubs correspondientes se minimice. Los problemas de localización de hubs se abordan por primera vez en [66, 67]. En el ámbito discreto, el primer trabajo se atribuye a [68], donde se propone un modelo cuadrático para el problema de P -hub mediana de asignación única, mientras que la primera formulación basada en MIP se propone en [69]. Existen múltiples trabajos dedicados a redes discretas. Entre ellos, destacan las revisiones [70, 71] y las referencias incluidas en ellas. Muchos de los trabajos en el estado del arte no consideran limitaciones de capacidad en la red y trabajan con hipótesis de asignación única, como por ejemplo [72] y [73]. El primer modelo para el problema de asignación múltiple se encuentra en [74]. En [75] se presenta una formulación lineal que requiere menos variables y restricciones y en [76] se emplean otras formulaciones que mejoran el esquema de linealización tanto para asignaciones individuales como múltiples.

Tratando de hacer que el número de hubs sea una parte endógena del problema, se puede hacer explícito el coste operativo o considerar un presupuesto limitado para la localización. En [77] se introduce por primera vez el modelado de costes fijos y costes de localización de los nodos hub, mientras que [69] también incorpora costes fijos para los enlaces. En [73] se proponen formulaciones MIP mejoradas donde se consideran costes fijos para los enlaces entre los nodos hubs. En casos donde el número de hubs no está fijado *a priori*, además de los esquemas de asignación única y múltiple, también pueden existir políticas de capacidad. Estos casos se estudian en [69, 78, 79, 80, 81].

5.2. Datos básicos y descripción de la red

Tal y como se menciona en el párrafo introductorio del capítulo, el objetivo del problema de optimización que se aborda es el diseño de una RdT aéreo que opera en el mercado doméstico de EE.UU. Los nodos de la red corresponden a ciudades estadounidenses, cuyo tamaño vendrá dado por la capacidad dimensionada para dichos aeropuertos. Además, se determinará si dichos aeropuertos actúan como hub de la aerolínea o no. Los enlaces indican la existencia de vuelos directos entre dos ciudades y la capacidad de los mismos representa la cantidad de vuelos ofertados o frecuencia.

A continuación, se describen las fuentes de información utilizadas para establecer los distintos parámetros y variables de estado que sirven de entrada al problema de optimización, así como las suposiciones realizadas en el modelo.

- El conjunto de ciudades (aeropuertos) analizadas corresponde a la selección realizada en [82], que incluía los 25 núcleos urbanos principales en el momento de la publicación de su estudio (véase tabla 5.1).
- Los valores de demanda entre origen y destino se han tomado también del estudio realizado por [82], actualizándolos de forma proporcional al crecimiento del tráfico de pasajeros en el mercado doméstico de EE. UU. entre el momento de publicación del estudio y el tercer trimestre de 2023. Los datos utilizados para realizar el cálculo del crecimiento se han obtenido de la base de datos del Bureau of Transportation Statistics (BTS) de EE.UU. [83].
- Los valores de distancia entre aeropuertos se han tomado de las bases de datos geográficas del servicio de cartografía estadounidense [84].
- Para dimensionar la capacidad de la infraestructura, se considera una hora pico típica, es decir, el periodo de 3 horas que concentra el mayor porcentaje de vuelos diarios. Para estimar dicho volumen de carga, se ha realizado un análisis sobre los datos procedentes del aeropuerto de

Madrid-Barajas en el mes de noviembre de 2023 [85], obteniendo un factor de carga máximo del 18 %. Es decir, en el periodo más cargado tienen lugar el 18 % de los vuelos planeados para ese día.

- Se asume un factor de ocupación medio de las aeronaves del 85 % [86], de cara a dotar al problema de mayor realismo, puesto que, debido a los sistemas de gestión de ingresos, en la mayoría de casos es muy improbable vender todos los billetes ofertados.
- Los precios para viajar entre los distintos pares origen-destino se han ajustado de acuerdo con la información obtenida del BTS [83], que desagrega las tarifas por distintos trayectos, considerando tanto billetes de solo ida (que representan en torno al 45 % del total) como la mitad del coste de los billetes de ida y vuelta (que representan en torno al 55 % restante). La base de datos, además, proporciona los precios de distintas tarifas, habiéndose tomado la media entre las múltiples tarifas asociadas a un mismo par origen-destino.
- Los costes de operación de los vuelos se han estimado teniendo en cuenta los costes laborales, el coste de combustible, los seguros, impuestos y las tasas incurridas por la aerolínea. Varios de estos costes están asociados a la distancia, otros al tiempo y otros son costes fijos. Para relacionar los costes asociados a tiempo y distancia, se ha asumido que un vuelo opera a una velocidad media de 450 millas por hora y que se pierden 40 minutos adicionales asociados a las fases de despegue, aproximación y aterrizaje a los aeropuertos de origen y destino. Conforme a los datos suministrados por las aerolíneas con las que se colabora en el GISAT-URJC, este coste se ha fijado a 5000\$ por hora de vuelo.
- Los costes fijos de despliegue (construcción) en un aeropuerto se han tomado teniendo en cuenta la información suministrada por AENA en [87]. Por otra parte, en el caso de una red aérea, los costes fijos de despliegue de un trayecto son relativamente pequeños, estando asociados básicamente a estudios de mercado, disponibilidad de recursos y viabilidad técnica. En los experimentos realizados, dichos costes se han fijado a un 5 % del coste fijo de despliegue en el aeropuerto.
- Los costes variables de despliegue se han determinado mediante la combinación de información de AENA en [87] y las aerolíneas colaboradoras. Estos costes incluyen el pago de tasas de uso de puerta y aparcamiento, así como los gastos relacionados con el personal de tierra y otros costes operativos menores. Según la información proporcionada por el operador aeroportuario y la aerolínea, se ha establecido el coste variable de despliegue en 3500\$ por aeronave en la hora cargada, *i.e.*, la hora punta en lo que respecta a tráfico.
- Finalmente, el análisis que se realiza en este capítulo considera que la red a diseñar no se desplegará de forma inmediata, sino en un horizonte temporal de tres años. Este intervalo permite a la potencial aerolínea negociar los permisos de operación, firmar contratos de compra o alquiler de aeronaves, y contratar a las tripulaciones y al personal de tierra. Por lo tanto, todos los parámetros relacionados con precios, costes y demanda se actualizan teniendo en cuenta el crecimiento económico del producto interior bruto estadounidense previsto por el Fondo Monetario Internacional (FMI) para ese periodo [88].

Comentario relativo a los potenciales errores en las estimaciones consideradas. Todos los valores indicados anteriormente deben interpretarse como estimaciones que presentan distintos niveles de incertidumbre. El diseño de la red puede y debe realizarse considerando distintas variaciones en estos parámetros. En este sentido, debe destacarse que una de las ventajas de la metodología propuesta es que, debido a su baja carga computacional, los algoritmos de optimización pueden ejecutarse en paralelo considerando un amplio rango de escenarios. Esto no solo permite una caracterización más exhaustiva del comportamiento de la red, sino también, si se estimase necesario, un diseño robusto de la misma.

1	Atlanta	6	Cleveland	11	Kansas City	16	Nueva Orleans	21	St. Louis
2	Baltimore	7	Dallas	12	Los Angeles	17	Nueva York	22	San Francisco
3	Boston	8	Denver	13	Memphis	18	Philadelphia	23	Seattle
4	Chicago	9	Detroit	14	Miami	19	Phoenix	24	Tampa
5	Cincinnati	10	Houston	15	Minneapolis	20	Pittsburgh	25	Washington, DC

Tabla 5.1: Enumeración de las ciudades consideradas en el estudio del diseño de la red de una aerolínea. El número que las acompaña indica el índice del nodo en la red.

5.3. Modificaciones en la formulación del problema

Después de introducir la nueva red de interés, es necesario reformular el modelo matemático de optimización presentado en capítulos anteriores para adaptarlo al problema de diseño de red de una aerolínea con estructura *hub & spoke*. Para tal fin, se explican las principales modificaciones en la formulación, para luego presentar el modelo detallado que se utilizará para diseñar la red. Se definen tanto las variables de estado como las de diseño, así como los diversos parámetros utilizados. Una vez que se conocen todos los elementos, se presenta el problema de optimización a resolver, analizando cada término de la función objetivo y las distintas restricciones.

5.3.1. Motivación de las modificaciones

Aunque los cambios de la formulación se detallan en las siguientes páginas, se proporcionan ahora una descripción de los dos principales cambios:

1. Dado que se ataca el diseño de la red de operación de una aerolínea privada, el objetivo del operador consiste en priorizar los mercados con mayor rentabilidad. Para ello, en la función objetivo se multiplican las variables de flujo asociadas a cada mercado por el precio del billete asociado al mismo. Así, se tratará de minimizar la expresión

$$- \sum_{\forall (o,d) \in \mathcal{K}} w^{od} p^{od} (\omega_p p^{od} F^{od} + \sum_{\forall (i,j) \in \mathcal{A}} \omega_t t_{ij} F_{ij}^{od}), \quad (5.1)$$

donde recordemos que w^{od} es la demanda nominal del mercado $(o, d) \in \mathcal{K}$ y p^{od} el precio del billete. En la expresión anterior, ω_t y ω_p son los coeficientes asociados al tiempo de viaje t_{ij} y al precio del billete p^{od} en la función de utilidad de los pasajeros. De esta forma, el operador prioriza maximizar la función de utilidad de los pasajeros pertenecientes a mercados más rentables, puesto que tendrá un mayor impacto en el valor objetivo. Sin embargo, la expresión (5.1) no recoge todos los términos de la función objetivo en los que aparecen las variables de flujo. Recordemos que para obtener el reparto *logit* de la demanda, es necesario introducir un término de entropía asociado a la variable de flujo F^{od} en cada mercado $(o, d) \in \mathcal{K}$. Por ello, para que la *logit* dependa únicamente de la función de utilidad de los pasajeros, también es necesario multiplicar los términos de entropía de la variable de flujo F^{od} por el precio p^{od} . Así, el término de entropía que se debe introducir en la función objetivo queda como

$$\sum_{\forall (o,d) \in \mathcal{K}} w^{od} p^{od} F^{od} (\log(F^{od}) - 1). \quad (5.2)$$

2. El segundo cambio hace referencia al hecho de que la red tenga una configuración de *hub & spoke*. La propuesta para modelar esta configuración manteniendo la convexidad del problema requiere una serie de modificaciones. La primera consiste en definir cuatro variables de capacidad por aeropuerto: las capacidades diseñada y operada de tráfico de tránsito (o capacidad de *hub*, denotadas como $S_i'^h$ y S_i^h , respectivamente) y las capacidades diseñada y operada de tráfico

propio (denotadas como S'_i y S_i). Si el aeropuerto i no es un *hub*, entonces $S_i^h = S_i^h = 0$. Una vez definidas estas capacidades, se modifican las restricciones de capacidad del modelo, de forma que, para cada aeropuerto $i \in \mathcal{S}$ ha de cumplirse

$$\sigma_i \sum_{\forall j \in \mathcal{N}_{out}(i)} \sum_{\forall (o,d) | o \neq i \in \mathcal{K}} \frac{w^{od} F_{ij}^{od}}{a_{nom} \tau_{ij}} \leq S_i^h, \quad (5.3)$$

donde σ_i , τ_{ij} y a_{nom} son parámetros que se definen en la sección 5.3.2, mientras que $\mathcal{N}_{out}(i)$ es el conjunto de posibles aeropuertos a los que se puede volar desde i . Así, el tráfico en tránsito que pasa por el nodo i (i.e., tráfico entrante cuyo destino no sea i) ha que ser menor que S_i^h . De esta manera, el nodo i solo puede reencaminar tráfico si se trata de un *hub* y, por lo tanto, si $S_i^h > 0$. El último paso en el diseño consiste en hacer que la red cuente únicamente con unos pocos *hubs*. Esto se consigue utilizando dos niveles de *sparsity* diferentes. Uno para todos los aeropuertos, que penaliza la activación de S'_i y S_i^h conjuntamente, dado por la expresión

$$\sum_{\forall i \in \mathcal{S}} \frac{c_{S_i}(S'_i + S_i^h)}{S_i^{h^{k-1}} + S_i^{h^{k-1}} + \varepsilon_S}, \quad (5.4)$$

donde c_{S_i} es el coste fijo asociado a operar en el aeropuerto $i \in \mathcal{S}$. El otro nivel, que tendrá un coste asociado mayor, solo aplica para los *hubs*, penalizando la activación de S_i^h mediante la expresión

$$\sum_{\forall i \in \mathcal{S}} \frac{c_{H_i} S_i^h}{S_i^{h^{k-1}} + \varepsilon_S}, \quad (5.5)$$

donde c_{H_i} es el coste fijo por establecer el aeropuerto i como *hub* de la aerolínea. Además de su motivación teórica, el hecho de que el coste asociado al establecimiento de un *hub* sea más alto tiene un claro sentido práctico, puesto que es común que las aerolíneas dispongan de recursos adicionales en sus *hubs* (estacionamientos, mantenimiento, personal...). Por último, para modelar mejor los efectos de los *hubs*, también se ha introducido una variable de tiempo de transbordo $t_{ijtrans}^{od}$ en la función de utilidad de los pasajeros, de forma que se modele un comportamiento más realista en el que el pasajero penaliza el vuelo con escala frente al vuelo directo.

Una vez que se han descrito verbalmente los principales cambios en la formulación, es momento de comenzar la definición formal del problema, introduciendo todas las variables y parámetros, para luego presentar el problema de optimización.

5.3.2. Variables de estado

Las variables de estado, que se extraen de las bases de datos listadas en la sección 5.2, se modelan como deterministas (constantes conocidas) a lo largo de una simulación. Las variables de estado consideradas en el problema de transporte aéreo que se presenta en este capítulo son:

- c_{S_i} es el coste fijo asociado a la operación de una aerolínea en el aeropuerto $i \in \mathcal{S}$.
- c_{H_i} es el coste fijo extra asociado a la operación de una aerolínea en el aeropuerto $i \in \mathcal{S}$ si este sirve como *hub* de la aerolínea.
- $c_{A_{ij}}$ es el coste fijo asociado a la operación del vuelo en el enlace $(i, j) \in \mathcal{A}$, relacionado con los estudios de viabilidad de dicho vuelo.
- $\bar{c}_{S_i}$ es el coste por unidad de capacidad dimensionada en el aeropuerto $i \in \mathcal{S}$ por la nueva aerolínea.

- c_{oij} es el coste de operación por unidad de capacidad o frecuencia del vuelo en el enlace $(i, j) \in \mathcal{A}$.
- w^{od} es la demanda diaria esperada entre las ciudades asociadas a los nodos $(o, d) \in \mathcal{K}$.
- p^{od} es el precio promedio que pagan los pasajeros del mercado $(o, d) \in \mathcal{K}$ en la nueva aerolínea por hacer su viaje, independientemente de la ruta seguida.
- p_{ext}^{od} es el precio promedio que pagan los pasajeros del mercado $(o, d) \in \mathcal{K}$ en las aerolíneas competidoras para realizar su viaje.
- t_{ij} es la duración del vuelo entre los aeropuertos situados en los nodos $(i, j) \in \mathcal{A}$.
- $t_{ijtrans}^{od}$ es el tiempo de transbordo asociado al vuelo entre los aeropuertos de los nodos $(i, j) \in \mathcal{A}$ para el mercado $(o, d) \in \mathcal{K}$. Este tiempo es nulo cuando el origen del vuelo i coincide con el origen del mercado o .
- t_{ext}^{od} es el tiempo medio requerido por las aerolíneas competidoras para cubrir la ruta entre los aeropuertos situados en los nodos $(o, d) \in \mathcal{K}$.
- $n_{airlines}$ es el número de aerolíneas que representan la competencia.
- σ_i el porcentaje de vuelos sobre el total diario que opera un aeropuerto en hora punta.
- τ_{ij} es el factor de carga u ocupación medio sobre la capacidad nominal de una aeronave.
- s_{max_i} es la máxima capacidad que se permite desplegar en el aeropuerto $i \in \mathcal{S}$.

5.3.3. Parámetros

El siguiente paso consiste en listar los parámetros utilizados (la mayoría coincidentes con los de capítulos anteriores):

- ω_t es el coeficiente asociado al tiempo de viaje en la función de utilidad de los pasajeros.
- ω_p es el coeficiente asociado al precio del billete en la función de utilidad de los pasajeros.
- $\varepsilon_S, \varepsilon_A$ y ε_H son números muy pequeños.
- h_{S_i} es el coeficiente asociado a la holgura entre la capacidad dimensionada y operada en el aeropuerto del nodo $i \in \mathcal{S}$. Sirve para penalizar la congestión de la nueva aerolínea en dicho aeropuerto.
- $h_{S'_i}$ es el coeficiente asociado a la holgura entre la capacidad máxima en el aeropuerto $i \in \mathcal{S}$ y la capacidad total dimensionada. Sirve para penalizar la congestión de dicho aeropuerto.
- β es el hiperparámetro / multiplicador de Lagrange asociado a la restricción de presupuesto. Si β se incrementa, el presupuesto destinado al despliegue de la red decrecerá.
- λ es el coeficiente que multiplica a los costes fijos asociados al establecimiento de un aeropuerto como *hub* en el nodo $i \in \mathcal{S}$. Sirve para modelar distintos escenarios de coste que se pueden plantear, sobre todo cuando hay incertidumbre en la estimación de la variable de estado a la que acompaña.

5.3.4. Variables de diseño

Las variables de diseño son aquellas cuyos valores se obtienen mediante la solución del modelo de optimización. Al igual que en capítulos anteriores, pueden agruparse, según su naturaleza, en variables de construcción, variables de flujo y variables auxiliares. En el caso de la red aérea que se desea diseñar en este capítulo, la mayor diferencia es que en las variables de construcción asociadas a los aeropuertos deben considerarse dos tipos de variables: uno para modelar la capacidad de tráfico propio y otro para modelar la capacidad de tráfico de tránsito.

Variables de construcción

- $S_i'^h \in \mathbb{R}^{\geq 0}$ es la capacidad construida en el horario de máxima carga en el aeropuerto $i \in \mathcal{S}$ para vuelos en los que el aeropuerto actúa como *hub*. Si i no es un *hub*, entonces $S_i'^h = 0$.
- $S_i' \in \mathbb{R}^{\geq 0}$ indica la capacidad construida en el horario de máxima carga en el aeropuerto $i \in \mathcal{S}$ para vuelos en los que el aeropuerto no actúa como *hub*. Si i no es un *hub*, S_i' representa toda la capacidad de la aerolínea en el aeropuerto. Si i es un *hub*, la capacidad de la aerolínea es la suma de $S_i'^h$ y S_i' .
- $A_{ij} \in \mathbb{R}^{\geq 0}$ es la capacidad o frecuencia con que se opera el enlace entre los aeropuertos del arco $(i, j) \in \mathcal{A}$.

Variables de flujo

- $F_{ext}^{od} \in [0, 1]$ es la proporción de pasajeros del mercado entre las ciudades del par $(o, d) \in \mathcal{K}$ que viaja con alguna de las aerolíneas competidoras.
- $F^{od} \in [0, 1]$ es igual a la proporción de pasajeros del mercado entre las ciudades del par $(o, d) \in \mathcal{K}$ que viaja con la nueva aerolínea.
- $F_{ij}^{od} \in [0, 1]$ indica la proporción de pasajeros del mercado entre las ciudades del par $(o, d) \in \mathcal{K}$ que realiza el vuelo entre los aeropuertos de los nodos $(i, j) \in \mathcal{A}$ con la nueva aerolínea.

