PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
ESCUELA DE INGENIERIA

RESTRICCIÓN DE CAPACIDAD EN UN MODELO DE ASIGNACIÓN ESTOCÁSTICA EN TRANSPORTE PÚBLICO

BELÉN ANDREA BARRIGA ROSALES

Tesis para optar al grado de
Magíster en Ciencias de la Ingeniería

Profesores Supervisores:
JUAN CARLOS MUÑOZ ABOGABIR
SEBASTIÁN RAVEAU FELIÚ

Santiago de Chile, Enero, 2021

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
ESCUELA DE INGENIERIA

RESTRICCIÓN DE CAPACIDAD EN UN MODELO DE ASIGNACIÓN ESTOCÁSTICA EN TRANSPORTE PÚBLICO

BELÉN ANDREA BARRIGA ROSALES

Tesis presentada a la Comisión integrada por los profesores:

JUAN CARLOS MUÑOZ ABOGABIR

SEBASTIÁN RAVEAU FELIÚ

ALEXANDRA SOTO OGUETA

SEBASTIÁN TAMBLAY MOËNNE

JOSÉ MIGUEL DEL VALLE LLADSER

Para completar las exigencias del grado de
Magíster en Ciencias de la Ingeniería

Santiago de Chile, Enero, 2021

(A mis Padres, hermanos, profesores y amigos, que me acompañaron en este largo camino.)

INDICE GENERAL

Pág.

INDICE DE TABLAS

Pág.

INDICE DE FIGURAS

Pág.

RESUMEN

En redes de transporte público altamente congestionadas, la capacidad de los vehículos juega un rol determinante en las elecciones, ya que no todos los usuarios podrán subir al primer bus o tren, generando aumento en los tiempos de espera percibidos y ampliando el conjunto de líneas atractivas que consideran los viajeros.

En esta tesis se propone un modelo de asignación que considere la restricción de capacidad de los vehículos y que respete el S.U.E. en las predicciones de flujos. Para esto se consideró un modelo de elección de rutas de transporte publico logit, en el que la generación de rutas es exógena al problema y la capacidad de los vehículos definida es estricta, es decir la densidad pasajeros/m^2 que se modela para el problema es la máxima y no puede sobrepasarse.

La metodología de trabajo consistió en definir un modelo de asignación de dos etapas que incluye los efectos de la congestión en la definición del conjunto de líneas atractivas y en los aumentos de tiempos de espera. Luego, se diseñaron redes ficticias reducidas con demandas y comportamiento esperado de los usuarios, con el objetivo de evaluar el modelo en su capacidad predictiva y desempeño.

Posteriormente del análisis y mejoras del modelo, se realizaron pruebas en una red real, altamente congestionada, como la del Metro de Santiago (Fuente SECTRA año 2013) . En esta etapa, se comparó su desempeño con otros dos modelos, en que la restricción de capacidad se modela como una función de la ocupación, Raveau et al. (2011) y De Cea y Fernández (1993).

Los resultados obtenidos, arrojaron que para cierta oferta y demanda existen soluciones de equilibrio que respetan la capacidad de los vehículos y que se obtienen al asignar con este modelo y no con los otros modelos analizados.

Palabras claves: Modelo de Asignación, Transporte Público, Restricción de Capacidad, Redes de Transporte

ABSTRACT

In highly congested public transport networks, vehicle capacity plays a determining role in the choices, since not all users will be able to board the first bus or train, generating an increase in perceived waiting times and widening the set of attractive lines considered by travelers.

This thesis proposes an assignment model that considers the vehicle capacity constraint and respects the S.U.E. in the flow predictions. For this purpose, a logit public transport route choice model was considered, in which route generation is exogenous to the problem and the defined vehicle capacity is strict, i.e. the passenger density/m^2 that is modeled for the problem is the maximum and cannot be exceeded.

The methodology consisted of defining a two-stage assignment model that includes the effects of congestion in the definition of the set of attractive lines and in the increases in waiting times. Then, reduced fictitious networks were designed with demands and expected user behavior, with the objective of evaluating the model in its predictive capacity and performance.

