



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

MODELO DE OPTIMIZACIÓN ROBUSTO PARA EL DISEÑO DE REDES BAJO
CONDICIONES DE EQUILIBRIO

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN
GESTIÓN DE OPERACIONES

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL INDUSTRIAL

AZUCENA MARCELA ORELLANA INSUNZA

PROFESOR GUÍA:
FERNANDO ORDÓÑEZ PIZARRO

PROFESOR CO-GUÍA:
VÍCTOR BUCAREY LÓPEZ

COMISIÓN:
ANDRÉS WEINTRAUB POHORILLE

SANTIAGO DE CHILE
2022

RESUMEN DE LA MEMORIA PARA OPTAR AL
TÍTULO DE INGENIERA CIVIL INDUSTRIAL Y AL
GRADO DE MAGÍSTER EN GESTIÓN DE OPERACIONES
POR: AZUCENA MARCELA ORELLANA INSUNZA
FECHA: 2022
PROF. GUÍA: FERNANDO ORDÓÑEZ PIZARRO

MODELO DE OPTIMIZACIÓN ROBUSTO PARA EL DISEÑO DE REDES BAJO CONDICIONES DE EQUILIBRIO

Los problemas de diseño de red se utilizan en muchos contextos, como por ejemplo, para modelar el transporte de pasajeros que se mueven desde y hacia distintos lugares dentro de una ciudad como también para la transmisión de mensajes en una red de comunicación o transmisión de energía en una red eléctrica. En este trabajo de tesis, se busca enfocar el problema de diseño de red para modelar el tránsito de personas de un punto a otro dentro de un área urbana incorporando su comportamiento en este tipo de situaciones.

Para ello, se considera el problema de diseño de redes modelado a través de una programación binivel, donde un nivel de decisión determina el comportamiento egoísta de los usuarios de la red que buscan escoger el camino más corto entre sus orígenes y destinos, mientras que en el otro nivel, un planificador central decide reducir los costos de la red completa. Además de esto, se considera que la demanda del problema es incierta y se incorpora al modelo utilizando una metodología de optimización robusta. Se presentan dos formulaciones de modelos que se diferencian principalmente por el tipo de función de latencia que se considera. En una se utiliza una función de latencia lineal y se incorporan restricciones de capacidad, mientras que en la otra se ingresa la capacidad en una función de latencia estrictamente convexa, pudiendo remover las restricciones de capacidad.

Con el fin de aproximar el problema robusto para que pueda ser computacionalmente tratable, se demuestra que si el conjunto de incertidumbre es un poliedro y, además, la función de costos es convexa, entonces resolver el problema de diseño de redes con demandas que sean vértices del conjunto de incertidumbre permite determinar cotas superiores e inferiores al valor óptimo del problema de diseño de redes robusto. Por lo que en definitiva, se encuentra una solución robusta que es una aproximación a la solución óptima del problema robusto original.

Los resultados computacionales muestran que en instancias pequeñas el modelo funciona eficientemente con bajos tiempos de ejecución, sin embargo, si el tamaño del problema crece, la resolución en forma exacta incurre en altos costos computacionales. No obstante, en las instancias pequeñas se observa que la solución robusta aproximada tiene una modesta sub-optimalidad en escenarios deterministas cerca del promedio mientras que reduce significativamente el costo del peor caso. En particular, a medida que la incertidumbre crece, la protección del modelo robusto en el peor caso es mayor.

En síntesis, en esta tesis se plantea, resuelve e implementa un modelo de optimización de diseño de redes que es robusto respecto a demandas inciertas y que minimiza costos de tiempos de viaje sujeto a restricciones de presupuesto, donde los usuarios se comportan satisfaciendo condiciones de equilibrio.

"No puedes conectar los puntos mirando hacia adelante; sólo puedes hacerlo mirando hacia atrás. Así que tienes que confiar en que los puntos se conectarán de alguna forma en el futuro."

Steve Jobs

Agradecimientos

Cuando uno llega hasta acá y mira el pasado, es imposible no recordar y reconocer a todas esas personas y situaciones que te ayudaron y construyeron el camino que recorriste para poder llegar a este destino. Este trabajo, este logro en mi vida, fue realizado por muchas más personas que yo. Ellos son los puntos que conecto hoy cuando miro hacia atrás y no me queda más que agradecerles.

Agradezco a Dios, que me hizo y me hace ver todos los días lo que soy capaz de lograr aún cuando tengo miedo a fracasar. Su amor es siempre fiel y estoy segura que Él me bendice a través de mi familia, amigos y todas esas personas que pone en mi camino. A papá, porque siempre deseó darme lo mejor regalándome una vida próspera, y esta carrera es prueba de su esfuerzo por nosotras. A mamá, que siempre está en los detalles de mi vida, y esos detalles son mi soporte y mi seguridad. A mi hermana mayor, Amapola, por reconocermme en el interior y por descifrar exactamente lo que necesito. A mi hermana menor, Jazmín, por amarme a través de esas cosas que no se ven, pero que son las que más aprecio (sí, ella me cocinaba y me alimentaba, cuando yo estaba estudiando). A Matías, por ser ese compañero sincero, por empezar este camino juntos y ayudarme a crecer y a mejorar a pesar de las dificultades que, sí, fueron varias.

En este logro académico, agradezco a mi profesores guías Fernando Ordoñez y Víctor Bucarey, por darme la oportunidad de trabajar con ellos y de aprender de ellos. A Fernando, gracias por confiar en mis habilidades y darme oportunidades para crecer profesionalmente. A Víctor, por mostrarme lo genial que es esta área de modelos matemáticos y optimización, que decidí especializarme, a través de su enseñanza cercana en ese primer curso de industrias.

Sin duda, agradezco a esas personas que me eligieron para formar parte de su vida: a mis amigos. Gracias Karen y Karín, por haber hecho del colegio el lugar más divertido, nuestra amistad formada allí no se rompe, aunque pase el tiempo. A mis amigos de Plan Común, las Ratitas, que llenaron de recuerdos esos dos primeros años de estudio. En especial, agradezco a Vale, porque su amistad ha perdurado desde el primer día que ingresé a mi carrera. A Cynthia, Belén y Magda, las mejores amigas que la universidad me permitió tener, una palabra las describe y es que son incondicionales. Y también, a Lucca, por ser mi mejor amigo. Gracias por apoyarme, ayudarme, soportarme y entenderme. Por último agradezco a Simpliroute, mi equipo de Data Science, a Javiera, mi mejor amiga en este equipo y a Víctor, por ser el mejor primer jefe.

Azucena Orellana Insunza

Tabla de Contenido

| | |
|--|----------|
| 1. Introducción | 1 |
| 1.1. Definición del problema | 1 |
| 1.2. Objetivos | 6 |
| 1.2.1. Objetivo general | 6 |
| 1.2.2. Objetivos específicos | 6 |
| 1.3. Alcances | 7 |
| 1.4. Estructura de la tesis | 8 |
| 2. Marco Conceptual | 9 |
| 2.1. Formulación del problema de diseño de redes | 10 |
| 2.1.1. Funciones de latencia | 12 |
| 2.2. Problemas de programación binivel | 14 |
| 2.3. Equilibrio en juegos de red | 17 |
| 2.3.1. Equilibrio de Wardrop | 17 |
| 2.3.2. Precio de la anarquía | 20 |
| 2.4. Optimización robusta | 21 |
| 2.4.1. Contraparte robusta | 23 |
| 2.4.2. Conjuntos de incertidumbre | 25 |
| 2.4.3. Métodos de resolución | 28 |
| 2.5. Trabajos relacionados | 29 |

| | |
|--|-----------|
| 3. Modelos | 34 |
| 3.1. Modelos deterministas | 36 |
| 3.1.1. Modelo 1: Función de costo lineal y con restricciones de capacidad . . | 37 |
| 3.1.2. Modelo 2: Función de latencia BPR y sin restricciones de capacidad . | 42 |
| 3.2. Modelos robustos | 44 |
| 3.2.1. Extensión del modelo determinista a escenarios inciertos de demanda | 44 |
| 3.2.2. Aproximación de la contraparte robusta ajustable | 45 |
| | |
| 4. Resultados Computacionales | 54 |
| 4.1. Implementación | 55 |
| 4.1.1. Conjunto de incertidumbre | 55 |
| 4.1.2. Métricas de evaluación y tipos de análisis | 55 |
| 4.2. Resultados en redes pequeñas | 57 |
| 4.2.1. Soluciones gráficas y simulaciones | 59 |
| 4.2.2. Análisis variando presupuesto | 62 |
| 4.2.3. Análisis variando nivel de incertidumbre | 64 |
| 4.3. Resultados en red Sioux Falls | 66 |
| 4.3.1. Análisis variando la discretización de valores enteros | 68 |
| | |
| 5. Conclusiones | 72 |
| 5.1. Conclusiones generales | 72 |
| 5.2. Trabajo futuro | 75 |
| | |
| Bibliografía | 76 |
| | |
| Anexos | 82 |
| A . Aproximación de la distribución normal para simular demandas discretas . . | 82 |
| B . Costos de viaje óptimos para distintos presupuestos | 84 |
| B .1. Red pequeña | 84 |
| B .2. Red Sioux Falls | 84 |

Índice de Tablas

| | |
|---|----|
| 2.1. Precio de la anarquía para clases comunes de funciones de latencia por arco. Los coeficientes polinomiales son no negativos (Roughgarden, 2003). | 21 |
| 4.1. Costos y diferencia relativa de la solución robusta en cada caso, variando el presupuesto B . Modelo 1 . Grafo 1.1, con incertidumbre en el par $OD = (2,3)$. | 63 |
| 4.2. Costos y diferencia relativo de la solución robusta en cada caso, variando el presupuesto B . Modelo 2 . Grafo 1.1, con incertidumbre en el par $OD = (2,3)$. | 63 |
| 4.3. Costos y cambio relativo al usar la solución robusta en cada caso. Se considera dos discretizaciones diferentes del aumento de capacidad. Modelo 1 . Grafo Sioux Falls. | 69 |

Índice de Ilustraciones

| | |
|---|----|
| 1.1. Grafo de una red simple. Cada arco posee una capacidad (c), un costo de inversión (g) y un tiempo de viaje (t). | 4 |
| 1.2. Solución factible para el grafo de la Figura 1.1 considerando una demanda para el par $OD = (1, 4)$ igual a 10 y para el par $OD = (2, 3)$ igual a 15. La variable x representa la cantidad de flujo por cada arco. | 5 |
| 2.1. Función <i>Bureau of Public Roads</i> (BPR) que relaciona el tiempo de viaje con el flujo del arco. Está basada en el tiempo de viaje T_{ff} bajo condiciones de libre flujo y en la capacidad c del arco (Maerivoet and De Moor, 2005). . . . | 14 |
| 4.1. Aumento de capacidad y flujos óptimos x por cada solución en el escenario promedio. Modelo 1. Grafo 1.1, con incertidumbre en el par $OD = (2, 3)$. . . | 59 |
| 4.2. Aumento de capacidad y flujos óptimos x por cada solución en el escenario promedio. Modelo 2. Grafo 1.1, con incertidumbre en el par $OD = (2, 3)$. . . | 59 |
| 4.3. Simulación de escenarios de demanda y su distribución de costos. La línea punteada muestra el promedio de la distribución y el área sombreada, su desviación estándar. Modelo 1. Grafo 1.1, con incertidumbre en el par $OD = (2, 3)$ | 61 |
| 4.4. Simulación de escenarios de demanda y su distribución de costos. La línea punteada muestra el promedio de la distribución y el área sombreada, su desviación estándar. Modelo 2. Grafo 1.1, con incertidumbre en el par $OD = (2, 3)$ | 62 |

| | | |
|------|---|----|
| 4.5. | Diferencia relativa del costo de la solución robusta en el caso promedio y en el peor caso, como función del tamaño (sd) del conjunto de incertidumbre. Modelo 2. Grafo 1.1, con incertidumbre en el par OD = (2,3). | 64 |
| 4.6. | Diferencia relativa del costo de la solución robusta en el caso promedio y en el peor caso, como función del tamaño (sd) del conjunto de incertidumbre . Modelo 2. Grafo 1.1, incertidumbre en ambos pares OD y $\Gamma = 1$ | 65 |
| 4.7. | Diferencia relativa del costo de la solución robusta en el caso promedio y en el peor caso, como función del tamaño (sd) del conjunto de incertidumbre . Modelo 2. Grafo 1.1, incertidumbre en ambos pares OD y $\Gamma = 2$ | 66 |
| 4.8. | Red Sioux Falls (Long et al., 2015) | 67 |
| 4.9. | Soluciones z . Experimento 5. Grafo Sioux Falls. Modelo 1. El grosor de los arcos indica el aumento de capacidad. Los colores en los nodos indican los orígenes y destino de los pares OD críticos. | 70 |
| 5.1. | Curva de distribución normal. Se muestran la probabilidades para distintos intervalos. | 83 |
| 5.2. | Costo de viaje óptimo en el escenario determinista en función del presupuesto B. Modelo 1. Grafo 1.1. | 84 |
| 5.3. | Costo de viaje óptimo en el escenario determinista en función del presupuesto B. Modelo 1. Grafo <i>Sioux Falls</i> (4.8). | 85 |

Capítulo 1

Introducción

1.1. Definición del problema

Los problemas de diseño de red (o *Network Design Problem* en inglés) se utilizan en muchos contextos, como por ejemplo, para modelar el transporte de pasajeros que se mueven desde y hacia distintos lugares dentro de una ciudad, como también para la transmisión de mensajes en una red de comunicación entre diferentes *switching centers*, o incluso, abarcando el área de sistemas de energía, la red puede representar líneas eléctricas de mediana o alta tensión que permiten transmitir energía entre plantas eléctricas y puntos de demanda (Minoux, 1989). En la presente tesis, se busca enfocar el problema de diseño de red para modelar el tránsito de personas de un punto a otro dentro una área urbana incorporando su comportamiento en este tipo de situaciones. En la práctica, resolver un problema de estas características serviría para modelar el tráfico de vehículos, o de transporte público en un sector poblacional, así como también, modelar redes de telefonía o de internet, los cuales tienen congestión en las líneas de comunicación entre los distintos puntos de conexión.

Un problema de diseño de red se puede modelar como un problema de flujo a costo mínimo *multi-commodity* en donde se satisface la demanda y también se considera la inversión de aumentar la capacidad de los arcos (Minoux, 1989). En este tipo de problemas se evalúa el *trade-off* entre los costos de utilización de los arcos (por ejemplo, tiempos de viaje) y

los costos de inversión para la construcción de redes. Una red más grande tiene un costo de construcción más alto, pero puede reducir los costos de utilización e incluir rutas para distintos puntos de origen y destino más atractivos. Por el contrario, una red más pequeña puede aumentar los costos de utilización pero es más barata de construir. Encontrar una solución óptima para este problema permite poder equilibrar ambos costos.

Un enfoque más simple al modelar este problema es considerar que las personas no toman decisiones respecto a su trayecto, sino más bien, se les asigna un camino a seguir. Sin embargo, esto dista mucho de la realidad. En esencia, los usuarios se comportan de manera egoísta y toman decisiones buscando su propio beneficio. En definitiva, si se considera que las personas que ocupan la red son racionales, entonces, es posible distinguir tipos de decisiones diferentes en este problema. Por un lado están los usuarios que utilizan la red, quienes buscan encontrar el camino más corto para llegar a su destino, y por lo tanto, quieren disminuir sus propios tiempos de viajes. Y por otro lado se encuentra un planificador de la red, que tiene el poder de decidir incrementar la capacidad de los arcos (por ejemplo, construir una pista más a la calle o aumentar el número de buses para el transporte de pasajeros) y que busca principalmente reducir el tiempo total de viaje de todos los transeúntes a la vez de invertir lo menos posible en la construcción de capacidad.

En consecuencia, en este problema existen dos tipos de decisiones que se pueden considerar y una forma de modelar esto es a través de un problema binivel (es decir, un problema de dos niveles o etapas (Candler and Norton, 1977)). Así, el planificador puede considerar los resultados de las decisiones de los usuarios y utilizar esta información para diseñar la red de manera óptima. La decisión del planificador se encuentra en el primer nivel y busca típicamente minimizar los costos totales de transporte sujeto a una restricción presupuestaria para expandir la capacidad (nivel superior). Mientras que en el segundo nivel está la decisión de los usuarios. Cada persona escoge su propio camino minimizando sus costos, por lo que es necesario llegar a una situación de equilibrio donde se considere todas las decisiones individuales de los transeúntes y en la que ninguno tenga incentivo a desviarse (nivel inferior).

Este segundo nivel es posible modelarlo matemáticamente a través de juegos de red (*Network Games* en inglés) (Correa and Stier-Moses, 2011).

Hasta este punto de la descripción del problema se ha considerado un supuesto relevante: no hay variación en la cantidad de usuarios que utilizan la red como tampoco se considera que las personas pueden cambiar su punto de origen y de destino. En términos más concretos, la demanda de cada par de origen-destino, es decir, la cantidad de usuarios que tiene un mismo origen y destino, es determinista. ¿Es posible considerar que esta demanda sea aleatoria? ¿Cómo se puede abordar la incertidumbre de este tipo de datos en el problema? Considerar la incertidumbre de los parámetros permite enriquecer los modelos deterministas porque incorporan el dinamismo que está presente en las situaciones reales. Por ejemplo, los tiempos de viaje entre un punto y otro dependen de condiciones del tráfico que pueden ser inciertas, producto del clima o de accidentes, o los usuarios no siempre ocupan el mismo trayecto para desplazarse ni tampoco viajan todos los días al mismo lugar. Además, el problema de diseño de redes es una decisión del tipo estratégica, es decir, no se toma todos los días, por ende, la solución a este problema debería considerar escenarios que se pueden presentar en 5 años o más, y con lo único que se cuenta en el presente son con estimaciones de demanda y tiempos de viaje. Por lo que tener en cuenta estas variaciones inciertas requiere desarrollar modelos más complejos que se adapten a los cambios que se presenten.

Obsérvese el siguiente ejemplo presentado en la Figura 1.1. Este es un grafo simple de cuatro nodos y cinco arcos. Cada arco posee una capacidad (c), un costo por aumentar una unidad de capacidad (g) y un costo lineal de viaje por cada unidad de flujo (t).

Cada nodo de este grafo podría representar sectores de una ciudad a los que se puede viajar utilizando la red de calles. Pero es posible, también, cruzar de un punto a otro de la ciudad utilizando una autopista ($arco = (2, 3)$). Las calles tienen tiempos de viaje distintos, pero mayores a la de la autopista ($t = 2$), puesto que por ella se puede ir a mayor velocidad. También, la autopista permite una mayor capacidad (puede tener más de una pista por sentido). Sin embargo, aumentar la capacidad de la autopista es mucho más costoso que

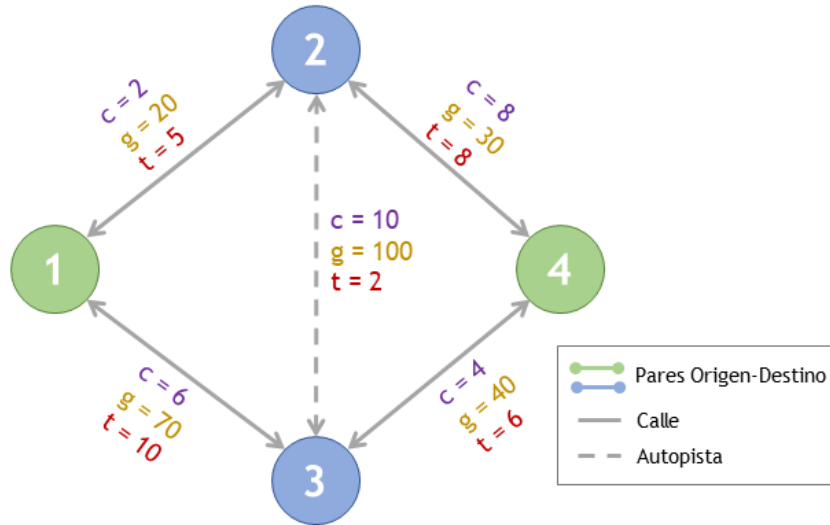


Figura 1.1: Grafo de una red simple. Cada arco posee una capacidad (c), un costo de inversión (g) y un tiempo de viaje (t).

cualquier otra calle ($g = 100$). La decisión de este problema consiste en determinar dónde y cuánto se debe aumentar la capacidad en los arcos, y por dónde transitan los usuarios de la red (la variable utilizada para esto es x , denominada flujo).

Supongamos que existen dos tipos de pasajeros: los que van desde el nodo 1 al nodo 4 y los que van del nodo 2 al nodo 3 (este tipo de clasificación es a través de pares de origen-destino, en adelante, OD). Lo más natural es que los usuarios que quieren ir desde el nodo 2 al nodo 3 utilicen la autopista, pero a su vez los otros tipos de usuarios también podrían utilizarla para ahorrar tiempo. El problema es que la capacidad de la autopista podría no ser suficiente. No obstante, podría ocurrir lo contrario, los usuarios que utilizan la autopista podrían preferir ocupar las calles debido a la congestión de la autopista, limitando la capacidad de las mismas.

Si suponemos que la cantidad de usuarios que utilizan el par $OD = (2, 3)$ es 15 y los que utilizan el par $OD = (1, 4)$ es 10, entonces una solución factible sería la que se observa en la Figura 1.2. Se privilegiaría que los del par $OD = (2, 3)$ utilicen la autopista, teniendo que aumentar la capacidad lo suficiente para que puedan transitar, y a los demás, se les permitiría ocupar las calles, que estarán más descongestionadas, teniendo sólo que incrementar capacidad a la calle más barata ($arco = (1, 2)$).

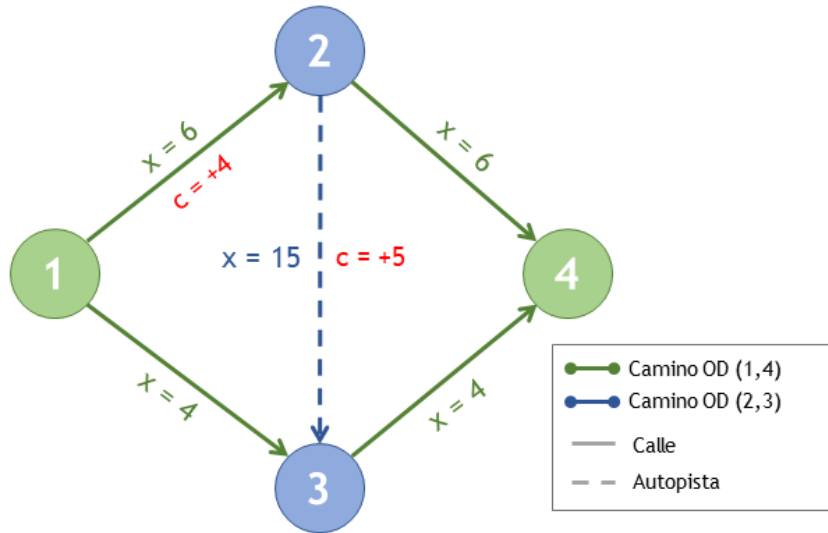


Figura 1.2: Solución factible para el grafo de la Figura 1.1 considerando una demanda para el par $OD = (1, 4)$ igual a 10 y para el par $OD = (2, 3)$ igual a 15. La variable x representa la cantidad de flujo por cada arco.

Ahora, supongamos que la cantidad de usuarios con origen en 2 y que van hacia 3 es incierta. Es decir, no se sabe con exactitud cuántos van a transitar por la autopista. ¿Se incrementaría las mismas unidades de capacidad en la autopista? ¿Será que la solución óptima sería privilegiar que estos usuarios sólo ocupen la autopista? Probablemente lo razonable sería aumentar la capacidad de las calles y que, dependiendo de la demanda, algunos circulen por ellas. Así, no se subutiliza la autopista cuando la demanda sea baja, más aún cuando la inversión en mejorar su capacidad es costosa.