Variables auxiliares

- $S_i^h \in \mathbb{R}^{\geq 0}$ es la capacidad operable en el horario de máxima carga en el aeropuerto $i \in \mathcal{S}$ para vuelos en los que el aeropuerto actúa como *hub*. Si i no es un *hub*, entonces $S_i^h = 0$.
- $S_i \in \mathbb{R}^{\geq 0}$ indica la capacidad operable en el horario de máxima carga en el aeropuerto $i \in \mathcal{S}$ para vuelos en los que el aeropuerto no actúa como *hub*. Si i no es un *hub*, S_i representa toda la capacidad de la aerolínea en el aeropuerto. Si i es un *hub*, la capacidad de la aerolínea es la suma de S_i^h y S_i .
- $\Delta S_i^h \in \mathbb{R}^{\geq 0}$ es la diferencia entre la capacidad dimensionada y operada en el aeropuerto $i \in \mathcal{S}$ para vuelos en los que el aeropuerto opera como *hub*. Sirve para modelar la congestión de la aerolínea en el aeropuerto i para este tipo de vuelos.
- $\Delta S_i \in \mathbb{R}^{\geq 0}$ es la diferencia entre la capacidad dimensionada y operada en el aeropuerto $i \in \mathcal{S}$ para vuelos en los que el aeropuerto no opera como *hub*. Sirve para modelar la congestión de la aerolínea en el aeropuerto i para este tipo de vuelos.
- $\Delta S_i' \in \mathbb{R}^{\geq 0}$ es igual a la diferencia entre la capacidad máxima disponible en aeropuerto $i \in \mathcal{S}$ y la capacidad dimensionada por la nueva aerolínea, y servirá para modelar la congestión en los aeropuertos.

Por último, para simplificar la notación, se definen tres conjuntos de variables de diseño:

$$\mathcal{X}_{cons} = \{S'_i\}_{\forall i \in \mathcal{S}} \cup \{S_i^{th}\}_{\forall i \in \mathcal{S}} \cup \{A_{ij}\}_{\forall (i,j) \in \mathcal{A}} \quad (5.6)$$

$$\mathcal{X}_{flow} = \{F_{ext}^{od}\}_{\forall (o,d) \in \mathcal{K}} \cup \{F^{od}\}_{\forall (o,d) \in \mathcal{K}} \cup \{F_{ij}^{od}\}_{\forall ((i,j),(o,d)) \in \mathcal{F}} \quad (5.7)$$

$$\mathcal{X}_{aux} = \{S_i\}_{\forall i \in \mathcal{S}} \cup \{S_i^h\}_{\forall i \in \mathcal{S}} \cup \{\Delta S_i\}_{\forall i \in \mathcal{S}} \cup \{\Delta S_i^h\}_{\forall i \in \mathcal{S}} \cup \{\Delta S'_i\}_{\forall i \in \mathcal{S}}. \quad (5.8)$$

5.3.5. Formulación del modelo

En este capítulo la formulación del modelo de optimización debe centrarse en la meta (propósito) del operador, que es fundamentalmente económico, sin términos que busquen satisfacer métricas sociales o medioambientales. Desde ese punto de vista, los términos principales de la función objetivo son los ingresos, provenientes de la compra de billetes por parte de los pasajeros y los gastos de operación de rutas. También deberán tenerse en cuenta los gastos de despliegue. En este caso, el valor del parámetro β puede entenderse como una limitación de presupuesto inicial o, alternativamente, como un coste de financiación. Finalmente, desde un punto de vista técnico, en la función objetivo del problema, además de los términos ya mencionados, deben incluirse las funciones de entropía sobre los flujos de pasajeros. Recordemos que estas funciones no tienen que entenderse como un objetivo *per se*, sino como un mecanismo para garantizar que los usuarios se comportan de acuerdo a un modelo de reparto *logit* de la demanda. Teniendo en cuenta todo lo anterior, se describe a continuación de forma más rigurosa los distintos términos que componen la función objetivo:

- **Costes de operación:** este término recoge los costes asociados a la operación de vuelos, incluyendo personal, combustible, o alquiler de servicios de *handling* en los aeropuertos de origen y destino. Aunque la metodología propuesta podría manejar cualquier función convexa, se asume por sencillez que se modelan de forma lineal como

$$\sum_{\forall (i,j) \in \mathcal{A}} c_{o_{ij}} A_{ij}. \quad (5.9)$$

- **Costes de despliegue:** son equivalentes a los costes de construcción de los problemas formulados en anteriores capítulos. Modelan los costes asociados al despliegue de infraestructura por parte de la nueva aerolínea en los aeropuertos en los que desea operar. Estos costes incluyen, por ejemplo el alquiler de hangares, el alquiler de mostradores o la obtención de permisos de operación. Algunos de estos costes son fijos, mientras que otros dependen de la capacidad dimensionada en cada aeropuerto. Además, también se incluyen los costes asociados a los estudios de viabilidad para operar vuelos entre aeropuertos. Es importante distinguir el coste fijo asociado a la operación de la aerolínea en un aeropuerto, modelado por c_{S_i} , del coste adicional que conlleva establecer como *hub* dicho aeropuerto, modelado por c_{H_i} . Este coste adicional se debe a que, para poder establecer un *hub*, la infraestructura necesaria es mayor (*e.g.*, es necesario disponer de hangares y personal para poder llevar a cabo tareas de mantenimiento en las aeronaves). Estos costes se modelan como

$$\begin{aligned} \beta \Big(\sum_{\forall (i,j) \in \mathcal{A}} \frac{c_{A_{ij}} A_{ij}}{A_{ij}^{k-1} + \varepsilon_A} + \sum_{\forall i \in \mathcal{S}} \frac{c_{S_i} (S'_i + S_i^{th})}{S_i'^{k-1} + S_i^{th^{k-1}} + \varepsilon_S} \\ + \sum_{\forall i \in \mathcal{S}} \lambda \frac{c_{H_i} S_i^{th}}{S_i^{th^{k-1}} + \varepsilon_S} + \sum_{\forall (i,j) \in \mathcal{A}} \bar{c}_{A_{ij}} A_{ij} + \sum_{\forall i \in \mathcal{S}} \bar{c}_{S_i} S'_i \Big). \end{aligned} \quad (5.10)$$

Nótese que, a diferencia de capítulos anteriores, en este caso el parámetro λ , que sirve para modular la importancia de determinados costes fijos, se aplica únicamente al término asociado a la activación de los *hubs*. Esta decisión es consistente con la discusión de la sección 5.3.1, donde se motiva el diseño de redes en el que el número de *hubs* sea reducido. Como se demostrará en

el apartado de simulaciones, el valor del parámetro puede utilizarse por parte del operador para variar el número de *hubs* que se construyen en la red.

- **Costes de los pasajeros:** en este término se introduce el negativo de la función de utilidad de los pasajeros procedentes de todos los mercados $(o, d) \in \mathcal{K}$. En el problema formulado en el capítulo 4 se pretendía maximizar la demanda atraída, por lo que el término de utilidad de los mercados origen-destino estaba ponderado por la demanda total del mercado. Sin embargo, en este capítulo el diseño se realiza para un operador privado y, por lo tanto, el objetivo es maximizar el beneficio. Por este motivo, en este caso el término de utilidad de los mercados no se multiplica únicamente por la demanda, sino también por el precio que pagan los pasajeros de cada mercado origen-destino. Esto prioriza pasajeros de mercados en los que se obtienen mayores ingresos. Por lo tanto, los términos asociados a la utilidad de los pasajeros vienen dados por

$$- \sum_{\forall(o,d) \in \mathcal{K}} w^{od} p^{od} \sum_{\forall(i,j) \in \mathcal{A}} \omega_t(t_{ij} + t_{ijtrans}^{od}) F_{ij}^{od} - \sum_{\forall(o,d) \in \mathcal{K}} w^{od} p^{od} (\omega_p p^{od}) F^{od}. \quad (5.11)$$

Nótese que en la función de utilidad, además de la multiplicación por el precio p^{od} , se ha incluido un nuevo término, $t_{ijtrans}^{od}$, en el tiempo de viaje. Este término sirve para diferenciar si el vuelo es directo o si el pasajero realiza escala. Puesto que se considera que la nueva aerolínea compite con otras ya presentes, es necesario indicarlo en el objetivo. Por lo tanto, hay que introducir en el objetivo la utilidad de las redes competidoras

$$- \sum_{\forall(o,d) \in \mathcal{K}} w^{od} p^{od} (\omega_t t_{ext}^{od} + \omega_p p_{ext}^{od}) F_{ext}^{od}. \quad (5.12)$$

- **Términos asociados al modelo de elección de la demanda tipo *logit*:** la metodología planteada en este TFG para obtener un reparto de la demanda tipo *logit* pasa por introducir en el objetivo funciones de entropía sobre los flujos por las distintas redes, además de la función de utilidad negativa para los pasajeros en todas ellas. Adicionalmente, para obtener un reparto *logit* de la demanda de acuerdo a la utilidad por las distintas redes, es necesario que tanto la utilidad negativa para todas las redes como la entropía vayan ponderadas por los mismos coeficientes. Por este motivo, hay que multiplicar la función de entropía asociada a los flujos de cada mercado $(o, d) \in \mathcal{K}$ por el precio p^{od} que pagan los pasajeros del mismo en la nueva red. Además, hay que añadir los términos de entropía asociados a las redes competidoras. Puesto que en este caso se considera que la competencia no consiste en una sola aerolínea, sino que viene dada por $n_{airlines}$, es necesario indicarlo en el objetivo. Por lo tanto, hay que introducir en el objetivo el término

$$\sum_{\forall(o,d) \in \mathcal{K}} w^{od} p^{od} F^{od} (\log(F^{od}) - 1) + \sum_{\forall(o,d) \in \mathcal{K}} w^{od} p^{od} F_{ext}^{od} \left(\log\left(\frac{1}{n_{airlines}} F_{ext}^{od}\right) - 1 \right). \quad (5.13)$$

- **Términos asociados a la congestión:** de nuevo, se considera la congestión en la nueva aerolínea mediante el modelo hiperbólico presentado en la sección 3.1. Al tratarse de una RdT aéreo, puesto que el número de asientos disponibles en las aeronaves es determinado y no se considera *overbooking*, solamente se modela la congestión en aeropuertos. Esta congestión nace, principalmente, de dos posibles fuentes. La primera tiene lugar si la nueva aerolínea opera en el aeropuerto $i \in \mathcal{S}$ al límite de su capacidad y, por lo tanto, los pasajeros perciben tiempos de espera prolongados, como bien podrían ser largas colas en los mostradores de facturación. Estos fenómenos están modelados por las variables ΔS_i y ΔS_i^h . La segunda es si el aeropuerto $i \in \mathcal{S}$ en el que opera la aerolínea está al límite de su capacidad, habiendo retrasos y congestión en tareas asociadas al aeropuerto, como pueden ser controles de seguridad. Este tipo de congestión se modela mediante la variable $\Delta S_i'$. Así, el término asociado a la congestión que se introduce en el objetivo viene dado por

$$\sum_{\forall i \in \mathcal{S}} \left(\frac{1}{h_{S_i} \Delta S_i + \varepsilon_H} + \frac{1}{h_{S_i} \Delta S_i^h + \varepsilon_H} + \frac{1}{h_{S_i'} \Delta S_i' + \varepsilon_H} \right). \quad (5.14)$$

Teniendo en cuenta todo lo anterior, el modelo matemático propuesto para el diseño de una red aérea con configuración en *hub & spoke* viene definido por la función objetivo en (5.15) y por las restricciones en (5.16)-(5.29).

$$\min_{\mathcal{X}_{cons}, \mathcal{X}_{flow}, \mathcal{X}_{aux}} (5.9) + (5.10) + (5.11) + (5.12) + (5.13) + (5.14) \quad (5.15)$$

$$S'_i = S_i + \Delta S_i \quad \forall i \in \mathcal{S} \quad (5.16)$$

$$S_i^{th} = S_i^h + \Delta S_i^h \quad \forall i \in \mathcal{S} \quad (5.17)$$

$$s_{max_i} \geq S_i^{th} + S'_i + \Delta S'_i \quad \forall i \in \mathcal{S} \quad (5.18)$$

$$\sigma_i \sum_{\forall j \in \mathcal{N}_{in}(i)} A_{ji} \leq S_i + S_i^h \quad \forall i \in \mathcal{S} \quad (5.19)$$

$$\sigma_i \sum_{\forall j \in \mathcal{N}_{out}(i)} A_{ij} \leq S_i + S_i^h \quad \forall i \in \mathcal{S} \quad (5.20)$$

$$\sigma_i \sum_{\forall j \in \mathcal{N}_{out}(i)} \sum_{\forall (o,d) | o \neq i \in \mathcal{K}} \frac{w^{od} F_{ij}^{od}}{a_{nom} \tau_{ij}} \leq S_i^h \quad \forall i \in \mathcal{S} \quad (5.21)$$

$$\sigma_i \sum_{\forall j \in \mathcal{N}_{in}(i)} \sum_{\forall (o,d) | d \neq i \in \mathcal{K}} \frac{w^{od} F_{ji}^{od}}{a_{nom} \tau_{ji}} \leq S_i^h \quad \forall i \in \mathcal{S} \quad (5.22)$$

$$\max(\sigma_o, \sigma_d) \frac{w^{od} F_{od}^{od}}{a_{nom} \tau_{od}} \leq S_o^h + S_d^h \quad \forall (o, d) \in \mathcal{K} \quad (5.23)$$

$$A_{ij} = A_{ji} \quad \forall (i, j) \in \mathcal{A} \quad (5.24)$$

$$\frac{\sum_{\forall (o,d) \in \mathcal{K}} F_{ij}^{od} w^{od}}{a_{nom}} \leq \tau_{ij} A_{ij} \quad \forall (i, j) \in \mathcal{A} \quad (5.25)$$

$$\sum_{\forall (i,j) \in \mathcal{A}} A_{ij} \leq a_{max} \quad (5.26)$$

$$\sum_{\forall j \in \mathcal{N}_{out}(i)} F_{ij}^{od} - \sum_{\forall j \in \mathcal{N}_{in}(i)} F_{ji}^{od} = \begin{cases} F^{od} & i = o, \\ 0 & i \neq o \neq d, \\ -F^{od} & i = d \end{cases} \quad \forall i, (o, d) \in \mathcal{S} \times \mathcal{K} \quad (5.27)$$

$$F^{od} + F_{ext}^{od} = 1 \quad \forall (o, d) \in \mathcal{K} \quad (5.28)$$

$$p^{od} w^{od} F^{od} \geq w^{od} \sum_{\forall (i,j) \in \mathcal{A}} \frac{c_{oij} F_{ij}^{od}}{a_{nom} \tau_{ij}} \quad \forall (o, d) \in \mathcal{K}. \quad (5.29)$$

Las restricciones (5.16)-(5.18) modelan la holgura en el dimensionamiento de la capacidad en aeropuertos, distinguiendo entre *hub* y *spoke*, por lo que su significado es ligeramente distinto a sus análogas en las formulaciones de capítulos anteriores, donde solo se percibe un tipo de congestión en las estaciones. Por una parte, (5.16) iguala ΔS_i a la diferencia entre la capacidad de operación dimensionada para la aerolínea y el flujo que se desea operar realmente en aeropuertos *spoke*, es decir, aeropuertos en los que la aerolínea solo opera vuelos que van a *hubs* o vienen de estos. Por otra parte, en las restricciones (5.17), la variable ΔS_i^h es la holgura entre capacidad dimensionada y flujo operado en aeropuertos que actúan como *hub* de la aerolínea, es decir, aeropuertos que pueden operar vuelos desde o hacia cualquier otro aeropuerto de la red. En cambio, (5.18) representa la diferencia entre la máxima capacidad disponible en el aeropuerto $i \in \mathcal{S}$ y la capacidad que dimensiona la aerolínea. Es decir, las restricciones (5.18) establecen la variable $\Delta S'_i$ como una cota inferior a la capacidad de operación sobrante en el aeropuerto $i \in \mathcal{S}$. En suma, las restricciones (5.16)-(5.18) modelan la congestión en los aeropuertos de la red.

En el modelo *hub & spoke*, si un aeropuerto $i \in \mathcal{S}$ no es designado como *hub*, solo puede servir vuelos que cumplan simultáneamente los dos requisitos siguientes: 1) el origen o el destino del mercado tiene que ser el nodo i y 2) el vuelo debe despegar o aterrizar de un aeropuerto que sea un *hub*. Por tanto, es necesario definir restricciones distintas para la capacidad de *hub* S_i^h y la capacidad S_i . Las restricciones (5.19)-(5.23) relacionan la capacidad en los enlaces con la capacidad del aeropuerto a través del parámetro σ_i , que modela el porcentaje máximo de vuelos diarios que el aeropuerto ha de tener en cuenta simultáneamente. Mientras que (5.19) y (5.20) aseguran que el número total de vuelos no sobrepase la capacidad de los aeropuertos de llegada y salida, respectivamente, (5.21) y (5.22) obligan a que todo el tráfico (número de vuelos) de conexión (se fuerza esto excluyendo el caso $o = i$ y $j = d$ del sumatorio) que sale y entra de un aeropuerto $i \in \mathcal{S}$ sea menor a la capacidad de *hub* de dicho aeropuerto. La restricción (5.23) permite vuelos directos entre el origen y el destino de un mercado $(o, d) \in \mathcal{K}$ únicamente si uno de los aeropuertos del par del mercado es un *hub*.

Las restricciones (5.24) imponen simetría en los vuelos operados por la nueva aerolínea. Nótese que esto refleja una regla de negocio común en las aerolíneas que operan redes *hub & spoke*, en las que un requisito habitual es que las aeronaves pernocten en aeropuertos designados como *hubs*, donde se suele disponer de la infraestructura necesaria para realizar las tareas de mantenimiento. Las restricciones (5.25) relacionan el número de pasajeros que viajan en un enlace (vuelo entre dos aeropuertos) con la capacidad nominal de las aeronaves. La restricción (5.26) impide superar el tamaño de flota disponible en la aerolínea. Las restricciones de conservación de flujo se dan en (5.27), mientras que el grupo de restricciones (5.28) asegura que todos los pasajeros realicen el viaje en alguna de las aerolíneas.

Por último, dado que en este capítulo se aborda el diseño de una RdT por parte de un operador privado, es necesario cerciorarse de operar mercados que resulten beneficiosos. Las restricciones (5.29) aseguran que la nueva aerolínea opere únicamente mercados rentables, es decir, que los ingresos superen a los costes de la ruta.

En las siguientes secciones se resuelve el modelo matemático propuesto para distintos escenarios, analizando las soluciones obtenidas bajo distintas métricas.

5.4. Presentación y análisis de los resultados numéricos

En esta sección se presentan los resultados de las simulaciones realizadas en este capítulo. Se comienza con el análisis detallado de un caso base inicial, que sirve como punto de partida. Se pretende diseñar la red de operación de una nueva aerolínea que entrará al mercado doméstico estadounidense en un horizonte de tres años. En este caso de estudio se consideran 25 nodos (aeropuertos) distribuidos por el país (véase la tabla 5.1 presentada en la sección 5.2), y se fijan unos valores de $\beta = 7$ (relacionado con la limitación presupuestaria) y $\lambda = 10$ (relacionado con los costes de establecimiento de *hubs*). Una vez estudiado este caso, se analiza cómo cambia el diseño de la red frente a distintas modificaciones en los parámetros del modelo, las cuales se listan a continuación.