After the analysis and improvements of the model, tests were carried out in a real, highly congested network, such as the Santiago Metro (Source SECTRA year 2013). At this stage, its performance was compared with two other models, in which the capacity constraint is modeled as a function of occupancy, Raveau et al. (2011) and De Cea and Fernandez (1993).

The results obtained showed that for certain supply and demand there are equilibrium solutions that respect vehicle capacity and that are obtained by allocating with this model and not with the other models analyzed.

Key words: Assignment Model, Public Transportation, Capacity Constraint, Transportation Networks.

1 INTRODUCCIÓN

1.1 Motivación y contexto

La necesidad de planificar y operar eficientemente un sistema de transporte público, tanto en largo como en el corto plazo, requiere contar con herramientas computacionales que permitan predecir con suficiente precisión cómo las personas eligen sus rutas de viajes, desde un origen a un destino, considerando los servicios disponibles en la red.

Esta predicción de cómo se asignan los viajes en una red de transporte público se realiza mediante Modelos de Asignación. Estos modelos predicen la elección de ruta de los viajeros considerando los costos de cada ruta que el modelador asume son relevantes para ellos, como el tiempo de viaje, tiempo de espera, número de transbordos, entre otros factores.

Debido a las costosas inversiones que es necesario realizar en infraestructura y vehículos de redes de transporte público, es frecuente que la demanda en algunos momentos y lugares bordee o incluso supere la oferta disponible. En estas condiciones se indica que la red está sometida a congestión.

La congestión en transporte público afecta directamente la experiencia de los usuarios, ya sea porque los viajes de otros usuarios en el sistema cambian sustancialmente el nivel de comodidad durante las etapas del viaje o porque el usuario incluso no puede abordar el primer vehículo que pasa, aumentando así su tiempo de espera. Es por esto que los Modelos de Asignación deben incluir en su formulación la restricción de capacidad de los vehículos.

La capacidad de los vehículos varía dependiendo del nivel máximo de ocupación que los viajeros consideran aceptable para abordar. Mientras en una ciudad los usuarios rehúsan subir a un vehículo que lleva un cierto número de pasajeros, en otra ciudad los pasajeros suben al mismo vehículo, aun cuando lleva un 30% más de

pasajeros que su capacidad de diseño. Incluso en una misma ciudad, distintas personas consideran distintos umbrales máximos aceptables para la ocupación al interior del vehículo.

El objeto de estudio de esta tesis corresponde a analizar el caso en que todos los usuarios rehúsan ingresar a un vehículo cuando su capacidad de pasajeros (definida por un determinado nivel de ocupación) ha sido alcanzada.

Cuando la capacidad del vehículo ha sido alcanzada los viajeros no pueden abordar el primer vehículo que llega a la parada del servicio que les interesa, experimentando un aumento en sus tiempos de espera. En los Modelos de Asignación publicados en la literatura, frecuentemente se modela este aumento en la espera como fruto de una caída en la frecuencia original (o frecuencia nominal) del servicio, debido a que algunos vehículos que atienden la parada no estarán disponibles para los usuarios. Esta frecuencia de vehículos con capacidad disponible para subir es conocida como la frecuencia efectiva. Es importante destacar que la frecuencia efectiva es siempre menor o igual que la frecuencia nominal.

Dependiendo de la magnitud del aumento en el tiempo de espera, debido a las frecuencias efectivas experimentadas, las decisiones de los viajeros podrían verse alteradas, ya sea esperando en la misma parada nuevas líneas atractivas para completar su viaje o simplemente cambiando su ruta.

En relación con las herramientas de planificación de transporte urbano utilizadas en Santiago de Chile, la más usada es ESTRAUS (SECTRA, 2019), que considera dos niveles de planificación: estratégica y táctica.

Para la planificación táctica, ESTRAUS ofrece un módulo que considera un Modelo de Asignación Determinística en redes de transporte público denominado MAITE. Este módulo contempla una restricción de capacidad de los vehículos que se traduce en un tiempo de espera adicional para un determinado viaje en función de los flujos

que restan capacidad para ese viaje, acrecentando la competencia por los cupos disponibles.