En resumen, el problema a resolver en esta tesis consiste en diseñar y balancear óptimamente los costos de transporte y los costos de inversión de una red sujeta a demandas inciertas, donde los usuarios se comportan buscando maximizar su utilidad y el problema satisface condiciones de equilibrio. En la literatura, se han desarrollado modelos de diseño de red con equilibrio (Davis, 1994; Luatkep et al., 2011; Wang et al., 2013) sin embargo este trabajo de tesis amplía el análisis a escenarios inciertos, específicamente, a una demanda incierta.

Por lo que en definitiva, considerando todos los aspectos mencionados, este trabajo de tesis consiste en responder las siguientes preguntas:

- ¿Es posible desarrollar un modelo de diseño de red que incorpore dos niveles de decisión, condiciones de equilibrio en el tráfico y además incertidumbre en los parámetros? ¿Cuán complejo es resolverlo?
- ¿Cómo se puede modelar la incertidumbre en la demanda?
- ¿Cómo cambia la solución al considerar incertidumbre en la demanda en comparación al escenario determinista? ¿En qué escenarios es mejor y en cuáles es peor?
- ¿Cómo varía la solución al invertir más (o menos) en la construcción de capacidad de los arcos al considerar incertidumbre?
- ¿Qué relación existe entre la magnitud de la incertidumbre considerada y los costos del problema?
- ¿Qué tan complejo se vuelve resolver el problema en contextos más realistas?

1.2. Objetivos

1.2.1. Objetivo general

El objetivo de esta tesis es plantear, resolver e implementar un modelo de optimización para diseñar una red sujeta a demandas inciertas, que minimice costos de tiempo de viaje y se limite a una restricción de presupuesto para la inversión, donde los usuarios se comportan satisfaciendo condiciones de equilibrio.

1.2.2. Objetivos específicos

1. Desarrollar un modelo de optimización determinista que diseñe una red que satisfaga condiciones de equilibrio y minimice costos de viaje sujeto a una restricción presupuestaria para la inversión.
2. Extender el modelo planteado considerando escenarios de incertidumbre en la demanda.

3. Evaluar la utilización de modelos de optimización robusto para abordar la incertidumbre en el modelo.
4. Implementar la resolución en forma exacta de los modelos de optimización determinista y robusto.
5. Evaluar los resultados entregados por los modelos utilizando tanto instancias creadas como instancias aplicadas.

1.3. Alcances

Como se ha mencionado en la sección anterior, el problema de diseño de red puede abarcar muchas áreas de estudio. Sin embargo, este proyecto de tesis se enfoca en desarrollar modelos que permitan resolver el problema para una red de transporte, donde los usuarios de la red podrían ser pasajeros o vehículos. Por lo tanto, estos modelos pueden no ser aplicables en otros tipos de redes, puesto que se analiza el comportamiento intrínseco de las personas para considerar un equilibrio en el tráfico.

Al incorporar escenarios inciertos al problema de diseño de red, es posible incorporar incertidumbre en los tiempos de viajes como en la demanda. En este proyecto de tesis se escoge considerar sólo uno de estos aspectos de manera de simplificar el modelo. Se opta por la demanda como incierta dado que es posible modelar los costos de transporte a través de funciones de latencia no lineales, incorporando los efectos de congestión.

Dentro de los objetivos de esta tesis no se encuentra investigar metodologías algorítmicas o desarrollar heurísticas que puedan resolver el problema planteado en tiempos razonables. Más bien, se decide resolver los modelos en forma exacta en instancias creadas y en instancias utilizadas en la literatura, con el objetivo de evaluar el comportamiento de las soluciones obtenidas frente a circunstancias inciertas.

1.4. Estructura de la tesis

La estructura de esta tesis sigue de esta manera: en el Capítulo 2 se revisa la literatura existente con respecto al problema de diseño de red y los distintos modelos que permiten incorporar niveles de decisión, condiciones de equilibrio en el tráfico y escenarios de incertidumbre. En el Capítulo 3 se plantean diferentes modelos de optimización que resuelven el problema propuesto y se identifican las condiciones que permiten ser resueltos de manera eficiente, demostrando algunas proposiciones. Se presentan los resultados computacionales y se analiza el comportamiento de las soluciones de los modelos en el Capítulo 4, para finalmente en el Capítulo 5, concluir y discutir los trabajos futuros que pueden enriquecer esta línea de investigación.

Capítulo 2

Marco Conceptual

Dentro de los problemas de optimización más estudiados en la literatura están los problemas de flujo en redes. Particularmente, en el trabajo de esta tesis se evalúa un tipo particular del problema de flujo en redes que busca minimizar los costos de viaje. El problema más simple de este tipo modela el flujo de un sólo producto a través de una red. Sin embargo, este problema se puede generalizar al considerar varios productos que necesitan utilizar la misma red subyacente y se denomina problema de flujo de múltiples productos (*Multicommodity Flow Problem* o MCFP por sus siglas en inglés).

Estos productos pueden ser diferenciados por sus características físicas o simplemente por donde parten y hacia donde tienen que llegar. De este modo, los distintos productos tienen diferentes orígenes y destinos, y cada producto tiene sus restricciones de conservación de flujo en cada nodo por separado. Sin embargo, lo que une a estos diferentes productos es que deben compartir las capacidades de los arcos que tienen en común. Particularmente, la decisión esencial que aborda este problema es la asignación de la capacidad de cada arco a los productos individuales de una manera que minimice los costos generales de la red (Ahuja et al., 1993).

Los problemas de flujo de múltiples productos surgen en muchas situaciones prácticas, como por ejemplo, el transporte de pasajeros de diferentes orígenes a diferentes destinos dentro de una ciudad, el ruteo de vehículos no homogéneos (no homogéneos en términos de

velocidad, capacidad de transporte y costos operacionales) o la transmisión de mensajes en una red de comunicación entre diferentes pares de orígenes-destinos, entre otras áreas.

Aunque existe mucha investigación relacionada al problema de flujo multi-producto durante las últimas cuatro décadas, este problema es aún considerado un problema de programación matemáticamente difícil, especialmente cuando hay un gran número de productos involucrados. Resolver el problema general de flujo multi-producto con variables enteras es NP-completo (Karp, 1975). De hecho, incluso resolver el problema de flujo entero con sólo dos productos es NP-completo (Even et al., 1976). Por otra parte, el problema lineal de flujo multi-producto puede teóricamente ser resuelto en tiempo polinomial (Tardos, 1986). Sin embargo, computacionalmente, los problemas de MCF, especialmente los problemas de MCF enteros, pueden requerir mucho tiempo de cálculo para resolverlos (Wang, 2018).

En la sección 2.1 se profundiza la formulación matemática del problema de flujo a costo mínimo como la formulación particular del problema de diseño de redes y cómo se puede modelar su función objetivo. Como el problema abordado en esta tesis extiende el problema de diseño de redes a más de un nivel de decisión, se estudian los problemas de optimización binivel en la sección 2.2. Considerando que las decisiones de los usuarios son egoístas, en la sección 2.3 se investigan las condiciones que permiten determinar un equilibrio en el tráfico de la red. En la sección 2.4, se estudia la metodología de optimización robusta que permite modelar y resolver la incertidumbre considerada en el problema, y finalmente, en la sección 2.5 se realiza una revisión de la literatura de los trabajos que relacionan todos los aspectos mencionados.

2.1. Formulación del problema de diseño de redes

La formulación del problema de flujo multi-producto a costo mínimo se define formalmente de la siguiente manera. Sean N el conjunto de todos los nodos en un grafo G , A el conjunto de todos los arcos, y K el conjunto de todos los productos. Para cada producto k con origen s_k y destino t_k , existe un costo por unidad de flujo c_{ij}^k en el arco (i, j) . Además, se define

como x_{ij}^k al flujo en el arco (i, j) del producto k . Dentro de los parámetros del problema, se denota como b_i^k a la oferta/demanda en el nodo i , y d^k a la demanda total del producto k . Por último, u_{ij} representa la capacidad del arco (i, j) . Sin pérdida de generalidad, se asume que cada unidad de cada producto ocupa una unidad de capacidad en cada arco donde transita.

$$\text{(MCFP) } \min \sum_{k \in K} \sum_{(i,j) \in A} c_{ij}^k x_{ij}^k \quad (2.1a)$$

$$\text{s.a. } \sum_{(i,j) \in A} x_{ij}^k - \sum_{(j,i) \in A} x_{ji}^k = b_i^k \quad \forall i \in N, \forall k \in K \quad (2.1b)$$

$$\sum_{k \in K} x_{ij}^k \leq u_{ij} \quad \forall (i, j) \in A \quad (2.1c)$$

$$x_{ij}^k \geq 0 \quad \forall (i, j) \in A, \forall k \in K \quad (2.1d)$$

donde $b_i^k = d^k$ si $i = s_k$, $b_i^k = -d^k$ si $i = t_k$, y $b_i^k = 0$ si $i \in N \setminus \{s_k, t_k\}$. Esta formulación tiene $|K||A|$ variables y $|N||K| + |A|$ restricciones no triviales.

Como se mencionó, una aplicación particular del problema de flujo de múltiples productos es el problema de diseño de redes (*Network Design Problem* en inglés). Este problema involucra, además de la decisión de la cantidad de flujo que atraviesa cada arco por cada producto utilizando restricciones de conservación de flujo (2.1b) y de capacidad (2.1c), algunas restricciones secundarias relacionadas con la decisión de aumentar la capacidad de los arcos o crear nuevos arcos en la red, todo esto minimizando los costos de viaje y acotando los costos de inversión. La función objetivo de este tipo problema puede ser no lineal (Wang, 2018).

Podemos definir este problema considerando que existe una variable adicional z_{ij} que corresponde al aumento (o disminución) de la capacidad del arco (i, j) y que se añade a las restricciones de capacidad (2.1c). También es posible considerar que la función de costos que se minimiza es no lineal y que depende del flujo total que transita por los arcos. Sea $g_{ij}(\cdot)$ una función de costos (que puede ser no lineal), el problema de diseño de red se define como:

$$\text{(NDP) } \min \sum_{(i,j) \in A} g_{ij}(x_{ij}) \quad (2.2a)$$

$$\text{s.a. } \sum_{(i,j) \in A} x_{ij}^k - \sum_{(j,i) \in A} x_{ji}^k = b_i^k \quad \forall i \in N, \forall k \in K \quad (2.2b)$$

$$x_{ij} = \sum_{k \in K} x_{ij}^k \quad \forall (i,j) \in A \quad (2.2c)$$

$$x_{ij} \leq u_{ij} + z_{ij} \quad \forall (i,j) \in A \quad (2.2d)$$

$$x_{ij}^k \geq 0 \quad \forall (i,j) \in A, \forall k \in K \quad (2.2e)$$

Otra formulación que se utiliza para el problema de diseño de redes es considerar la capacidad u_{ij} como una variable de decisión, con lo cual se toman decisiones de crear arcos cuando $u_{ij} \geq 0$ para poder cumplir con la demanda de los productos.

2.1.1. Funciones de latencia

Como ya fue mencionado, el modelo clásico de diseño de redes de transporte asigna a todos los viajeros con pares OD en común un camino en la red, en otras palabras, se les asigna espacio en los arcos para que transiten a un mínimo costo. Un aspecto fundamental de este enfoque es que asume que todos los flujos de la red están en equilibrio, en particular, que los flujos son constantes, independientes del tiempo y del tráfico. Esto ocurre principalmente porque el modelo clásico considera que el costo en cada arco es lineal con respecto al flujo que transita por él.

Más precisamente el modelo clásico asume que (1) los usuarios escogen la misma ruta a costo mínimo entre un mismo par de origen y destino, (2) que todos tienen la misma información acerca de los tiempos que demora transitar por las vías y (3) que ese tiempo de viaje es constante, es decir, es independiente de la cantidad de usuarios que ocupan la vía (Maerivoet and De Moor, 2005). En consecuencia, este modelo no toma en cuenta los efectos de congestión que se producen en las calles.

Sin embargo, lo común en las redes de transporte es considerar que cuanto mayor es el tráfico en las vías, mayor es el tiempo de viaje, porque existe una capacidad que no permite el libre flujo por la red. Ignorar los efectos de congestión, implicaría suponer que no existen efectos de espera cuando hay una mayor cantidad de usuarios utilizando las vías (lo que posiblemente ocasionará cambiar la ruta a un arco con costo más barato). En concreto, una variación al modelo que permita incorporar esto es considerar que el costo del arco será mayor ante un aumento del flujo, lo que implica utilizar un costo no lineal en los arcos.

Este aumento en los tiempos de viaje se especifican mediante funciones de latencia (o funciones de congestión), las cuales son usadas para expresar el tiempo de viaje (o costo) (T) como función del flujo en un arco (q). Usualmente estas funciones son expresadas como el producto entre un tiempo de viaje libre de flujo (*free flow time* en inglés) y una función de congestión $f(x)$.

$$T(q) = T_{ff} f\left(\frac{q}{c}\right)$$

donde el argumento de la función de congestión es el ratio q/c , con c la capacidad del arco (Spiess, 1990). Algo importante a destacar es que las restricciones de capacidad se reemplazan por las funciones de latencia.

Existen muchas funciones de latencia que han sido usadas en la práctica en el pasado (para una revisión detallada ver (Branston, 1976)). Pero últimamente, la función de latencia más utilizada es la función *Bureau of Public Roads* (BPR) (1964), la cual está definida como:

$$T(q) = T_{ff} \left(1 + \alpha \left(\frac{q}{c}\right)^\beta\right)$$

En la ecuación de BPR, los coeficientes α y β determinan la forma de la función. Con altos valores de β , la aparición de los efectos de congestión se vuelve cada vez más repentina. En la Figura 2.1 se presenta un ejemplo de la función BPR. Se puede observar que para flujos bajos, la función es bastante plana y el tiempo de viaje corresponde al *free flow time* (T_{ff}). Cuando

existen flujos altos en los arcos, el coeficiente β determina el umbral en el que la función BPR aumenta significativamente (en algunas formulaciones se aproxima asintóticamente a la capacidad del arco (c)).

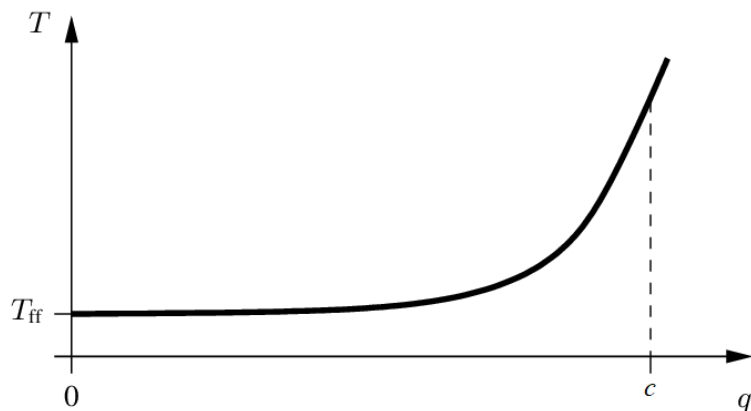


Figura 2.1: Función *Bureau of Public Roads* (BPR) que relaciona el tiempo de viaje con el flujo del arco. Está basada en el tiempo de viaje T_{ff} bajo condiciones de libre flujo y en la capacidad c del arco (Maerivoet and De Moor, 2005).

A pesar de la simplicidad de la función BPR, esta presenta algunas desventajas. Cuando los arcos están muy congestionados, la función sobreestima los tiempos de viaje, lo que podría ocasionar una tasa de convergencia lenta. Además, subestima los costos de los arcos que tienen un bajo ratio de *flujo/capacidad* y no proporciona una garantía de solución única cuando β es grande (Spiess, 1990).

2.2. Problemas de programación binivel

La optimización binivel es un campo de la programación matemática en la cual hay un problema de optimización (el superior, también llamado “líder”) con problemas de optimización anidados (los inferiores, también llamados “seguidores”) en las restricciones. En la práctica, estos problemas se utilizan para modelar procesos de decisión jerárquicos. Primero, el “líder” tiene que decidir sobre un subconjunto de las variables de decisión, las cuales afectan la región factible del “seguidor”. Después de eso, el “seguidor” tiene que decidir sobre el otro subconjunto de variables de decisión, las cuales afectan al valor de la función objetivo del “líder”.

Este tipo de programación es atractiva para modelar problemas más realistas, pero también resulta en modelos de optimización difíciles de resolver en teoría y en la práctica. Ben-Ayed y Blair (1990) mostraron que incluso para problemas biniveles lineales con variables de decisión continuas tanto para el líder como los seguidores son *NP-hard*.

El problema general de optimización binivel es de la forma

$$\min_{x \in X, y} F(x, y) \quad (2.3a)$$

$$\text{s.a. } G(x, y) \geq 0 \quad (2.3b)$$

$$y \in S(x) \quad (2.3c)$$

donde $S(x)$ es el conjunto de soluciones óptimas del siguiente problema parametrizado en x .

$$\min_{y \in Y} f(x, y) \quad (2.4a)$$

$$\text{s.a. } g(x, y) \geq 0 \quad (2.4b)$$

El problema (2.3) es el problema del nivel superior o el del “líder” y el problema (2.4) es el problema del nivel inferior o el del “seguidor”. Además, las variables $x \in \mathbb{R}^n$ son las variables del nivel superior y las variables $y \in \mathbb{R}^m$ son las del nivel inferior. Las funciones objetivos están dadas por $F, f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ y las funciones de restricciones por $G : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ y $g : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$. X y Y son el conjunto factible de las variables x e y respectivamente. A las restricciones del nivel superior se les llaman restricciones de acoplamiento si dependen explícitamente del vector variable del nivel inferior y . Además, todas las variables del nivel superior que aparecen en las restricciones del nivel inferior se denominan variables de enlace (Kleinert et al., 2021).

Los problemas de optimización binivel se remontan a las publicaciones fundamentales sobre los juegos de líder-seguidor de von Stackelberg (1934; 1952). La formulación presentada en 2.3 y 2.4 fue usada en Bracken y McGill (1973) en un contexto de aplicación militar que

considera combinaciones de armas a costo mínimo. Otra discusión preliminar de problemas multi-niveles, o en particular, de dos niveles se pueden encontrar en (Candler and Norton, 1977). A través de los años, la optimización binivel ha sido reconocida como una importante herramienta de modelamiento ya que permite formalizar procesos de decisión jerárquicos que a menudo aparecen en aplicaciones tales como energía, seguridad, o gestión de operaciones.

Se ha estudiado en la literatura que el desarrollo de algoritmos de solución para problemas de optimización binivel depende fuertemente de la estructura y las propiedades del problema del nivel inferior como también del acoplamiento entre el nivel superior e inferior a través de las variables de decisión. Por ejemplo, las técnicas de solución son muy diferentes dependiendo de si el problema del “seguidor” es continuo y convexo o no es convexo, debido a, por ejemplo, variables enteras.

Por esta razón, muchos métodos han sido propuestos para resolver modelos de dos niveles. Para problemas lineales, Bard y Moore (1990) presentó un enfoque de *Branch-and-Bound* basado en las condiciones de Kuhn-Tucker del problema inferior y Hansen et al. (1992) propone también un algoritmo de *Branch-and-Bound* con ramificaciones en las restricciones del “seguidor”. Por otro lado, Saharidis y Lerapetritou (2009) sugirieron un enfoque de descomposición de Benders para resolver problemas binivel entero mixto con variables discretas para el “líder” y variables continuas para el “seguidor”.

En el contexto de esta tesis, una aplicación de este tipo de programación son los problemas de diseño de redes. Todos estos tipos de problemas pueden ser modelados como problemas de dos niveles porque el objetivo del planificador de la red (“líder”) -reducir la congestión de la red completa- y el objetivo de los viajeros (“seguidores”) -encontrar el camino más rápido entre su origen y destino- no tienen que ser el mismo (Braess et al., 2005). En esta tesis en particular, se estudia el caso de programación binivel con variables del nivel superior enteras y variables del nivel inferior continuas.

2.3. Equilibrio en juegos de red

Es lógico que todos los viajeros se esfuercen por tomar la ruta más corta entre sus respectivos orígenes y destinos, o dicho de otra forma, la ruta más rápida cuando la función objetivo representa el tiempo. Si suponemos que cada usuario elige una ruta que percibe como la más corta con las condiciones de tráfico predominantes, la situación resultante de estas decisiones individuales es una en la cual los viajeros no pueden reducir sus tiempos de viaje eligiendo unilateralmente otra ruta (la única manera de lograrlo, al decir esto, es que deben ponerse de acuerdo para poder disminuir sus tiempos de viaje, pero suponiendo un comportamiento egoísta esto no ocurriría). Knight (1924) llamó a los patrones de flujos de tráfico resultantes una situación de equilibrio.

Wardrop desarrolló algunos principios básicos que subordinan el comportamiento de cada individuo al escoger una ruta, que se utilizan hasta el día de hoy, formalizando un equilibrio del tráfico. A continuación, se presentan aquellos principios y se define matemáticamente el conocido *Equilibrio de Wardrop*.

2.3.1. Equilibrio de Wardrop

La contribución fundamental que realizó Wardrop (1952) fue declarar dos principios que formalizan la noción de equilibrio y que rigen la distribución del tráfico en rutas alternativas:

Equilibrio de los usuarios (*User Equilibrium*): "*Los tiempos de viaje en todas las rutas realmente utilizadas son iguales y menores que los que experimentaría un solo vehículo en cualquier ruta no utilizada*".

Óptimo del sistema (*System optimum*): "*El tiempo promedio de viaje es mínimo*".

Además, en otras palabras, el primer principio de Wardrop se puede definir como: "*una distribución del tráfico es un equilibrio cuando ningún usuario tiene una ruta alternativa menos costosa*" (Smith, 1979).

En el primer criterio se asume que todas las decisiones de los individuos tienen un efecto insignificante en el desempeño de los demás. Es importante destacar, además, que se está asumiendo dos principios fundamentales aquí: (1) el hecho de que en la situación de equilibrio, no se asume la cooperación entre los individuos, y (2) que todos los individuos toman sus decisiones de una manera egoísta y racional. En la vida real se espera que todos sigan el primer criterio, así todo el sistema se establece en un equilibrio en el que nadie esté mejor eligiendo una ruta alternativa. Sin embargo, el segundo criterio es menos probable que ocurra espontáneamente. La única manera es que la percepción de un usuario del costo de una ruta cambie a un costo generalizado o marginal (es decir, se consideran los costos de los efectos provocados por agregar un vehículo adicional a la demanda de viajes). Entonces, así, se alcanza el óptimo en el sistema. Cabe señalar que como algunas personas estarán mejor, otras estarán peor, pero el sistema de transporte en su conjunto estará mejor (Maerivoet and De Moor, 2005).

Para que el equilibrio tenga sentido es necesario que la función que buscan minimizar los usuarios integre efectos de congestión, de lo contrario, todos elegirían el mismo camino sin importar cuántos lo hayan hecho, porque su propio tiempo de viaje no cambia ya que no depende de la cantidad de flujo.

La diferencia entre el equilibrio de Wardrop y Nash es sutil pero importante. Haurie y Marcotte (1985) probaron que un equilibrio de Nash en un juego de red con un número finito de jugadores, converge al equilibrio de Wardrop cuando el número de jugadores se incrementa. Por esta razón, aunque los conceptos de solución son diferentes, un equilibrio de Wardrop puede ser visto como una instancia de equilibrio de Nash en un juego con un número infinito de jugadores.