1. Modificaciones en los costes de despliegue, fijos y variables. En este apartado, se modifican los costes fijos asociados al establecimiento de *hubs*, así como los costes variables de despliegue.
2. Restricción de la capacidad máxima de despliegue en *hubs*. En estos experimentos se estudia el cambio de la solución óptima para el caso en el que haya limitaciones en el dimensionamiento de capacidad de los aeropuertos designados como *hubs*.
3. Restricción del tamaño de flota. En este escenario se limita la cantidad máxima de aeronaves de las que puede disponer la aerolínea, observando cómo afecta esta limitación a la solución obtenida.
4. Consideración de distintos escenarios económicos. En este apartado se analiza el impacto de la incertidumbre económica sobre la solución óptima. Para ello, se realizan simulaciones considerando dos escenarios distintos, generados a partir de las previsiones económicas del FMI.

5. Modificación de patrones de precios. En estos experimentos se desea medir el impacto que tienen en las soluciones proporcionadas por el modelo los cambios en los precios entre pares origen-destino (o, d) .

Los resultados obtenidos se analizan a través de distintas métricas, representadas en tablas y figuras de este capítulo, pudiendo visualizarse tanto la topología de la red obtenida así como el grado de conexión entre aeropuertos. Las métricas reportadas en las tablas son:

- Demanda atendida (“D.A.”): esta métrica hace referencia al porcentaje de pasajeros que la aerolínea es capaz de absorber.
- Presupuesto (“Pres.”): la inversión inicial para el dimensionamiento de la aerolínea es un elemento clave para entender el grado de despliegue de las soluciones propuestas.
- Beneficios (“Ben.”): se corresponden con la diferencia entre los ingresos percibidos por la venta de billetes y el coste operacional.
- Ingreso por asiento ofertado por milla (RASM, por sus siglas en inglés *Revenue per Available Seat and Mile*): esta métrica representa el ingreso por unidad de distancia y asiento ofertado que obtiene la aerolínea. En términos globales se calcula como el cociente entre el total de ingresos percibidos por la venta de billetes y el ASM, donde el ASM es el producto entre el número de asientos ofrecidos por la aerolínea multiplicado por la distancia total volada. De esta forma, el RASM reporta el ingreso por unidad de pasajero y distancia y se calcula como

$$RASM = \frac{\sum_{\forall(o,d) \in \mathcal{K}} p^{od} w^{od} F^{od}}{\sum_{\forall(o,d) \in \mathcal{K}} \sum_{\forall(i,j) \in \mathcal{A}} w^{od} d_{ij} F_{ij}^{od}}. \quad (5.30)$$

- Coste por asiento ofertado por milla (CASM, por sus siglas en inglés *Cost per Available Seat and Mile*): el valor de esta métrica será igual al coste de operación que supone ofrecer un asiento por unidad de distancia. Es decir, el CASM en términos globales es igual al cociente entre el coste de operación total y el producto entre los asientos ofertados y la distancia total volada. Así, el CASM se computa como

$$CASM = \frac{\sum_{\forall(o,d) \in \mathcal{K}} \sum_{\forall(i,j) \in \mathcal{A}} c_{oij} \frac{w^{od} F_{ij}^{od}}{a_{nom}}}{\sum_{\forall(o,d) \in \mathcal{K}} \sum_{\forall(i,j) \in \mathcal{A}} w^{od} d_{ij} F_{ij}^{od}}. \quad (5.31)$$

- Número de enlaces (Arcos): esta métrica reporta la cantidad de conexiones entre aeropuertos, es decir, la cantidad de vuelos diferentes que ofrece la nueva aerolínea.
- Número de aeropuertos (Nodos): es igual al número de aeropuertos en los que la nueva aerolínea opera.
- Número de *hubs* (Hubs): indica el número de aeropuertos elegidos para operar como *hub* en el diseño propuesto para la red.
- Tiempo de computación (t_{comp}): es el tiempo requerido por el *solver* para proporcionar la solución al problema planteado en el escenario correspondiente, medido en segundos. Las simulaciones cuyos resultados se reportan se han lanzado en un ordenador con 64 GB de RAM y un procesador de 8 núcleos a 3 GHz de frecuencia de reloj, utilizando el entorno CVX (versión 2.2) desde Matlab (versión 2023b), utilizando MOSEK (versión 10.1.13) como *solver* de optimización. Así mismo, se ha simulado también el caso base en un ordenador portátil de prestaciones intermedias (MacOS con 1 procesador de 2 núcleos de frecuencia de 2.8 GHz cada uno y 8GB de memoria RAM), incurriendo en un tiempo de cómputo que es, aproximadamente, el doble que el del servidor.

Debido a las limitaciones de espacio, la presentación y el análisis de los resultados en esta sección se centran en los aspectos más relevantes, remitiendo al lector al repositorio de código y a los artículos relacionados para consultar detalles adicionales.

5.4.1. Análisis del caso base

En primer lugar, se presenta la topología de la red construida para un caso base. La configuración exacta de parámetros elegida se muestra en la tabla 5.2.

Parámetro	ω_t	ω_p	$\varepsilon_S = \varepsilon_A = \varepsilon_H$	h_{S_i}	$h_{S'_i}$
Valor	-0.3	-0.03	10^{-4}	0.1	0.1

Tabla 5.2: Parámetros para el diseño de la red para una aerolínea.

En la tabla 5.3 se muestran distintas métricas del modelo matemático de optimización para el caso de estudio propuesto. En la tabla 5.4 se presentan los resultados de demanda atraída, presupuesto, beneficios obtenidos, RASM, CASM, número de enlaces, número de nodos (aeropuertos en los que se opera) y número de *hubs*. Por otro lado, en la figura 5.1 se visualiza la red construida, donde:

- Los aeropuertos construidos se representan mediante círculos sólidos, y el tamaño (área) del círculo refleja la capacidad dimensionada para el mismo ($S'_i + S''_i \forall i \in \mathcal{S}$).
- Los aeropuertos que funcionan como *hubs* se identifican en color rojo.
- Las conexiones entre aeropuertos se representan mediante líneas sólidas, cuyo color y tamaño dependerán de la frecuencia de esa conexión, $A_{ij} \forall (i, j) \in \mathcal{A}$.

Núm. V.C.	Núm. V.B.	Núm. Restr.	Núm. NZE
391375	0	849676	1612921

Tabla 5.3: Métricas del problema de optimización para la red aérea, para el caso concreto de $\beta = 7$ y $\lambda = 10$. Se reporta el número de variables continuas (“Núm. V.C.”), número de variables binarias (“Núm. V.B.”), número de restricciones (“Núm. Restr.”) y número de *non-zero elements* (“Núm. NZE”), entradas distintas de cero en el modelo matemático de optimización.

β	λ	D.A. [%]	Pres. [\$]	Ben. [\$]	RASM	CASM	Arcos	Nodos	Hubs	$t_{\text{comp}}[\text{s}]$
7	10	8.55	2.73×10^6	2.53×10^7	0.25	0.06	180	25	4	3780

Tabla 5.4: Métricas reportadas para la solución obtenida en el caso base, donde $\lambda = 10$ y $\beta = 7$.

En general, los resultados demuestran la efectividad del método propuesto. Se construye una red relativamente *sparse*, con 4 *hubs*, y se asignan más recursos a los pares $(o, d) \in \mathcal{K}$ que resultan más rentables. Además, la metodología de diseño es exitosa: los únicos nodos con más de 8 enlaces son los propios *hubs*, mientras que el resto de los aeropuertos tiene un máximo de 8 enlaces (conexiones desde/hacia los *hubs*). Es destacable que algunos aeropuertos, como Baltimore, no se conectan a todos los *hubs*, lo que se debe a la existencia de rutas no rentables, dado que el precio viene dado en función de la distancia y operar una ruta supone unos costes ineludibles en términos de capacidad aeroportuaria y costes operativos.

En lo que respecta a la selección de *hubs*, se observa que, en general, se han elegido aeropuertos con un alto nivel de demanda, como Nueva York, Los Ángeles, Chicago y San Francisco. Nueva York

destaca como el *hub* con mayor tráfico, y Chicago se posiciona como el que presenta un mayor nivel de tráfico de interconexión, tal como muestra la figura 5.2. La designación de Nueva York como aeropuerto principal era esperada, dado que es el aeropuerto con mayor demanda y, además, tiene la capacidad de ofrecer servicios de interconexión a otros aeropuertos de la costa este. Así mismo, la ubicación geográfica central de Chicago dentro del territorio estadounidense lo convierte en un candidato ideal para vuelos de interconexión. Para muchos trayectos, la distancia recorrida al hacer escala en Chicago no difiere significativamente de la distancia de un vuelo directo. Por ejemplo, Chicago absorbe una proporción considerable de la demanda de vuelos con origen o destino en ciudades como Filadelfia o Baltimore, debido a que la escala no aumenta mucho la distancia de la ruta. En contraste, los *hubs* situados en San Francisco y Los Ángeles presentan un comportamiento opuesto al de Chicago. Apenas redirigen tráfico debido a su posición geográfica, donde destaca la ausencia de otros aeropuertos con menor demanda a su alrededor. A pesar de esto, San Francisco y Los Ángeles se seleccionan como *hub* debido a su alta demanda como origen o destino. Designarlos como *hubs* permite ofrecer vuelos directos con origen o destino en San Francisco y Los Ángeles, que, debido a la considerable distancia que estos vuelos recorren, generan mayores ingresos.

En cuanto a otras métricas, como el RASM y el CASM, cabe destacar la diferencia entre ambas, existiendo un margen de beneficio bastante holgado para la mayoría de mercados entre pares $(o, d) \in \mathcal{K}$. Este es otro de los motivos por los que la solución óptima pasa por construir *hubs* en Los Ángeles y San Francisco, puesto que las rutas más largas tienen un precio mayor asociado, generando así, por lo general, un mayor beneficio, dado el hecho que los costes de operación son de una magnitud menor.

Finalmente, en lo que al tiempo de cómputo se refiere, observamos que el se tarda en torno a 1 hora en encontrar la solución óptima. En comparación, un modelo MIP ejecutándose en el mismo ordenador durante 48 horas no pudo encontrar una solución para la red de trabajo. En cualquier caso, debe señalarse que, si la complejidad computacional fuera la métrica más crítica, el hecho de trabajar con una formulación convexa permite el diseño de algoritmos de descenso adaptados a la estructura del problema, lo que reduciría el tiempo de cómputo de una forma sustancial.

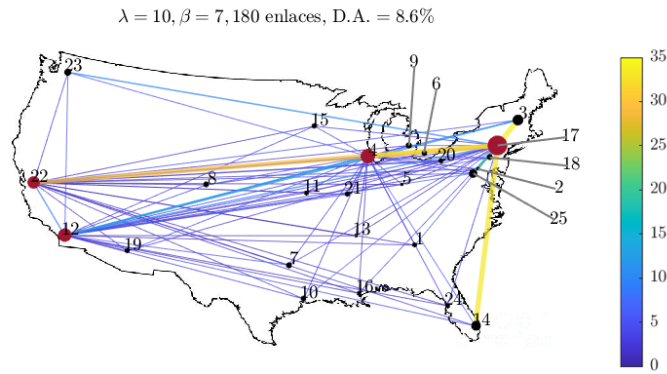


Figura 5.1: Topología de la red para el caso base con $\beta = 7$ y $\lambda = 10$. En color rojo, se distinguen 4 *hubs*, a saber: Nueva York, Chicago, Los Ángeles y San Francisco. El grosor y color de los enlaces son proporcionales a la frecuencia de vuelo asociada.

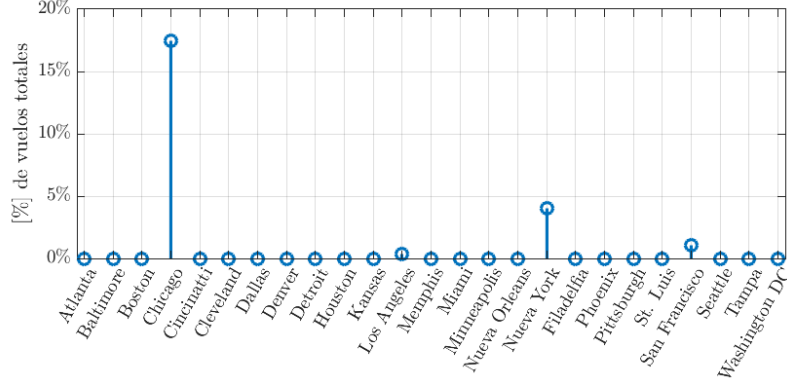


Figura 5.2: Porcentaje de vuelos que hacen escala respecto al total de vuelos en cada aeropuerto. Caso base, con $\beta = 7$ y $\lambda = 10$. Cabe destacar que solo se permiten conexiones en aeropuertos designados como *hub*.

5.4.2. Modificación de costes fijos y variables

Para continuar con la validación de la metodología propuesta, al igual que en capítulos anteriores, se realiza un conjunto de experimentos para analizar la red construida considerando distintos valores de costes fijos y variables. La figura 5.3 muestra nueve gráficas, cada una correspondiente a una configuración de costes diferente. Específicamente, cada fila representa un nivel presupuestario, y cada columna representa un nivel de coste fijo asociado a la construcción de *hubs*. Al igual que en la gráfica anterior, la capacidad construida para cada aeropuerto se representa mediante el área del círculo, y los aeropuertos que funcionan como *hubs* se identifican con color rojo.

Los resultados demuestran que, a medida que se restringe más presupuesto (dado por el parámetro β), se obtienen redes con menos enlaces y aeropuertos. Al analizar los resultados con más detalle, se observa que en la mayoría de los casos se opera en todos los aeropuertos. Por otra parte, analizando la variación de la topología propuesta para diferentes valores de λ , se observa que el número de *hubs* construidos disminuye gradualmente a medida que aumentan los costes fijos de estos. Además, se aprecia cómo la red evoluciona hacia una configuración en estrella con un nodo central en Nueva York. Para más detalle, la tabla 5.5 contiene las métricas definidas anteriormente. Respecto a los beneficios de la aerolínea, cabe destacar cómo, a medida que se construye una red más *sparse*, mejoran las métricas RASM y CASM, es decir, se operan solamente las rutas más rentables respecto a configuraciones con mayor despliegue. Estos resultados no solo proporcionan intuición sobre el comportamiento de la red, sino que también pueden ayudar a la aerolínea a elegir el punto de equilibrio para su despliegue inicial.

β	λ	D.A. [%]	Pres. [\$]	Ben. [\$]	RASM	CASM	Arcos	Nodos	Hubs	t_{comp} [s]
5	5	12.28	3.74×10^6	3.414×10^7	0.2496	0.0619	256	25	6	5203
5	10	12.28	3.89×10^6	3.415×10^7	0.2497	0.0619	256	25	6	3826
5	20	11.24	3.78×10^6	3.198×10^7	0.2503	0.0620	180	25	4	3688
7	5	9.05	2.78×10^6	2.636×10^7	0.2456	0.0612	220	25	5	5330
7	10	8.55	2.74×10^6	2.529×10^7	0.2461	0.0612	180	25	4	3780
7	20	7.05	2.38×10^6	2.163×10^7	0.2509	0.0614	94	25	2	3616
10	5	5.29	1.64×10^6	1.690×10^7	0.2384	0.0599	128	24	3	4930
10	10	4.76	1.55×10^6	1.527×10^7	0.2464	0.0605	86	23	2	3762
10	20	3.55	1.23×10^6	1.165×10^7	0.2631	0.0615	42	23	1	3586

Tabla 5.5: Métricas reportadas para la solución obtenida en distintos escenarios (β, λ) .

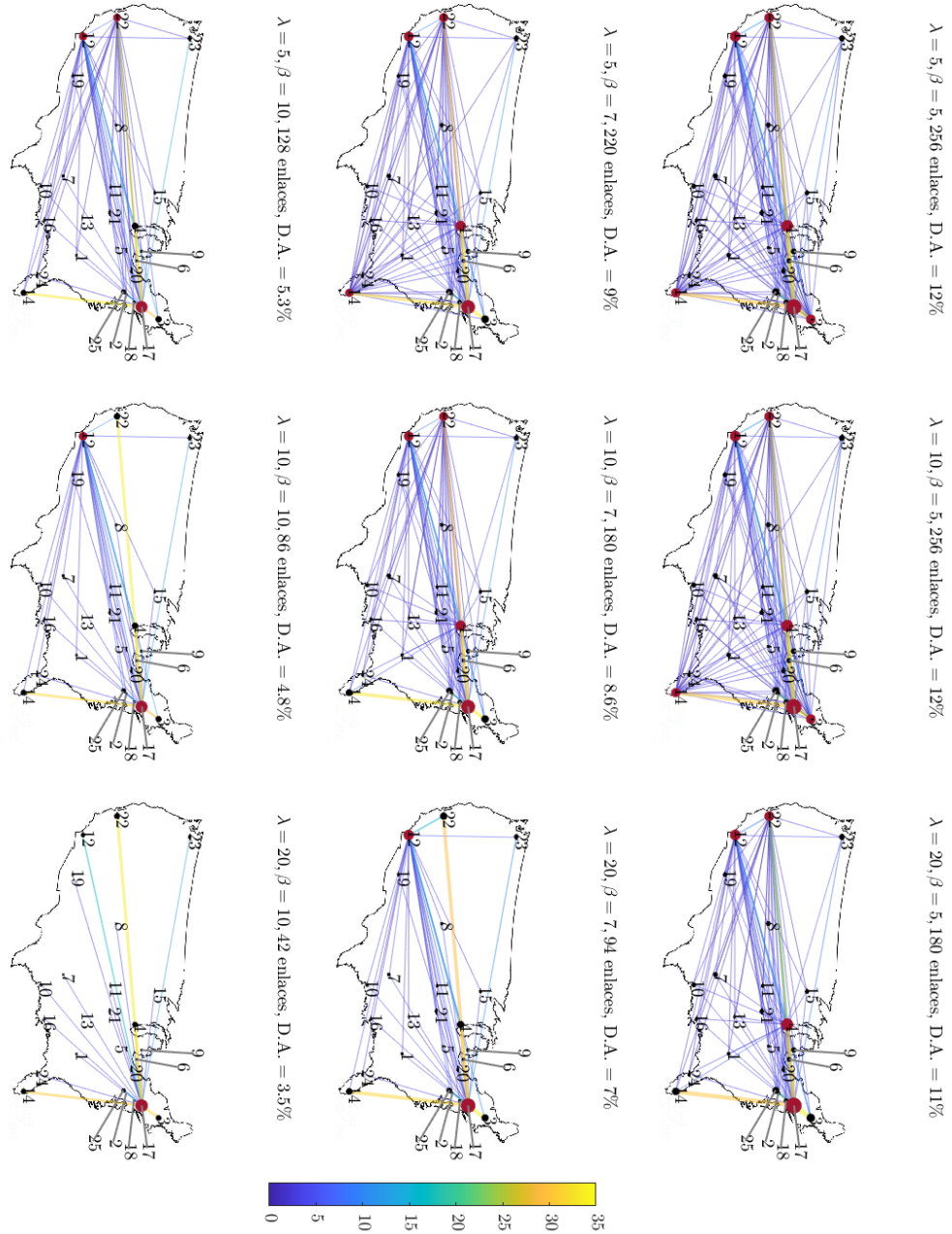


Figura 5.3: Topología de la red para distintos escenarios de costes fijos y presupuestos, representados por los valores de β y λ . En color rojo se denotan los aeropuertos designados como *hub*. El grosor de los enlaces es proporcional a la frecuencia de vuelo asociada, así como el color. Con mayores valores de costes fijos (λ) se designan menos *hubs*, mientras que para mayores valores de β se obtienen configuraciones más *sparse*.