Adicionalmente, se ha desarrollado un nuevo modelo de planificación táctica para sistemas de transporte público urbano denominado STEP. Este es un Modelo de Asignación Estocástica basada en el modelo de Raveau et al. (2011), que no incluye explícitamente la restricción de capacidad de los vehículos, sino que tiene un costo asociado a la ocupación de los servicios. Así, servicios que presentan más congestión tienen una penalización mayor en la utilidad que servicios menos congestionados.

En ambos modelos, los efectos del elevado uso de los vehículos se presentan como un costo que es creciente con el nivel de ocupación. Ambas formulaciones no garantizan que la capacidad de un arco sea siempre mayor o igual al flujo asignado al arco. Si la asignación supera a la capacidad este resultado podría interpretarse como una densidad (pasajeros/m^2) al interior del vehículo superior a la que define la capacidad. Por ejemplo, en Santiago, a menudo se planifica considerando que la capacidad de los pasajeros de pie en un vehículo equivale a 6 pax/m^2; sin embargo, Metro de Santiago ha reportado densidades promedio sistemáticamente superiores a los 6 pax/m^2 durante ciertos períodos de la punta mañana y en ciertos tramos de la red. En consecuencia, un buen Modelo de Asignación debiera ser capaz de predecir que en estos casos la asignación superará la capacidad, pero en otros contextos también podría considerarse que una asignación de flujo que excede la capacidad pierde realismo, si efectivamente se considera que la capacidad de los vehículos no puede superarse. Este último caso será el caso abordado por esta tesis.

Si la capacidad realmente no puede superarse entonces, no definir un límite máximo puede generar asignaciones en donde se sobreestimen los flujos en algunos arcos y se subestimen en otros. Como consecuencia, los costos asociados a los viajes no estarán correctamente estimados y por lo tanto la predicción de flujo de pasajeros se alejará de la realidad.

1.2 Hipótesis, objetivos y alcances

A continuación, se presenta la hipótesis de trabajo con los objetivos asociados y se indican los alcances que tendrá la tesis.

1.2.1 Hipótesis

Es posible elaborar un Modelo de Asignación que genere una predicción de flujos consistente con un Equilibrio de Usuarios Estocástico y una restricción estricta de capacidad de los vehículos.

1.2.2 Objetivos

El objetivo general es construir un Modelo de Asignación de viajes en una red de transporte público basado en el concepto de Equilibrio de Usuarios Estocástico al que se le incorpora una restricción de capacidad estricta de los vehículos.

Objetivos específicos:

1) Proponer y validar un algoritmo de solución para el Modelo de Asignación Estocástica de viajes respetando la restricción de capacidad estricta de los vehículos.

2) Analizar el desempeño del modelo desarrollado para la predicción de las rutas realizadas por cada usuario considerando distintos escenarios. Primero se probarán redes ficticias y luego una red real, altamente congestionada, como es la del Metro de Santiago.

3) Analizar el desempeño del algoritmo en números de iteraciones para red de prueba.

1.2.3 Alcances

En esta tesis, se considera un Equilibrio de Usuario Estocástico por sobre un Equilibrio de Usuario Determinístico, ya que el primero se ajusta más a la realidad al modelar un efecto probabilístico asociado a las percepciones individuales de cada usuario para distintos niveles de congestión.

En relación con cómo los usuarios escogen dentro de un conjunto de rutas disponibles, se considera un Modelo de Asignación tipo Logit en el que la generación de rutas es un dato exógeno al problema.

La capacidad de los vehículos definida se considera estricta, es decir, la densidad medida en pax/m^2 que se modela para el problema es la máxima y por lo tanto la asignación de flujos no la podrá sobrepasar.

El modelo asume que la oferta disponible en la red es suficiente para transportar a través de alguna de las rutas todos los viajes de la matriz origen destino que se asignará.

1.3 Estructura

La estructura de la tesis se organiza en cinco capítulos. El segundo capítulo corresponde al marco teórico donde se explican los Modelos de Asignación en redes de transporte, algunos conceptos básicos asociados a las redes de transporte público, cómo se ha modelado en la literatura la restricción de capacidad y el tiempo de espera adicional. El tercer capítulo muestra el Modelo de Asignación propuesto que respeta el Equilibrio de Usuarios Estocástico y la restricción de capacidad de los vehículos estricta, evaluando su capacidad predictiva en redes ficticias. En el cuarto capítulo se muestran los resultados en la red de Metro de Santiago, y se analiza su capacidad predictiva y desempeño en tiempos de ejecución. Finalmente, se termina

con el quinto capítulo en donde se muestran las conclusiones de este trabajo y se discuten nuevas líneas de investigación.