A continuación se formula matemáticamente el primer principio de Wardrop. Sea una red dirigida $\mathcal{G} = (N, A)$ y un conjunto $C \subseteq N \times N$ de productos representados por pares OD . Para cada $k \in C$, sea R_k el conjunto de rutas en G conectando el origen y destino correspondiente. El flujo por arco es un vector no negativo $f = (f_a)_{a \in A}$ que describe la cantidad de tráfico en

cada arco, además existe un flujo de ruta que es un vector no negativo $h = (h_r)_{r \in R}$ tal que cumple la demanda de cada par OD . Para un flujo f , el costo de viaje a lo largo de una ruta r es $c_r(f) := \sum_{a \in r} t_a(f_a)$. Con $t_a(\cdot)$ una función de costo de viaje por arco no negativa, no decreciente y continua. Por último, sea X_f el conjunto de los flujos factibles. Como el primer principio de Wardrop requiere que todo el flujo transite por las rutas más cortas, un flujo h se denomina equilibrio de Wardrop si y sólo si:

$$c_r(h) \leq c_q(h) \quad \forall k \in C \quad \forall \text{ rutas } r, q \in R_k \text{ tal que } h_r > 0 \quad (2.5)$$

Beckmann, McGuire y Winsten (1955) probaron que tal flujo siempre existe considerando el siguiente problema de flujo multi-producto a costo mínimo con función objetivo separable:

$$\min \left\{ \sum_{a \in A} \int_0^{f_a} t_a(z) dz : f \in X_f \right\} \quad (2.6)$$

El problema anterior es convexo porque la función objetivo es la integral de una función no decreciente y, dado que su dominio es un conjunto compacto, alcanza su óptimo.

Junto con ello, es posible reformular la ecuación (2.5) como un problema de desigualdad variacional propuesto por Smith (1979). Un flujo f es un equilibrio de los usuarios si y sólo si:

$$\sum_{a \in A} t_a(f_a) f_a \leq \sum_{a \in A} t_a(f_a) x_a \quad \forall x \in X_f \quad (2.7)$$

Nótese que esta desigualdad es una consecuencia directa del hecho de que, en equilibrio, los usuarios viajan por las rutas más baratas con respecto a los costos del arco $t_a(f_a)$ (Correa and Stier-Moses, 2011). Además, esta reformulación se centra en los costos de los arcos en vez de los costos de las rutas posibles para pares OD , lo cual resulta útil para la formulación del problema de diseño de redes.

2.3.2. Precio de la anarquía

De acuerdo a lo establecido por Wardrop, se pueden apreciar dos estados en una red de tráfico al considerar minimizar los costos de viaje: el equilibrio del tráfico y el óptimo del sistema. Como ya fue mencionado, estos casos difieren porque, al considerar un modelo de ruteo en el cual el costo experimentado por el tráfico de la red en un arco es una función de la congestión de éste, en el primer criterio se asume que no existe una regulación central de la red y que los usuarios actúan de manera puramente egoísta. Esta manera egoísta en que actúan establece un equilibrio en la red, en el que cada usuario no puede estar mejor, sin embargo, en su totalidad, la red está peor. Esto es porque está estudiado que los equilibrios de Nash en la teoría de juegos (que en este caso, es el equilibrio de Wardrop) no optimizan en general el bienestar social (Dubey, 1986). La pregunta que surge entonces es ¿cómo se puede cuantificar esta ineficiencia en el sistema? ¿Cuál es la relación entre el costo del equilibrio de Wardrop y el costo de la asignación óptima del tráfico en la red?

Esta pérdida de desempeño de la red es cuantificada a través del peor ratio posible entre los costos del bienestar social logrado por un equilibrio en la red y el ruteo óptimo del tráfico, el cual fue nombrado *Precio de la Anarquía* por Papadimitriou (2001) y definido por primera vez en (Koutsoupias and Papadimitriou, 1999). También es conocido como ratio de coordinación. Considerando $c(\cdot)$ una función de costos, \bar{x} como el flujo en equilibrio, x^* el flujo óptimo y α el valor del precio de la anarquía, entonces se establece que:

$$\frac{c(\bar{x})}{c(x^*)} \leq \alpha$$

El precio de la anarquía fue evaluado en diferentes clases de funciones de latencia. Los primeros en evaluar este ratio en los modelos de ruteo de tráfico fueron Roughgarden y Tardos (2002), quienes probaron que el precio de la anarquía en redes con funciones de latencia lineales es exactamente $4/3$. Además, Roughgarden evaluó el precio de la anarquía para funciones de latencia polinomiales arbitrarias con coeficientes no negativos y grado

máximo d , entre otras funciones (Roughgarden, 2003). Los diferentes valores que calculó se pueden observar en la Tabla 2.1.

| Tipo de función de latencia | Representación típica | Precio de la anarquía |
|--------------------------------|------------------------|--|
| Lineal | $ax + b$ | $\frac{4}{3} \approx 1,333$ |
| Cuadrática | $ax^2 + bx + c$ | $\frac{3\sqrt{3}}{3\sqrt{3}-2} \approx 1,626$ |
| Cúbica | $ax^3 + bx^2 + cx + d$ | $\frac{4\sqrt[3]{4}}{4\sqrt[3]{4}-3} \approx 1,896$ |
| Polinomiales de grado $\leq p$ | $\sum_{i=0}^p a_i x^i$ | $\frac{(p+1)\sqrt[p]{p+1}}{(p+1)\sqrt[p]{p+1}-p} = \Theta\left(\frac{p}{\ln p}\right)$ |

Tabla 2.1: Precio de la anarquía para clases comunes de funciones de latencia por arco. Los coeficientes polinomiales son no negativos (Roughgarden, 2003).

Básicamente, Roughgarden estableció que el precio de la anarquía es pequeño a menos que las funciones de costo sean extremadamente “empinadas” (Roughgarden, 2005), es decir, para grados polinomiales muy grandes, el precio de la anarquía aumenta bastante. Sin embargo, Dumrauf y Gairing (2006) demostraron límites superior e inferior para el precio de la anarquía en funciones de latencia polinomiales de grado máximo p y grado mínimo s , incluso para casos en los que el grado máximo p es grande y por lo tanto, la función de costos es “empinada”.

2.4. Optimización robusta

La metodología de optimización robusta busca encontrar una solución que es insensible a la incertidumbre considerada en el problema y, por tanto, es una solución eficaz en la práctica. La principal característica de una solución robusta es que se comporta bien sobre todos los escenarios posibles de incertidumbre aunque sólo se acerque al óptimo para cada uno de los escenarios específicos. Además, en muchos problemas encontrar una solución robusta no es más difícil que resolver el problema determinista. El enfoque de optimización robusta fue introducido para problemas de optimización convexa por Ben-Tal y Nemirovski (1998) y para problemas de programación definida por El-Ghaoui y Lebret (1997).

La solución robusta de un problema de optimización bajo incertidumbre es definida como la solución que alcanza el mejor valor objetivo para el peor caso de escenario incierto. Otra definición, que permite comprender mejor el concepto de robustez, es que una solución robusta

es aquella solución que es factible de aplicar a cualquier escenario posible de parámetros y que además, permanece cercana a la condición de optimalidad (García and Peña, 2018).

Esta metodología de optimización robusta ha generado interesantes investigaciones en un gran número de campos tales como optimización de portafolio (El-Ghaoui et al., 2003; Goldfarb and Iyengar, 2003), problemas de gestión de la cadena de suministros (Ben-Tal et al., 2005; Bertsimas and Thiele, 2003), programación entera y redes de flujo (Bertsimas and Sim, 2003; Atamtürk and Zhang, 2007), ruteo de vehículos (Agra et al., 2013), machine learning (Xu et al., 2008), políticas públicas (Ben-Tal et al., 2011), entre otros.

Como se menciona anteriormente, se desea que la solución robusta sea factible para cualquier valor de parámetros a analizar. Pero para lograr esto, es necesario tener en cuenta algunas implicancias de este enfoque. En primer lugar, cuando se considera todos los posibles valores de los parámetros se tiende a incurrir en significativos costos computacionales por la gran cantidad de escenarios probables. Además, existe un deterioro en la calidad de la solución, pues mientras mayor sea la exigencia en cuanto a la factibilidad, mayor será la probabilidad de alejarse de la optimalidad. Por lo tanto, existe un *trade-off* que depende del problema a resolver. En este punto, se destaca que la elección del conjunto de incertidumbre (el conjunto en el cual vive el parámetro del problema que está bajo incertidumbre) juega un papel fundamental en la factibilidad de resolver un problema y en la calidad de la solución obtenida.

A continuación se formaliza el problema de optimización robusto. Primero, se considera el siguiente problema de optimización incierto:

$$\begin{aligned} \text{(P)} \quad & \min_{u,v} f(u, v, w) \\ & \text{s.a. } g(u, v, w) \leq 0 \end{aligned}$$

donde la variable w es un parámetro incierto y en este caso, corresponde a un escenario probable para el problema definido.

En la mayoría de los problemas, no se sabe con exactitud cuál es el valor del parámetro w , pero se establece que el parámetro recae en un conjunto de incertidumbre \mathcal{U} , que contiene todos los escenarios suficientes para considerar en el problema. Luego, se tiene una familia de problemas de optimización dados por el par $(P(w), w \in \mathcal{U})$. Uno de los principales objetivos de la optimización robusta es poder convertir esta familia de problemas en un solo problema de optimización, donde la elección del conjunto de incertidumbre es fundamental para el resultado y complejidad del problema.

2.4.1. Contraparte robusta

Como parte de los objetivos de la optimización robusta, que busca poder unificar en un sólo problema de optimización la familia de problemas que corresponden a cada escenario de incertidumbre, Ben-Tal y Nemirovski (1998) proponen una formulación considerando problemas que poseen las siguientes características:

- (1) Existe un conocimiento básico de los datos, los cuales pueden ser parcial o totalmente “inciertos”, y todo lo que se sabe sobre el vector de datos w es que pertenece a un conjunto de incertidumbre \mathcal{U} .
- (2) Todas las restricciones del problema deben cumplirse cualquiera sea la realización de $w \in \mathcal{U}$.

Con estas condiciones, es posible reformular el problema de optimización incierto (P) como un problema de optimización robusto que se denomina contraparte robusta de (P) (*Robust Counterpart Problem* o RC por sus siglas en inglés).

$$\begin{aligned}
 \text{(RC)} \quad & \min_{u,v,\gamma} \quad \gamma \\
 \text{s.a.} \quad & f(u, v, w) \leq \gamma \quad \forall w \in \mathcal{U}, \\
 & g(u, v, w) \leq 0 \quad \forall w \in \mathcal{U}
 \end{aligned}$$

De esta manera, las soluciones factibles de (RC) se denominan soluciones robustas del

problema de optimización incierto (P), y el valor óptimo de (RC) es denominado valor óptimo robusto del problema incierto (P).

Una cualidad de este enfoque es que la complejidad de resolver el problema (RC) es, para casos muy generales, la misma complejidad que el problema original (P). Por ejemplo, cuando el problema original es un problema lineal, de (Ben-Tal and Nemirovski, 1999) se sabe que el problema (RC) es equivalente a un problema lineal cuando \mathcal{U} es un poliedro, y es un programa convexo con restricciones cuadráticas cuando \mathcal{U} es un conjunto elipsoidal acotado. Además, el tamaño del problema (RC) es acotado por un polinomio en las dimensiones del problema original, el cual implica un método polinomial para resolver la solución robusta.

Contraparte robusta ajustable

Uno de los supuestos considerados en el enfoque anterior es que todas las decisiones del problema se toman antes de conocer las realizaciones de los parámetros inciertos. Es decir, las variables de decisión toman valores sin saber cuál es el escenario real que ocurre.

Sin embargo, una de las características principales del problema a resolver en esta tesis es que posee dos tipos de decisiones, y una de esas decisiones son determinadas posteriormente al resultado de los datos que son considerados inciertos. Específicamente, la decisión de escoger un camino por parte de los usuarios se toma después de conocer la demanda por cada par origen-destino.

Por lo tanto, existe una extensión a la contraparte robusta que se denomina contraparte robusta ajustable (*Adjust Robust Counterpart* o ARC por sus siglas en inglés). En ella, algunas variables de decisión se pueden ajustar en un momento posterior a la resolución del problema. En particular, en esta variante se consideran dos tipos de variables de decisión: (1) las variables que representan decisiones “*a priori*” y que obtienen valores numéricos específicos al resolver el problema antes de que los parámetros “se conozcan”; y (2) las variables que representan decisiones “*a posteriori*” que pueden ser tomadas cuando los datos reales se revelen parcialmente, o son variables de análisis que no representan decisiones (ej., variables

de holgura que son añadidas para convertir el problema original a su forma general, esto es, a un problema lineal).

En definitiva, la contraparte robusta ajustable (ARC) está dada por:

$$\begin{aligned}
 \text{(ARC)} \quad & \underset{u, \gamma}{\text{mín}} \quad \gamma \\
 \text{s.a.} \quad & \forall w \in \mathcal{U} \text{ exists } v : \begin{cases} f(u, v, w) \leq \gamma \\ g(u, v, w) \leq 0 \end{cases}
 \end{aligned}$$

donde u son las variables de decisión “*a priori*” y v denota las variables de decisión “*a posteriori*”.

Para el caso particular del problema de esta tesis, u representa la decisión del planificador central, que decide invertir en la capacidad de los arcos antes de observar la demanda de los usuarios y el flujo del tráfico v se adapta posteriormente cuando se sabe la cantidad de usuarios que utilizarán la vía.

Cabe destacar que la optimización robusta ajustable es menos conservadora que el enfoque clásico de optimización robusta dado que conlleva a decisiones más flexibles que pueden ser ajustadas de acuerdo a los datos que se conocen en una determinada etapa. Más precisamente, (ARC) produce valores objetivos óptimos que son al menos tan buenos que los de (RC) ($z_{ARC} \leq z_{RC}$) (Gorissen et al., 2015). En el aspecto de complejidad, este tipo de problemas puede no resultar eficiente computacionalmente, al menos que se consideren ciertas condiciones en el problema (para más detalles, véase *Robust Optimization* (Ben-Tal et al., 2009)).

2.4.2. Conjuntos de incertidumbre

Como se puede observar en el problema (RC), existe una restricción para cada realización del parámetro w . Como el objetivo de la optimización robusta es encontrar las soluciones de menor costo entre aquellas que tienen que satisfacer las restricciones para todas la

realizaciones de $w \in \mathcal{U}$, esto implica tener un número infinito de restricciones.

Dada la infinidad de restricciones, no es claro que el problema (RC) pueda ser resuelto eficientemente. Y es posible pensar que agregar robustez a un problema de optimización conlleva a buscar una solución a expensas de una complejidad computacional significativamente mayor, lo cual es cierto, puesto que la contraparte robusta de cualquier problema de optimización convexo en general es difícil de resolver. A pesar de esto, existen muchos problemas robustos que pueden ser resueltos de una forma eficiente y mucha de la literatura está enfocada en especificar los conjuntos de incertidumbre junto con las clases de restricciones que permiten obtener contrapartes robustas manejables.

En definitiva, darle estructura al conjunto de incertidumbre y elegir el apropiado para el tipo de problema a solucionar es esencial para obtener problemas robustos que se puedan resolver eficientemente, ya que es posible aplicar técnicas de reformulación robustas que permiten reescribir el problema con un número finito de restricciones.

En la literatura se han estudiado y utilizado los siguientes conjuntos de incertidumbre para problemas robustos lineales que conllevan a problemas robustos que son tratables computacionalmente. Para un revisión detallada de los conjuntos de incertidumbre y la reformulación de la contraparte robusta referirse a (Bertsimas et al., 2010).

- **Incetidumbre elipsoidal:** Los conjuntos de incertidumbre elipsoidales fueron utilizados por primera vez por El-Ghaoui y Lebret (1997) y Ben-Tal y Nemirovski (1999). Estos conjuntos elipsoidales controlan el tamaño de la variabilidad a través de un parámetro (ρ) denominado Presupuesto de Incertidumbre (*Budget of Uncertainty* en inglés). El tomador de decisión selecciona este parámetro con el objetivo de hacer fácilmente el *trade-off* entre robustez y desempeño. La forma del conjunto elipsoidal es:

$$\mathcal{U} = \{w \mid \|w\|_2 \leq \rho\}$$

Para el caso de un conjunto elipsoidal, la reformulación del problema en el peor caso

implica un problema que se optimiza sobre una restricción cuadrática. Utilizando el dual que involucra funciones cuadráticas, conduce a un problema de optimización de cono de segundo orden.

- **Incertidumbre poliedral:** Este tipo de conjunto de incertidumbre también se puede encontrar en (Ben-Tal and Nemirovski, 1999). Su forma matemática es:

$$\mathcal{U} = \{w \mid \mathbf{D}w \leq q\}$$

Cuando \mathcal{U} es un poliedro, la reformulación del problema en el peor caso es lineal y la contraparte robusta resultante es equivalente a un problema lineal. Por lo tanto, el tamaño de estos problemas crece polinomialmente en el tamaño del problema original y en las dimensiones del conjunto de incertidumbre.

- **Incertidumbre restringida de cardinalidad:** En (Bertsimas and Sim, 2004) se define un conjunto de incertidumbre que crea una cota superior en términos de restricciones de cardinalidad, esto es, un número máximo de parámetros del problema que pueden variar de sus valores nominales. Considerando una matriz incierta $\mathbf{A} = (\tilde{a}_{ij})$ suponga que cada componente incierto \tilde{a}_{ij} se encuentra en el intervalo $[a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}]$, donde a_{ij} es el valor nominal de \tilde{a}_{ij} . En lugar de considerar el caso en que cada parámetro puede desviarse, se permite que a lo más Γ_i coeficientes de la fila i pueden desviarse. Por lo tanto, el número positivo Γ_i denota el presupuesto de incertidumbre para la i -ésima restricción. Al igual que con el tamaño elipsoidal, este parámetro controla el equilibrio entre la optimización de la solución y su robustez a la perturbación de parámetros.

Este tipo de incertidumbre se puede ver como un caso particular de la incertidumbre poliedral, al relajar el parámetro que restringe la cardinalidad. Por lo mismo, al utilizar este conjunto, se puede derivar una formulación lineal del problema incierto.

- **Incertidumbre con norma:** En (Bertsimas et al., 2004) se demuestra que los problemas de optimización robustos lineales con conjuntos de incertidumbre descritos por normas más generales conllevan a problemas convexos con restricciones relacionadas

con la norma dual. Por lo tanto, un modelo original basado en normas conlleva a un problema equivalente robusto con restricciones de norma dual correspondientes. En particular, como se mencionó, la norma $\|\cdot\|_1$ y la norma $\|\cdot\|_\infty$ resultan en problemas lineales y la norma $\|\cdot\|_2$ en un problema de cono de segundo orden.

2.4.3. Métodos de resolución

Una vez que se deriva la reformulación de la contraparte robusta del problema de optimización incierto, éste se puede resolver de forma exacta o a través de algoritmos eficientes de resolución. Por ejemplo, algunos algoritmos utilizados para resolver estos problemas son heurísticas de generación de columnas, métodos de punto interior, métodos de descomposición, entre otros. Por otra parte, si no es posible escribir la contraparte robusta o realizar una reformulación aproximada que determine un problema manejable computacionalmente se puede utilizar el método denominado *Enfoque adversario* (Gorissen et al., 2015).

Este enfoque permite comenzar con un subconjunto finito de escenarios del conjunto donde pertenece el parámetro incierto. Luego, al utilizar este subconjunto finito en el problema de optimización robusto se tiene un número finito de restricciones y por lo tanto es posible resolverlo. Si la solución resultante es factible y robusta entonces se ha llegado a la solución óptima, si por el contrario no es robusta, entonces se busca un nuevo escenario que haga esta solución infactible, y este escenario se agrega al subconjunto finito con el que se está trabajando, para volver a resolver el problema. Se realiza este procedimiento iterativamente hasta converger a una solución óptima robusta y factible.

Este método de resolución fue utilizado por Bienstock y Özay (2008), y se demuestra que se puede converger al óptimo en unas pocas iteraciones, incluso podría ser más eficiente que utilizar la reformulación de la contraparte robusta para conjuntos de incertidumbre poliedrales. Cabe notar que una ventaja de este método es que se mantiene la estructura original del problema de optimización robusto, puesto que sólo cambia en cada iteración el

número de escenarios a considerar, y con ello, se agregan restricciones del mismo tipo que deben ser válidas para todos los escenarios que se encuentran en el subconjunto finito.

2.5. Trabajos relacionados

Los problemas de diseño de red (NDP por sus siglas en inglés) son problemas que se han analizado con muchas variantes y extensiones en la literatura. Generalmente los NDP son categorizados en tres clases. Están los problemas de diseño de redes continuos (CNDP) y discretos (DNDP) en que se utilizan variables de expansión de capacidad continua y discreta respectivamente, además, los problemas que mezclan estos dos tipos de variables. En esta tesis se desarrolla un modelo de diseño de redes mixto, con variables discretas de aumento de capacidad y variables de flujo continuas.

Hasta hace muy poco tiempo se desconocía la complejidad de los problemas CNDP, no obstante, en (Gairing et al., 2017) se mostró que esta clase de problemas es fuertemente *NP-hard* y *APX-hard*, tanto en redes dirigidas como no dirigidas e incluso en instancias con funciones de latencias afines.

Por otro lado, debido a la propiedad inherente de no-convexidad del problema de diseño de redes discreto, es difícil obtener una solución óptima global para este tipo de problema. Incluso, la mayoría de las investigaciones de DNDP no desarrollaron métodos de solución de optimización global. En los últimos estudios, los métodos de solución de optimización global para DNDP han atraído cada vez más la atención de los investigadores en el campo del diseño de redes de transporte.

Wang y Lo (2010) transformaron el modelo binivel de diseño de redes continuo a un problema de un sólo nivel, que luego aproxima mediante una programación general lineal entera mixta (MILP por sus siglas en inglés) y resuelve el óptimo global a través de un algoritmo de solución MILP tradicional. Luathep et al. (2011), por otra parte, presentó una formulación lineal entera mixta para el NDP discreto formulando el problema de decisión de

los usuarios (segundo nivel) con el problema de la desigualdad variacional. Este problema aproxima la función de tiempos de viajes con una función lineal por partes y relaja los términos no lineales introduciendo variables auxiliares binarias.

Farvaresh y Sepehri (2011), sin embargo, aplicó condiciones de Karush-Kuhn-Tucker para transformar el DNDP de dos niveles a uno de un sólo nivel que posteriormente es linealizado igualmente con variables auxiliares binarias y Wang et al. (2013) propone un enfoque de optimización de solución global usando la relación entre el óptimo del sistema y el equilibrio óptimo de los usuarios. En este trabajo se presenta un DNDP con múltiples niveles de capacidad para decidir cuántas pistas deberían agregarse a las carreteras congestionadas existentes.

Por último, Fontaine y Minner (2014) evitan linealizar el problema de diseño de un sólo nivel que resulta de las metodologías anteriores con variables binarias auxiliares. Gracias a esto, se obtienen muy pocas variables binarias y por lo tanto el problema se vuelve computacionalmente eficiente. Se termina resolviendo el problema con un algoritmo de descomposición de Benders para alcanzar el óptimo global.

Como se ha visto, en los problemas de diseño de redes binivel, el nivel inferior es modelado con ecuaciones que permiten llegar a un equilibrio en el tráfico considerando la decisión de los usuarios. El desafío de resolver este problema radica en cómo se considera el comportamiento de elección de ruta de los usuarios de la red, por ejemplo, se puede plantear que el equilibrio de usuario es determinista (*Deterministic user equilibrium* en inglés o DUE) o estocástico (*Stochastic user equilibrium* o SUE). El equilibrio determinista más utilizado es el equilibrio de Wardrop (Wardrop, 1952) debido a su simplicidad y en este se asume que todas las rutas usadas para el mismo par de origen-destino tienen el mismo costo mínimo de viaje. Beckmann et al. (1955) fueron los primeros en probar que un equilibrio de Wardrop siempre existe y que esencialmente es único, y es el que ha sido utilizado en las investigaciones recientemente mencionadas.