5.4.3. Restricción de la capacidad de *hubs*

En los resultados anteriores, se ha asumido que los aeropuertos no tenían restricciones en la capacidad que podían ofrecer a la nueva aerolínea. Sin embargo, en este caso de estudio, se analiza el impacto en la red diseñada si los aeropuertos pueden ofertar únicamente el 60 % de la capacidad inicialmente demandada, obtenida como resultado de la optimización en el caso base. Esta restricción viene dada por el grupo de restricciones (5.18). Para ello, se resuelve el problema de diseño óptimo de la red para los mismos escenarios de coste fijo y valor de β , pero esta vez imponiendo la capacidad máxima, s_{max_i} , para los tres aeropuertos considerados como *hub* y más utilizados (*i.e.*, con mayor capacidad) en la simulación original, de tal forma que $s_{max_i} = 0.6(S_i^{h*} + S_i^{*})$, siendo $S_i^{h*} + S_i^{*}$ la capacidad

dimensionada en el aeropuerto $i \in \mathcal{S}$ en el caso de resolver el problema sin límite de capacidad. Este análisis permite comprender cómo las limitaciones de capacidad afectan a la estructura de la red y cómo se deben adaptar las estrategias de despliegue de la aerolínea en un escenario más realista.

Los resultados obtenidos se presentan en la figura 5.4 y en la tabla 5.6. Pese a que no hay cambios topológicos significativos, sí es posible apreciar que la capacidad diseñada en los *hubs* se reduce notablemente, disminuyendo también el número de rutas operadas y la demanda atendida. También cabe mencionar que el RASM reporta unos valores más bajos. Debido a que el grupo de restricciones (5.18) es considerablemente más restrictivo en estos experimentos, se encuentra una solución subóptima respecto al problema original, con lo que el presupuesto invertido para un mismo valor de β se reduce. Puesto que el valor de β indica la importancia de los costes de inversión iniciales para la aerolínea, al encontrar una solución de menor beneficio económico, la importancia relativa de los costes de inversión aumenta. Por esta razón, se decide disminuir el valor de β para realizar una comparación más justa, de tal forma que la solución propuesta con la restricción que limita la capacidad en los *hubs* tenga asociada una inversión inicial muy similar a la de los escenarios originales, donde dicha restricción no existe.

Se presentan experimentos para tres escenarios distintos de costes fijos, en los cuales se emplea un presupuesto con una variación máxima del 2 % con respecto a sus análogos de la figura 5.3 y tabla 5.5 para $\beta = 7$. Las soluciones se reportan en la figura 5.5 y en la tabla 5.7. En este caso, sí que es bastante notorio el cambio de topología para todos los escenarios, siempre con la construcción de una red menos *sparse*, con un mayor número de rutas operadas y el establecimiento de más *hubs*, que pasan de ser 5, 4 y 2 a 8, 6 y 4 para $\lambda = 5$, $\lambda = 10$ y $\lambda = 20$, respectivamente. Pese a ello, son soluciones subóptimas (comparativamente respecto al caso base), con lo que los ingresos y la demanda atraída son menores que para los experimentos de la tabla 5.5. Esto se pone de manifiesto al visualizar el RASM obtenido, siendo hasta un 15 % inferior.

β	λ	D.A. [%]	Pres. [\$]	Ben. [\$]	RASM	CASM	Arcos	Nodos	Hubs	$t_{\text{comp}}[\text{s}]$
5	5	8.43	2.63×10^6	2.35×10^7	0.2347	0.0610	254	25	6	2962
5	10	8.45	2.78×10^6	2.36×10^7	0.2352	0.0610	252	25	6	3755
5	20	7.21	2.55×10^6	2.13×10^7	0.2373	0.0607	178	25	4	3609
7	5	6.04	1.91×10^6	1.78×10^7	0.2316	0.0602	216	25	5	3951
7	10	5.49	1.83×10^6	1.67×10^7	0.2326	0.0600	176	25	4	3813
7	20	4.27	1.50×10^6	1.38×10^7	0.2428	0.0602	86	23	2	3664
10	5	3.19	1.02×10^6	1.05×10^7	0.2335	0.0593	120	23	3	3859
10	10	2.86	0.97×10^6	0.96×10^7	0.2375	0.0594	74	22	2	3870
10	20	2.14	0.77×10^6	0.75×10^7	0.2529	0.0601	36	19	1	3752

Tabla 5.6: Métricas reportadas para la solución obtenida limitando la capacidad en los *hubs* principales.

β	λ	D.A. [%]	Pres. [\$]	Ben. [\$]	RASM	CASM	Arcos	Nodos	Hubs	$t_{\text{comp}}[\text{s}]$
3.6	5	8.85	2.84×10^6	2.31×10^7	0.2281	0.0614	324	25	8	5985
3.6	10	8.06	2.71×10^6	2.17×10^7	0.2255	0.0609	254	25	6	5897
5.9	20	6.73	2.44×10^6	1.97×10^7	0.2332	0.0607	168	24	4	6323

Tabla 5.7: Métricas reportadas para la solución obtenida limitando la capacidad en los *hubs* principales, utilizando el mismo presupuesto que en la tabla 5.5.

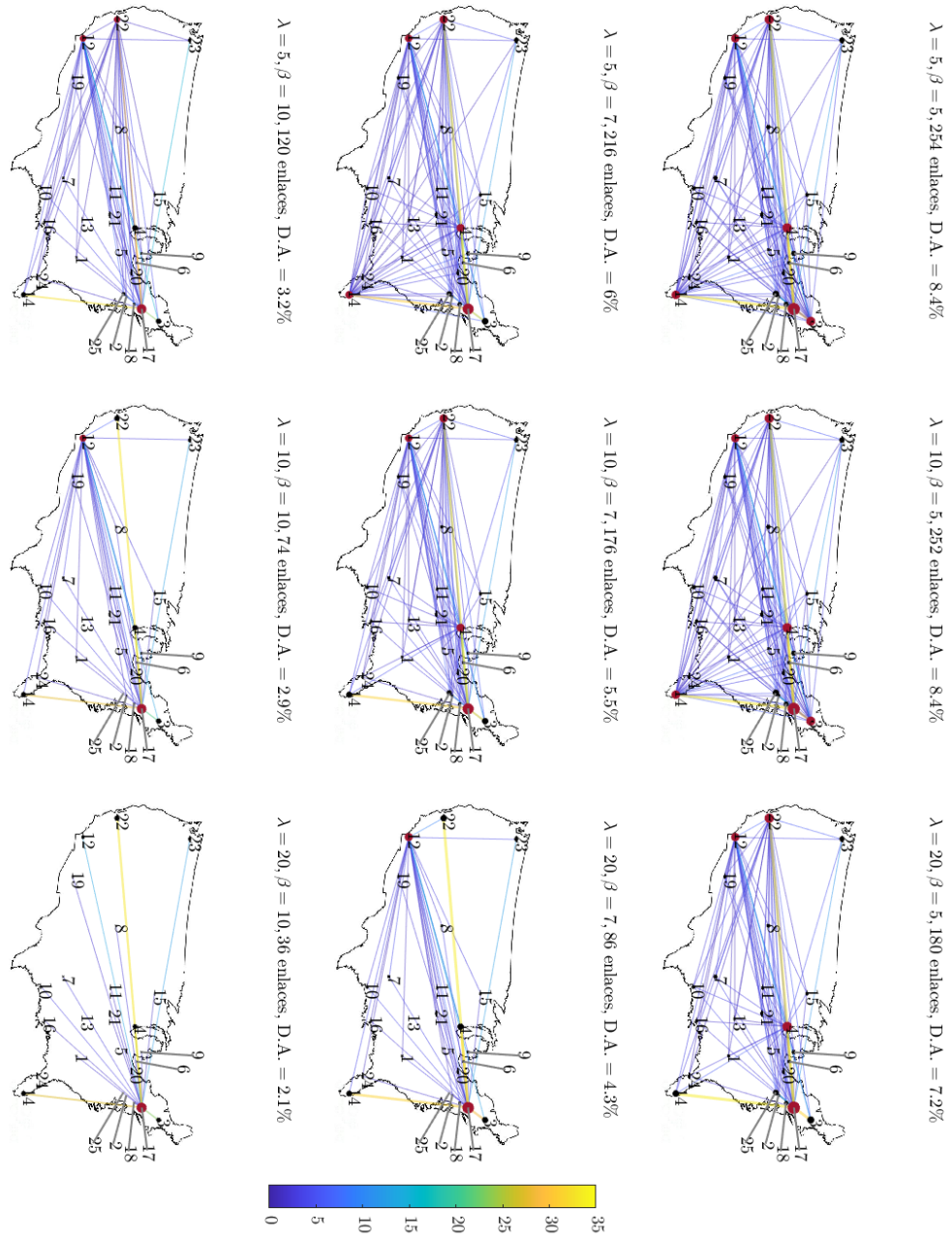


Figura 5.4: Topología de la red para distintos escenarios de costes fijos y presupuestos, limitando la capacidad de los *hubs* al 60 % respecto al resultado original. Nótese como la demanda atraída es menor, al igual que el presupuesto utilizado (compárense los valores de presupuesto en este caso, disponibles en tabla 5.6, con los valores de presupuesto sin limitar la capacidad disponible en los *hubs*, disponibles en la tabla 5.5).

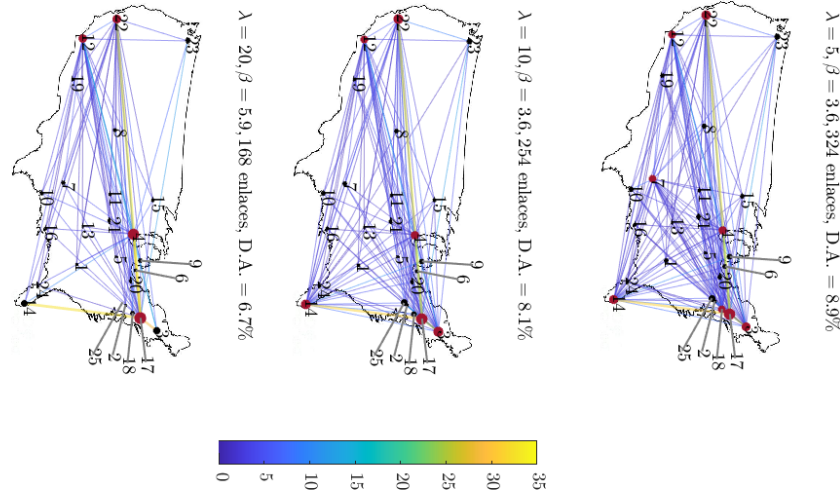


Figura 5.5: Topología de la red para distintos escenarios de costes fijos y presupuestos, limitando la capacidad de los *hubs* al 60 % con el mismo presupuesto que el escenario original. Nótese como el número de *hubs* aumenta considerablemente y, sin embargo, se obtienen menores beneficios que en el caso original, tal y como se observa al comparar las tablas 5.7 (este caso) y 5.5 (caso original).

5.4.4. Restricción del tamaño de flota

β	λ	D.A. [%]	Pres. [\$]	Ben. [\$]	RASM	CASM	Arcos	Nodos	Hubs	t_{comp} [s]
5	5	9.90	3.03×10^6	2.84×10^7	0.2561	0.0613	256	25	6	3775
5	10	9.81	3.13×10^6	2.83×10^7	0.2570	0.0614	220	25	5	3844
5	20	9.12	3.11×10^6	2.67×10^7	0.2572	0.0613	180	25	4	3738
7	5	7.29	2.26×10^6	2.18×10^7	0.2519	0.0606	220	25	5	3786
7	10	6.91	2.24×10^6	2.10×10^7	0.2525	0.0606	180	25	4	3819
7	20	5.73	1.95×10^6	1.80×10^7	0.2585	0.0609	94	25	2	3644
10	5	4.26	1.34×10^6	1.38×10^7	0.2446	0.0595	132	24	3	3839
10	10	3.85	1.27×10^6	1.26×10^7	0.2526	0.0600	90	24	2	3757
10	20	2.88	1.01×10^6	0.96×10^7	0.2712	0.0611	42	22	1	3629

Tabla 5.8: Métricas reportadas para la solución obtenida en distintos escenarios(β, λ), restringiendo el tamaño de flota.

Hasta este punto, se había considerado que la única restricción por parte de la aerolínea era el presupuesto disponible para la inversión, independientemente de la capacidad o tamaño de flota resultante de la optimización. En este caso de estudio, se analiza el impacto que tiene en la solución una potencial limitación en el tamaño de la flota disponible. Dicha limitación puede estar asociada, por ejemplo, a restricciones en el suministro de aeronaves o a la capacidad de contratar tripulaciones que las operen. Desde un punto de vista de modelado, las limitaciones en términos del tamaño de flota se traducen en la restricción dada por (5.26), donde se da una cota superior a la capacidad máxima, a_{max} . Para el análisis numérico, en cada experimento se fija el tamaño de la flota al 80 % del tamaño obtenido en el caso base análogo (mismo valor de β y λ) analizado al principio de esta sección para los distintos escenarios de costes simulados, $a_{max} = \sum_{(i,j) \in \mathcal{A}} A_{ij}^*$.

Los resultados obtenidos se reflejan en la figura 5.6 y la tabla 5.8. Como es de esperar, la limitación en el tamaño de flota da lugar a un diseño de red más selectivo, donde se diseñan menos vuelos y se atrae una menor demanda. Pese a que el beneficio total disminuye, el beneficio por pasajero, visible en la diferencia entre el RASM y CASM (tabla 5.8), reporta mejores resultados respecto a los presentados por la tabla 5.5. Puesto que el tamaño de flota es limitado, la aerolínea decide operar un subconjunto de las rutas operadas en el caso base, manteniéndose únicamente las más rentables.

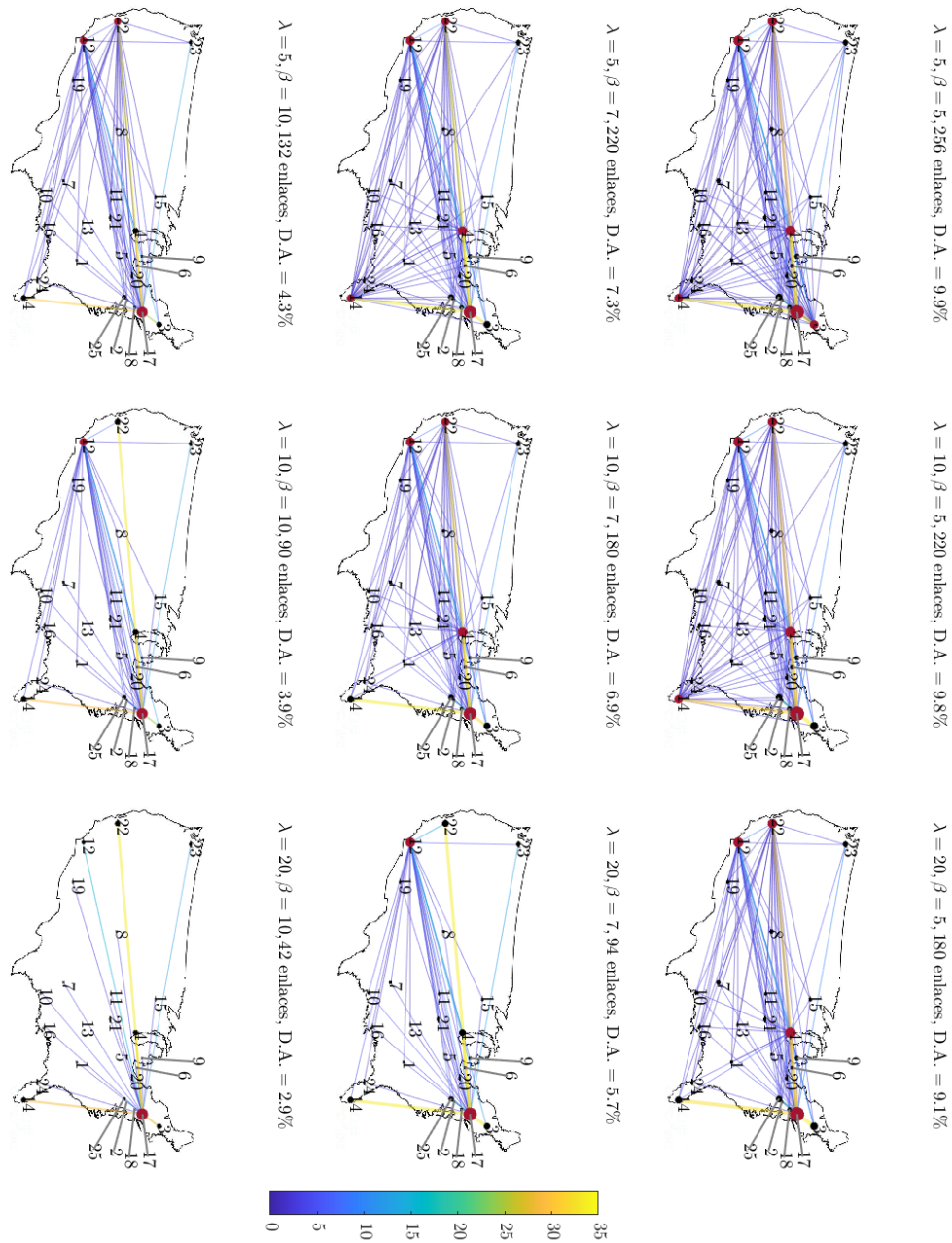


Figura 5.6: Topología de la solución propuesta para distintos escenarios de costes fijos y presupuesto, considerando un tamaño de flota máximo del 80 % respecto al escenario original.

5.4.5. Consideración de distintos escenarios económicos

En este caso de estudio, se analizan los cambios en la red frente a distintos escenarios económicos. Recordemos que la red se había dimensionado para ser desplegada en un horizonte temporal de tres años, teniendo en cuenta las previsiones de crecimiento del FMI. En este punto, se consideran dos escenarios: 1) escenario optimista (en el que el FMI ha infraestimado el crecimiento económico en un 2 % anual) y 2) escenario pesimista (en el que el FMI ha sobreestimado el crecimiento en un 2 % anual). Es importante señalar que las previsiones del FMI suelen ser bastante acertadas, con un error anual menor al 1 % [89]. Por lo tanto, los dos escenarios a analizar recogerían las colas de la distribución en términos de incertidumbre económica.

Para simplificar el análisis, se supone que el crecimiento económico tiene un impacto directo y

proporcional en la demanda. Los costes de construcción se mantienen fijos, ya que se supone que se negocian con antelación al inicio del horizonte temporal y no se ven afectados por cambios futuros.

Los resultados obtenidos para el escenario optimista se detallan en la figura 5.7 y en la tabla 5.9, y para el escenario pesimista en la figura 5.8 y en la tabla 5.10. Si bien la incertidumbre económica acarrea pequeñas modificaciones en la solución, no supone un cambio cualitativo en el diseño de la red. Esto es un efecto esperado, puesto que el diseño inicial solo abría rutas en las que existiese un margen en términos de ingresos frente a costes de operación.