2 MARCO TEÓRICO

En este capítulo inicialmente se presentan los principales Modelos de Asignación utilizados en redes de transporte. A continuación, se describe cómo se modelan las rutas de transporte público mediante distintas formulaciones de grafos. Finalmente se explica cómo se han abordado en la literatura los problemas de congestión asociados a la restricción de capacidad de los vehículos en redes de transporte público.

2.1 Modelos de Asignación

Los Modelos de Asignación de viajes en redes de transporte buscan determinar el Equilibrio de Tráfico en una red. Este equilibrio consiste en identificar los flujos resultantes en la red de transporte (público o privado) considerando la elección de ruta de cada uno de estos usuarios, una demanda fija entre los distintos pares orígenes–destinos y un comportamiento para elegir rutas previamente definido (De Cea, 2012).

Estos modelos se pueden clasificar bajo múltiples criterios, pero para efectos de esta tesis interesan especialmente dos. El primero corresponde a incluir o no los efectos de la congestión en las elecciones de rutas de los viajeros. La congestión se genera producto de la capacidad limitada de la oferta en redes de transporte, ya sea porque la capacidad de los vehículos en transporte público es restringida o porque la capacidad de las vías en transporte privado es fija. Esta oferta limitada implica que los costos de las rutas dependan de sus flujos. Y al aumentar el costo de una ruta otras rutas antes descartadas, se vuelven atractivas. Así, para distintos niveles de congestión pueden existir más o menos rutas atractivas.

El segundo criterio se asocia con suponer o no que todos los usuarios comparten una misma valoración de los atributos de viajes y que esta valoración es conocida por el modelador. Si se asume que este supuesto no se cumple equivale a reconocer que no todos los usuarios se comportan de la misma manera al valorar una ruta para su viaje. Modelar

esta cualidad equivale a suponer, a ojos del modelador, que existe un efecto estocástico en la valoración del viaje por parte del usuario.

Los modelos asociados a la clasificación bidimensional recién descrita se muestran en la Tabla 2-1: .

Tabla 2-1: Clasificación de Modelos de Asignación

		¿Incluye los efectos estocásticos?	
		No	Sí
¿Considera los efectos de la congestión?	No	Asignación Todo o Nada	Asignación Estocástica
	Sí	Asignación considerando Equilibrio de Usuarios Determinístico (D.U.E.)	Asignación considerando Equilibrio de Usuarios Estocástico (S.U.E.)

Fuente: (Herrera & Sommariva, 2013)

Debido a que esta tesis se enmarca en los problemas que enfrentan congestión, interesa profundizar en los modelos de Equilibrio de Usuarios mencionados en la segunda fila de la Tabla 2-1: .

2.1.1 Equilibrio de Usuarios Determinístico

Los Modelos de Asignación que consideran un Equilibrio de Usuarios Determinístico (D.U.E.) consideran los siguientes supuestos de comportamiento (Sheffi, 1985):

1) Todos los usuarios tienen información perfecta de la red de transporte.
2) Los individuos son racionales y siempre toman las decisiones correctas respecto a las elecciones de rutas.
3) Todos los viajeros tienen el mismo comportamiento, es decir, valoran los atributos de cada alternativa de forma idéntica y el modelador conoce esta valoración.

Estos supuestos significan que cada uno de los usuarios escoge sus rutas minimizando sus costos totales de viaje de acuerdo con una valoración de los atributos de los viajes que coincide exactamente con la consideración del modelador. Si se considera este comportamiento en una red que no está afecta a la congestión, entonces la asignación buscada consistiría en que todos los usuarios elegirían la ruta de menor costo para su viaje (en ausencia de congestión, este costo es independiente de la asignación de los viajes por la red). Como todos los usuarios que realizan un viaje elegirían una mista ruta, entonces la asignación esperada equivaldría a un Modelo de Asignación Todo o Nada.