Por otro lado, en el equilibrio de usuarios estocástico se asume que todos los transeúntes pueden no conocer perfectamente la información del tráfico de la red y, por lo tanto, eligen su mejor ruta percibida (McCord, 1987). Los primeros modelos de tráfico con equilibrio estocástico fueron propuestos por Dial (1971) y Danganan et al. (1977) que incorporan variaciones inciertas en la percepción de los costos por diferentes agentes.

Davis (1994) propuso dos enfoques de programación no lineal para resolver el CNDP con SUE, pero, no pueden resolver la versión discreta, como también Liu y Wang (2015) que se enfocan en el mismo problema continuo con decisiones estocásticas en los usuarios. Sin embargo en (Liu and Wang, 2016) desarrollan una programación no lineal entera mixta para resolver de manera global el problema de diseño de redes discreto con SUE, que es inherentemente no convexo e incluso un desafío resolverlo en su óptimo local.

Otro método para considerar el equilibrio de usuarios es con un enfoque robusto y en (Ordóñez and Stier-Moses, 2010) se considera que los agentes seleccionan caminos resolviendo un problema de optimización robusto que impone un límite en el número de arcos que pueden desviarse de la media. Y por último, Lo y Tung (2003) introduce un modelo probabilístico de equilibrio de usuarios para redes con capacidad estocástica. Este modelo de equilibrio requiere que las rutas utilizadas no solo tengan el mismo valor de tiempo medio de viaje, sino que su varianza esté limitada por las garantías de desempeño dadas.

Ahora bien, los trabajos de diseño de redes incorporando incertidumbre en los parámetros del problema se remontan a (Ferguson and Dantzig, 1956).

Anterior a la aplicación de la metodología robusta para tratar problemas de optimización inciertos, un método estándar para representar la incertidumbre utilizada en la literatura era a través de escenarios inciertos y discretos, que es un enfoque usado en optimización estocástica (Birge and Louveaux, 1997). Mulvey et al. (1995) utiliza este enfoque para resolver problemas de optimización en varios campos de gestión de operaciones, como también Ahmed et al. (2003) lo usa en planeación de capacidad.

Estos métodos basados en escenarios enfrentan las siguientes dificultades: (1) se asume que la discretización de la incertidumbre en escenarios es conocida, lo cual puede ser una aproximación muy simple de la realidad, (2) la gran cantidad de escenarios utilizados para representar con precisión la incertidumbre puede conducir a desafiantes problemas computacionales y (3) la solución obtenida puede ser sensible a los posibles resultados de la incertidumbre, una dificultad ya abordada por Mulvey et al. (1995).

En este sentido, la metodología robusta se hace cargo de estas problemáticas y considera que los parámetros inciertos pertenecen a un conjunto de incertidumbre, obteniendo una solución que es insensible a la variabilidad de los escenarios. Además, la resolución del problema robusto puede llegar a ser de complejidad comparable con la resolución del problema determinista en casos particulares.

En (Gutiérrez et al., 1996) se desarrollan métodos algorítmicos que consideran escenarios de incertidumbre para resolver un problema robusto de diseño de redes sin capacidad. Bertsimas y Sim (2003) consideran soluciones robustas para problemas de flujo en redes con un conjunto de incertidumbre de intervalo para los coeficientes de los costos. En tanto, Atamtürk y Zhang (2007) investigan soluciones robustas para redes de flujo y diseño de redes, mostrando que para ciertas redes y para conjuntos de incertidumbre de demanda específicos, el problema robusto es eficiente.

Ordóñez y Zhao (2007) muestran que el problema de diseño de redes robusto es tratable para redes con un único par origen-destino por producto y bajo incertidumbre general de costos de transporte y un conjunto de incertidumbre de intervalo en la demanda. Mientras que Mudchanatongsuk et al. (2008) extiende los resultados anteriores considerando un problema de diseño de redes bajo supuestos más generales y realistas.

Otros estudios de metodologías robustas se encuentran en (Ferris and Ruszczyński, 2000) para el ruteo de una red con fallas y en (Paraskevopoulos et al., 1991) para la planificación de la capacidad con aplicación en la fabricación. Para conocer más acerca de los problemas

robustos discretos se recomienda referirse a (Gerodimos, 1998).

En resumen, el trabajo de esta tesis aborda los problemas de diseño de redes con aumento de capacidad, a través de una programación binivel y un equilibrio de usuarios determinístico, junto con considerar incertidumbre a través de una metodología robusta. Es decir, mezcla tres aspectos distintos: expansión de capacidad discreta, equilibrio en el tráfico e incertidumbre, en un problema de diseño de redes de transporte, que en la literatura, se han desarrollado por separado o combinando sólo dos aspectos de los tres.

Capítulo 3

Modelos

En este capítulo se formulan dos modelos diferentes que incorporan todos los aspectos que integra el problema a resolver en esta tesis. En el primer modelo, el costo por arco es lineal con respecto al flujo que transita por él, además considera restricciones de capacidad para cada arco. En cambio, el segundo modelo ocupa la función de latencia BPR como costo por arco, por lo que al incorporar la capacidad dentro de la función, no hay restricciones de capacidad por arco involucradas.

En primer lugar, en la Sección 3.1 se modela el escenario cuando la demanda es determinista, es decir, conocida e invariable. Posteriormente, en la Sección 3.2 se realizan reformulaciones al modelo determinista para extenderlo a que considere escenarios inciertos de demanda. Finalmente, se analizan las condiciones que permiten abordar las reformulaciones del problema para así resolverlo eficientemente de manera exacta, definiendo un modelo robusto aproximado.

Para este problema se considera un grafo dirigido $\mathcal{G} = (N, A)$, con n nodos y m arcos. También, se tiene un conjunto $\Omega \subset N \times N$ de pares de nodos orígenes y destino. Existen costos de inversión para mejorar cada arco $g_a(\cdot)$ y costos de transporte $t_a(\cdot)$. Cada arco tiene una capacidad c_a . A continuación se explicita todas las notaciones y variables que se ocupan para formular los modelos.

Notación

N = conjunto de nodos. $|N| = n$

A = conjunto de arcos. $|A| = m$

Ω = conjunto de pares origen-destino.

I = matriz de incidencia del grafo \mathcal{G} .

d^k = demanda para cada par origen-destino $k \in \Omega$.

B = presupuesto de inversión.

c_a = capacidad del arco $a \in A$

$g_a(\cdot)$ = función de costos lineal para aumentar la capacidad en el arco a que depende de las unidades de aumento.

$t_a(\cdot)$ = función de costo de viaje por cada arco a que depende del flujo del arco.

Variables de decisión

x_a^k = flujo en el arco $a \in A$ del par $OD = k \in \Omega$.

$x_a = \sum_{k \in \Omega} x_a^k :=$ flujo en el arco $a \in A$.

$x^k = (x_a^k)_{a \in A} :$ vector de flujo del par $OD = k \in \Omega$.

z_a = aumento de capacidad del arco $a \in A$

El problema general a resolver es un modelo de diseño de redes de dos etapas, con dos variables de decisión: la decisión de aumentar la capacidad de los arcos (z) y la decisión del patrón de flujos por cada arco (x).

Problema binivel de diseño de redes general

$$\text{mín} \quad \sum_{a \in A} t_a(x_a, z_a) x_a \quad (3.1a)$$

$$\text{s.a} \quad \sum_{a \in A} g_a(z_a) \leq B \quad (3.1b)$$

$$z \geq 0 \quad (3.1c)$$

$$x \in \text{argmin} \left\{ \sum_{a \in A} \int_0^{v_a} t_a(y, z_a) dy : Iv^k = d^k \quad \forall k \in \Omega, \quad v \geq 0 \right\} \quad (3.1d)$$

La función objetivo en la ecuación (3.1a) busca minimizar el costo total de todos los arcos. El costo por arco depende del flujo de tráfico y también, generalizando para la función de latencia, podría depender de la capacidad total de cada arco (considerando entonces la variable de aumento de capacidad z_a). La ecuación (3.1b) indica la restricción presupuestaria de la inversión en capacidad, siendo el aumento de capacidad una variable no negativa (3.1c). Estas primeras ecuaciones indican la decisión en el primer nivel: la del planificador central.

En el segundo nivel, la decisión de los usuarios (3.1d), se expresa el equilibrio de Wardrop reformulada en la versión de Beckmann et al. (1955). De esta manera, al resolver este problema se obtiene un flujo x en equilibrio donde los traseúntes se comportan siguiendo el primer principio de Wardrop (ver acápite 2.3.1).

3.1. Modelos deterministas

En lo que sigue se presentan los modelos deterministas de las dos formulaciones mencionadas anteriormente. El modelo determinista considera una demanda d^k nominal, que es constante, para cada par origen-destino $k \in \Omega$.

Inicialmente, se propone el primer modelo determinista, que considera una función de latencia lineal, y las reformulaciones para poder obtener un modelo entero mixto que se pueda resolver de manera exacta. Luego, se formula el segundo modelo determinista entero

mixto, el que considera una función de latencia no lineal, el cual se construye de la misma forma que el modelo anterior.

3.1.1. Modelo 1: Función de costo lineal y con restricciones de capacidad

En este primer modelo la función de costo por arco es lineal con respecto al flujo que transita por él. Además, cabe mencionar que la capacidad no forma parte de la función objetivo, por lo tanto:

$$t_a(x_a, z_a) = t_a(x_a) = time_a x_a$$

con $time_a$ una constante no negativa.

Problema de una etapa

Esta reformulación del problema permite transformar el problema de dos etapas (3.1) a un problema de una etapa (3.2), utilizando desigualdades variacionales (Smith, 1979) para el problema de decisión de los usuarios (nivel inferior).

$$\text{mín} \quad \sum_{k \in \Omega} \sum_{a \in A} t_a(x_a) x_a^k \quad (3.2a)$$

$$\text{s.a} \quad \sum_{a \in A} g_a(z_a) \leq B \quad (3.2b)$$

$$\sum_{a \in A} t_a(x_a) x_a \leq \sum_{a \in A} t_a(x_a) v_a \quad \forall v \text{ flujo factible} \quad (3.2c)$$

$$\sum_{a \in \delta^+(i)} x_a^k - \sum_{a \in \delta^-(i)} x_a^k = \begin{cases} d^k, & \text{if } i = k_s, \\ -d^k, & \text{if } i = k_t, \\ 0, & \text{en otro caso,} \end{cases} \quad \forall k \in \Omega, i \in N \quad (3.2d)$$

$$0 \leq \sum_{k \in \Omega} x_a^k \leq c_a + z_a \quad \forall a \in A \quad (3.2e)$$

$$x_a^k, z_a \in \mathbb{R}^+ \quad \forall a \in A, \forall k \in \Omega \quad (3.2f)$$

El costo óptimo (3.2a) y la restricción (3.2b) son iguales al modelo anterior. La ecuación (3.2c) representa la desigualdad variacional que permite capturar el flujo en equilibrio. Sin embargo, esta desigualdad variacional implica un número no numerable de restricciones, ya que se tiene una restricción por cada variable v flujo factible, que podrían ser incontables dependiendo del tamaño del problema. La restricción (3.2d) modela la conservación de flujo y la satisfacción de la demanda por cada par origen-destino $k \in \Omega$. En la restricción (3.2e) se establece el cumplimiento de la capacidad de los arcos y por último, en (3.2f) se presentan la naturaleza de las variables.

Reformulación del problema de una etapa usando dualidad

Dada la gran cantidad de restricciones asociadas a la segunda ecuación (3.2c) del modelo anterior, en esta reformulación del problema se utiliza el dual del siguiente subproblema para aproximar el equilibrio de flujo de Wardrop.

$$\min_v \left\{ \sum_{a \in A} t_a(x_a) v_a : I v^k = d^k \quad \forall k \in \Omega, \quad x_a \leq c_a + z_a \quad \forall a \in A, \quad v \geq 0 \right\} \quad (3.3)$$

Considerando $\beta_a \quad \forall a \in A$ y $\alpha_i^k \quad \forall k \in \Omega, \quad \forall i \in N$ como variables duales de las restricciones de capacidad y conservación de flujo respectivamente, entonces la función objetivo del dual se escribe como:

$$\sum_{k \in \Omega} \alpha_{k_s}^k d^k - \sum_{a \in A} \beta_a z_a - \sum_{a \in A} \beta_a c_a \quad (3.4)$$

Se destaca que se fija como $\alpha_{k_t}^k = 0$ para obtener restricciones linealmente independientes.

De esta manera bajo el teorema de dualidad débil, si se impone que el costo del dual del subproblema (3.3) es mayor o igual al lado izquierdo de la ecuación (3.2c), entonces se tiene que:

$$\sum_{a \in A} t_a(x_a) x_a \leq \sum_{k \in \Omega} \alpha_{k_s}^k d^k - \sum_{a \in A} \beta_a z_a - \sum_{a \in A} \beta_a c_a \leq \sum_{a \in A} t_a(x_a) v_a \quad (3.5)$$

y la restricción del equilibrio de Wardrop se cumple. Por lo tanto, la reformulación del problema quedaría como:

$$\text{mín} \quad \sum_{k \in \Omega} \sum_{a \in A} t_a(x_a) x_a^k \quad (3.6a)$$

$$\text{s.a} \quad \sum_{a \in A} g_a(z_a) \leq B \quad (3.6b)$$

$$\sum_{a \in A} \sum_{k \in \Omega} t_a(x_a) x_a^k \leq \sum_{k \in \Omega} \alpha_{k_s}^k d^k - \sum_{a \in A} \beta_a z_a - \sum_{a \in A} \beta_a c_a \quad (3.6c)$$

$$\alpha_i^k - \alpha_j^k - \beta_{ij} \leq t_{ij}(x_{ij}) \quad \forall k \in \Omega, \forall a = (i, j) \in A \quad (3.6d)$$

$$\sum_{a \in \delta^+(i)} x_a^k - \sum_{a \in \delta^-(i)} x_a^k = \begin{cases} d^k, & \text{if } i = k_s, \\ -d^k, & \text{if } i = k_t, \\ 0, & \text{en otro caso,} \end{cases} \quad \forall k \in \Omega, \forall i \in N \quad (3.6e)$$

$$0 \leq \sum_{k \in \Omega} x_a^k \leq c_a + z_a \quad \forall a \in A \quad (3.6f)$$

$$x_a^k, \beta_a, z_a \in \mathbb{R}^+ \quad \alpha_i^k \in \mathbb{R} \quad \forall a \in A, \forall k \in \Omega, \forall i \in N \quad (3.6g)$$

Esta reformulación del problema solo presenta cambios en dos restricciones. En la restricción (3.6c) se presenta la ecuación que se obtiene al considerar dualidad débil en (3.5). Además, en la ecuación (3.6d) se establece la factibilidad de las variables duales y su relación con la variable primal. Finalmente, la naturaleza de estas variables se mencionan en (3.6g).

Problema entero mixto (MIP) aproximado

En la reformulación anterior existen funciones no lineales con respecto a las variables de decisión, por ejemplo, en la función objetivo (3.6a) y en la restricción (3.6c). Por lo tanto, en esta parte, se construyen aproximaciones lineales por partes de los costos de viaje de los arcos y para la capacidad de los mismos.

Para ello, se definen los siguientes parámetros y variables:

s = valores enteros posibles para la cantidad de flujo a transportarse en cada arco.

p = valores enteros posibles para la capacidad de cada arco.

$$\lambda_a^l = \begin{cases} 1 & \text{si se elige la componente } l \text{ del vector } s, \text{ para cada arco } a \\ 0 & \text{si no} \end{cases}$$

$$\gamma_a^r = \begin{cases} 1 & \text{si se elige la componente } r \text{ del vector } p, \text{ para cada arco } a \\ 0 & \text{si no} \end{cases}$$

Considerando esto, al linealizar el costo de un arco a por su flujo se escribe como:

$$t_a(x_a)x_a = \sum_l t_a(s_l)s_l\lambda_a^l$$

siendo s_l la componente l del vector s , $\sum_l \lambda_a^l = 1$ y $\lambda_a^l \geq 0$. Así también, la linealización de la multiplicación entre la variable dual β y la variable de aumento de la capacidad, quedaría como:

$$\beta_a z_a = \beta_a \sum_r \gamma_a^r p_r$$

con p_r la componente r del vector p , $\sum_r \gamma_a^r = 1$ y $\gamma_a^r \geq 0$.

Además, se realiza el siguiente cambio de variable: $\omega_a^r = \beta_a \gamma_a^r$ y se utilizan ecuaciones que permiten linealizar la multiplicación entre variables diferentes. Finalmente, el problema entero mixto queda como:

$$\text{mín} \quad \sum_{a \in A} \sum_l \lambda_a^l t_a(s_l) s_l \tag{3.7a}$$

$$\text{s.a} \quad \sum_{a \in A} g_a(z_a) \leq B \tag{3.7b}$$

$$\sum_l \lambda_a^l s_l = \sum_{k \in \Omega} x_a^k \quad \forall a \in A \tag{3.7c}$$

$$\sum_r \gamma_a^r p_r = z_a \quad \forall a \in A \tag{3.7d}$$

$$\sum_{a \in A} \sum_l \lambda_a^l t_a(s_l) s_l \leq \sum_{k \in \Omega} \alpha_{k_s}^k d^k - \sum_{a \in A} \sum_r \omega_a^r p_r - \sum_{a \in A} \beta_a c_a \quad (3.7e)$$

$$\alpha_i^k - \alpha_j^k - \beta_{ij} \leq \sum_l \lambda_{ij}^l t_{ij}(s_l) \quad \forall k \in \Omega, \forall a = (i, j) \in A \quad (3.7f)$$

$$\sum_{a \in \delta^+(i)} x_a^k - \sum_{a \in \delta^-(i)} x_a^k = \begin{cases} d^k, & \text{if } i = k_s, \\ -d^k, & \text{if } i = k_t, \\ 0, & \text{en otro caso,} \end{cases} \quad \forall k \in \Omega, i \in N \quad (3.7g)$$

$$0 \leq \sum_{k \in \Omega} x_a^k \leq c_a + z_a \quad \forall a \in A \quad (3.7h)$$

$$\sum_l \lambda_a^l = 1 \quad \forall a \in A \quad (3.7i)$$

$$\sum_r \gamma_a^r = 1 \quad \forall a \in A \quad (3.7j)$$

$$\sum_{l=1}^{n-1} q_a^l = 1 \quad \forall a \in A \quad (3.7k)$$

$$\lambda_a^{l+1} \leq q_a^{l+1} + q_a^l \quad \forall a \in A, \forall l = 1, \dots, n-2 \quad (3.7l)$$

$$\omega_a^r \leq \beta_a \quad (3.7m)$$

$$\omega_a^r \leq \beta^U \gamma_a^r \quad (3.7n)$$

$$\omega_a^r \geq \beta_a + \beta^U (\gamma_a^r - 1) \quad (3.7ñ)$$

$$\omega_a^r \geq 0 \quad (3.7o)$$

$$\lambda_a^l, \beta_a, \omega_a^r, x_a^k \in \mathbb{R}^+ \quad \alpha_i^k \in \mathbb{R} \quad z_a \in \mathbb{Z}_+ \quad q_a^l, \gamma_a^r \in \{0, 1\} \quad (3.7p)$$

En este problema las restricciones (3.7b), (3.7g) y (3.7h) no presentan cambios. En las ecuaciones (3.7a), (3.7e) y (3.7f) se aplica la linealización de la función de costos, que ocupa el vector s que contiene flujos por cada arco. También en esta última ecuación, se realiza la aproximación lineal de la capacidad de los arcos, que utiliza por su parte el vector p , junto con el cambio de variable propuesto. En las restricciones (3.7c) y (3.7d) se establece la relación entre las nuevas variable enteras y las variables de decisión originales. Las restricciones (3.7i) - (3.7l) permiten generar correctamente la linealización, siendo para las variables x

una aproximación lineal para valores reales y para las variables z , una aproximación para valores enteros. Las ecuaciones (3.7m) - (3.7o) son conocidas como ecuaciones de McCormick y permiten generar restricciones lineales para la multiplicación entre variables enteras y continuas (McCormick, 1972). Al final, se identifica la naturaleza de las variables en (3.7p).

3.1.2. Modelo 2: Función de latencia BPR y sin restricciones de capacidad

La gran diferencia de este segundo modelo con respecto al anterior es que la función de costos por arco es no lineal con respecto al cantidad de flujo, y además, la capacidad ingresa dentro de esta función, por lo que se remueven las restricciones relacionadas a ésta.

La función de latencia a utilizar es la función BPR (ver acápite 2.1.1), con t_{ff}^a el tiempo de viaje libre de flujo, α' y β' los parámetros de la función:

$$t_a(x_a, z_a) = t_{ff}^a \left(1 + \alpha' \left(\frac{x_a}{c_a + z_a} \right)^{\beta'} \right)$$

Este modelo es un problema de programación entero mixto, y se desarrolló con la misma metodología que el modelo anterior. Aquí se presenta el modelo con las aproximaciones duales y lineales. Por otra parte, se utiliza un cambio de variable diferente para linealizar la multiplicación entre variables: $f_a^{l,r} = \gamma_a^r \lambda_a^l$.

$$\text{mín} \quad \sum_{a \in A} \sum_l \sum_r f_a^{l,r} t_a(s_l, p_r) s_l \quad (3.8a)$$

$$\text{s.a} \quad \sum_{a \in A} g_a(z_a) \leq B \quad (3.8b)$$

$$\sum_l \lambda_a^l s_l = \sum_{k \in \Omega} x_a^k \quad \forall a \in A \quad (3.8c)$$

$$\sum_r \gamma_a^r p_r = z_a \quad \forall a \in A \quad (3.8d)$$

$$\sum_{a \in A} \sum_l \sum_r f_a^{l,r} t_a(s_l, p_r) s_l \leq \sum_{k \in \Omega} \alpha_{k_s}^k d^k \quad (3.8e)$$

$$\alpha_i^k - \alpha_j^k \leq \sum_l \sum_r f_{ij}^{l,r} t_{ij}(s_l, p_r) \quad \forall k \in \Omega, \quad \forall a = (i, j) \in A \quad (3.8f)$$

$$\sum_{a \in \delta^+(i)} x_a^k - \sum_{a \in \delta^-(i)} x_a^k = \begin{cases} d^k, & \text{if } i = k_s, \\ -d^k, & \text{if } i = k_t, \\ 0, & \text{en otro caso,} \end{cases} \quad \forall k \in \Omega, \quad i \in N \quad (3.8g)$$

$$\sum_l \lambda_a^l = 1 \quad \forall a \in A \quad (3.8h)$$

$$\sum_r \gamma_a^r = 1 \quad \forall a \in A \quad (3.8i)$$

$$\sum_{l=1}^{n-1} q_a^l = 1 \quad \forall a \in A \quad (3.8j)$$

$$\lambda_a^{l+1} \leq q_a^{l+1} + q_a^l \quad \forall a \in A, \quad \forall l = 1, \dots, n-2 \quad (3.8k)$$

$$f_a^{l,r} \leq \lambda_a^l \quad (3.8l)$$

$$f_a^{l,r} \leq \gamma_a^r \quad (3.8m)$$

$$f_a^{l,r} \geq \lambda_a^l + \gamma_a^r - 1 \quad (3.8n)$$

$$f_a^{l,r} \geq 0 \quad (3.8ñ)$$

$$\lambda_a^l, f_a^{l,r}, x_a^k \in \mathbb{R}^+ \quad \alpha_i^k \in \mathbb{R} \quad z_a \in \mathbb{Z}_+ \quad q_a^l, \gamma_a^r \in \{0, 1\} \quad (3.8o)$$

Las principales diferencias de formulación en este modelo es que al considerar que el costo de viaje depende tanto del flujo como de la capacidad del arco, entonces, para linealizar esta función por partes se debe considerar una variable que multiplique las decisiones de elección de flujo (λ_a^l) y de capacidad (γ_a^r). Esta variable $f_a^{l,r}$, por lo tanto, ingresa en la función objetivo (3.8a) y en las restricciones (3.8e) y (3.8f). Otra diferencia es que se eliminan las restricciones de capacidad por arco y, por esta misma razón, en estas últimas restricciones, no se presenta la variable dual β que se relaciona con la capacidad. Por último, las restricciones (3.8l)-(3.8ñ) indican las ecuaciones de McCormick para, en este caso, la variable auxiliar $f_a^{l,r}$, que al multiplicar una variable binaria y una variable continua en el rango $[0, 1]$ presenta algunos sutiles cambios. Todas las demás restricciones se mantienen iguales. De todas maneras, la

naturaleza de las variables se observan en la (3.8o).