De forma más específica, se aprecia un pequeño cambio en los pasajeros atraídos para el escenario optimista (véase figura 5.7 para $\lambda = 20$ y $\beta = 10$), además de una pequeña disminución en la métrica del RASM, puesto que al operar más rutas, las nuevas rutas generan menores beneficios que las que estaban presentes también en la solución original. No obstante, estas diferencias no son significativas, dado que todas las aerolíneas se ven afectadas por los cambios de indicadores económicos. Como línea de trabajo futura, también se investigará el impacto de contingencias económicas disruptivas, como la pandemia del COVID-19, que afectan de manera heterogénea a precios, costes y demanda.

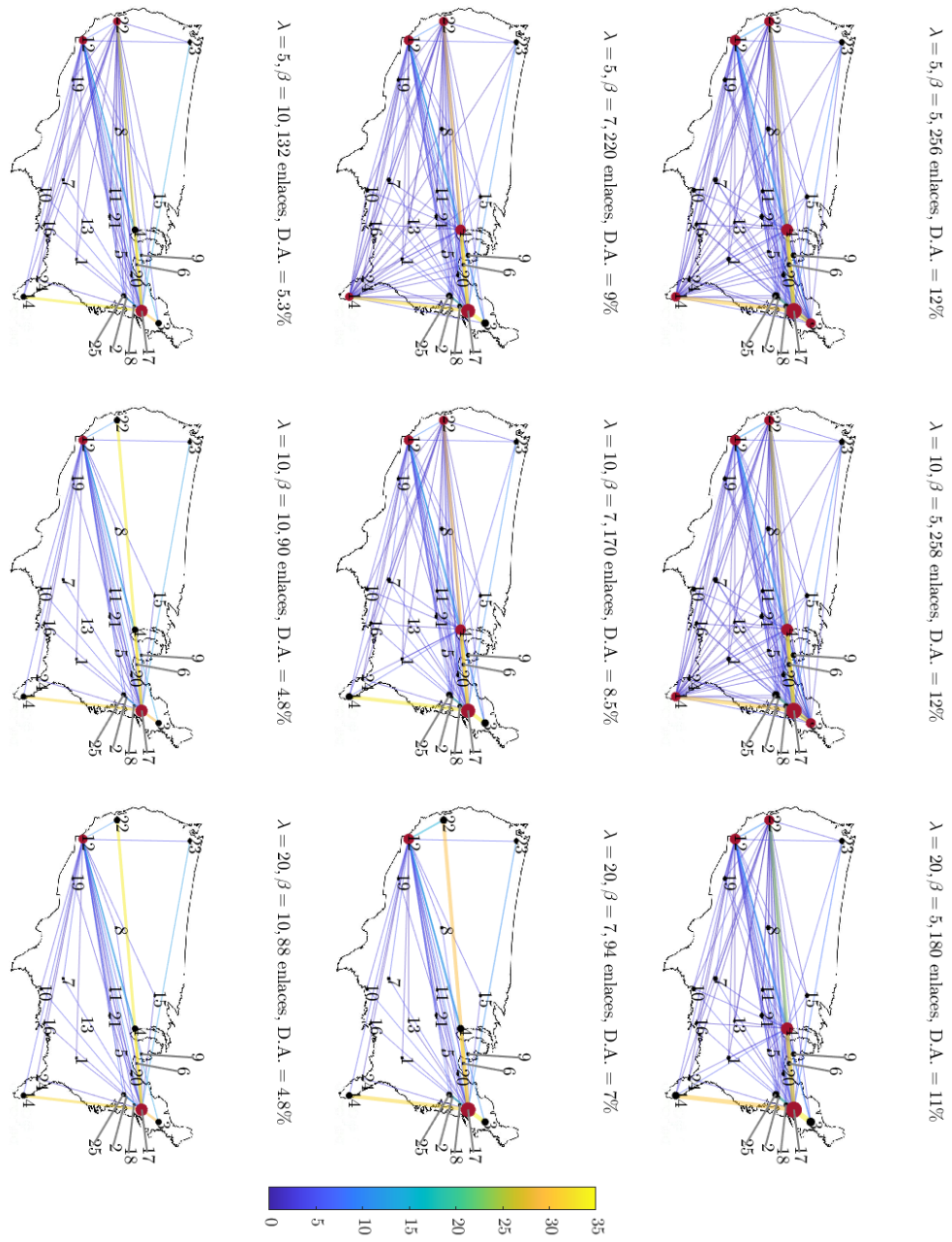


Figura 5.7: Topología de la solución propuesta para distintos escenarios de costes fijos y presupuesto, considerando una previsión económica optimista

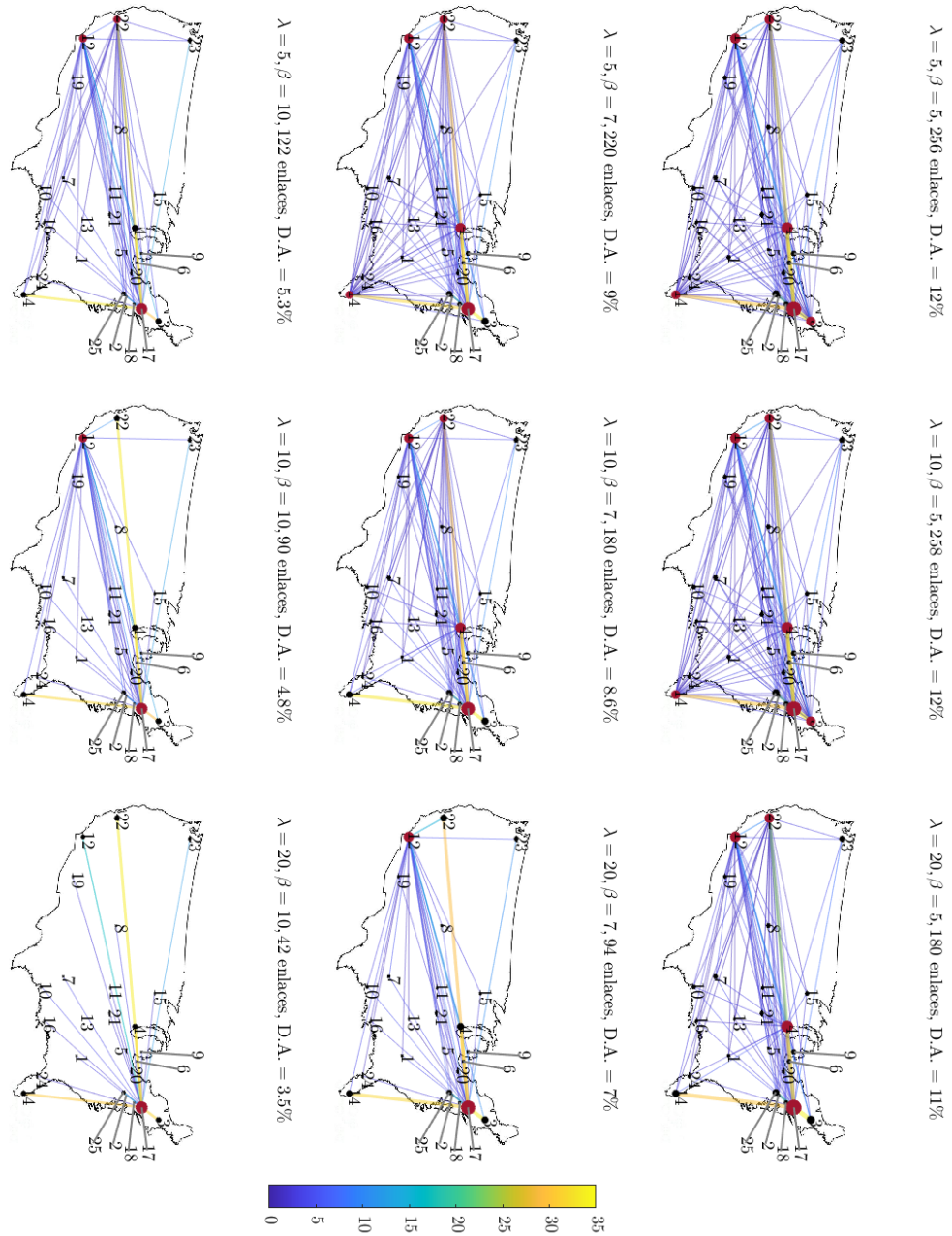


Figura 5.8: Topología de la solución propuesta para distintos escenarios de costes fijos y presupuesto, considerando una previsión económica pesimista.

5.4.6. Modificación de patrones de precios

Para finalizar con el análisis de resultados numéricos, se cambia el modelo de precios para comprobar la sensibilidad de la solución obtenida respecto a esta variable de estado. En este caso, se establece el precio para una ruta en base a la distancia entre los aeropuertos pero sobre todo a la demanda entre dichas ciudades, garantizando que la media de los precios entre mercados sea igual que en los experimentos anteriores. De esta forma, los mercados con un mayor precio asociado serán aquellos entre grandes ciudades, reduciendo considerablemente la rentabilidad de mercados entre ciudades con menor población. Esto lleva a soluciones más *sparse* que favorecen la topología de una red con unos pocos *hubs*, puesto que en ellos se centrará la mayor parte de los beneficios. Los resultados se visualizan en la figura 5.9 y en la tabla 5.11. En esta ocasión, es notorio el aumento de demanda absorbida para un

β	λ	D.A. [%]	Pres. [\$]	Ben. [\$]	RASM	CASM	Arcos	Nodos	Hubs	t_{comp} [s]
5	5	12.28	3.92×10^6	3.59×10^7	0.2496	0.0619	256	25	6	3596
5	10	12.29	4.07×10^6	3.59×10^7	0.2497	0.0619	258	25	6	3847
5	20	11.24	3.95×10^6	3.37×10^7	0.2503	0.0620	180	25	4	3723
7	5	9.05	2.91×10^6	2.77×10^7	0.2456	0.0612	220	25	5	3707
7	10	8.51	2.85×10^6	2.65×10^7	0.2462	0.0611	170	25	4	3849
7	20	7.05	2.49×10^6	2.28×10^7	0.2509	0.0614	94	25	2	3730
10	5	5.31	1.73×10^6	1.78×10^7	0.2384	0.0599	132	24	3	3658
10	10	4.80	1.63×10^6	1.61×10^7	0.2461	0.0605	90	24	2	3816
10	20	4.78	1.73×10^6	1.61×10^7	0.2460	0.0605	88	24	2	3630

Tabla 5.9: Métricas reportadas para la solución obtenida en distintos escenarios, para valores de β y λ con una previsión económica optimista.

β	λ	D.A. [%]	Pres. [\$]	Ben. [\$]	RASM	CASM	Arcos	Nodos	Hubs	t_{comp} [s]
5	5	12.28	3.54×10^6	3.22×10^7	0.2496	0.0619	256	25	6	3803
5	10	12.29	3.69×10^6	3.22×10^7	0.2497	0.0619	258	25	6	3737
5	20	11.24	3.59×10^6	3.02×10^7	0.2503	0.0620	180	25	4	3760
7	5	9.05	2.63×10^6	2.49×10^7	0.2456	0.0612	220	25	5	3885
7	10	8.55	2.60×10^6	2.39×10^7	0.2461	0.0612	180	25	4	3785
7	20	7.05	2.26×10^6	2.04×10^7	0.2509	0.0614	94	25	2	3604
10	5	5.26	1.54×10^6	1.59×10^7	0.2387	0.0599	122	23	3	3843
10	10	4.80	1.48×10^6	1.45×10^7	0.2461	0.0605	90	24	2	3691
10	20	3.55	1.17×10^6	1.10×10^7	0.2631	0.0615	42	22	1	3635

Tabla 5.10: Métricas reportadas para la solución obtenida en distintos escenarios, para valores de β y λ con una previsión económica pesimista.

número de rutas operadas similar, si bien el número de vuelos es sustancialmente mayor. Como se ha comentado, este modelo de precios favorece topologías de red con unos pocos *hubs* que consigan atraer la demanda entre las principales ciudades, garantizando la operación de las rutas más beneficiosas. Cabe mencionar que San Francisco deja de ser uno de los *hubs* en escenarios más restrictivos. Esto se debe a que en el modelo de precios utilizado influye más la demanda que la distancia, mientras que los costes de operación siguen dependiendo de la distancia. Así, establecer un *hub* en grandes ciudades con múltiples aeropuertos cercanos reporta mayores beneficios, como es el caso de Boston, ya que la mayoría de ciudades se concentran en la costa este. Observando la tabla 5.11, se comprueba que la medida del RASM aumenta hasta un 38 % respecto a las soluciones obtenidas con el modelo de precios original, dadas en la tabla 5.5, manteniendo el CASM unos valores muy similares.

β	λ	D.A. [%]	Pres. [\$]	Ben. [\$]	RASM	CASM	Arcos	Nodos	Hubs	t_{comp} [s]
5	5	27.63	7.99×10^6	7.75×10^7	0.3263	0.0661	290	25	7	3664
5	10	27.33	8.07×10^6	7.70×10^7	0.3267	0.0660	256	25	6	3745
5	20	26.07	7.70×10^6	7.48×10^7	0.3354	0.0667	178	25	4	3647
7	5	24.83	7.18×10^6	7.15×10^7	0.3374	0.0663	248	25	6	3714
7	10	23.89	7.02×10^6	6.93×10^7	0.3445	0.0667	170	24	4	3749
7	20	23.29	7.01×10^6	6.81×10^7	0.3473	0.0669	132	25	3	3725
10	5	21.23	6.12×10^6	6.32×10^7	0.3600	0.0670	166	24	4	3715
10	10	20.77	6.07×10^6	6.22×10^7	0.3629	0.0672	130	24	3	3674
10	20	20.78	6.23×10^6	6.23×10^7	0.3628	0.0672	132	24	3	3616

Tabla 5.11: Métricas reportadas para la solución obtenida en distintos escenarios, para valores de β y λ , modificando el modelo de precios.

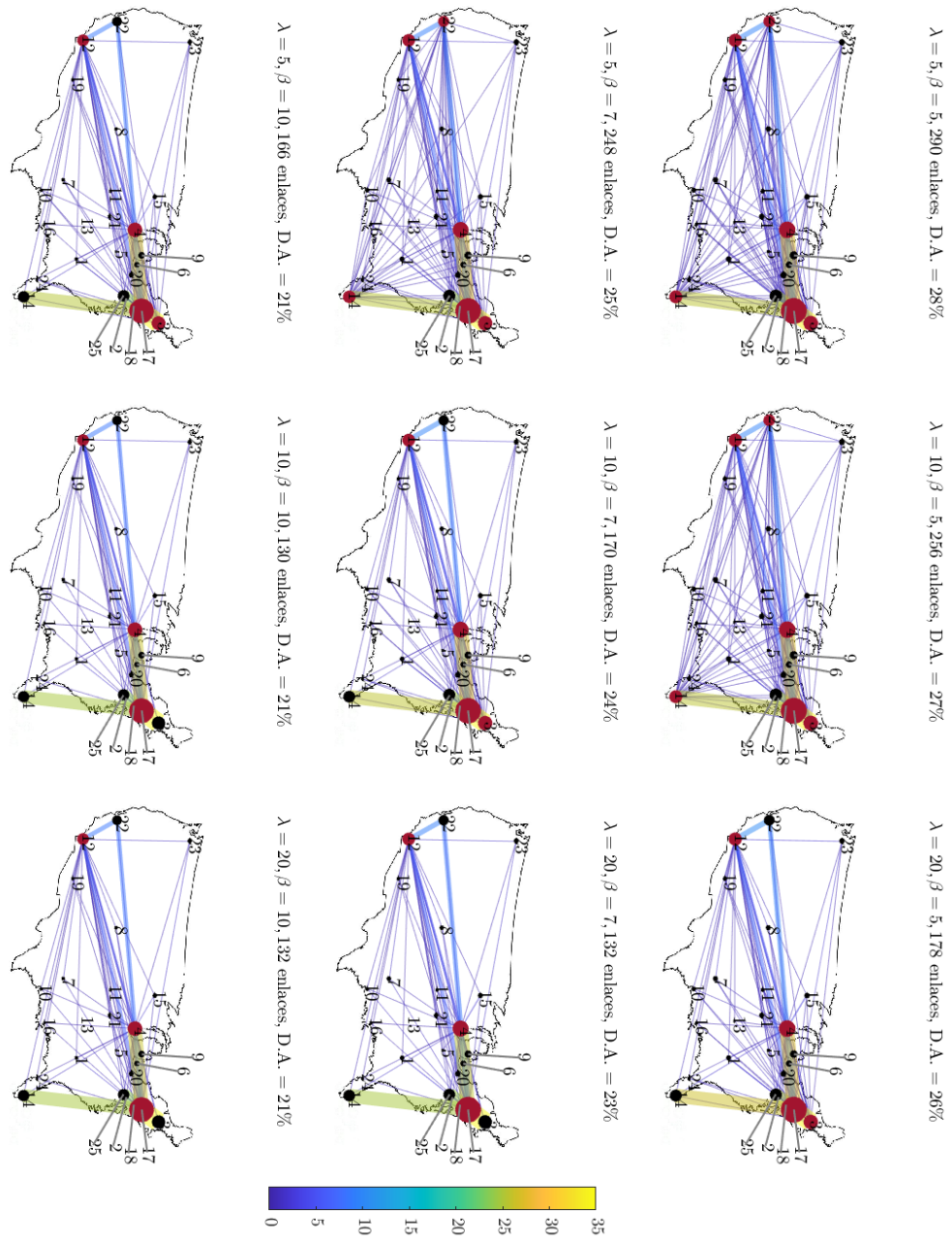


Figura 5.9: Topología de la solución propuesta para distintos escenarios de costes fijos y presupuesto, modificando el modelo de precios.

Capítulo 6

Conclusiones y líneas futuras

Este capítulo tiene un doble propósito. En primer lugar, en la sección 6.1 se realiza un breve resumen de las tareas y contribuciones realizadas, enfatizando tanto los aspectos más novedosos como los que tienen un mayor impacto. En segundo lugar, en la sección 6.2 se identifican vías de trabajo futuro relevantes que permitirían continuar con la senda iniciada por este TFG.

6.1. Contribuciones, resultados e impacto

La contribución principal de este TFG es la introducción de nuevas formulaciones para el diseño óptimo de RdT, permitiendo obtener modelos convexos cuyo tiempo de cómputo es sustancialmente menor que el de las formulaciones MIP clásicamente planteadas en el estado del arte. El menor coste computacional es clave desde dos puntos de vista complementarios. En primer lugar, permite el diseño de RdT de mayor tamaño, pudiendo aplicarse a escenarios realistas con decenas o centenas de estaciones. En segundo lugar, permite obtener el diseño óptimo de la red no solo para los valores nominales de las variables de estado, sino para muchos otros escenarios. Esto permite a los operadores y reguladores considerar distintas condiciones de contorno (modificando, por ejemplo, las estimaciones de demanda, los costes de construcción, los precios del transporte o los coeficientes de utilidad de los pasajeros), obtener el diseño óptimo de la RdT para cada uno de estos escenarios y, comparando las distintas soluciones, tomar la decisión del diseño final de la red de una forma más rigurosa e informada.