Por el contrario, si se considera este comportamiento en una red afecta a congestión, es decir, a que los costos medios de los arcos son crecientes con sus flujos, todos los usuarios que realizan un mismo viaje no necesariamente elegirían la misma ruta. Por ejemplo, si se considera una red de dos nodos y dos arcos con las funciones de costos que se muestran en la Figura 2-1, una demanda de 10 viajes desde del nodo 1 al nodo 2 y dos rutas paralelas, la primera corresponde al arco con costo c_1 y la segunda al arco con costo c_2.

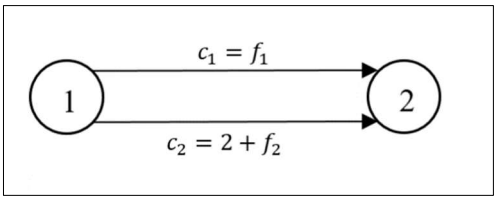


Figura 2-1: Red ejemplo 1
Fuente: Elaboración propia

En ausencia de congestión, la ruta 1 parece más conveniente, pues tiene un costo de 0 si nadie la usa en comparación de la ruta 2 que tiene un costo de 2. Pero si los 10 usuarios eligieran la primera ruta, su costo aumentaría a 10, mientras que el de la segunda ruta permanecería en 2. Esto implicaría que algunos usuarios podrían cambiarse a la segunda ruta y disminuir su costo total de viaje esperado, por lo tanto, esta asignación no podría ser considerada de equilibrio. Lo mismo ocurre si todos los usuarios eligieran la segunda ruta, el costo total esperado correspondería a 12 y el de la primera ruta sería de 0, lo que

implicaría que algunos usuarios se debieran cambiar a la primera ruta con el objetivo de obtener un menor tiempo total esperado de viaje.

La única solución de equilibrio, en la que todos los usuarios minimizan su costo total esperado de viaje, exige en este caso que ambas rutas sean utilizadas. Para determinar la asignación de viajes de la red de transporte (cuántos viajeros eligen la primera y segunda ruta), se utiliza el Primer Principio de Wardrop:

"*Los tiempos de viajes de todas las rutas utilizadas son iguales y las rutas no utilizadas tienen tiempos de viajes iguales o mayores a las rutas utilizadas*." (Wardrop, 1952, pág. 345)

Es decir, la solución de equilibrio del ejemplo tiene que cumplir con las siguientes condiciones: (1) que el costo de la primera ruta sea igual al costo de la segunda ruta (ecuación 2.1), ya que ambas rutas se utilizan y (2) que la asignación de viajes para ambas rutas sume la demanda total para el par origen-destino (ecuación 2.2). Esto conduce al siguiente sistema de ecuaciones.

$$f_1 = f_2 + 2 \tag{2.1}$$

$$f_1 + f_2 = 10 \tag{2.2}$$

Al resolver este sistema de ecuaciones, se obtiene que la solución de equilibrio corresponde a 6 viajeros en la primera ruta y 4 viajeros en la segunda ruta. El costo de equilibrio será 6.

Es importante mencionar que para una red genérica con múltiples nodos, arcos y pares origen-destino, los Modelos de Asignación que arrojan un Equilibrio de Usuarios Determinístico (D.U.E.) se resuelven mediante un conjunto de condiciones de equilibrio conocido como desigualdad variacional (ecuación 2.3), además de un conjunto de restricciones que exigen que los flujos resultantes del equilibrio (ya sea en términos de flujos en rutas h_r^* como en términos de los arcos f_a^*) satisfagan la condición de factibilidad de los flujos por la red (mediante las ecuaciones 2.4, 2.5 y 2.6) (De Cea & Fernández, 1993).

$$\sum_{w \epsilon W} \sum_{r \epsilon R_w} c_r^* \cdot (h_r - h_r^*) \geq 0 \quad \forall\, r \ \in R_w, w \ \in \ W \tag{2.3}$$

$$\sum_{r \,\epsilon\, R_w} h_r = \ T_w \quad \forall\, w \in W \tag{2.4}$$

$$\sum_{w \epsilon W} \sum_{r \,\epsilon\, R_w} \alpha_{ar} h_r = f_a \quad \forall\, a \ \in A \tag{2.5}$$

$$h_r \geq 0 \quad \forall\, r \in R_w, w \,\epsilon\, W \tag{2.6}$$

La notación se describe a continuación:

1) **Conjuntos:**

   W: Es el conjunto de pares orígenes-destinos w de la red.