3.2. Modelos robustos

En esta sección, ya construidos dos modelos deterministas distintos para el problema binivel de diseño de redes, se busca desarrollar un modelo que permita incorporar la demanda en un contexto de incertidumbre.

3.2.1. Extensión del modelo determinista a escenarios inciertos de demanda

En esta primera parte, se desarrolla la extensión a escenarios inciertos en el primer modelo determinista, para luego, generalizar la metodología al segundo modelo propuesto.

Para introducir incertidumbre al modelo, se utilizará la metodología de optimización robusta por las ventajas expuestas en la Sección 2.4. Dada esta metodología, como no se conoce cómo es el comportamiento de la demanda, permite considerar que el parámetro incierto forme parte de un conjunto cerrado y convexo.

En vista de que el problema que se aborda en esta tesis posee dos tipos de decisiones: decisiones de inversión que deben tomarse antes de conocer la demanda (decisiones *a priori*) y decisiones de flujo que se toman con el conocimiento de las condiciones reales del tráfico (decisiones *a posteriori*), entonces es posible expresar la contraparte robusta ajustable del problema de una etapa (3.2).

Sea \mathcal{U} el conjunto incierto al que pertenece la demanda d^k para cada par origen-destino $k \in \Omega$. Entonces, la contraparte robusta ajustable (ARC) se plantea como:

$$\text{(ARC) } \min \quad \gamma \tag{3.9a}$$

$$\text{s.a. } \sum_{a \in A} g_a(z_a) \leq B \tag{3.9b}$$

$$z \in \mathbb{Z}_+^m \tag{3.9c}$$

$$\forall d \in \mathcal{U} \text{ existe } x^d : \begin{cases} \sum_{a \in A} t_a(x_a^d) x_a^d \leq \gamma \\ \sum_{a \in A} t_a(x_a^d) x_a^d \leq \sum_{a \in A} t_a(x_a^d) v_a^d \quad \forall v \text{ flujo factible} \\ Ix^{d,k} = d^k \quad \forall k \in \Omega \\ 0 \leq x_a^d \leq c_a + z_a \quad \forall a \in A \end{cases} \tag{3.9d}$$

Este problema robusto se puede interpretar de la siguiente manera. Por cada escenario $d \in \mathcal{U}$, se puede encontrar un flujo que cumple dos condiciones: (1) dado la segunda restricción de equilibrio de Wardrop, este flujo está en equilibrio con respecto a la decisión de los usuarios y (2) es un flujo factible de la red puesto que cumple la tercera y cuarta restricción que implica la conservación de flujo, el cumplimiento de la demanda y la restricción de la capacidad (para el Modelo 2, esto último no se considera). Por lo que en definitiva, el flujo x^d es una respuesta directa del tipo de escenario que se presente en la demanda.

Ahora, lo que se busca al resolver el problema robusto es encontrar una solución z que pueda satisfacer todos los escenarios de flujos al mismo tiempo que minimice el costo máximo permitido para estos flujos (γ). Por lo que en concreto, la solución robusta de este problema es z , la decisión de incrementar la capacidad de los arcos.

3.2.2. Aproximación de la contraparte robusta ajustable

En general las contrapartes robustas ajustables son difíciles de resolver, y es necesaria una reformulación del problema para que pueda ser tratable. Más aún en este caso en que se considera flujos factibles, en equilibrio y de costo mínimo. Una forma de considerar el problema (ARC) anterior es generar un conjunto de restricciones de la segunda etapa por cada uno de los escenarios $d \in \mathcal{U}$, lo cual resulta inviable ya que pueden existir infinitos escenarios.

La forma clásica de reformular el problema anterior, según la metodología robusta, es resolver el problema para el peor caso de escenario de demanda, de esta manera, al minimizar el peor (máximo) costo que tiene un escenario, entonces por consecuencia, se estaría minimizando los costos de otros escenarios posibles. Pero entonces surge una pregunta importante ¿cuál es el peor caso de demanda en este problema? ¿Cómo se puede encontrar? Además, dado que aquí se mezclan flujos factibles que deben estar en equilibrio y cuyo costo es el que se desea minimizar ¿será que los costos de los flujos factibles y en equilibrio de cualquier escenario de demanda están acotados por un costo de algún flujo factible y en equilibrio de un escenario en particular?

En lo que sigue, se demuestran algunas proposiciones que serán útiles para encontrar la relación entre los costos de los escenarios de demanda como también la definición de un peor caso, para así desarrollar una reformulación que se resuelva eficientemente.

Proposición 3.1 *Sea $d \in \mathcal{U}$ la combinación convexa de los puntos extremos del conjunto, es decir, $d = \sum_j \theta_j d^j$, con d^j punto extremo del conjunto \mathcal{U} , $\sum_j \theta_j = 1$ y $\theta_j \geq 0$. Entonces x^d , un flujo factible de la demanda d , es una combinación convexa de flujos factibles de los puntos extremos x^{d^j} , es decir, $x^d = \sum_j \theta_j x^{d^j}$.*

DEMOSTRACIÓN. Primero, se realiza esta demostración por cada par origen-destino, para luego generalizar. Sea un flujo factible $x^{d^j,k}$ del par OD k con respecto a la demanda de un punto extremo d^j . Este es un vector en el que cada componente corresponde al flujo de k en el arco a , es decir, $x^{d^j,k} = (x_a^{d^j,k})_{a \in A}$. Dado que $x^{d^j,k}$ es factible, entonces cumple la tercera restricción de (3.9d) de conservación de flujo. Por lo tanto, s.p.g, para el nodo $i = k_s \in N$ se tiene que:

$$d^{j,k} = \sum_{a \in \delta^+(k_s)} x_a^{d^j,k} - \sum_{a \in \delta^-(k_s)} x_a^{d^j,k} \quad (3.10a)$$

$$\Rightarrow d^k = \sum_j \theta_j \left(\sum_{a \in \delta^+(k_s)} x_a^{d^j,k} - \sum_{a \in \delta^-(k_s)} x_a^{d^j,k} \right) \quad (3.10b)$$

$$d^k = \sum_{a \in \delta^+(k_s)} \sum_j \theta_j x_a^{dj,k} - \sum_{a \in \delta^-(k_s)} \sum_j \theta_j x_a^{dj,k} \quad (3.10c)$$

De (3.10a) a (3.10b) se considera el hecho de que d es combinación convexa de los puntos extremos del conjunto \mathcal{U} y gracias a esto, cada componente d^k también cumple que $d^k = \sum_j \theta_j d^{j,k}$.

Ahora, si se considera que $x^{d,k}$ flujo factible cumple que $x_a^{d,k} = \sum_j \theta_j x_a^{dj,k}$, es decir, que el flujo factible de la demanda d en el par origen-destino k es combinación convexa de los flujos factibles de los puntos extremos de \mathcal{U} , reemplazando en (3.10c), se obtiene que:

$$d^k = \sum_{a \in \delta^+(k_s)} x_a^{d,k} - \sum_{a \in \delta^-(k_s)} x_a^{d,k}$$

Lo que satisface la restricción de conservación de flujo pues lo anterior se cumple $\forall k \in \Omega$.

Con esto, y dado que $x_a^d = \sum_{k \in \Omega} x_a^{d,k}$, es posible derivar que:

$$\Rightarrow \sum_{k \in \Omega} x_a^{d,k} = \sum_{k \in \Omega} \sum_j \theta_j x_a^{dj,k} \quad (3.11a)$$

$$\Rightarrow x_a^d = \sum_j \theta_j x_a^{dj} \quad (3.11b)$$

$$\Rightarrow x^d = \sum_j \theta_j x^{dj} \quad (3.11c)$$

Del mismo modo, para la restricción de capacidad, dado los flujos factibles x_a^{dj} , se cumple que x_a^d también satisface la capacidad.

$$x_a^d = \sum_{k \in \Omega} x_a^{d,k} = \sum_{k \in \Omega} \sum_j \theta_j x_a^{dj,k} = \sum_j \theta_j \sum_{k \in \Omega} x_a^{dj,k} \leq \sum_j \theta_j c_a = c_a \quad (3.12a)$$

□

Proposición 3.2 Sea $g_a(x_a) = t_a(x_a)x_a$ una función convexa que representa el costo del flujo factible x_a para cada arco a . Si $x^d = \sum_j \theta_j x^{dj}$, con x^d un flujo factible de la demanda d , d^j un punto extremo de \mathcal{U} , $\sum_j \theta_j = 1$ y $\theta_j \geq 0$. Entonces, $c(x^d) = \sum_a g_a(x_a^d) \leq \sum_j \theta_j c(x^{dj})$

DEMOSTRACIÓN. Sea $x^d = \sum_j \theta_j x^{dj}$ un flujo factible para la demanda d , con $\sum_j \theta_j = 1$ y $\theta_j \geq 0$. Sea $g_a(\cdot)$ una función convexa. Dada la definición de $c(x^d)$, se tiene que:

$$c(x^d) = \sum_a g_a(x_a^d) \quad (3.13a)$$

$$= \sum_a g_a\left(\sum_j \theta_j x_a^{dj}\right) \quad (3.13b)$$

$$\leq \sum_a \sum_j \theta_j g_a(x_a^{dj}) \quad (3.13c)$$

$$= \sum_j \theta_j \sum_a g_a(x_a^{dj}) \quad (3.13d)$$

$$= \sum_j \theta_j c(x^{dj}) \quad (3.13e)$$

De (3.13b) a (3.13c) se utiliza la propiedad de convexidad de la función $g_a(\cdot)$. □

Proposición 3.3 Sea $d \in \mathcal{U}$ la combinación convexa de los puntos extremos del conjunto, es decir, $d = \sum_j \theta_j d^j$, con d^j punto extremo del conjunto \mathcal{U} , $\sum_j \theta_j = 1$ y $\theta_j \geq 0$. Sea \bar{x}^d un flujo factible para la demanda $d \in \mathcal{U}$ y que está en equilibrio con respecto a la decisión de los usuarios. Sea además, $c(x^d) = \sum_a g_a(x_a^d) = \sum_a t_a(x_a^d)x_a^d$ con $g_a(\cdot)$ una función convexa.

Entonces existen cotas inferiores y superiores tal que

$$\max_j \{c(\bar{x}^{d^j})\} \leq \max_{d \in \mathcal{U}} \{c(\bar{x}^d)\} \leq \alpha \max_j \{c(\bar{x}^{d^j})\}$$

con d^j punto extremo del conjunto \mathcal{U} y $\alpha \geq 1$ el coeficiente del precio de la anarquía del juego de congestión.

DEMOSTRACIÓN. Sea x^d un flujo factible, \bar{x}^d un flujo factible y que, además, está en equilibrio con respecto a la decisión de los usuarios, y x^{*d} el flujo factible de costo mínimo, todos con respecto a la demanda $d \in \mathcal{U}$.

Además, para cada d^j , punto extremo del conjunto \mathcal{U} , encontramos los flujos factibles de costo mínimo x^{*d^j} .

Considerando un flujo factible \hat{x}^d de la demanda d , por la Proposición 3.1 se puede escribir como $\hat{x}^d = \sum_j \theta_j x^{*d^j}$, con $\sum_j \theta_j = 1$ y $\theta_j \geq 0$. Por lo tanto, se tiene que:

$$c(\hat{x}^d) \leq \sum_j \theta_j c(x^{*d^j}) \quad (3.14a)$$

$$c(x^{*d^j}) \leq c(\bar{x}^{d^j}) \quad \forall j \quad (3.14b)$$

Con (3.14a) por la Proposición 3.2 y (3.14b) por el hecho de que x^{*d^j} es el flujo de costo mínimo.

Ahora, considerando la formulación del precio de la anarquía, con α el valor de este ratio determinado por Roughgarden (2003) para funciones de latencia lineales y polinomiales.

$$\frac{c(\bar{x}^d)}{c(x^{*d})} \leq \alpha \quad (3.15)$$

Se obtiene que,

$$c(\bar{x}^d) \leq \alpha c(x^{*d}) \quad (3.16a)$$

$$\leq \alpha c(\hat{x}^d) \quad (3.16b)$$

$$\leq \alpha \sum_j \theta_j c(x^{*d^j}) \quad (3.16c)$$

$$\leq \alpha \sum_j \theta_j c(\bar{x}^{d^j}) \quad (3.16d)$$

$$\leq \alpha \sum_j \theta_j \max_k \{c(\bar{x}^{d^k})\} \quad (3.16e)$$

$$\Rightarrow c(\bar{x}^d) \leq \alpha \max_j \{c(\bar{x}^{dj})\} \quad (3.16f)$$

La primera equivalencia hace referencia a que $c(x^{*d})$ siempre es menor que cualquier otro flujo factible. Además, de (3.16b) a (3.16c) se utilizó (3.14a) y de (3.16c) a (3.16d) se utilizó (3.14b).

Como lo anterior se cumple para cualquier $d \in \mathcal{U}$ entonces particularmente se cumple para $\max_{d \in \mathcal{U}} \{c(\bar{x}^d)\}$.

Por otro lado, el peor costo de los flujos \bar{x} de demandas extremas puede pertenecer o no al peor caso de demanda del conjunto. Es decir,

$$\max_j \{c(\bar{x}^{dj})\} \leq \max_{d \in \mathcal{U}} \{c(\bar{x}^d)\} \quad (3.17)$$

Por lo tanto,

$$\Rightarrow \max_j \{c(\bar{x}^{dj})\} \leq \max_{d \in \mathcal{U}} \{c(\bar{x}^d)\} \leq \alpha \max_j \{c(\bar{x}^{dj})\} \quad (3.18)$$

□

La conclusión más importante de todas estas proposiciones es la siguiente: si \mathcal{U} , el conjunto de incertidumbre, es un conjunto en el que cualquier punto se puede escribir como una combinación convexa de los puntos extremos del conjunto y, además, la función de costos $g_a(x_a) = t(x_a)x_a$ por cada arco es convexa, entonces, el costo máximo, sobre todas las demandas $d \in \mathcal{U}$, de los flujos factibles y en equilibrio está acotado inferiormente por el costo de un flujo factible y en equilibrio correspondiente a un punto extremo de \mathcal{U} y superiormente por ese mismo costo multiplicado por el valor del precio de la anarquía. Dicho en otras palabras, el peor escenario del problema se encuentra bien determinado por el peor caso de las demandas extremas del conjunto \mathcal{U} y, al calcular el peor costo de esos puntos extremos, sólo es necesario multiplicar por una constante para obtener una cota superior. Cabe destacar

que los conjuntos que tienen la propiedad de que cualquier elemento dentro de él son una combinación convexa de sus puntos extremos son los poliedros.

Sin embargo, también es posible obtener una mejor cota superior para el peor caso de la demanda considerando los flujos óptimos de las demandas extremas. Si considera la ecuación (3.16b) de la demostración de la Proposición 3.3, y de ahí se obtuviera el máximo sobre todos los puntos extremos d^j entonces se tiene que:

$$\max_{d \in \mathcal{U}} \{c(\bar{x}^d)\} \leq \alpha \max_j \{c(x^{*d^j})\} \leq \alpha \max_j \{c(\bar{x}^{d^j})\} \quad (3.19)$$

Ahora, como se pudo demostrar, se puede obtener soluciones factibles robustas muy buenas y que se acerquen al óptimo considerando los puntos extremos del conjunto de demanda. Por lo tanto, una buena aproximación del problema robusto (ARC) es considerar el peor caso de los puntos extremos de \mathcal{U} y utilizar un conjunto de incertidumbre \mathcal{U} poliedral. Lo que en definitiva, permite convertir el problema (ARC) en un problema finito y, por lo tanto, computacionalmente tratable.

Por consiguiente, la aproximación de (ARC) se describe como:

$$\text{(ARCA) } \min \quad \gamma \quad (3.20a)$$

$$\text{s.a. } \sum_{a \in A} g_a(z_a) \leq B \quad (3.20b)$$

$$\max_{j=1, \dots, p} \left\{ \sum_{a \in A} t_a(x_a^{d^j}) x_a^{d^j} \right\} \leq \gamma \quad (3.20c)$$

$$\sum_{a \in A} t_a(x_a^{d^j}) x_a^{d^j} \leq \sum_{a \in A} t_a(x_a^{d^j}) v_a^{d^j} \quad \forall v \text{ factible } \quad \forall j = 1, \dots, p \quad (3.20d)$$

$$I x^{d^j, k} = d^k \quad \forall k \in \Omega \quad \forall j = 1, \dots, p \quad (3.20e)$$

$$0 \leq x_a^{d^j} \leq c_a + z_a \quad \forall a \in A \quad \forall j = 1, \dots, p \quad (3.20f)$$

$$z \in \mathbb{Z}_+^m, \quad x \in \mathbb{R}^{+m \times |\Omega| \times p} \quad (3.20g)$$

Es preciso señalar que este modelo permite obtener soluciones robustas factibles, pero no se sabe si éstas soluciones son óptimas para problema robusto (ARC) pues se desconoce si algún d^j es el peor escenario de demanda. Sin embargo, son buenos escenarios para poder encontrar cotas del costo óptimo del problema robusto (ARC), puesto que el costo óptimo de (ARCA) (C_{ARCA}^*) es una cota inferior del costo óptimo de (ARC) y $\alpha \cdot C_{ARCA}^*$ es una cota superior, considerando α como el precio de la anarquía. Además, en lugar de considerar restricciones de factibilidad y equilibrio por cada d , sólo se considera para las demanda de puntos extremos d^j , que son posibles de calcular.

En la práctica, la restricción del máximo (3.20c), que es no lineal, se puede linealizar considerando que el costo de cada flujo de un punto extremo es acotado por γ . Finalmente, las contrapartes robustas de ambos modelos quedan de la siguiente manera.

Modelo robusto 1: Función de costo lineal y con restricciones de capacidad

$$\text{mín } \gamma \tag{3.21a}$$

$$\text{s.a } \sum_{a \in A} g_a(z_a) \leq B \tag{3.21b}$$

$$\sum_{a \in A} t_a(x_a^{dj})x_a^{dj} \leq \gamma \quad \forall j = 1, \dots, p \tag{3.21c}$$

$$\sum_{a \in A} t_a(x_a^{dj})x_a^{dj} \leq \sum_{a \in A} t_a(x_a^{dj})v_a^{dj} \quad \forall v \text{ factible } \forall j = 1, \dots, p \tag{3.21d}$$

$$Ix^{dj,k} = d^k \quad \forall k \in \Omega \quad \forall j = 1, \dots, p \tag{3.21e}$$

$$0 \leq x_a^{dj} \leq c_a + z_a \quad \forall a \in A \quad \forall j = 1, \dots, p \tag{3.21f}$$

$$z \in \mathbb{Z}_+^m, x \in \mathbb{R}^{+m \times |\Omega| \times p} \tag{3.21g}$$

Modelo robusto 2: Función de latencia BPR y sin restricciones de capacidad

$$\text{mín } \gamma \tag{3.22a}$$

$$\text{s.a } \sum_{a \in A} g_a(z_a) \leq B \tag{3.22b}$$

$$\sum_{a \in A} t_a(x_a^{d^j}, z_a) x_a^{d^j} \leq \gamma \quad \forall j = 1, \dots, p \quad (3.22c)$$

$$\sum_{a \in A} t_a(x_a^{d^j}, z_a) x_a^{d^j} \leq \sum_{a \in A} t_a(x_a^{d^j}, z_a) v_a^{d^j} \quad \forall v \text{ factible} \quad \forall j = 1, \dots, p \quad (3.22d)$$

$$I x^{d^j, k} = d^k \quad \forall k \in \Omega \quad \forall j = 1, \dots, p \quad (3.22e)$$

$$z \in \mathbb{Z}_+^m, x \in \mathbb{R}^{+m \times |\Omega| \times p} \quad (3.22f)$$

Cabe destacar que tanto para el primer modelo robusto como para el segundo, la restricción de equilibrio de Wardrop se reemplaza por las restricciones construidas para los modelos deterministas. Además, los costos de transporte de ambos modelos son funciones convexas. De esta manera se construyen modelos robustos que se pueden resolver eficientemente de forma exacta.

Capítulo 4

Resultados Computacionales

En este capítulo se presentan experimentos computacionales que buscan comparar las soluciones deterministas y las soluciones robustas de los dos tipos de modelos formulados, analizando la mejora (o pérdida) relativa de la solución robusta por sobre la determinista. El objetivo es analizar el comportamiento de estos costos, bajo distintos escenarios de demanda, para poder determinar en qué casos es mejor utilizar un modelo sobre otro, considerando las dos formulaciones y la incorporación de incertidumbre.

Se presentan resultados en dos tipos de redes. En primer lugar, se ocupa una red pequeña que permite analizar de manera concreta cómo cambian los flujos de los arcos con ambos modelos y también permite variar, obteniendo resultados más rápidos, los distintos parámetros del problema, estableciendo una intuición del comportamiento de las soluciones. En segundo lugar, se utiliza una red más grande, que se acerca más a la configuración de una red de calles reales, con la cual se permitirá verificar la intuición anteriormente desarrollada, observar los cambios que se presentan al aumentar el tamaño del problema y evaluar la dificultad de resolución. En específico, en esta oportunidad se utiliza la red de *Sioux Falls*, la cual se utiliza ampliamente en la literatura (los datos de esta red se obtienen de (LeBlanc et al., 1975)).

4.1. Implementación

4.1.1. Conjunto de incertidumbre

Para la realización de los experimentos de esta sección se define un conjunto de incertidumbre para la demanda. Por lo demostrado en el capítulo anterior, para que la aproximación del problema robusto (ARCA) tenga sentido, el conjunto debe ser un poliedro. El conjunto de incertidumbre restringido de cardinalidad propuesto por Bertsimas y Sim (2003) es un conjunto poliedral.

Se escoge este conjunto de incertidumbre por la simplicidad de su formulación, además de que es posible controlar la variabilidad de la demanda a través del valor del presupuesto de incertidumbre (Γ). Como se explica en el acápite 2.4.2, este presupuesto de incertidumbre indica un máximo en la cantidad de coeficientes que pueden variar. De forma intuitiva, es poco probable que todos los parámetros cambien al mismo tiempo en los diferentes escenarios de demanda, por eso, al utilizar este tipo de conjunto se considera que sólo un subconjunto de demandas cambiará. Por último, se tiene la ventaja de que con este conjunto es simple calcular los puntos extremos, los cuales son claves en la resolución del problema.

La forma del conjunto de incertidumbre se define a continuación:

$$\mathcal{U} = \left[d \mid \underline{d}_k \leq d_k \leq \underline{d}_k + \rho z_k, \sum_{k=1}^n z_k \leq \Gamma, 0 \leq z_k \leq 1 \forall k = 1, \dots, n \right] \quad (4.1)$$

Donde ρ es un parámetro que permite variar el tamaño del intervalo en el que se mueve cada componente d_k , con \underline{d}_k un valor fijo de demanda, y Γ el presupuesto de incertidumbre.

4.1.2. Métricas de evaluación y tipos de análisis

Por cada experimento a realizar se calcula la solución z , que es el incremento de capacidad en los arcos, de ambos o de un modelo propuesto. Se define z_{det} como la solución del problema determinista y z_{rob} como la solución del problema robusto, que corresponden,

respectivamente, al caso promedio de demanda (o demanda nominal) y al peor caso de demanda.