Desde un punto de vista científico-técnico las dos propuestas críticas para la formulación del problema como una optimización convexa han sido: i) el uso de regularizadores logarítmicos para potenciar sparsity en la infraestructura (estaciones y arcos) construida y ii) el uso de regularizadores de entropía y utilidades lineales para dar lugar a soluciones óptimas en las que los pasajeros toman decisiones en función de un modelo *logit*. La primera propuesta permite eliminar las variables binarias de construcción del modelo MIP. La segunda permite eliminar la restricción no lineal (y no convexa) asociada al reparto *logit* de la demanda. Pese a que ambas técnicas habían sido utilizadas con anterioridad en otras disciplinas científicas, su uso en el ámbito del transporte y, más específicamente, en el diseño de RdT capacitadas es novedoso. Para validar la formulación propuesta, se ha comparado de forma exhaustiva con las implementaciones clásicas, obteniendo resultados prácticamente idénticos en términos de optimalidad y reduciendo con creces el tiempo necesario para hallar la solución. Debe también destacarse que, más allá del ahorro computacional, la propuesta de una metodología de diseño basada en optimización convexa permite, además, integrar de forma sencilla condiciones de operación y requisitos de servicios que puedan modelarse de forma convexa enriqueciendo la expresividad del modelo y aumentando el rango de problemas que pueden atacarse.

Le segunda mitad del TFG se ha dedicado a adaptar y aplicar la nueva formulación a dos problemas reales diferentes: el desarrollo de una red de metro y la creación de una nueva aerolínea.

- Para la aplicación al diseño y despliegue de la red de metro, se ha realizado un estudio conside-

rando varios escenarios de costes y presupuesto, analizando también los cambios en la topología de la red diseñada en función de los pesos de los diferentes términos en la función objetivo, como puede ser la priorización de la población procedente de barrios deprimidos o la predicción de crecimiento por distrito de la ciudad.

- Por la aplicación al diseño y despliegue de una nueva aerolínea, se ha realizado una serie de cambios asociados a la naturaleza del medio de transporte, promoviendo topologías de *hub & spoke*. En este caso, se consideran varios escenarios de coste y presupuesto, distintas limitaciones adicionales (de capacidad aeroportuaria o tamaño de flota), y modificaciones en la previsión económica (cambios en el modelo de precios, etc.). Los resultados obtenidos dan lugar a topologías de *hub & spoke* y, en situaciones extremas, dan lugar a redes en estrella, que es uno de los tipos de red más comunes en el mundo aéreo.

Debe destacarse, una vez más, que la capacidad de llevar a cabo este tipo de análisis (considerando decenas de casos de estudio que reflejan distintos escenarios sociales, condiciones económicas y restricciones de operación) en los dos tipos RdT reales consideradas, pone en valor el impacto económico y social de la metodología de diseño propuesta en este TFG.

Por último, y desviándome un poco de la línea argumental establecida, me gustaría resaltar también el impacto que el trabajo que he llevado a cabo durante los 9 meses que he dedicado a este TFG ha tenido en mi desarrollo académico, profesional y personal. En primer lugar, creo que el TFG me ha permitido integrar de manera natural conocimientos adquiridos a lo largo de la carrera, en asignaturas tan distintas como Introducción a la Ingeniería Aeroespacial, Optimización del Espacio Aéreo o Programación, así como hacer un uso práctico de muchas de las herramientas y tecnologías introducidas en dichas asignaturas. En segundo lugar, me ha permitido la adquisición de nuevos conocimientos, especialmente en los ámbitos de la optimización, las RdT y el diseño y operación de las redes aéreas, reforzando mi formación y haciéndome más consciente de las oportunidades de aplicación de dichos conocimientos. También me ha servido para aprender el uso de nuevas tecnologías y herramientas de desarrollo, desde aspectos relacionados con el uso de servidores de simulación y de computación distribuida, la consulta y extracción de información de bases de datos de organismos (AENA, INE, FMI o BTS) o la toma de decisiones por parte de operadores de transporte privados. Quizás de forma más inesperada, mi trabajo en este TFG me ha permitido también la adquisición de habilidades transversales de investigación, incluyendo la lectura y escritura de artículos de investigación, la colaboración con grupos internacionales de prestigio, la realización de presentaciones en inglés o la redacción de documentos científico-técnicos. Creo sinceramente que todo esto me ha hecho crecer como ingeniero y como persona y lo percibo como un final sobresaliente que me hace valorar todavía más la formación que he recibido a lo largo de estos años.

6.2. Líneas de trabajo futuro

Esta sección pretende identificar, de forma no exhaustiva, algunas de las potenciales líneas de trabajo futuro que han surgido tras la realización de este TFG. Si bien el interés, plazo de ejecución o nivel de detalle de las líneas identificadas no es homogéneo, se intenta proporcionar un listado representativo. En primer lugar, se identifican vías de trabajo futuro relacionadas con los escenarios de aplicación de la metodología de optimización propuesta:

- **Particularización de la metodología de diseño para otro tipo de RdT.** Los dos casos analizados en los capítulos 4 y 5 se han escogido no solo por su relevancia y su impacto socio-económico, sino también por representar dos ejemplos de aplicación diferentes. En uno de ellos el operador era público y los costes de construcción de nuevas rutas es muy elevado, mientras que en el otro el operador era privado y el coste de construcción era mucho más moderado y, habitualmente, menor que el que supone los costes de operación. No obstante, existen otros muchos

tipos de movilidad y RdT donde, con las modificaciones oportunas, la metodología propuesta podría aplicarse. De forma general, esto podrá requerir la adaptación de las funciones objetivo, la potencial consideración de nuevas variables de estado y, principalmente, la inclusión de restricciones que capturen las condiciones de operación o requisitos de servicio de dichas redes. Casos específicos que se pretenden estudiar en el futuro próximo incluyen el diseño de redes de carretera y el diseño de servicios en una RdT. Este último representa un problema perteneciente a un nivel jerárquico inferior, y con él se busca ilustrar la aplicabilidad de la metodología desarrollada a problemas relacionados con la gestión de RdT.

- **Generalización a otros tipos de redes.** De forma más general, varios de los principios de diseño planteados en este TFG son también de aplicación a otro tipo de redes, incluyendo redes de vehículos eléctricos (donde debería incorporarse también el dimensionamiento de las estaciones de carga o de estacionamiento de vehículos compartidos), redes de transmisión y distribución de potencia, o redes de comunicaciones.

Desde un punto de vista científico-tecnológico, existen multitud de líneas de trabajo futuro. Algunas de las más relevantes son:

- **Incorporación de incertidumbre y variables de estado estocásticas a la optimización.** Uno de los mayores retos que afrontan los esquemas de optimización entera es la consideración de incertidumbre en los parámetros que se utilizan como punto de partida para la optimización. Las RdT son inherentemente distribuidas y en su diseño y operación deben integrarse distintos horizontes temporales. Esto provoca que muchos de los parámetros que sirven de entrada a la optimización tengan una naturaleza espacio-temporal que provoca que, en caso de que se deban modelar de una forma aleatoria, existan lugar a fuertes dependencias estadísticas. Estas dependencias complican tremendamente no solo la generación de escenarios que sean representativos, sino también el número de escenarios que deben generarse, haciendo en ocasiones inviable el uso de técnicas de programación entera. Por el contrario, la incorporación de incertidumbre y variables aleatorias espacio-temporales es un ámbito mucho más maduro en el campo de la optimización convexa, existiendo distintas familias de algoritmos que permiten garantizar (bien en esperanza o bien en probabilidad) determinadas prestaciones, así como robustez en términos de cumplimiento de las restricciones. Estos aspectos son críticos para garantizar la sostenibilidad económica y social de la inversión en distintos escenarios, así como para el diseño robusto de RdT que permita hacer frente a disrupciones y problemas en la red.
- **Postulación de funciones de entropía generalizadas que permitan manejar de forma integral la división de flujos, transbordos y categorías de usuarios.** Por sencillez, en este TFG se han presentado resultados suponiendo que los usuarios toman decisiones de acuerdo con un modelo *logit* que tiene en cuenta las características de la ruta origen-destino de la red a construir, así como de la competencia. No obstante, dicho modelado puede complicarse (enriquecerse) desde múltiples puntos de vista, incluyendo la división de los flujos de pasajeros, la existencia de transbordos o la toma de decisiones basadas en modelos *logit* anidados. Replicar la metodología expuesta en este TFG para cubrir estos escenarios requeriría, entre otros, la postulación de nuevas funciones objetivo cuyas condiciones de optimalidad dieran lugar a los comportamientos deseados. Esta es una vía que ya se está explorando y que se perseguirá de forma activa tras la finalización de este TFG.
- **Desarrollo de algoritmos de optimización ajustados a la estructura del problema de diseño de RdT.** En este TFG se ha adoptado un enfoque híbrido para el diseño de los algoritmos que resuelven el problema de optimización planteado. Más concretamente, se ha propuesto un algoritmo iterativo que: i) en cada iteración obtiene los costes fijos y lineales de construcción a través de un esquema de MM asociado a una función logarítmica y ii) una vez obtenidos dichos costes, resuelve el resto de variables a través de un *solver* genérico de optimización convexa. Si

bien esto reduce el tiempo que ha llevado desarrollar el algoritmo, es igualmente cierto que el tiempo de ejecución es mayor que el que podría obtenerse. En este sentido, debe destacarse que la taxonomía de algoritmos de optimización convexa es muy amplia (mucho mayor que la de los MIP) y que en los últimos años se han experimentado unos avances muy significativos en el diseño de algoritmos. Desde este punto de vista, si la metodología de diseño propuesta quiere aplicarse a redes mucho mayores o pretende incorporarse una mayor resolución temporal o espacial en las variables, cobra especial sentido desarrollar algoritmos específicos que tengan en cuenta las particularidades del problema, incluidas la presencia de costes y restricciones lineales que hacen que el problema no sea diferenciable, así como el modelado de infraestructura de la RdT como un grafo, lo que permite el uso de esquemas de optimización distribuida.

- **Formulación del problema de diseño de RdT como una optimización binivel con restricciones y desarrollo de algoritmos de optimización eficientes para su solución aproximada.** Si bien la práctica totalidad de la literatura ataca el diseño de RdT como un problema de optimización de un único nivel en el que se integran los objetivos de los distintos actores involucrados en el diseño y operación de la red (pasajeros, operadores, reguladores...), una formulación binivel en la que los pasajeros actúan de forma totalmente autónoma en función de las decisiones de diseño realizadas por el operador resulta más realista. El problema en este caso es que los algoritmos de optimización binivel son tremendamente complejos y sus condiciones de convergencia no son triviales. En este sentido, gracias a la relevancia de la optimización binivel en el ámbito del aprendizaje máquina, se han producido recientes avances en el desarrollo y caracterización de algoritmos de optimización binivel por descenso para problemas no lineales (véase, *e.g.*, [90]). Aunque los algoritmos existentes en el estado del arte no consideran la presencia de restricciones (un aspecto fundamental en el diseño de RdT), resultados iniciales obtenidos durante el transcurso de este TFG indican que, mediante ciertas modificaciones, es posible generalizar los resultados existentes a problemas con restricciones. La generalización de dichos resultados y su aplicación al diseño de RdT es otra de las líneas de trabajo futuro de este TFG.

Naturalmente, hay numerosas áreas de interés para futuras investigaciones que podrían haber sido abordadas, pero que, debido a restricciones de espacio, no han sido incluidas en la discusión. Algunas de estas áreas, particularmente relevantes, incluyen: i) la integración de los métodos de optimización propuestos con enfoques de predicción Bayesiana, aprendizaje automático y aprendizaje por refuerzo; ii) la profundización en el diseño e implementación de métricas que, de forma rigurosa, permitan incorporar aspectos relacionados con el impacto social, la reducción de la desigualdad y la accesibilidad en el diseño de RdT; o iii) la incorporación de funciones y restricciones relacionadas con la sostenibilidad y el impacto ecológico en el diseño de RdT. Dada la complejidad, no linealidad y naturaleza eminentemente continua (analógica) de las variables y métricas mencionadas, su integración en formulaciones clásicas basadas en MIP es poco viable. Sin embargo, disponer de la metodología de optimización convexa presentada en este TFG, sugiere que es factible incorporar estos objetivos en nuestros algoritmos en un futuro cercano. En este sentido, los tutores de este trabajo han decidido integrar los resultados obtenidos durante la realización de este TFG en varios proyectos de investigación que, en los campos de la optimización y el transporte, dirigen.

En resumen, los resultados obtenidos en este TFG se consideran prometedores, ya que ofrecen una metodología de optimización de RdT que es novedosa y presenta ventajas sustanciales con respecto a las técnicas predominantes en el estado del arte, basadas en programación entera. Por consiguiente, estos hallazgos pueden servir como un sólido punto de partida para la formulación de problemas y el desarrollo de algoritmos que contribuyan al diseño flexible, eficiente, sostenible y robusto de las RdT en los próximos años.

Bibliografía

- [1] D. Teodorovic and M. Janić, *Transportation Engineering: Theory, Practice and Modeling*. Butterworth-Heinemann, 2016.
- [2] V. R. Vuchic, *Urban transit systems and technology*. John Wiley & Sons, 2007.
- [3] C. Barnhart and A. Cohn, “Airline schedule planning: Accomplishments and opportunities,” *Manufacturing & Service Operations management*, vol. 6, no. 1, pp. 3–22, 2004.
- [4] L. Cadarso, V. Vaze, C. Barnhart, and Á. Marín, “Integrated airline scheduling: Considering competition effects and the entry of the high speed rail,” *Transportation Science*, vol. 51, no. 1, pp. 132–154, 2017.
- [5] P. Belobaba, “The airline planning process,” *The Global Airline Industry*, pp. 153–181, 2009.
- [6] J.-W. Goossens, S. Van Hoesel, and L. Kroon, “A branch-and-cut approach for solving railway line-planning problems,” *Transportation Science*, vol. 38, no. 3, pp. 379–393, 2004.
- [7] A. Schöbel, “Line planning in public transportation: models and methods,” *OR Spectrum*, vol. 34, no. 3, pp. 491–510, 2012.
- [8] V. Guihaire and J.-K. Hao, “Transit network design and scheduling: A global review,” *Transportation Research Part A: Policy and Practice*, vol. 42, no. 10, pp. 1251–1273, 2008.
- [9] R. Church and C. ReVelle, “The maximal location covering problem,” *Papers in Regional Science*, vol. 32, pp. 101–118, 1974.
- [10] M. E. Ben-Akiva and S. R. Lerman, *Discrete choice analysis: theory and application to travel demand*, vol. 9. MIT press, 1985.
- [11] E. Cascetta, *Transportation systems analysis: models and applications*, vol. 29. Springer Science & Business Media, 2009.
- [12] T. L. Magnanti and R. T. Wong, “Network design and transportation planning: Models and algorithms,” *Transportation Science*, vol. 18, no. 1, pp. 1–55, 1984.
- [13] G. Laporte, J. A. Mesa, and F. A. Ortega, “Optimization methods for the planning of rapid transit systems,” *European Journal of Operational Research*, vol. 122, no. 1, pp. 1–10, 2000.
- [14] G. Bruno, M. Gendreau, and G. Laporte, “A heuristic for the location of a rapid transit line,” *Computers & Operations Research*, vol. 29, no. 1, pp. 1–12, 2002.
- [15] G. Laporte, J. A. Mesa, F. A. Ortega, and I. Sevillano, “Maximizing trip coverage in the location of a single rapid transit alignment,” *Annals of Operations Research*, vol. 136, pp. 49–63, 2005.
- [16] G. Laporte, Á. Marín, J. A. Mesa, and F. A. Ortega, “An integrated methodology for the rapid transit network design problem,” in *Algorithmic Methods for Railway Optimization: International Dagstuhl Workshop, Dagstuhl Castle, Germany, June 20-25, 2004, 4th International Workshop, ATMOS 2004, Bergen, Norway, September 16-17, 2004, Revised Selected Papers*, pp. 187–199, Springer, 2007.
- [17] G. Gutiérrez-Jarpa, C. Obreque, G. Laporte, and V. Marianov, “Rapid transit network design for optimal cost and origin–destination demand capture,” *Computers & Operations Research*, vol. 40, no. 12, pp. 3000–3009, 2013.
- [18] G. Laporte, A. Marín, J. A. Mesa, and F. Perea, “Designing robust rapid transit networks with alternative routes,” *Journal of Advanced Transportation*, vol. 45, no. 1, pp. 54–65, 2011.
- [19] F. Perea, M. B. Menezes, J. A. Mesa, and F. Rubio-Del-Rey, “Transportation infrastructure network design in the presence of modal competition: computational complexity classification and a genetic algorithm,” *Top*, vol. 28, no. 2, pp. 442–474, 2020.
- [20] J. Zheng, Q. Meng, and Z. Sun, “Liner hub-and-spoke shipping network design,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 75, pp. 32–48, 2015.
- [21] B. García-Archilla, A. J. Lozano, J. A. Mesa, and F. Perea, “GRASP algorithms for the robust railway network design problem,” *Journal of Heuristics*, vol. 19, pp. 399–422, 2013.

- [22] D. Canca, A. De-Los-Santos, G. Laporte, and J. A. Mesa, “An adaptive neighborhood search metaheuristic for the integrated railway rapid transit network design and line planning problem,” *Computers & Operations Research*, vol. 78, pp. 1–14, 2017.
- [23] D. Canca, A. De-Los-Santos, G. Laporte, and J. A. Mesa, “Integrated railway rapid transit network design and line planning problem with maximum profit,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 127, pp. 1–30, 2019.
- [24] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media, 2009.
- [25] K. Lange, *Karush-Kuhn-Tucker Theory*, pp. 107–135. New York, NY: Springer New York, 2013.
- [26] L. Cadarso and Á. Marín, “Improved rapid transit network design model: considering transfer effects,” *Annals of Operations Research*, vol. 258, no. 2, pp. 547–567, 2017.
- [27] L. Cadarso and Á. Marín, “Combining robustness and recovery in rapid transit network design,” *Transportmetrica A: Transport Science*, vol. 12, no. 3, pp. 203–229, 2016.
- [28] Á. G. Marín and P. Jaramillo, “Urban rapid transit network design: accelerated benders decomposition,” *Annals of Operations Research*, vol. 169, no. 1, pp. 35–53, 2009.
- [29] S. T. Waller, K. C. Mouskos, D. Kamaryiannis, and A. K. Ziliaskopoulos, “A linear model for the continuous network design problem,” *Computer-Aided Civil and Infrastructure Engineering*, vol. 21, no. 5, pp. 334–345, 2006.
- [30] N. Ghaffarinasab, T. Van Woensel, and S. Minner, “A continuous approximation approach to the planar hub location-routing problem: Modeling and solution algorithms,” *Computers & Operations Research*, vol. 100, pp. 140–154, 2018.
- [31] T. Thomadsen and T. Stidsen, “The generalized fixed-charge network design problem,” *Computers & Operations Research*, vol. 34, no. 4, pp. 997–1007, 2007.
- [32] Z. Yang, B. Yu, and C. Cheng, “A parallel ant colony algorithm for bus network optimization,” *Computer-Aided Civil and Infrastructure Engineering*, vol. 22, no. 1, pp. 44–55, 2007.
- [33] J. Nocedal and S. J. Wright, *Numerical optimization*. Springer, 1999.
- [34] F. S. Hillier and G. J. Lieberman, *Introduction to Operations Research*. McGraw-Hill, 2015.
- [35] L. A. Wolsey, *Integer programming*. John Wiley & Sons, 2020.
- [36] E.-G. Talbi, *Metaheuristics: from design to implementation*. John Wiley & Sons, 2009.
- [37] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to algorithms*. MIT press, 2022.
- [38] M. Hong, M. Razaviyayn, Z.-Q. Luo, and J.-S. Pang, “A unified algorithmic framework for block-structured optimization involving big data: With applications in machine learning and signal processing,” *IEEE Signal Processing Magazine*, vol. 33, no. 1, pp. 57–77, 2016.
- [39] L. Garrow, *Customer Modeling*, pp. 1–33. Boston, MA: Springer US, 2012.
- [40] D. McFadden, “Econometric models of probabilistic choice,” *Structural Analysis of Discrete Data with Econometric Applications*, vol. 198272, 1981.
- [41] J. d. D. Ortúzar, “On the development of the nested logit model,” *Transportation Research Part B: Methodological*, vol. 35, no. 2, pp. 213–216, 2001.
- [42] M. Fischetti, F. Glover, and A. Lodi, “The feasibility pump,” *Mathematical Programming*, vol. 104, p. 91–104, Mar. 2005.
- [43] E. J. Candes, M. B. Wakin, and S. P. Boyd, “Enhancing sparsity by reweighted ℓ_1 minimization,” *Journal of Fourier Analysis and Applications*, vol. 14, pp. 877–905, 2008.
- [44] Y. Sun, P. Babu, and D. P. Palomar, “Majorization-minimization algorithms in signal processing, communications, and machine learning,” *IEEE Transactions on Signal Processing*, vol. 65, no. 3, pp. 794–816, 2016.
- [45] N. Oppenheim, “Equilibrium trip distribution/assignment with variable destination costs,” *Transportation Research Part B: Methodological*, vol. 27, no. 3, pp. 207–217, 1993.
- [46] D. Branston, “Link capacity functions: A review,” *Transportation Research*, vol. 10, no. 4, pp. 223–236, 1976.
- [47] S. Elhedhli and H. Wu, “A lagrangean heuristic for hub-and-spoke system design with capacity selection and congestion,” *INFORMS Journal on Computing*, vol. 22, no. 2, pp. 282–296, 2010.
- [48] V. Bayram, B. Yıldız, and M. S. Farham, “Hub network design problem with capacity, congestion, and stochastic demand considerations,” *Transportation Science*, vol. 57, no. 5, pp. 1276–1295, 2023.
- [49] M. Labbé and P. Marcotte, “Bilevel network design,” *Network Design with Applications to Transportation and Logistics*, pp. 255–281, 2021.
- [50] S. Elhedhli and F. X. Hu, “Hub-and-spoke network design with congestion,” *Computers & Operations Research*, vol. 32, no. 6, pp. 1615–1632, 2005.