   R_w: Es el conjunto de rutas que conectan el par w.

   A: Es el conjunto de arcos de la red.

2) **Variables:**

   c_r^*: Es el costo de la ruta r en la asignación de equilibrio.

   h_r^*: Es el flujo de la ruta r en la asignación de equilibrio.

   h_r: Es el flujo de la ruta r asociado a cualquier asignación factible de la demanda sobre la red (no necesariamente la de equilibrio).

   f_a: Es el flujo del arco a, asociado a cualquier asignación factible de la demanda sobre la red (no necesariamente la de equilibrio).

3) **Parámetros:**

   T_w: Demanda de viajes del par w

   α_{ar}: Elemento representativo de la matriz de incidencia arco-ruta que toma el valor 1 si el arco a pertenece a la ruta r, o el valor de 0 en otro caso.

2.1.2 Equilibrio de Usuarios Estocástico

El Equilibrio de Usuarios Estocástico (S.U.E.) se sustenta en dos supuestos de los tres considerados para el Equilibrio de Usuarios Determinístico (D.U.E): 1) los individuos

poseen información perfecta de la red y 2) los individuos son seres racionales que quieren minimizar sus tiempos de viajes. El tercer supuesto del D.U.E., que define a todos los usuarios con igual comportamiento, se relaja distinguiendo entre el costo de viaje que cada usuario percibe y el costo de viaje que el modelador asume (Sheffi, 1985). Así el Primer Principio de Wardrop, se puede volver a enunciar como:

"*Ningún usuario* ***cree*** *que puede mejorar su tiempo de viaje cambiando unilateralmente su ruta*" (Daganzo & Sheffi, 1977, pág. 255).

Las condiciones de equilibrio de este problema prescinden de la desigualdad variacional (2.3) y se enuncian en vez de acuerdo con la ecuación (2.7) y manteniendo las restricciones de factibilidad de flujo del problema anterior (ecuaciones 2.4 -2.6):

$$h_r = T_w \cdot P_r^w \quad \forall\, r \in R_w, w \in W \tag{2.7}$$

Donde P_r^w, representa la probabilidad de que un usuario viajando en el par w elija la ruta r (ecuación 2.8). Esta probabilidad de elección consiste en la probabilidad de que el costo de la ruta r percibido por un usuario (C_r) sea menor o igual que los costos percibidos de todas las otras rutas alternativas, y por lo tanto sea la ruta escogida por ese usuario.

Sin embargo, el modelador solo conoce una parte del costo real percibido por el usuario, aquella que puede medir. El modelador considera que este costo observado es la media del costo percibido por cualquier usuario. Se asume que el costo percibido responde a una distribución aleatoria que el modelador conoce. Así, la probabilidad P_r^w depende de los costos de la ruta r y de todas las otras rutas que sirven al mismo par w con $c_w = \left(c_1 \dots c_{|R_w|}\right)$ (Sheffi, 1985, págs. 309-310).

$$P_r^w = P_r^w(c_w) = \Pr\{C_r \leq C_k \;\forall\, k \neq r, k \in R_w | \, c_w\} \quad \forall\, r \in R_w, w \in W \tag{2.8}$$

La formulación del costo percibido (ver ecuación 2.9) se sustenta en la teoría de utilidad aleatoria que agrega a cada costo medido de una ruta *r* un error aleatorio (ξ_r), relacionado a todo aquello que no puede ser observado por el modelador, pero que determina la decisión del individuo.

$$C_r = c_r + \xi_r \quad \forall\, r \in R_w, w \in W \tag{2.9}$$

Distintos modelos de asignación surgen de la distribución de probabilidad de los errores aleatorios como, Logit, GEV, Probit y Logit Mixto. En esta tesis se usará los modelos Tipo Logit y serán descritos en el capítulo 3.

Cabe mencionar, que otra forma de incluir el efecto estocástico de que no todos los usuarios consideran los costos de viajes igual, se puede hacer mediante la modelación de los efectos considerados al subir a un vehículo en transporte público. En este caso los costos cambian según el paradero de subida como lo hace Cortés et al (2013), este enfoque queda fuera del alcance de esta tesis.