En los experimentos con redes pequeñas se presentan y se comparan los resultados de ambos escenarios para diferentes niveles de incertidumbre (que se obtienen variando el valor del parámetro ρ del conjunto de incertidumbre (4.1) como también el valor del presupuesto de incertidumbre Γ) y para diferentes valores de presupuesto B . Por otro lado, para la red de *Sioux Falls*, se analiza cómo varían los resultados al cambiar el intervalo de discretización de los valores que linealizan las funciones no lineales.

Para evaluar el *trade-off* de utilizar la solución robusta, junto con obtener la inversión total en capacidad de ambas soluciones, se calculan los siguientes costos:

C_d^* costo óptimo del problema determinista.

C_{ar}^* costo óptimo de la aproximación del modelo robusto, que es cota inferior del costo óptimo robusto. Como se mencionó, al multiplicar por una constante este costo, se obtiene una cota superior del costo óptimo robusto.

C_{cp}^* valor de la función objetivo de la solución robusta en el caso promedio.

C_{pc}^* valor de la función objetivo de la solución determinista en el peor caso.

Con estos costos, se establecen las siguientes métricas que miden el incremento y la disminución relativa de la solución robusta frente a la solución determinista en los diferentes casos:

r_{cd} incremento relativo del costo de la solución robusta en el caso promedio. Este valor es positivo.

$$r_{cp} = \frac{C_{cp}^* - C_d^*}{C_d^*}$$

r_{pc} disminución relativa del costo de la solución robusta en el peor caso. Este valor es negativo.

$$r_{pc} = \frac{C_{ar}^* - C_{pc}^*}{C_{pc}^*}$$

Adicionalmente, se realizaron análisis con escenarios de demandas simuladas. Para ello, se generaron 1000 simulaciones de escenarios bajo una distribución normal que, debido a que la demanda es discreta, es aproximada por una distribución Bernoulli dentro del intervalo considerado en el conjunto de incertidumbre (4.1) (para ver la aproximación realizada revisar Anexo A). Para cada escenario simulado y para cada tipo de modelo, se calcula el costo de la solución factible del problema robusto y el costo óptimo determinista, graficando la distribución obtenida de estos costos.

Por último, los experimentos fueron programados en el lenguaje Python y se utilizó Gurobi v9.0.2 para solucionar los modelos en forma exacta. El computador utilizado para ejecutar el programa tiene sistema operativo Windows 10 Pro con un procesador Intel Core i5-6200U, 2.30GHz (4 CPUs) y una memoria RAM de 8 GB.

4.2. Resultados en redes pequeñas

Los primeros experimentos se realizan en la red que se presenta a la Figura 1.1, usada para ejemplificar el problema a resolver y como motivación para el desarrollo de esta tesis. Esta red tiene 4 nodos, 5 arcos y 2 pares origen-destino. Como se ha mencionado, los parámetros del problema y que se muestran en la Figura 1.1 son los costos de viaje (t_a), costos de inversión (g_a) y la capacidad (c_a) para cada arco a . Además, de acuerdo a lo comentado, existe un arco que tiene un tiempo de viaje menor a todos los demás, pero que su costo de construcción de capacidad es muy alto (autopista, $arco = (2, 3)$).

La demanda nominal \hat{d}_k es de 10 para el par $OD = (1, 4)$ y 15 para el par $OD = (2, 3)$. Para el modelo robusto se considera que la demanda pertenece al conjunto de incertidumbre (4.1), con $\underline{d}_k = \hat{d}_k - sd$ y $\rho = 2sd$ (siendo sd la desviación estándar y el parámetro que se varía para los distintos niveles de incertidumbre).

Para el primer tipo de modelo MIP aproximado, en el que es necesario discretizar los flujos por arco (s) como el aumento de capacidad de los arcos (p), se determina que $s = [0, 25]$ y $p = [0, 15]$, ambos con intervalos discretos de tamaño 1. Como en este modelo la función de latencia es lineal el precio de la anarquía es igual a $4/3$.

Para el segundo tipo de modelo, por su parte, tiene los mismos rangos que el modelo anterior pero con un intervalo discretos de tamaño 2 en vez de 1. Además, la función BPR en este modelo tiene como parámetros $\alpha' = 0,15$ y $\beta' = 4$, valores que comúnmente se utilizan en este tipo de función (LeBlanc et al., 1975) y también se considera que $t_{ff}^a = t_a$. Por su parte, la función BPR es un caso particular de una función de latencia polinomial, por lo que el precio de la anarquía equivale a $\frac{(\beta'+1)^{1+1/\beta'}}{(\beta'+1)^{1+1/\beta'} - \beta'}$, puesto que β' es el máximo grado de la función (véase el acápite 2.3.2).

A continuación se mencionan los experimentos que se realizan en esta sección:

Experimento 1: Se considera el grafo de la red simple con incertidumbre sólo en el par $OD = (2, 3)$. El par $OD = (1, 4)$ tiene demanda determinista en ambos modelos.

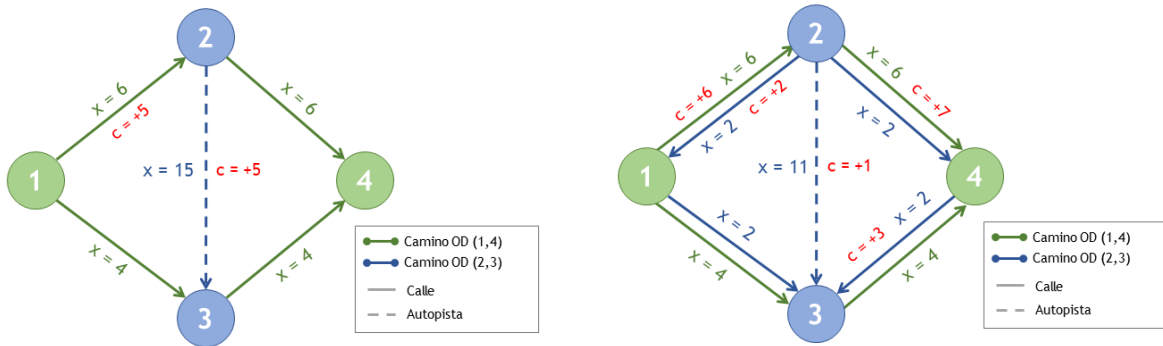
En este experimento se muestran las soluciones deterministas y robustas de ambos modelos gráficamente en las Figuras 4.1 y 4.2 respectivamente. Además, se llevan a cabo simulaciones que se muestran en las Figuras 4.3 y 4.4.

Junto con ello, se muestran los costos y las métricas de ambos modelos para distintos presupuestos B , en las Tablas 4.1 y 4.2; y también se muestra el cambio relativo de las soluciones variando la desviación estándar sd en el gráfico de la Figura 4.5.

Experimento 2: Se considera el grafo de la red simple con incertidumbre en ambos pares OD . Con este experimento sólo se analizan las métricas variando la incertidumbre a través de los parámetros sd y Γ , usando $\Gamma = 1$ en la Figura 4.6 y $\Gamma = 2$ en la Figura 4.7.

4.2.1. Soluciones gráficas y simulaciones

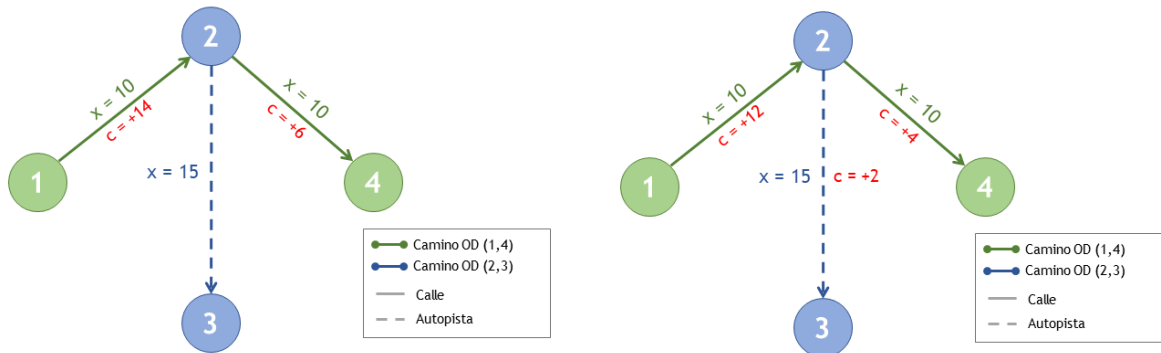
Los siguientes resultados forman parte del Experimento 1, fijando los parámetros $B = 600$ y $sd = 7$. El presupuesto $B = 600$ es el presupuesto en el cual se llega a un costo de viaje mínimo, es decir, con presupuestos mayores, el costo de viaje sigue manteniéndose igual (véase Anexo B .1). Por otra parte, considerar $sd = 7$ significa considerar un desviación estándar de un 50% de la demanda nominal del par $OD = (2, 3)$.



(a) Solución determinista

(b) Solución robusta

Figura 4.1: Aumento de capacidad y flujos óptimos x por cada solución en el escenario promedio. **Modelo 1.** Grafo 1.1, con incertidumbre en el par $OD = (2,3)$.



(a) Solución determinista

(b) Solución robusta

Figura 4.2: Aumento de capacidad y flujos óptimos x por cada solución en el escenario promedio. **Modelo 2.** Grafo 1.1, con incertidumbre en el par $OD = (2,3)$.

Se puede observar que los resultados del Modelo 1 (Figura 4.1) son intuitivos, tal como se dedujo en la introducción del problema, puesto que al utilizar una demanda determinista

se privilegia a la demanda del par $OD = (2, 3)$ que transite por el arco central, aumentando la capacidad lo necesario para que cubra la demanda. Del mismo modo, para satisfacer la demanda de los transeúntes del par $OD = (1, 4)$ sólo se aumenta la capacidad en el arco más barato lo suficiente para que viajen aquellos usuarios que no caben por el camino de abajo a través del camino de arriba.

Por otro lado, en la versión robusta del Modelo 1, como no se sabe con certeza cuál es el valor de la demanda, hay que considerar todos los escenarios incluyendo el peor caso. Por eso mismo, no vale la pena construir demasiada capacidad en la autopista porque pueden ocurrir situaciones de demandas pequeñas en el par $OD = (2, 3)$ que terminarían ocupando un bajo porcentaje de la capacidad. Por esa razón, la solución robusta busca, para no incurrir en mayores costos de inversión, que cuando existan demandas grandes, éstas viajen por arcos más baratos. En definitiva, por lo mismo se aumenta capacidad en las calles del grafo, y en consecuencia, en la solución robusta la demanda determinista del par $OD = (2, 3)$ se divide en dos caminos.

Un aspecto interesante a destacar es que la restricción presupuestaria busca ser activa en este modelo, porque era igualmente factible aumentar sólo 4 unidades de capacidad en el arco $(1, 2)$ pero se aumentaron 5 unidades que permiten alcanzar el presupuesto de 600.

En el Modelo 2 (Figura 4.2) se aprecian resultados diferentes por cómo se desempeña la función de BPR y puesto que también no existen restricciones de capacidad. En este caso la inversión de capacidad tiene directa relación con la disminución de los costos de viaje, porque la capacidad es parte de la función objetivo. De esta manera, la demanda del par $OD = (1, 4)$ no necesita dividirse en caminos diferentes, sino que transita por el camino con arcos más baratos (que es el camino de arriba), aumentando la capacidad para disminuir los costos de transporte por ese trayecto. Ahora, no aumenta capacidad en el arco central porque es muy costoso, entonces proyecta la disminución de costos en los arcos más económicos.

Sin embargo, cuando existe incertidumbre en la demanda del arco central, es necesario

aumentar capacidad en ese enlace, puesto que pueden existir altas demandas en algunos casos y es necesario disminuir los tiempos de viajes para evitar la congestión. Se destaca también, que para considerar los peores casos de demanda, la solución robusta invierte mucho más en capacidad.

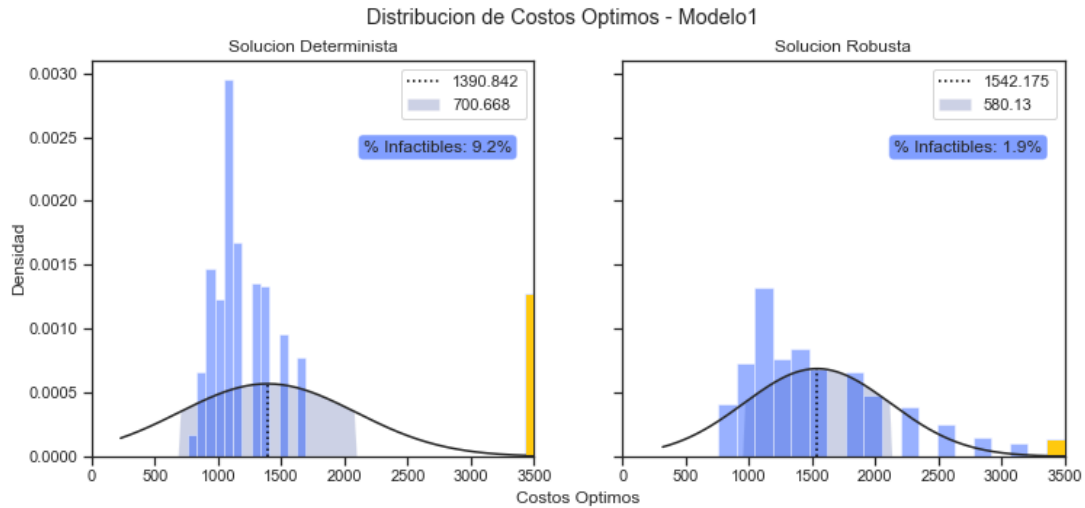


Figura 4.3: Simulación de escenarios de demanda y su distribución de costos. La línea punteada muestra el promedio de la distribución y el área sombreada, su desviación estándar. **Modelo 1.** Grafo 1.1, con incertidumbre en el par $OD = (2,3)$.

En la Figura 4.3 se puede ver que existen algunas instancias infeasibles en el Modelo 1 y esto ocurre por la restricción de capacidad que no es suficiente para acomodar todos los flujos. Sin embargo, se observa que la solución determinista tiene más instancias infeasibles (9,2%) que la solución robusta (1,9%), por lo que el modelo robusto logra resolver más escenarios que el modelo determinista.

En otro aspecto, la solución robusta muestra un costo de viaje promedio mayor (1542,18 vs. 1390,84) debido a que en general los costos de algunas instancias son peores con la solución robusta, aunque la desviación estándar es mayor en el determinista (700,67 vs. 580,13) porque alcanzan peores costos para las instancias alejadas del promedio.

En los resultados del Modelo 2, que se encuentran en la Figura 4.4, no hay infeasibilidades, observando claramente el comportamiento de ambas soluciones y cómo se diferencian. Se muestra que la solución determinista efectivamente llega a costos más altos que la solución

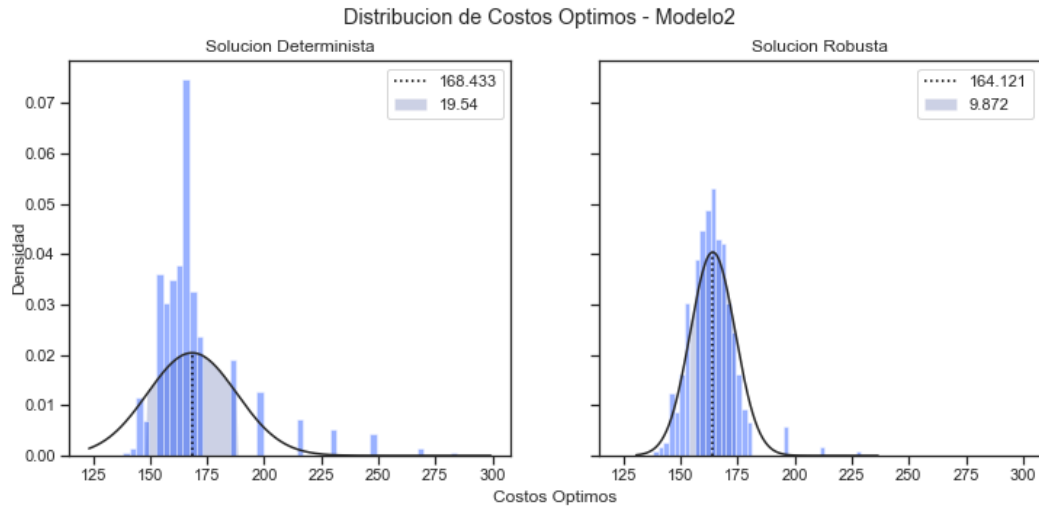


Figura 4.4: Simulación de escenarios de demanda y su distribución de costos. La línea punteada muestra el promedio de la distribución y el área sombreada, su desviación estándar. **Modelo 2.** Grafo 1.1, con incertidumbre en el par OD = (2,3).

robusta, y además, tiene una desviación estándar mayor (19,5 vs. 9,9). Es decir, el modelo robusto tiene costos más conservadores porque trata de hacerse cargo de todos los escenarios que pertenecen al conjunto de incertidumbre. En particular, para este modelo robusto, que es una aproximación del problema robusto original que considera todos los escenarios, se enfoca en un subconjunto de escenarios y permite que todos estos casos sean eficientes.

Por otro lado, en la solución determinista se observan muchas más instancias cerca del costo promedio. Esto quiere decir que cerca del promedio el modelo determinista logra obtener instancias con menores costos que el robusto no puede alcanzar, por lo que en ese sentido, en estos casos, el modelo robusto es sub-óptimo.

4.2.2. Análisis variando presupuesto

En lo que sigue, se muestran resultados del Experimento 1 variando el presupuesto y fijando la desviación estándar $sd = 7$.

Lo primero que se destaca en estos resultados es que, como se observaba en las simulaciones, el Modelo 1 sigue obteniendo instancias infactibles. En este caso, para presupuestos menores a 600, la versión robusta no tiene suficiente inversión para atender

| Presupuesto | Tipo de solución | Costo Inversión | Tiempo de ejecución (s) | Caso Promedio | | Peor Caso |
|-------------|------------------|-----------------|-------------------------|----------------|--------------------------------|----------------|
| | | | | Costo de viaje | % aumento con solución robusta | Costo de viaje |
| B = 350 | Determinista | 280 | 0,3637 | 1342 | - | Infactible |
| | Robusta | INFACTIBLE | | | | |
| B = 600 | Determinista | 600 | 0,5156 | 1174 | 22,15 % | Infactible |
| | Robusta | 590 | 0,8123 | 1434 | | 3120 |
| B = 900 | Determinista | 900 | 0,7108 | 1174 | 0 % | 1840 |
| | Robusta | 880 | 0,6783 | 1174 | | 1840 |

Tabla 4.1: Costos y diferencia relativa de la solución robusta en cada caso, variando el presupuesto B . **Modelo 1.** Grafo 1.1, con incertidumbre en el par $OD = (2,3)$.

| Presupuesto | Tipo de solución | Costo Inversión | Tiempo de ejecución (s) | Caso Promedio | | Peor Caso | |
|-------------|------------------|-----------------|-------------------------|----------------|--------------------------------|----------------|------------------------------------|
| | | | | Costo de viaje | % aumento con solución robusta | Costo de viaje | % disminución con solución robusta |
| B = 100 | Determinista | 80 | 1,63 | 178,75 | 0,00 % | 252,8 | 0,00 % |
| | Robusta | 80 | 2,81 | 178,75 | | 252,8 | |
| B = 200 | Determinista | 200 | 3,33 | 172 | 25,23 % | 245,3 | -5,63 % |
| | Robusta | 200 | 3,70 | 215,4 | | 231,5 | |
| B = 350 | Determinista | 320 | 0,73 | 164,5 | 4,92 % | 236,9 | -18,70 % |
| | Robusta | 340 | 1,81 | 172,6 | | 192,6 | |
| B = 600 | Determinista | 460 | 121,78 | 164,5 | 0,00 % | 230,9 | -21,78 % |
| | Robusta | 520 | 2,64 | 164,5 | | 180,6 | |

Tabla 4.2: Costos y diferencia relativo de la solución robusta en cada caso, variando el presupuesto B . **Modelo 2.** Grafo 1.1, con incertidumbre en el par $OD = (2,3)$.

a todos los escenarios y más aún al peor caso (Tabla 4.1).

Por el contrario, en el Modelo 2 sí es posible obtener soluciones factibles para presupuestos menores. Además, se evidencia claramente el *trade-off* entre las soluciones a medida que se aumenta el presupuesto. En primer lugar, se observa que cuando se cuenta con un presupuesto menor, el modelo robusto no logra disminuir más el costo de viaje en el peor caso porque no le alcanza el presupuesto para aumentar la capacidad, lo que conlleva a que la solución robusta sea igual a la determinista.

Ahora, si aumentamos el presupuesto, la versión robusta del Modelo 2 es capaz de construir capacidad en ciertos arcos para disminuir el costo en el peor caso. Por ejemplo, para $B = 200$, utiliza los recursos para mejorar el peor costo, pero como no tiene suficiente, no se preocupa del caso promedio, y el costo de viaje en ese escenario es un poco mayor que el

costo determinista óptimo. Pero cuando puede invertir más, es capaz de buscar una solución que disminuya el peor caso y que también no empeore el caso promedio. Es más, con $B = 600$ la solución robusta iguala el costo en el caso promedio y mejora significativamente el peor caso (Tabla 4.2).

Finalmente se analiza que, en general, el Modelo 1 necesita más inversión para alcanzar soluciones óptimas que el Modelo 2, y en este último, su complejidad no influye en mayor medida en el tiempo de resolución en este tipo de redes.

4.2.3. Análisis variando nivel de incertidumbre

A continuación se analiza el aumento de la incertidumbre de la demanda sólo con el Modelo 2, incluyendo resultados del Experimento 1 y del Experimento 2.

Se utiliza sólo el Modelo 2 ya que se comporta mejor y no deja instancias infactibles. Dado que en el Modelo 1 la mayoría de las instancias generan infactibilidad, no se permitirá apreciar bien el comportamiento de la solución robusta ante la variación de la desviación estándar. En estos resultados se fija $B = 600$, porque es el valor del presupuesto en que el Modelo 2 tiene un mayor impacto de mejora en el peor caso.

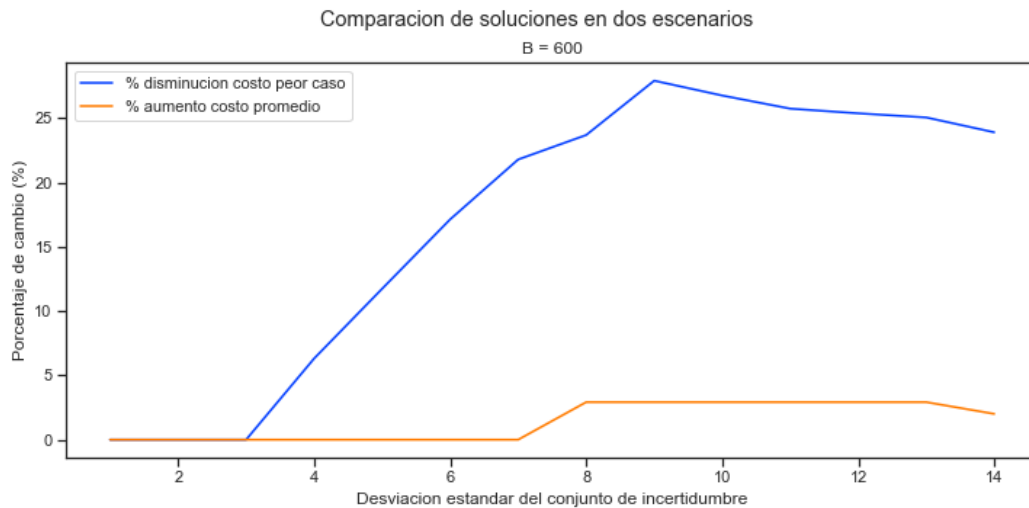


Figura 4.5: Diferencia relativa del costo de la solución robusta en el caso promedio y en el peor caso, como función del tamaño (sd) del conjunto de incertidumbre. **Modelo 2.** Grafo 1.1, con incertidumbre en el par $OD = (2,3)$.