- [51] Y. Collette and P. Siarry, *Multiobjective optimization: principles and case studies*. Springer Science & Business Media, 2013.
- [52] K. Deb, K. Sindhya, and J. Hakanen, “Multi-objective optimization,” in *Decision Sciences*, pp. 161–200, CRC Press, 2016.
- [53] L. Escudero and S. Muñoz, “An approach for solving a modification of the extended rapid transit network design problem,” *Top*, vol. 17, pp. 320–334, 12 2009.
- [54] Á. Marín and R. García-Ródenas, “Location of infrastructure in urban railway networks,” *Computers & Operations Research*, vol. 36, no. 5, pp. 1461–1477, 2009.
- [55] A. Marin and P. Jaramillo, “Urban rapid transit network capacity expansion,” *European Journal of Operational Research*, vol. 191, no. 1, pp. 45–60, 2008.
- [56] J.-Y. Wang and C.-M. Lin, “Mass transit route network design using genetic algorithm,” *Journal of the Chinese Institute of Engineers*, vol. 33, no. 2, pp. 301–315, 2010.
- [57] G. Laporte, J. A. Mesa, and F. Perea, “A game theoretic framework for the robust railway transit network design problem,” *Transportation Research Part B: Methodological*, vol. 44, no. 4, pp. 447–459, 2010.
- [58] S. Ngamchai and D. J. Lovell, “Optimal time transfer in bus transit route network design using a genetic algorithm,” *Journal of Transportation Engineering*, vol. 129, no. 5, pp. 510–521, 2003.
- [59] F. Zhao and I. Ubaka, “Transit network optimization—minimizing transfers and optimizing route directness,” *Journal of Public Transportation*, vol. 7, no. 1, pp. 63–82, 2004.
- [60] E. Klotz and A. M. Newman, “Practical guidelines for solving difficult linear programs,” *Surveys in Operations Research and Management Science*, vol. 18, no. 1, pp. 1–17, 2013.
- [61] “Atlas de distribución de renta de los hogares,” octubre 2023. Fecha de consulta: 23 de marzo de 2024. Enlace.
- [62] N. Adler and K. Smilowitz, “Hub-and-spoke network alliances and mergers: Price-location competition in the airline industry,” *Transportation Research Part B: Methodological*, vol. 41, no. 4, pp. 394–409, 2007.
- [63] P. Belobaba, A. Odoni, and C. Barnhart, *The global airline industry*. John Wiley & Sons, 2015.
- [64] S. Nickel, A. Schöbel, and T. Sonneborn, *Hub Location Problems in Urban Traffic Networks*, pp. 95–107. Boston, MA: Springer US, 2001.
- [65] A. Goldman, “Optimal locations for centers in a network,” *Transportation Science*, vol. 3, no. 4, pp. 352–360, 1969.
- [66] M. E. O’Kelly, “The location of interacting hub facilities,” *Transportation Science*, vol. 20, no. 2, pp. 92–106, 1986.
- [67] M. E. O’Kelly, “Activity levels at hub facilities in interacting networks,” *Geographical Analysis*, vol. 18, no. 4, pp. 343–356, 1986.
- [68] M. E. O’Kelly, “A quadratic integer program for the location of interacting hub facilities,” *European Journal of Operational Research*, vol. 32, no. 3, pp. 393–404, 1987.
- [69] J. F. Campbell, “Integer programming formulations of discrete hub location problems,” *European Journal of Operational Research*, vol. 72, no. 2, pp. 387–405, 1994.
- [70] J. F. Campbell, A. T. Ernst, and M. Krishnamoorthy, “Hub location problems,” *Facility Location*, pp. 373–407, 2002.
- [71] S. Alumur and B. Y. Kara, “Network hub location problems: The state of the art,” *European Journal of Operational Research*, vol. 190, no. 1, pp. 1–21, 2008.
- [72] D. Skorin-Kapov, J. Skorin-Kapov, and M. O’Kelly, “Tight linear programming relaxations of uncapacitated p-hub median problems,” *European Journal of Operational Research*, vol. 94, no. 3, pp. 582–593, 1996.
- [73] J. Sohn and S. Park, “Efficient solution procedure and reduced size formulations for p-hub location problems,” *European Journal of Operational Research*, vol. 108, no. 1, pp. 118–126, 1998.
- [74] J. F. Campbell, “Location and allocation for distribution systems with transshipments and transportation economies of scale,” *Annals of Operations Research*, vol. 40, no. 1, pp. 77–99, 1992.
- [75] A. T. Ernst and M. Krishnamoorthy, “Efficient algorithms for the uncapacitated single allocation p-hub median problem,” *Location Science*, vol. 4, no. 3, pp. 139–154, 1996.
- [76] M. E. O’Kelly, D. Bryan, D. Skorin-Kapov, and J. Skorin-Kapov, “Hub network design with single and multiple allocation: A computational study,” *Location Science*, vol. 4, no. 3, pp. 125–138, 1996.
- [77] M. E. O’Kelly, “Hub facility location with fixed costs,” *Papers in Regional Science*, vol. 71, no. 3, pp. 293–306, 1992.
- [78] T. Aykin, “Lagrangian relaxation based approaches to capacitated hub-and-spoke network design problem,” *European Journal of Operational Research*, vol. 79, no. 3, pp. 501–523, 1994.
- [79] M. Labbé, H. Yaman, and E. Gourdin, “A branch and cut algorithm for hub location problems with single assignment,” *Mathematical Programming*, vol. 102, no. 2, pp. 371–405, 2005.

- [80] H. Yaman and G. Carello, “Solving the hub location problem with modular link capacities,” *Computers & Operations Research*, vol. 32, no. 12, pp. 3227–3245, 2005.
- [81] A. Marín, “Formulating and solving splittable capacitated multiple allocation hub location problems,” *Computers & Operations Research*, vol. 32, no. 12, pp. 3093–3109, 2005.
- [82] M. E. O’kelly, “A quadratic integer program for the location of interacting hub facilities,” *European Journal of Operational Research*, vol. 32, no. 3, pp. 393–404, 1987.
- [83] “Bureau of transportation statistics. Transtats.” Fecha de consulta: 20 de enero de 2024. Enlace.
- [84] “Usgs national transportation dataset,” 2023. Fecha de consulta: 25 de marzo de 2024.
- [85] “Flight radar database.” Fecha de consulta: 20 de enero de 2024. Enlace.
- [86] “Air passenger market analysis,” diciembre 2023. Fecha de consulta: 29 de mayo de 2024. Enlace.
- [87] “Tarifas AENA mayo 2024.” Fecha de consulta: 30 de mayo de 2024. Enlace.
- [88] “World economic outlook,” marzo 2024. Fecha de consulta: 23 de marzo de 2024. Enlace.
- [89] R. Gatti, D. Lederman, A. M. Islam, H. Nguyen, R. Lotfi, and M. Emam Mousa, “Data transparency and GDP growth forecast errors,” *Journal of International Money and Finance*, vol. 140, p. 102991, 2024.
- [90] H. Shen, Q. Xiao, and T. Chen, “On penalty-based bilevel gradient descent method,” in *International Conference on Machine Learning*, 2023.

Apéndice A

Linealización por trozos de la función *logit*

En este apéndice se explica el procedimiento para linealizar una función a través de una aproximación por regiones, con el fin de formular un problema de optimización de diseño de RdT. Este problema incorpora un modelo de reparto de demanda basado en la función *logit*. La aproximación de la función *logit* se realiza con el propósito de obtener un modelo que pueda ser resuelto mediante un *solver* MIP. Para comenzar, consideremos una función genérica $f : \mathcal{D} \rightarrow \mathbb{R}$ que se quiere linealizar a trozos mediante la función $\hat{f} : \mathcal{D} \rightarrow \mathbb{R}$. El primer paso consiste en dividir el dominio $\mathcal{D}$ en un conjunto de regiones no solapadas. Concretamente, si $\mathcal{R}$ denota el conjunto de índices de dichas regiones, se tiene que $\mathcal{D} = \mathcal{D}_1 \cup \dots \cup \mathcal{D}_{|\mathcal{R}|}$ y que $\mathcal{D}_r \cap \mathcal{D}_{r'} = \emptyset$ para todo $r \neq r'$. El segundo paso consiste en definir una función lineal $\hat{f}_r$ para cada región $\mathcal{D}_r$ que, para las entradas $x \in \mathcal{D}_r$, aproxime f de forma fidedigna. Finalmente, si $\bar{r} : \mathcal{D} \rightarrow \mathcal{R}$ es una función que, para un valor de entrada, devuelve el índice de la región a la que pertenece la entrada, se tiene que $\hat{f}(x) = \hat{f}_{\bar{r}(x)}(x)$. Si bien la definición de dicha función es sencilla, su incorporación a un problema de optimización es algo más tediosa. Para describir cómo se puede conseguir esto, se supone que el dominio de la función f es escalar, de forma que: a) las distintas regiones son intervalos y, por lo tanto, es posible escribir $\mathcal{D}_r = [D_r^{\min}, D_r^{\max})$ y b) la función $\hat{f}_r(x)$ puede escribirse como $\hat{f}_r(x) = f(D_r^{\min}) + \hat{\xi}_r(x - D_r^{\min})$, donde $\hat{\xi}_r$ es la pendiente estimada para la región r . Así mismo, suele definirse un valor inicial de la función, denotado como $\hat{f}_0$, que se utiliza para aquellas entradas $x < D_1^{\min}$ y que, salvo que se indique lo contrario, se fija a $\hat{f}_0 = f(D_1^{\min})$. Teniendo en cuenta lo anterior, en la figura A.1 se representa la linealización de la función logística

$$f(x) = \frac{e^{-x}}{e^{-x} + e^{-x_{ext}}}, \quad (\text{A.1})$$

donde x_{ext} es un valor constante. En el ejemplo representado, el dominio $\mathcal{D} = [0, \infty)$ se ha dividido en $|\mathcal{R}| = 6$ regiones, el valor inicial se aproxima a 1 y los valores de $\hat{\xi}_r$ para las 6 regiones presentadas son negativos.

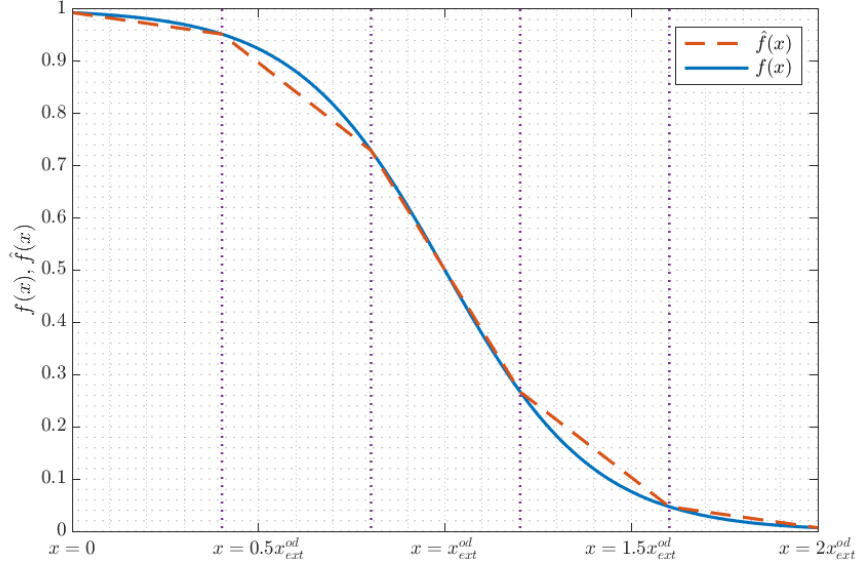


Figura A.1: Representación de la función logística dada en (A.1) y su equivalente linealizada, considerando $|\mathcal{R}| = 6$ regiones. Nótese que en la última región -no representada en la figura y correspondiente a $[2x_{ext}^{od}, \infty)$ - la función $\hat{f}(x)$ tiene pendiente constante de valor nulo.

Para explicar cómo tener en cuenta la función definida a trozos en el problema de optimización, deben considerarse tanto las variables binarias $\{\alpha_r\}_{r=1}^{|\mathcal{R}|}$ como las variables continuas $\{\delta_r\}_{r=1}^{|\mathcal{R}|} \in \mathbb{R}^{\geq 0}$, cuyo significado se describe a continuación:

- α_r es una variable binaria que sirve para identificar la región que está activa. En concreto, para una entrada x determinada, $\alpha_r = 0$ si $x < D_r^{\max}$ y $\alpha_r = 1 \forall x > D_r^{\min}$. De esta forma, x pertenece a la región $\mathcal{D}_r$ si $\alpha_{r-1} = 1$ y $\alpha_r = 0$. Por conveniencia de notación, se define α_0 (que en este caso no es una variable de optimización sino un parámetro fijo) como $\alpha_0 = 1$.
- $\delta_r \in [0, D_r^{\max} - D_r^{\min}]$ es una variable lineal asociada a la región r . Para un valor determinado de x , la variable δ_r toma el siguiente valor:
 - Si $x < D_r^{\min}$, entonces $\delta_r = 0$.
 - Si $x > D_r^{\max}$, entonces $\delta_r = D_r^{\max} - D_r^{\min}$.
 - Si $x \in [D_r^{\min}, D_r^{\max}]$, entonces $\delta_r = x - D_r^{\min}$.

Con esta notación, se tiene que $\alpha_{r-1} \geq \alpha_r$, $x = D_1^{\min} + \sum_{r=1}^{|\mathcal{R}|} \delta_r \alpha_{r-1}$ y que $\hat{f}(x) = \hat{f}_0 + \sum_{r=1}^{|\mathcal{R}|} \hat{\xi}_r \delta_r \alpha_{r-1}$. De esta manera, el modelo descrito en (2.3)-(2.11), (2.17)-(2.19) se puede aproximar como se muestra a continuación.

$$\min_{\mathcal{X}_{cons}, \mathcal{X}_{flow}} \sum_{\forall (o,d) \in \mathcal{K}} w^{od} F_{ext}^{od} \quad (\text{A.2})$$

s. a:

$$\frac{\sum_{\forall (o,d) \in \mathcal{K}} F_{ij}^{od} w^{od}}{a_{nom}} \leq A_{ij} \quad \forall (i,j) \in \mathcal{A} \quad (\text{A.3})$$

$$\sigma_i \sum_{\forall j \in \mathcal{N}_{in}(i)} A_{ji} \leq S_i \quad \forall i \in \mathcal{S} \quad (\text{A.4})$$

$$\sigma_i \sum_{\forall j \in \mathcal{N}_{out}(i)} A_{ij} \leq S_i \quad \forall i \in \mathcal{S} \quad (\text{A.5})$$

$$\sum_{\forall (i,j) \in \mathcal{A}} A_{ij} \leq a_{max} \quad (\text{A.6})$$

$$A_{ij} \leq M A_{ij}^b \quad \forall (i,j) \in \mathcal{A} \quad (\text{A.7})$$

$$S_i \leq M S_i^b \quad \forall i \in \mathcal{S} \quad (\text{A.8})$$

$$\sum_{\forall i \in \mathcal{S}} (c_{s_i} S_i^b + \bar{c}_i S_i) + \sum_{\forall (i,j) \in \mathcal{A}} (c_{A_{ij}} A_{ij}^b + \bar{c}_{A_{ij}} A_{ij}) \leq b \quad (\text{A.9})$$

$$F^{od} + F_{ext}^{od} = 1 \quad \forall (o,d) \in \mathcal{K} \quad (\text{A.10})$$

$$F^{od} \leq M Z^{od} \quad \forall (o,d) \in \mathcal{K} \quad (\text{A.11})$$

$$\sum_{\forall j \in \mathcal{N}_{out}(i)} F_{ij}^{od} - \sum_{\forall j \in \mathcal{N}_{in}(i)} F_{ji}^{od} = \begin{cases} Z^{od} & i = o, \\ 0 & i \neq o \neq d, \\ -Z^{od} & i = d \end{cases} \quad \forall i, (o,d) \in \mathcal{S} \times \mathcal{K} \quad (\text{A.12})$$

$$F^{od} \leq 1 + \sum_{r=1}^{|\mathcal{R}|} \hat{\xi}_r^{od} \delta_r^{od} \quad \forall (o,d) \in \mathcal{K} \quad (\text{A.13})$$

$$\sum_{\forall (i,j) \in \mathcal{A}} F_{ij}^{od} t_{ij} = \sum_{r=1}^{|\mathcal{R}|} \delta_r^{od} \quad \forall (o,d) \in \mathcal{K} \quad (\text{A.14})$$

$$\alpha_r^{od} \leq \alpha_{r-1}^{od} \quad \forall r, (o,d) \in \mathcal{R} \times \mathcal{K} \quad (\text{A.15})$$

$$\alpha_r^{od} (D_r^{od,max} - D_r^{od,min}) \leq \delta_r^{od} \quad \forall r, (o,d) \in \mathcal{R} \times \mathcal{K} \quad (\text{A.16})$$

$$\delta_r^{od} \leq \alpha_{r-1}^{od} (D_r^{od,max} - D_r^{od,min}) \quad \forall r, (o,d) \in \mathcal{R} \times \mathcal{K}, \quad (\text{A.17})$$

donde la restricción no lineal (2.18) ha sido sustituida por las restricciones (A.13)-(A.17), que implementan la aproximación lineal a trozos. Así por ejemplo, las restricciones (A.14) relacionan el tiempo de viaje percibido por los pasajeros de cada mercado $(o,d) \in \mathcal{K}$ con el sumatorio de las variables auxiliares δ_r^{od} . Las restricciones (A.15) obligan a que, para que una variable α_r^{od} se active, las variables $\alpha_{r'}^{od}$ asociadas a las regiones $r' \in \mathcal{R}$ tales que $r' < r$ han de estar activas. Las restricciones (A.16)-(A.17) establecen conjuntamente el valor de δ_r^{od} de acuerdo con tres casos posibles:

- Según (A.17), si la variable α_{r-1}^{od} asociada a la región anterior no está activa, el valor de δ_r^{od} ha de ser nulo.
- Según (A.16) y (A.17), si la variable α_{r-1}^{od} está activa pero α_r^{od} es igual a cero, el valor de δ_r^{od} estará comprendido entre cero y su valor máximo. Este valor viene dado por la longitud del intervalo donde se define el dominio de la región r , es decir, $D_r^{od,max} - D_r^{od,min}$, siendo $D_r^{od,max}$ y $D_r^{od,min}$ los valores máximo y mínimo del dominio de la región r .