Definido el Equilibrio de Usuarios Estocástico se puede mostrar que el Equilibrio de Usuarios Determinístico es un caso particular del caso estocástico. Para más detalle ver (Sheffi, 1985, págs. 310-312).

Se han descrito brevemente los Modelos de Asignación determinístico y estocástico, que arrojan una predicción para los flujos que se observan en una red considerando el efecto de la congestión en la modelación de los costos de las rutas. Una fortaleza importante del D.U.E. es que en muchos casos es posible generar una predicción de flujos para la red sin la necesidad de identificar las rutas disponibles para viajar entre cada par de nodos. Por el contrario, en el S.U.E. es necesario listar las rutas que serán consideradas por cada usuario y que podrían recibir una probabilidad no nula para viajar entre cada par origen-destino.

Así al trabajar en esta tesis con un S.U.E. es necesario determinar cómo los viajeros caracterizan las rutas que podrían ser usadas para sus viajes. En la siguiente sección se describirá distintos niveles de comprensión de la red de transporte público para determinar las rutas alternativas. En cada caso se define un grafo de transporte y se describen los supuestos considerados.

2.2 Redes de Transporte Público

Las redes de transporte público pueden modelarse a través de grafos en los cuales es posible identificar las rutas consideradas por los usuarios al viajar. La forma en que los viajeros caracterizan sus rutas de transporte público depende de la complejidad del razonamiento del individuo y su capacidad para comprender sus opciones de viajes. Estos distintos comportamientos difieren en cómo los usuarios estructuran su viaje en función de qué servicios llegan primero al paradero de cada una de sus etapas de viaje.

A continuación, se presentan los tres tipos de comportamiento de los usuarios para viajar en redes de transporte público que comúnmente son considerados en la modelación de rutas: desde un comportamiento muy simple, en una red que solo considera un servicio atractivo por etapa de viaje, hasta un comportamiento más complejo en que el usuario construye una híper-ruta que combina transbordos y servicios. Para cada uno de los comportamientos se describe la red que permite asignar los viajes suponiendo ese comportamiento.

Es importante mencionar, que estos tres tipos de redes representan sistemas de transporte público sin presencia de congestión. Es decir, se considera que los usuarios reciben un nivel de servicio que es independiente de lo que los demás usuarios hacen. En el caso de redes de transporte público en que los demás usuarios causan congestión al usar los vehículos, esta congestión puede impedir que algunos usuarios aborden el primer vehículo que están esperando debido a la capacidad limitada de estos sistemas, aumentando sus tiempos de espera y haciendo más incómodo el viaje en un vehículo con muchos pasajeros. Si se quisiera incluir los efectos de la congestión se podría hacer mediante un problema de punto fijo en que los costos asociados a cada ruta generada dependan de la asignación de flujos en la red. Este es el enfoque seguido en esta tesis.

2.2.1 Itinerarios mínimos

La construcción de rutas mediante itinerarios mínimos es la más simple que pueden considerar los viajeros, ya que estos predisponen de una secuencia de etapas de viaje para alcanzar el destino y completar su viaje eligiendo el servicio o los servicios de menor costo. Para representar este comportamiento mediante un grafo se construye una red G (N, L), con N el conjunto de nodos y L el conjunto de arcos llamados secciones de línea que representan todos los pares de nodos (o paraderos) conectados directamente por cada servicio. Al resolver un algoritmo de rutas mínimas sobre esta red G (N, L) se identifican los itinerarios de menor costo que unen cada par origen-destino w.

La construcción de la red G (N, L) se realiza a partir de la red G (N, A), con A el conjunto de arcos conocidos como segmentos de líneas. La red G (N, A), representa el trazado del recorrido a través de los segmentos de líneas (a_i), estos conectan un par de nodos (paraderos) consecutivos del recorrido. Por ejemplo, una red de 4 nodos con tres servicios: servicio verde que pasa por nodo 1, 2 y 3; servicio rojo que conecta los nodos 2, 3 y 4; y servicio azul que conecta el nodo 3 y 4, se puede representar mediante la red G(N,A) de la Figura 2-2.