En la Figura 4.5 se puede apreciar efectivamente cómo la solución robusta se protege del peor caso y disminuye significativamente el costo de éste a medida que aumenta la incertidumbre de la demanda. Más aún, sólo incurre en una modesta pérdida de optimalidad ($< 5\%$) en el costo del caso promedio y sólo cuando la incertidumbre es mayor.

Del gráfico también se desprende que a medida que aumenta la incertidumbre, la solución robusta protege muchísimo el peor caso hasta un punto (cerca de $sd = 10$) en que comienza a disminuir la mejora relativa, y en ese momento, también la solución robusta comienza a perder optimalidad frente al escenario nominal. De esto se deduce que en una situación con mucha incertidumbre (y por lo tanto con casos muy extremos), el modelo robusto sólo se preocupa de mejorar el peor caso pero perdiendo de vista la optimalidad de los escenarios cerca del promedio.

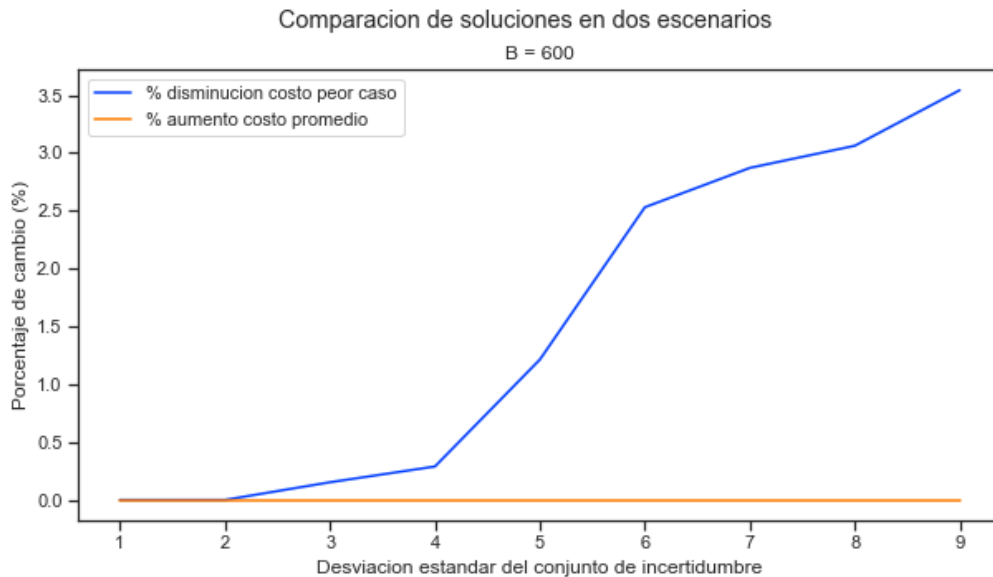


Figura 4.6: Diferencia relativa del costo de la solución robusta en el caso promedio y en el peor caso, como función del tamaño (sd) del conjunto de incertidumbre. **Modelo 2.** Grafo 1.1, incertidumbre en ambos pares OD y $\Gamma = 1$.

En estos últimos gráficos (Figuras 4.6 y 4.7) se considera que toda la demanda de la red es incierta y, además, se analiza agrandando el tamaño del conjunto de incertidumbre. Los principales resultados muestran que la solución robusta sigue protegiéndose de las demandas grandes sin perder demasiada optimalidad (de hecho cuando $\Gamma = 1$ la solución robusta nunca pierde optimalidad en el caso promedio). Sin embargo, al aumentar el presupuesto de

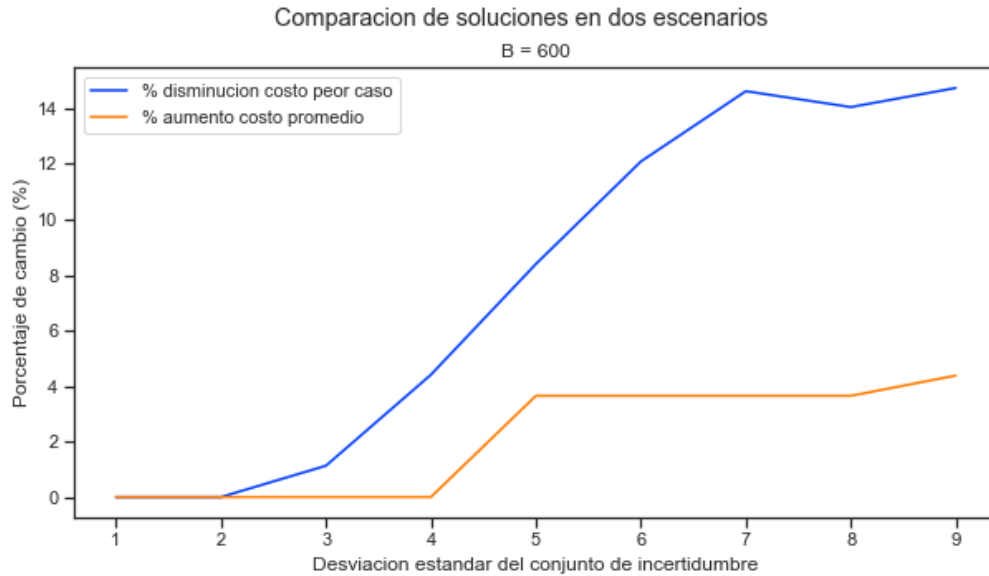


Figura 4.7: Diferencia relativa del costo de la solución robusta en el caso promedio y en el peor caso, como función del tamaño (sd) del conjunto de incertidumbre . **Modelo 2.** Grafo 1.1, incertidumbre en ambos pares OD y $\Gamma = 2$.

incertidumbre a $\Gamma = 2$ se comienza a notar un deterioro en el costo del caso promedio en la solución robusta. En este sentido, se observa que la diferencia entre el aumento relativo del costo en el caso promedio y la disminución relativa del costo en el peor caso se acorta. Pero aún así sigue siendo un resultado positivo, puesto que la pérdida de optimalidad del costo nominal es de un 4% aproximadamente mientras que la ganancia en el peor costo es mayor a 14%.

4.3. Resultados en red Sioux Falls

Los siguientes experimentos se realizan en la red *Sioux Falls* que se presenta en la Figura 4.8, que es un grafo más realista y bastante utilizado en la literatura. Esta red tiene 24 nodos y 76 arcos. Para estos análisis el tamaño del problema es de 100 pares origen-destinos, seleccionados aleatoriamente del conjunto de pares origen-destino que tiene la instancia *Sioux Falls*.

Los datos de este grafo como la demanda de cada par origen-destino, la capacidad y el tiempo de viaje por arco se obtuvieron de LeBlanc et al. (1975). La capacidad fue modificada

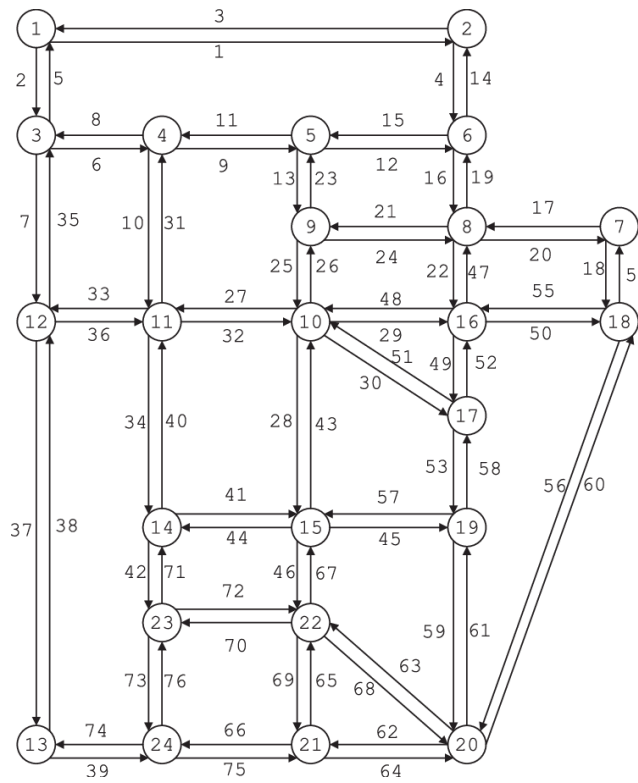


Figura 4.8: Red Sioux Falls (Long et al., 2015)

con el fin de reducir el espacio disponible para los flujos y permitir estresar el modelo para que optimice los costos de inversión, así, se redujo en un 90 % la capacidad de cada arco.

Los parámetros que no se encuentran en la data disponible de *Sioux Falls* son construidos. Primeramente, los costos de inversión por arco se obtienen aleatoriamente del intervalo $[1, 10]$. Se discretizan los flujos por arco (s), para linealizar la función de costos, dentro del rango $[0, \sum_{k \in K} d_k]$ con un intervalo de tamaño 200, de este modo, se establece que el flujo máximo que puede pasar por un arco es toda la demanda disponible. Por último, los valores discretos del aumento de capacidad de los arcos (p) están en el rango $[0, \frac{1}{2} \sum_{k \in K} d_k]$ fijando el tamaño del intervalo en 50, de esta manera sólo se permite aumentar la capacidad de un arco hasta la mitad de la demanda disponible y así restringir los recursos disponibles para la red.

Para el modelo robusto se evalúa un conjunto de incertidumbre para la demanda de la forma (4.1). Este conjunto tiene valores tal que $\underline{d}_k = \hat{d}_k$ y $\rho = sd$ (con sd la desviación estándar), lo que implica que la demanda nominal forma parte de los puntos extremos del conjunto.

Junto con ello, se considera un conjunto KC de pares origen-destino críticos a los cuales se aplica incertidumbre. Esto quiere decir que $\forall k \in KC, d_k \in \mathcal{U}$, con \mathcal{U} un conjunto de incertidumbre y $\forall k \in K \setminus KC, d_k = \hat{d}_k$. Este conjunto KC , para estos experimentos, se compone de 5 pares OD, los cuales son $KC = \{(1, 9), (13, 8), (6, 21), (12, 18), (4, 17)\}$ y se escogieron dado que son nodos lejanos entre sí y que no tienen camino directo dentro de la red (véase Figura 4.8).

A continuación se mencionan los experimentos que se realizan con el grafo de *Sioux Falls*:

Experimento 3: Se considera el grafo de la red *Sioux Falls* junto al conjunto de incertidumbre que considera la demanda nominal como punto extremo del conjunto, ya que las demandas inciertas se mueven dentro del intervalo $[\hat{d}_k, \hat{d}_k + sd]$. Los valores de los parámetros son los descritos anteriormente.

Experimento 4: Se considera el mismo tipo de problema anterior, pero, se cambia el tamaño del intervalo de la discretización del aumento de capacidad a 25, es decir, se decide achicar el tamaño de la capacidad que se aumenta en los arcos (recordando que es necesario para linealizar los costos no lineales involucrados en el problema).

Como resultado de estos experimentos se obtienen los costos de la comparación de los modelos determinista y robusto. Además, para el Experimento 4 se muestran las soluciones gráficas del aumento de capacidad.

Finalmente, en los dos experimentos se utiliza un presupuesto $B = 700000$, que es aproximadamente el presupuesto al que se mantiene con un costo de viaje mínimo (véase Anexo B .2). Además, se usa una desviación estándar (sd) igual al 75% de la demanda más pequeña de los arcos críticos (KC).

4.3.1. Análisis variando la discretización de valores enteros

Los resultados de los Experimentos 3 y 4 se contrastan en la Tabla 4.3. En estos experimentos se ha utilizado sólo el Modelo 1 que posee restricciones de capacidad.

| Tamaño del intervalo de capacidad | Tipo de solución | Costo de inversión | Tiempo de ejecución (s) | GAP | Caso Promedio | | Peor Caso | |
|-----------------------------------|------------------|--------------------|-------------------------|---------|----------------|--------------------------------|----------------|------------------------------------|
| | | | | | Costo de viaje | % aumento con solución robusta | Costo de viaje | % disminución con solución robusta |
| Intervalo = 50 | Determinista | 671.000 | 61,91 | 0,000 % | 1,86602e+09 | 0,0005 % | 1,8747e+09 | -0,0216 % |
| | Robusta | 699.050 | 15025,5 | 0,001 % | 1,86603e+09 | | 1,8742e+09 | |
| Intervalo = 25 | Determinista | 698.925 | 60,99 | 0,000 % | 1,86602e+09 | 0,0003 % | Infactible | - |
| | Robusta | 698.400 | 3897,4 | 0,001 % | 1,86603e+09 | | 1,8743e+09 | |

Tabla 4.3: Costos y cambio relativo al usar la solución robusta en cada caso. Se considera dos discretizaciones diferentes del aumento de capacidad. **Modelo 1.** Grafo Sioux Falls.

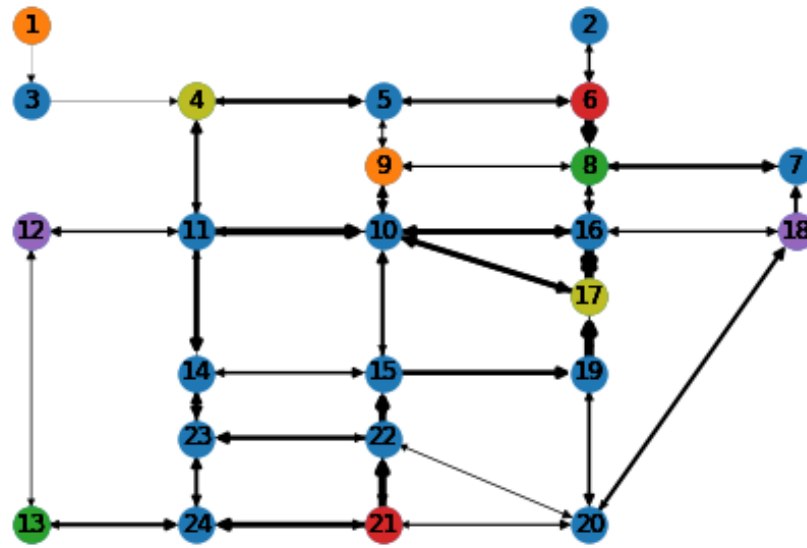
Un primer aspecto importante a mencionar es que efectivamente al crecer el tamaño de la red éste impacta en el tiempo de resolución. Los tiempos de ejecución de ambos modelos difieren bastante; si se considera el intervalo de discretización de 50, cerca de 6 horas se demora en resolver el modelo robusto comparado al minuto que lo hace el modelo determinista.

Asimismo, se puede observar que cuando se utiliza una discretización más gruesa, el modelo robusto no tiene un impacto de mejora notable para el peor caso y los costos del caso promedio de ambos modelos se mantienen igual. Esto motiva a evaluar algunos cambios en parte de la configuración de los parámetros con el fin de encontrar una solución robusta con más influencia en el peor escenario. Se puede deducir que al considerar más opciones de valores discretos para incrementar la capacidad, es posible llegar a soluciones mejores.

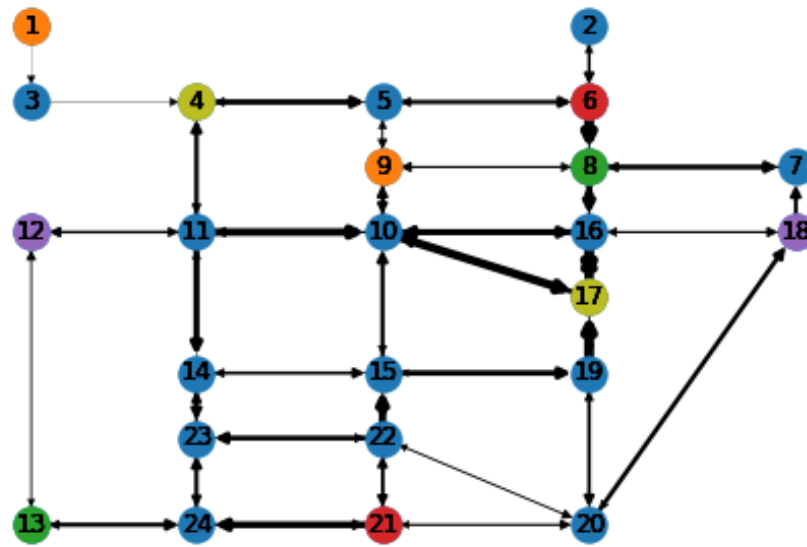
Dicho esto, al experimentar con un intervalo de tamaño 25, se puede apreciar en la tabla de resultados, que el modelo robusto logra obtener una solución para el peor caso mientras que la solución determinista resulta ser infactible. Se observa, además, que para el modelo determinista esta disminución del tamaño del intervalo no resulta ser favorable (aumenta el costo de inversión pero sin mejorar los tiempos de viaje que se mantienen igual). Sin embargo, para el modelo robusto se ve un impacto de esta variación: disminuye su tiempo de ejecución cerca de 5 horas, y permite obtener soluciones que se protegen contra el peor caso y que no empeoran el caso promedio, mientras que invierte un poco menos en capacidad.

En definitiva, se obtienen resultados que son notablemente distintos con sólo cambiar un tipo de configuración de parámetros. Esto nos dice de los efectos que se pueden producir en estos modelos ante la variación de los datos y de que éstos son capaces de encontrar soluciones

para diferentes discretizaciones de valores.



(a) Solución determinista



(b) Solución Robusta

Figura 4.9: Soluciones z . Experimento 5. Grafo Sioux Falls. **Modelo 1.** El grosor de los arcos indica el aumento de capacidad. Los colores en los nodos indican los orígenes y destino de los pares OD críticos.

Ahora bien, es posible observar gráficamente las soluciones z robusta y determinista del Experimento 4 en la Figura 4.9. En éstas se aprecia que las soluciones son parecidas salvo en dos arcos. Se identifica que la solución robusta incrementó más la capacidad en los arcos $(10, 17)$ y $(8, 16)$, que son arcos que están en el centro de la red y que poseen un nodo de origen o destino. Además, el par $(10, 17)$ es el par OD que tiene la mayor demanda, que en

este experimento es determinista. El aumento en estos arcos permiten obtener una solución factible para el peor caso.

Por último, es importante mencionar que los análisis realizados en la red *Sioux Falls* fueron realizados sólo con el Modelo 1, debido a que el Modelo 2 incurre en costos computacionales significativamente mayores. De lo observado en los resultados de la sección anterior, es posible deducir que con el Modelo 2 se pueda llegar a soluciones mejores, sin embargo, tiene la deficiencia del tiempo de resolución. En estos escenarios con problemas más grandes, el Modelo 1 tiene una ventaja por sobre el Modelo 2.

Capítulo 5

Conclusiones

5.1. Conclusiones generales

En este trabajo de tesis, se considera el problema de diseño de redes modelado a través de una programación binivel, dado que se asume un comportamiento egoísta en los usuarios de la red que buscan escoger su propio camino más corto entre sus orígenes y destinos, mientras que un planificador central decide reducir los costos de la red completa. Este comportamiento de los transeúntes es formalizado a través del equilibrio de Wardrop, estableciendo ecuaciones de equilibrio del tráfico en el nivel inferior del problema. Además de esto, se considera que la demanda del problema es incierta y se incorpora al modelo utilizando una metodología robusta.

En primer lugar, se desarrollan dos tipos de modelos de optimización deterministas de diseño de redes que satisfacen condiciones de equilibrio y minimizan los costos de viaje, limitándose a una restricción presupuestaria para la inversión de capacidad. Uno de los modelos utiliza una función de latencia -la función que determina el costo de viaje por arco que depende del flujo en él y que permite incorporar congestión- con una forma lineal y posee restricciones de capacidad, en cambio, el otro modelo posee una función polinomial, en particular, la función BPR, y se eliminan las restricciones de capacidad. Se desarrolla una metodología que permite transformar el problema binivel no lineal de estos dos modelos en

un problema de un sólo nivel con variables enteras y continuas.

Para modelar la incertidumbre en la demanda se usa la metodología robusta para encontrar una solución en la cual se reduce el costo para el peor caso, y por lo tanto, se comporta eficientemente para todos los escenarios inciertos de demanda. Para ello, se considera que estos escenarios pertenecen a un conjunto de incertidumbre. Debido a la naturaleza del problema de diseño de redes, es natural considerar un problema robusto con dos tipos de decisiones de acuerdo al momento en que se toman: las variables de expansión de capacidad son decididas antes de la incertidumbre mientras que las variables de flujo se ajustan al escenario resultante.

Con el fin de reformular el problema robusto para que pueda ser computacionalmente tratable, se demuestra que si el conjunto de incertidumbre es un poliedro y, además, la función de costos es convexa, entonces existen cotas que permiten determinar el peor caso de demanda. Específicamente, el peor escenario de demanda se encuentra bien determinado por el peor caso de las demandas de los puntos extremos del conjunto de incertidumbre, puesto que este peor caso es una cota inferior, y al multiplicar este costo por una constante, en particular, el precio de la anarquía, se logra obtener una cota superior.

De esta manera, se formula un modelo considerando las demandas extremas, con un conjunto de incertidumbre restringido de cardinalidad, el cual es eficiente y encuentra una solución robusta aproximada para el problema robusto original.

Este modelo fue posible implementarlo y resolverlo de forma exacta en instancias pequeñas y en instancias más realistas. En instancias pequeñas, el modelo funciona eficientemente con bajo tiempo de ejecución, sin embargo, si el tamaño del problema crece, la resolución en forma exacta incurre en altos costos computacionales.

Los resultados computacionales en redes pequeñas muestran que la solución robusta aproximada tiene una modesta sub-optimalidad en escenarios deterministas cerca del promedio mientras que reduce significativamente el costo del peor caso. En particular, a medida que la incertidumbre crece, la diferencia entre el costo determinista y robusto en el

peor caso se vuelve mayor.

Cuando se analiza el presupuesto, los resultados evidencian que cuando se cuenta con un presupuesto menor, el modelo robusto no es mejor que el modelo determinista, puesto que no tiene inversión suficiente para aumentar más la capacidad y permitir una disminución en el costo del peor caso. Sin embargo, con presupuesto mayor, el modelo robusto es capaz de construir capacidad en ciertos arcos para disminuir la congestión en el peor escenario y, por ende, los tiempos de viaje. Además, cuando se cuenta con una menor inversión, el modelo robusto prioriza protegerse del peor caso por sobre reducir los costos de los casos promedio, por lo que pierde eficiencia en esos escenarios.

Con respecto a los análisis realizados en el grafo de *Sioux Falls*, se puede apreciar que los modelos tienen la capacidad de resolver el problema para diferentes discretizaciones de los valores que linealizan las funciones no lineales. Sin embargo, estas soluciones difieren significativamente. Esto supone que estos modelos enfrentan desafíos algorítmicos al resolver estas instancias con heurísticas, puesto que un cambio en los valores de los parámetros produce efectos distintos en las redes diseñadas.

Por cierto, en los alcances de esta tesis no fue posible determinar la existencia de algún tipo de conjunto de incertidumbre en donde se cumpla que efectivamente el peor escenario se encuentra en las demandas extremas, como tampoco se pudo hallar un contraejemplo que muestre que el peor caso se encuentra fuera de los puntos extremos del conjunto de incertidumbre, lo cual queda pendiente para un análisis posterior.

Finalmente, al comparar las dos formulaciones en problemas con pocos nodos, el modelo con restricción de capacidad genera instancias infactibles con presupuestos menores, en cambio el modelo con función BPR y sin restricción de capacidad, llega a soluciones robustas con bajos presupuestos. Ahora bien, la formulación con restricción de capacidad se comporta mejor con problemas más grandes, puesto que su tiempo de resolución es significativamente menor que el modelo con la función BPR.