- Si la variable α_r^{od} está activa, el valor de δ_r^{od} ha de ser necesariamente $D_r^{od} - D_{r-1}^{od}$, puesto que es el único valor que satisface conjuntamente las restricciones (A.16) y (A.17).

En total, la aproximación de la función lineal a trozos y su incorporación al modelo de optimización suponen una cantidad adicional de $(3|\mathcal{R}|+1)|\mathcal{K}|$ restricciones y $|\mathcal{R}||\mathcal{K}|$ variables binarias, incrementando exponencialmente la complejidad asociada a resolver el modelo MIP.

Apéndice B

Análisis de sensibilidad de la formulación convexa

En este apéndice se realiza un análisis de sensibilidad de la función objetivo con respecto a distintas variables de estado. Los cambios de valor en las variables se analizan de forma individual y, entre las decenas de variables de estado y parámetros existentes, se escogen o bien las más significativas o bien aquellas en las que pueda existir una mayor incertidumbre con respecto a su valor real o a su evolución a lo largo del tiempo. Este análisis persigue dos objetivos principales: i) demostrar que la formulación convexa y el algoritmo de solución propuestos en este TFG son robustos frente a cambios en los datos y ii) estudiar la estabilidad de la solución. Concretamente:

- Se escogen como variables de estado a analizar: la demanda (w^{od}), el tiempo de viaje por la red alternativa (t_{ext}^{od}), el precio propio p_{ij} , el precio de la red alternativa p_{ext}^{od} , los costes fijos de construcción ($c_{A_{ij}}, c_{S_i}$), los costes lineales de construcción ($\bar{c}_{A_{ij}}, \bar{c}_{S_i}$) y el coste de operación ($c_{o_{ij}}$).
- En cada uno de los experimentos, se selecciona una variable de estado y se realiza un barrido sobre la misma, considerando variaciones desde el -20% al 20% del valor nominal utilizado en la simulación de referencia. Cuando la variable de estado esté formada por múltiples variables (e.g., la demanda es diferente para cada par (o, d)), el nivel de variación relativa se aplicará de forma simultánea a todas las variables. Así, por ejemplo, una subida de la demanda del 10% se traducirá en que todos los valores de la matriz de demanda se incrementan en un 10% .
- Para cada uno de los experimentos y variación, se resuelve el modelo convexo y se cuantifica la calidad de la solución obtenida evaluando la función objetivo del problema MIP, así como la satisfacción de las restricciones. Dicha evaluación se realiza de acuerdo a dos configuraciones de parámetros distintas: 1) utilizando los parámetros del caso de referencia y 2) utilizando los parámetros modificados con los que se resuelve el experimento. Esto supone que: a) podrán obtenerse soluciones que no sean factibles en 1), mientras que siempre lo son en 2) y b) si la solución es factible, el valor de la función objetivo será siempre peor que en el caso de referencia para 1), no siendo necesariamente así para 2).
- Cada uno de los experimentos se lleva cabo para tres configuraciones distintas de los parámetros (λ, β) , de forma que los resultados sean más representativos. Los tres valores considerados son $(\lambda = 5, \beta = 1.25)$, $(\lambda = 6, \beta = 1.25)$ y $(\lambda = 5, \beta = 1.50)$.

Las tablas B.1-B.6 proporcionan los resultados obtenidos tras los experimentos. Cada una de las tablas se corresponde con una de las tres configuraciones de los parámetros (λ, β) . En todas ellas, las columnas indican el porcentaje de variación de la variable de estado, la fila indica la variable de estado modificada y el valor dentro de cada celda indica el cambio (expresado en términos de porcentaje) con respecto al valor de la función objetivo obtenido para el caso de referencia. Porcentajes positivos indican un

mayor coste que el obtenido en el caso de referencia. “Inf.” indica soluciones que no son factibles. En las tablas B.1, B.3 y B.5, el valor de cada celda se calcula según el siguiente procedimiento: 1) se obtiene la solución generada por el algoritmo 2 utilizando como parámetros los asociados a la variación “Cambio” del parámetro indicado en la fila; 2) se evalúa la función objetivo del problema MIP con la configuración de parámetros original en la solución generada en 1); y 3) se compara el valor obtenido en el paso 2 con el valor de la función objetivo para el caso original (“Cambio” = 0 %). Por el contrario, en las tablas B.2, B.4 y B.6, el valor de cada celda se calcula siguiendo los siguientes pasos: 1) se obtiene la solución generada por el algoritmo 2 utilizando como parámetros los asociados a la variación “Cambio” del parámetro indicado en la fila; 2) se evalúa la función objetivo del problema MIP con la configuración de parámetros de 1) en la solución generada en 1); y 3) se compara el valor obtenido en el paso 2 con el valor de la función objetivo para el caso original (“Cambio” = 0 %).

A continuación se detallan, para cada una de las 7 variables de estado analizadas, las conclusiones más relevantes obtenidas a partir del análisis de sensibilidad mostrado en las tablas B.1-B.6:

- Demanda w^{od} . En todos los casos es la variable que da lugar a mayores variaciones de la función objetivo (mayor nivel de sensibilidad). En las tablas B.1, B.3 y B.5 la solución obtenida al reducir los valores estimados de demanda es infactible para la configuración de parámetros original. Si disminuye la demanda, con el mismo presupuesto se puede atraer una mayor proporción de pasajeros, dando lugar a flujos que no son factibles para el caso de referencia. Si se sobrestima la demanda, se observa que la función objetivo crece de forma lineal con la variación del parámetro. Para las tablas B.2, B.4 y B.6, la obtención de valores negativos (que indican un mejor valor de la función objetivo) es razonable puesto que se evalúa la solución obtenida en una función objetivo distinta a la del problema de referencia, donde los valores de los parámetros tienen una magnitud menor al caso original. Por ello, se observa una mayor variación de objetivo, pero, de nuevo, comportándose de forma lineal (con pendiente muy cercana a 1).
- Tiempo de viaje por la red alternativa t_{ext}^{od} . En las tablas B.1, B.3 y B.5 no se observan apenas cambios para variaciones del orden del 5 %. La variación no es simétrica, observándose un mayor impacto cuando se incrementa el tiempo (de manera que más usuarios escogen la nueva red). En las tablas B.2, B.4 y B.6, la variación es mayor, observándose un cambio en el objetivo que se comporta de forma simétrica y lineal (con pendiente muy cercana a 1), demostrando que el tiempo de viaje por la red alternativa es uno de los parámetros que más afectan al valor objetivo del modelo.
- Precio propio p_{ij} . La sensibilidad es bastante pequeña (dos órdenes de magnitud menor que con el cambio de demanda) y relativamente suave. Variaciones abultadas no hacen que la penalización en términos de coste sea mucho mayor. En promedio, se observa que la solución es menos sensible a la variación de precios que a la de tiempos de viaje, lo que en parte se justifica por los coeficientes de utilidad considerados. Es también interesante observar que la variación no es simétrica, observándose mayores diferencias cuando se reduce el precio para las tablas B.1, B.3 y B.5 y mayores diferencias cuando se aumenta el precio para las tablas B.2, B.4 y B.6.
- Precio de la red alternativa p_{ext}^{od} . En términos de magnitudes, en las tablas B.1-B.6 el nivel de sensibilidad es muy similar al de la modificación del precio propio. Las mayores diferencias se observan cuando se aumenta el precio (que es el escenario que lleva a que la nueva red gane más demanda).
- Costes fijos de construcción $(c_{A_{ij}}, c_{S_i})$. Para las tablas B.1, B.3 y B.5, la solución obtenida con menores valores de costes fijos no es factible en el caso original, puesto que es posible construir más capacidad con el mismo presupuesto. En caso de aumentar los precios, se puede observar que la variación de la función objetivo es positiva, puesto que la red se construye con un presupuesto que en realidad es menor al utilizado en el caso original. En las tablas B.2, B.4 y B.6 se observa que los valores positivos de la fila $c_{S_i}, c_{A_{ij}}$ (asociados a cambios positivos de los costes fijos) son

Cambio Variable	-20 %	-15 %	-10 %	-5 %	0 %	5 %	10 %	15 %	20 %
w^{od}	Inf.	Inf.	Inf.	Inf.	-	1.73	3.31	5.12	6.70
t_{ext}^{od}	0.97	0.59	0.30	-0.01	-	0.19	1.49	1.94	2.26
p_{ij}	0.23	0.18	0.13	0.07	-	0.05	0.08	0.09	0.06
p_{ext}^{od}	0.04	0.09	0.08	0.06	-	0.07	0.14	0.19	0.24
$c_{S_i}, c_{A_{ij}}$	Inf.	Inf.	Inf.	Inf.	-	0.46	1.52	2.17	2.12
$\bar{c}_{S_i}, \bar{c}_{A_{ij}}$	Inf.	Inf.	Inf.	Inf.	-	0.54	0.94	1.73	2.29
$c_{o_{ij}}$	1.08	0.53	0.20	0.29	-	0.20	0.16	0.21	0.13

Tabla B.1: Análisis de sensibilidad en distintas variables de estado para $\lambda = 5$ y $\beta = 1.25$. “Inf.” indica que la solución no es factible para el problema original. Para generar el valor del porcentaje mostrado en cada una de las celdas: 1) se obtiene la solución generada por el algoritmo 2 utilizando como parámetros los asociados a la variación “Cambio” del parámetro indicado en la fila; 2) se evalúa la función objetivo del problema MIP con la configuración de parámetros original en la solución generada en 1); y 3) se evalúa el cambio relativo entre el valor obtenido en el paso 2 y el valor de la función objetivo para el caso original (“Cambio” = 0 %).

Cambio Variable	-20 %	-15 %	-10 %	-5 %	0 %	5 %	10 %	15 %	20 %
w^{od}	-21.03	-15.73	-10.40	-5.03	-	5.55	11.04	16.88	22.58
t_{ext}^{od}	-24.18	-17.70	-11.55	-5.73	-	5.62	12.11	17.52	22.69
p_{ij}	-0.25	-0.18	-0.11	-0.05	-	0.17	0.32	0.44	0.54
p_{ext}^{od}	-0.9	-0.62	-0.39	-0.18	-	0.31	0.60	0.89	1.17
$c_{S_i}, c_{A_{ij}}$	-1.55	-1.13	-0.74	-0.13	-	0.46	1.52	2.17	2.12
$\bar{c}_{S_i}, \bar{c}_{A_{ij}}$	-1.3	-0.92	-0.46	-0.05	-	0.54	0.94	1.73	2.29
$c_{o_{ij}}$	-4.13	-3.33	-2.35	-0.95	-	1.40	2.53	3.65	4.61

Tabla B.2: Análisis de sensibilidad en distintas variables de estado para $\lambda = 5$ y $\beta = 1.25$. Para generar el valor del porcentaje mostrado en cada una de las celdas: 1) se obtiene la solución generada por el algoritmo 2 utilizando como parámetros los asociados a la variación “Cambio” del parámetro indicado en la fila; 2) se evalúa la función objetivo del problema MIP con la configuración de parámetros de 1) en la solución generada en 1); y 3) se evalúa el cambio relativo entre el valor obtenido en el paso 2 y el valor de la función objetivo para el caso original (“Cambio” = 0 %).

iguales que los valores reportados en dicha fila para las tablas B.1, B.3 y B.5. Esto se debe a que: a) los parámetros c_{S_i} y $c_{A_{ij}}$ no aparecen en la función objetivo y b) la solución obtenida es la misma que la obtenida en las tablas anteriores. La situación para los cambios negativos de los costes fijos es diferente puesto que, en este caso, las soluciones son factibles y, por lo tanto, dan lugar a porcentajes negativos.

- Costes lineales de construcción ($\bar{c}_{A_{ij}}, \bar{c}_{S_i}$). El comportamiento de la sensibilidad es muy parecido al de los costes fijos de construcción, con un nivel de sensibilidad algo menor que en el caso anterior.
- Costes de operación ($c_{o_{ij}}$). En las tablas B.1, B.3 y B.5 se observa que la sensibilidad es menor que la obtenida para los costes de construcción. Al tratarse de una variable que está en el objetivo pero que no forma parte de las restricciones, no presenta problemas de infactibilidad. Por el contrario, en las tablas B.2, B.4 y B.6 se observa que la variación del objetivo es mayor para los costes de operación que para los costes de construcción. Esto se debe a que, en estas tablas, la magnitud de los parámetros de optimización influye en el resultado. Al estar los costes operativos en la función objetivo y los costes de construcción no, la variación observada es mayor para los experimentos donde se modifica el valor de los costes operativos.

Cambio Variable	-20 %	-15 %	-10 %	-5 %	0 %	5 %	10 %	15 %	20 %
w^{od}	Inf.	Inf.	Inf.	Inf.	-	1.47	2.80	4.35	5.37
t_{ext}^{od}	0.43	-0.06	-0.03	0.43	-	0.30	0.89	1.59	2.41
p_{ij}	0.20	0.16	0.12	0.01	-	0.02	0.03	0.13	0.17
p_{ext}^{od}	0.18	0.14	0.02	0.03	-	0.01	0.13	0.17	0.21
c_{S_i} y $c_{A_{ij}}$	Inf.	Inf.	Inf.	Inf.	-	0.40	0.83	1.65	2.29
$\bar{c}_{S_i}$ y $\bar{c}_{A_{ij}}$	Inf.	Inf.	Inf.	Inf.	-	0.52	0.88	1.43	1.55
$c_{o_{ij}}$	0.59	0.42	0.13	0.24	-	0.22	0.39	0.41	0.28

Tabla B.3: Análisis de sensibilidad en distintas variables de estado para $\lambda = 5$ y $\beta = 1.50$. El significado del valor mostrado en cada celda es análogo al explicado en el subtítulo de la tabla B.1.

Cambio Variable	-20 %	-15 %	-10 %	-5 %	0 %	5 %	10 %	15 %	20 %
w^{od}	-21.50	-15.73	-10.12	-4.92	-	5.44	10.85	16.57	21.83
t_{ext}^{od}	-25.06	-18.70	-12.10	-5.46	-	5.97	11.98	17.92	23.96
p_{ij}	-0.22	-0.15	-0.09	-0.09	-	0.13	0.23	0.44	0.58
p_{ext}^{od}	-0.80	-0.59	-0.47	-0.22	-	0.25	0.61	0.90	1.18
c_{S_i} y $c_{A_{ij}}$	-2.16	-1.59	-0.69	0.07	-	0.40	0.83	1.65	2.28
$\bar{c}_{S_i}$ y $\bar{c}_{A_{ij}}$	-1.86	-0.97	-0.23	0.05	-	0.52	0.88	1.43	1.55
$c_{o_{ij}}$	-4.02	-2.94	-2.06	-0.84	-	1.27	2.46	3.46	4.27

Tabla B.4: Análisis de sensibilidad en distintas variables de estado para $\lambda = 5$ y $\beta = 1.50$. El significado del valor mostrado en cada celda es análogo al explicado en el subtítulo de la tabla B.2.

En resumen, los resultados numéricos ponen de manifiesto que la solución óptima generada por el algoritmo 2 se comporta de forma estable ante potenciales cambios de (o presencia de incertidumbre en) los valores de las variables de estado. Más específicamente, los resultados han revelado que: 1) la variación del valor objetivo con respecto a las variables de estado es relativamente pequeña para todas las variables consideradas; 2) la variable de estado que da lugar a una mayor sensibilidad es la demanda e , incluso para este caso, la curva de sensibilidad es lineal y tiene una pendiente unitaria; y 3) los resultados son muy similares en los tres escenarios de costes fijos y presupuesto (parametrizados por distintos valores de λ y β considerados). Todo esto indica que el método de diseño de RdT propuesto en este TFG es relativamente estable y, una vez más, justifica su uso en redes reales.

Cambio Variable	-20 %	-15 %	-10 %	-5 %	0 %	5 %	10 %	15 %	20 %
w^{od}	Inf.	Inf.	Inf.	Inf.	-	1.56	2.88	4.12	5.58
t_{ext}^{od}	0.69	0.15	-0.01	0.42	-	0.13	0.76	1.66	2.30
p_{ij}	0.11	0.06	0.01	0.02	-	0.04	0.08	0.11	0.15
p_{ext}^{od}	0.16	0.12	0.08	0.04	-	0.01	0.01	0.07	0.12
c_{S_i} y $c_{A_{ij}}$	Inf.	Inf.	Inf.	Inf.	-	0.38	0.89	1.74	2.02
$\bar{c}_{S_i}$ y $\bar{c}_{A_{ij}}$	Inf.	Inf.	Inf.	Inf.	-	0.45	0.76	1.06	1.56
$c_{o_{ij}}$	0.42	0.41	0.12	0.24	-	0.26	0.39	0.40	0.27

Tabla B.5: Análisis de sensibilidad en distintas variables de estado para $\lambda = 6$ y $\beta = 1.25$. El significado del valor mostrado en cada celda es análogo al explicado en el subtítulo de la tabla B.1.

Cambio Variable	-20 %	-15 %	-10 %	-5 %	0 %	5 %	10 %	15 %	20 %
w^{od}	-21.41	-15.64	-10.08	-4.88	-	5.49	10.83	16.16	21.78
t_{ext}^{od}	-24.77	-18.39	-12.00	-5.42	-	5.71	11.61	17.61	23.26
p_{ij}	-0.32	-0.26	-0.21	-0.09	-	0.15	0.29	0.43	0.57
p_{ext}^{od}	-0.81	-0.60	-0.40	-0.20	-	0.24	0.48	0.79	1.08
c_{S_i} y $c_{A_{ij}}$	-2.07	-1.65	-0.94	0.04	-	0.38	0.89	1.74	2.02
$\bar{c}_{S_i}$ y $\bar{c}_{A_{ij}}$	-1.82	-0.87	-0.19	0.09	-	0.45	0.76	1.06	1.56
$c_{o_{ij}}$	-4.37	-3.07	-2.16	-0.88	-	1.34	2.51	3.52	4.35

Tabla B.6: Análisis de sensibilidad en distintas variables de estado para $\lambda = 6$ y $\beta = 1.25$. El significado del valor mostrado en cada celda es análogo al explicado en el subtítulo de la tabla B.2.