5.2. Trabajo futuro

Un desarrollo inminente de este trabajo de tesis es poder resolver los modelos planteados de manera eficiente. Es necesario generar un algoritmo de resolución para el problema robusto, sobre todo para abarcar mejor las instancias de gran tamaño, y poder obtener más y mejores resultados en tiempos razonables. Uno de los algoritmos ideales para este tipo de problema, que busca ser factible para un conjunto de escenarios de demanda, son los algoritmos de descomposición. Por otro lado, es posible profundizar más en la formulación del modelo para lograr definir cómo encontrar el peor caso de demanda y reformular el problema robusto para obtener soluciones óptimas que sean factibles para cualquier escenario de demanda.

Para ampliar y precisar el desempeño de la demanda incierta, un buen enfoque sería incorporar otros tipos de conjuntos de incertidumbre y así permitir una mayor variabilidad en los diferentes casos de demanda. Además, sería bueno explorar el enfoque estocástico de incertidumbre en la demanda, estimando la distribución de la demanda con datos históricos de un sector poblacional para aplicar el modelo propuesto.

Por último, algunas extensiones interesantes que se pueden considerar para enriquecer el modelo son examinar otros supuestos de comportamiento de los transeúntes. Por ejemplo, se puede considerar que los usuarios tienen información imperfecta al momento de escoger el mejor camino, y por ende, éste es escogido de acuerdo a la percepción de cada uno. También, se puede eliminar el supuesto de racionalidad de los viajeros e incorporar algunos modelos de elección jerárquicos. Asimismo, se puede utilizar un enfoque estocástico para la decisión de los usuarios de la red, en donde se establece que cada persona, en lugar de escoger un sólo camino entre su origen y destino, se les asignan probabilidades de elección a diferentes trayectos que une ambos puntos.

Todos estos trabajos permitirán acercar cada uno de los problemas matemáticos de programación a un enfoque más realista, logrando modelar de mejor manera el comportamiento del transporte, de las personas y de las ciudades en general.

Bibliografía

- Agra, A., Christiansen, M., Figueiredo, R., Hvattum, L., Poss, M. and Requejo, C. (2013), ‘The robust vehicle routing problem with time windows’, *Computers and Operations Research* **40**(3), 856–866.
- Ahmed, S., King, A. and Parija, G. (2003), ‘A multi-stage stochastic integer programming approach for capacity expansion under uncertainty’, *Journal of Global Optimization* **26**, 3–24.
- Ahuja, R., Magnanti, T. and Orlin, J. (1993), *Network flows: Theory, algorithms, and applications*, pp. 649–694.
- Atamtürk, A. and Zhang, M. (2007), ‘Two-stage robust network flow and design under demand uncertainty’, *Operations Research* **55**(4), 662–673.
- Bard, J. and Moore, J. (1990), ‘A branch and bound algorithm for the bilevel programming problem’, *Siam Journal on Scientific and Statistical Computing* **11**(2), 281–292.
- Beckmann, M. J., McGuire, C. B. and Winsten, C. B. (1955), *Studies in the Economics of Transportation*, RAND Corporation, Santa Monica, CA.
- Ben-Ayed, O. and Blair, C. E. (1990), ‘Computational difficulties of bilevel linear programming’, *Operations Research* **38**(3), 556–560.
- Ben-Tal, A., Chung, B., Mandala, S. and Yao, T. (2011), ‘Robust optimization for emergency logistics planning: Risk mitigation in humanitarian relief supply chains’, *Transportation Research, Series B: Methodological* **45**(8), 1177–1189.
- Ben-Tal, A., El-Ghaoui, L. and Nemirovski, A. (2009), *Robust Optimization*, Princeton Series in Applied Mathematics, Princeton University Press.
- Ben-Tal, A., Golany, B., Nemirovski, A. and Vial, J.-P. (2005), ‘Retailer-supplier flexible commitments contracts: A robust optimization approach’, *Manufacturing & Service Operations Management* **7**(3), 248–271.
- Ben-Tal, A. and Nemirovski, A. (1998), ‘Robust convex optimization’, *Mathematics of Operations Research* **23**(4), 769–805.
- Ben-Tal, A. and Nemirovski, A. (1999), ‘Robust solutions of uncertain linear programs’,

- Operations Research Letters* **25**(1), 1–13.
- Bertsimas, D., Brown, D. and Caramanis, C. (2010), ‘Theory and applications of robust optimization’, *SIAM Review* **53**.
- Bertsimas, D., Pachamanova, D. and Sim, M. (2004), ‘Robust linear optimization under general norms’, *Operations Research Letters* **32**, 510–516.
- Bertsimas, D. and Sim, M. (2003), ‘Robust discrete optimization and network flows’, *Mathematical Programming* **98**, 49–71.
- Bertsimas, D. and Sim, M. (2004), ‘The price of robustness’, *Operations Research* **52**, 35–53.
- Bertsimas, D. and Thiele, A. (2003), ‘A robust optimization approach to supply chain management’, *Integer Programming and Combinatorial Optimization* pp. 86–100.
- Bienstock, D. and Özbay, N. (2008), ‘Computing robust basestock levels’, *Discrete Optimization* **5**(2), 389–414.
- Birge, J. R. and Louveaux, F. (1997), *Introduction to Stochastic Programming*, Springer-Verlag, New York, NY, USA.
- Bracken, J. and McGill, J. T. (1973), ‘Mathematical programs with optimization problems in the constraints’, *Operations Research* **21**(1), 37–44.
- Braess, D., Nagurney, A. and Wakolbinger, T. (2005), ‘On a paradox of traffic planning’, *Transportation Science* **39**, 446–450.
- Branston, D. (1976), ‘Link capacity functions: A review’, *Transportation Research* **10**(4), 223–236.
- Bureau of Public Roads, U. S. (1964), *Traffic Assignment Manual for Application with a Large, High Speed Computer*, number 2 in ‘Traffic Assignment Manual for Application with a Large, High Speed Computer’, U.S. Department of Commerce, Bureau of Public Roads, Office of Planning, Urban Planning Division.
- Candler, W. and Norton, R. D. (1977), *Multi-level programming*, Vol. 20, World Bank Development Research Center.
- Correa, J. R. and Stier-Moses, N. E. (2011), *Wardrop equilibria*, Wiley Encyclopedia of Operations Research and Management Science.
- Daganzo, C. F. and Sheffi, Y. (1977), ‘On stochastic models of traffic assignment’, *Transportation Science* **11**(3), 253–274.
- Davis, G. A. (1994), ‘Exact local solution of the continuous network design problem via stochastic user equilibrium assignment’, *Transportation Research, Series B: Methodological* **28**(1), 61–75.
- Dial, R. B. (1971), ‘A probabilistic multipath traffic assignment model which obviates path

- enumeration’, *Transportation Research* **5**(2), 83–111.
- Dubey, P. (1986), ‘Inefficiency of nash equilibria’, *Mathematics of Operations Research* **11**(1), 1–8.
- Dumrauf, D. and Gairing, M. (2006), Price of anarchy for polynomial wardrop games, in ‘Proceedings of the Second International Conference on Internet and Network Economics’, Springer-Verlag, Berlin, Heidelberg, p. 319–330.
- El-Ghaoui, L. and Lebret, H. (1997), ‘Robust solutions to least-squares problems with uncertain data’, *SIAM Journal on Matrix Analysis and Applications* **18**, 1035–1064.
- El-Ghaoui, L., Oks, M. and Oustry, F. (2003), ‘Worst-case value-at-risk and robust portfolio optimization: A conic programming approach’, *Operations Research* **51**, 543–556.
- Even, S., Itai, A. and Shamir, A. (1976), ‘On the complexity of timetable and multicommodity flow problems’, *SIAM J. Comput.* **5**, 691–703.
- Farvaresh, H. and Sepehri, M. M. (2011), ‘A single-level mixed integer linear formulation for a bi-level discrete network design problem’, *Transportation Research Part E: Logistics and Transportation Review* **47**(5), 623–640.
- Ferguson, A. R. and Dantzig, G. B. (1956), ‘The allocation of aircraft to routes—an example of linear programming under uncertain demand’, *Management Science* **3**(1), 45–73.
- Ferris, M. C. and Ruszczyński, A. (2000), *Robust path choice in networks with failures*, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II, Institut für Mathematik.
- Fontaine, P. and Minner, S. (2014), ‘Benders decomposition for discrete–continuous linear bilevel problems with application to traffic network design’, *Transportation Research Part B: Methodological* **70**, 163–172.
- Gairing, M., Harks, T. and Klimm, M. (2017), ‘Complexity and approximation of the continuous network design problem’, *SIAM Journal on Optimization* **27**, 1554–1582.
- García, J. and Peña, A. (2018), ‘Robust optimization: Concepts and applications’, *Nature-Inspired Methods for Stochastic, Robust and Dynamic Optimization* p. 7.
- Gerodimos, A. E. (1998), ‘Robust discrete optimization and its applications’, *Journal of the Operational Research Society* **49**, 1303–1304.
- Goldfarb, D. and Iyengar, G. (2003), ‘Robust portfolio selection problems’, *Mathematics of Operations Research* **28**, 1–38.
- Gorissen, B. L., İhsan Yanıkoğlu and den Hertog, D. (2015), ‘A practical guide to robust optimization’, *Omega* **53**, 124–137.
- Gutiérrez, G. J., Kouvelis, P. and Kurawarwala, A. A. (1996), ‘A robustness approach

- to uncapacitated network design problems’, *European Journal of Operational Research* **94**(2), 362–376.
- Hansen, P., Jaumard, B. and Savard, G. (1992), ‘New branch-and-bound rules for bilevel linear programming’, *SIAM Journal on Scientific and Statistical Computing* **13**(5), 273–292.
- Haurie, A. and Marcotte, P. (1985), ‘On the relationship between nash—cournot and wardrop equilibria’, *Networks* **15**(3), 295–308.
- Karp, R. (1975), ‘On the computational complexity of combinatorial problems’, *Networks* **5**, 45–68.
- Kleinert, T., Labbé, M., Ljubić, I. and Schmidt, M. (2021), ‘A survey on mixed-integer programming techniques in bilevel optimization’, *EURO Journal on Computational Optimization* **9**, 100007.
- Knight, F. H. (1924), ‘Some Fallacies in the Interpretation of Social Cost’, *The Quarterly Journal of Economics* **38**(4), 582–606.
- Koutsoupias, E. and Papadimitriou, C. (1999), Worst-case equilibria, in G. Meinel and S. Tison, eds, ‘Proceedings of the 16th Annual ACM Symposium on Theoretical Aspects of Computer Science’, Lecture Notes in Computer Science, Springer Verlag, pp. 404–413.
- LeBlanc, L. J., Morlok, E. K. and Pierskalla, W. P. (1975), ‘An efficient approach to solving the road network equilibrium traffic assignment problem’, *Transportation Research* **9**(5), 309–318.
- Liu, H. and Wang, D. Z. (2015), ‘Global optimization method for network design problem with stochastic user equilibrium’, *Transportation Research Part B: Methodological* **72**(C), 20–39.
- Liu, H. and Wang, D. Z. (2016), ‘Modeling and solving discrete network design problem with stochastic user equilibrium: Dndp with sue’, *Journal of Advanced Transportation* **50**, 1295–1313.
- Lo, H. K. and Tung, Y.-K. (2003), ‘Network with degradable links: capacity analysis and design’, *Transportation Research Part B: Methodological* **37**(4), 345–363.
- Long, J., Szeto, W., Shi, Q., Gao, Z. and Huang, H.-J. (2015), ‘A nonlinear equation system approach to the dynamic stochastic user equilibrium simultaneous route and departure time choice problem’, *Transportmetrica A: Transport Science* **11**, 388–419.
- Luathep, P., Sumalee, A., Lam, W. H., Li, Z.-C. and Lo, H. K. (2011), ‘Global optimization method for mixed transportation network design problem: A mixed-integer linear programming approach’, *Transportation Research Part B: Methodological* **45**(5), 808–827.
- Maerivoet, S. and De Moor, B. (2005), ‘Transportation planning and traffic flow models’.
- McCord, M. R. (1987), ‘Urban transportation networks: Equilibrium analysis with

- mathematical programming methods: Yosef sheffi. prentice-hall, inc., englewood cliffs, nj, u.s.a. 1985. 399 pp. + xvi. \$40.95', *Transportation Research Part A: General* **21**, 481–484.
- McCormick, G. P. (1972), 'Converting general nonlinear programming problems to separable nonlinear programming problems, Technical report, George Washington University, Washington DC, Program in Logistics.
- Minoux, M. (1989), 'Networks synthesis and optimum network design problems: Models, solution methods and applications', *Networks* **19**, 313–360.
- Morettini, M. (2013), 'Aproximaciones de distribuciones de probabilidad: Enfoque empírico'. Último acceso 30 Diciembre 2021.
URL: <http://nulan.mdp.edu.ar/>
- Mudchanatongsuk, S., Ordonez, F. and Liu, J. (2008), 'Robust solutions for network design under transportation cost and demand uncertainty', *Journal of the Operational Research Society* **59**, 652–662.
- Mulvey, J., Vanderbei, R. and Zenios, S. (1995), 'Robust optimization of large-scale systems', *Operations Research* **43**, 264–281.
- Ordóñez, F. and Stier-Moses, N. E. (2010), 'Wardrop equilibria with risk-averse users', *Transportation Science* **44**(1), 63–86.
- Ordóñez, F. and Zhao, J. (2007), 'Robust capacity expansion of network flows', *Networks* **50**(2), 136–145.
- Papadimitriou, C. (2001), 'Algorithms, games, and the internet', *Conference Proceedings of the Annual ACM Symposium on Theory of Computing* **2076**, 749–753.
- Paraskevopoulos, D., Karakitsos, E. and Rustem, B. (1991), 'Robust capacity planning under uncertainty', *Management Science* **37**(7), 787–800.
- Roughgarden, T. (2003), 'The price of anarchy is independent of the network topology', *Journal of Computer and System Sciences* **67**(2), 341–364. Special Issue on STOC 2002.
- Roughgarden, T. (2005), *Selfish Routing and the Price of Anarchy*, The MIT Press.
- Roughgarden, T. and Tardos, E. (2002), 'How bad is selfish routing?', **49**(2), 236–259.
- Saharidis, G. and Ierapetritou, M. (2009), 'Resolution method for mixed integer bi-level linear problems based on decomposition technique', *Journal of Global Optimization* **44**, 29–51.
- Smith, M. (1979), 'The existence, uniqueness and stability of traffic equilibria', *Transportation Research Part B: Methodological* **13**(4), 295–304.
- Spieß, H. (1990), 'Conical volume-delay functions', *Transportation Science* **24**.
- Tardos, E. (1986), 'A strongly polynomial algorithm to solve combinatorial linear programs', *Operations Research* **34**(2), 250–256.

- Von Stackelberg, H. (1934), 'Marktform und gleichgewicht. julius springer', *Julius Springer* .
- Von Stackelberg, H. (1952), 'The theory of market economy', *Oxford University Press* .
- Wang, D. Z. and Lo, H. (2010), 'Global optimum of the linearized network design problem with equilibrium flows', *Transportation Research Part B: Methodological* **44**, 482–492.
- Wang, I.-L. (2018), 'Multicommodity network flows: A survey, part i: Applications and formulations', *International Journal of Operations Research* **15**(4), 145–153.
- Wang, S., Meng, Q. and Yang, H. (2013), 'Global optimization methods for the discrete network design problem', *Transportation Research Part B: Methodological* **50**(C), 42–60.
- Wardrop, J. G. (1952), 'Some theoretical aspects of road traffic research.', *Proceedings of the Institution of Civil Engineers* **1**(3), 325–362.
- Xu, H., Caramanis, C. and Mannor, S. (2008), 'Robustness and regularization of support vector machines', *Journal of Machine Learning Research* **10**.

Anexos

A . Aproximación de la distribución normal para simular demandas discretas

Una distribución normal con media μ y varianza σ^2 puede ser aproximada por una distribución Bernoulli de parámetros n y p , cuando $n \geq 30$ y $0,1 \leq p \leq 0,9$, de la siguiente manera (Morettini (2013)).

$$B(n, p) \sim N(np, \sqrt{np(1-p)})$$

Con esto,

$$p = \left(1 - \frac{\sigma^2}{\mu}\right) \quad n = \frac{\mu}{p}$$

Dado que se busca realizar una simulación para obtener demandas discretas del par $OD = (2, 3)$, entonces se considera para la distribución normal que $\mu = \hat{d}_{(2,3)} = 15$. Por consiguiente, y respetando los valores para que la distribución normal pueda ser aproximada por una Bernoulli, se calcula que:

$$0,1 \leq p \leq 0,9 \tag{5.1a}$$

$$\Rightarrow 0,1 \leq \frac{\sigma^2}{\mu} \leq 0,9 \tag{5.1b}$$

$$\Rightarrow 1,5 \leq \sigma^2 \leq 13,5 \tag{5.1c}$$

$$\Rightarrow 1,22 \leq \sigma \leq 3,68 \tag{5.1d}$$

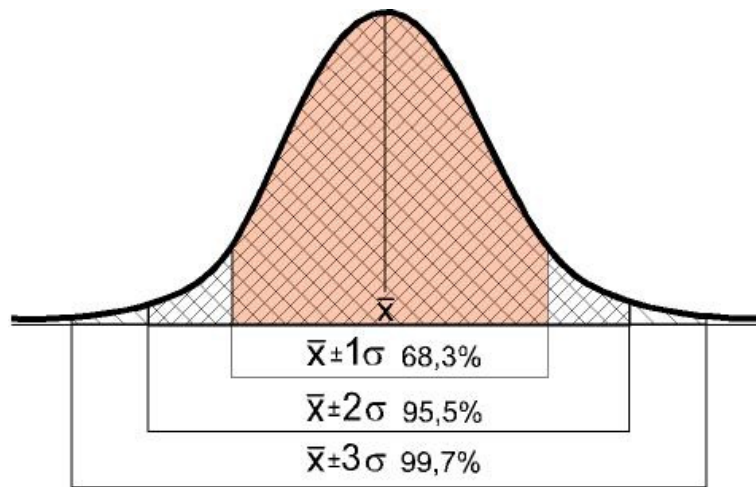


Figura 5.1: Curva de distribución normal. Se muestran la probabilidades para distintos intervalos.

Tomada de: <https://jesusgarciaj.com/2010/01/22/la-curva-de-distribucion-normal/>.

Ahora, en la Figura 5.1 se puede observar que, dada la distribución normal, con un 95 % de confianza aproximadamente los valores se encuentran en el intervalo $[\mu - 2\sigma, \mu + 2\sigma]$. Por lo tanto, si se desea que con un 95 % de probabilidad las demandas de las simulaciones estén dentro del conjunto de incertidumbre, entonces $sd = 2\sigma$, con sd la desviación estándar del conjunto de incertidumbre.

Entonces, para cumplir con lo obtenido para σ en la ecuación (5.1d), sd debe cumplir que:

$$2,44 \leq sd \leq 7,35$$

Como se considera $sd = 7$, que es el 50 % de la demanda nominal, y dado que se encuentra dentro de los valores calculados, entonces se pueden obtener demandas con un 95 % de confianza que están dentro del conjunto de incertidumbre.

B . Costos de viaje óptimos para distintos presupuestos

B .1. Red pequeña

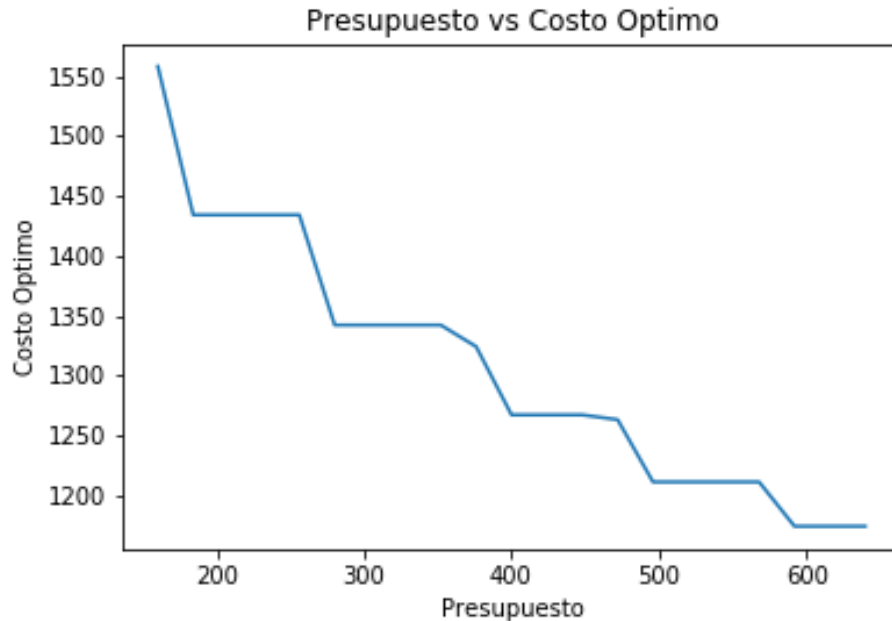


Figura 5.2: Costo de viaje óptimo en el escenario determinista en función del presupuesto B . Modelo 1. Grafo 1.1.

B .2. Red Sioux Falls

Si bien en este gráfico se observa que el costo de viaje mínimo se genera en un presupuesto $B = 667000$, se termina escogiendo, para realizar los análisis, un presupuesto $B = 700000$, porque era el presupuesto en que el modelo obtenía una solución en un tiempo razonable. Para los valores entre 667000 y 700000, el tiempo de ejecución del modelo era significativamente mayor. Además, el costo de viaje se mantenía en el mismo valor desde $B = 670000$.

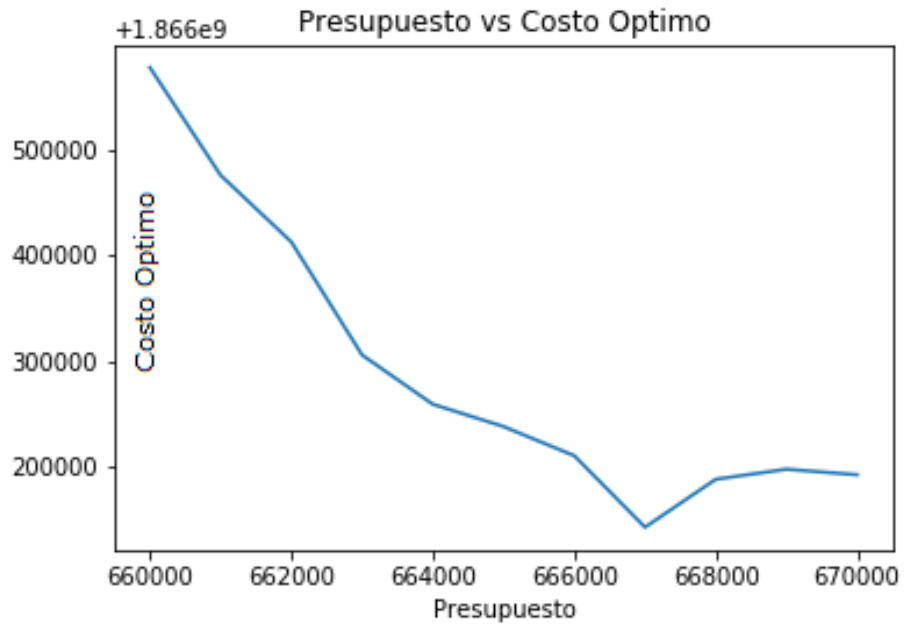


Figura 5.3: Costo de viaje óptimo en el escenario determinista en función del presupuesto B . **Modelo 1.** Grafo *Sioux Falls* (4.8).