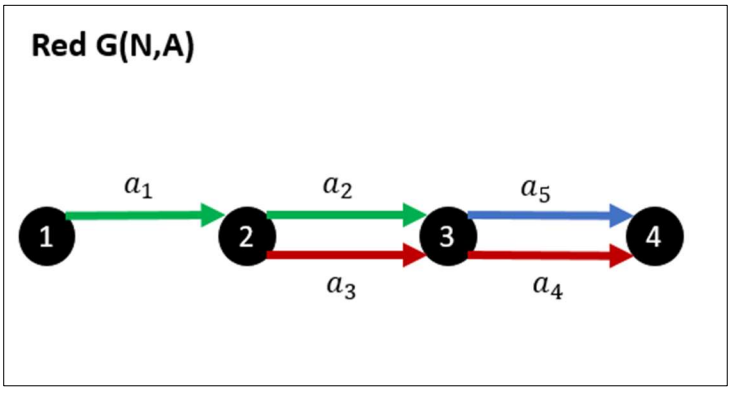


Figura 2-2: Ejemplo 2 y su red G(N,A)
Fuente: Elaboración Propia

A partir de la red G (N, A) se puede generar la red G (N, L), cuyos arcos representan todos los posibles viajes que se pueden completar con cada línea de la red en forma directa. Estos arcos se denominan secciones de línea (l_i), del conjunto de arcos L. En el ejemplo anterior, el recorrido verde puede ofrecer tres posibles viajes: desde el nodo 1 al nodo 2 (l_1), del nodo 1 al nodo 3 (l_2), y del nodo 2 al 3 (l_3). La línea roja también puede ofrecer tres conexiones: desde el nodo 2 al 3 (l_4), del nodo 2 al nodo 4 (l_5), y del nodo 3 al nodo 4 (l_6). Finalmente, la línea azul solo permite viajar desde 3 a 4 (l_7), ver Figura 2-3.

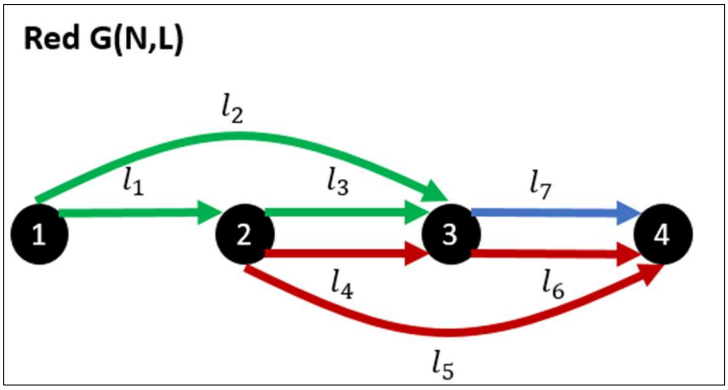


Figura 2-3: Ejemplo 2 y su red G (N, L)
Fuente: Elaboración propia

La red G (N, L) permite identificar las rutas mediante una secuencia de secciones de línea que unen un par w. La secuencia representa las etapas del viaje, por ejemplo, desde el nodo 1 al nodo 4 se puede viajar mediante dos etapas solo por las siguientes secciones de líneas: l_1 y l_5, l_2 y l_6 ó l_2 y l_7. Si se considera que los usuarios buscan minimizar el costo de su viaje, en esta red el itinerario de costo mínimo consistiría en la ruta de menor costo en la red G (N, L) de acuerdo con los Modelos de Asignación descritos en la sección 2.1.

2.2.2 Rutas mínimas

La siguiente red utilizada para asignar los viajes de los usuarios considera un comportamiento más complejo que el anterior, ya que para cada etapa de viaje en vez de esperar un solo servicio (que podría forzar demoras innecesarias si hay otro servicio de similar atractivo), los usuarios toman el primer vehículo que pasa de entre un conjunto de líneas que ellos consideran atractivas para esta etapa, pues les permite minimizar la esperanza del tiempo de viaje. Este comportamiento se modela mediante el algoritmo de Chriqui y Robillard (1975). Este algoritmo se aplica separadamente para cada posible etapa de viaje en la red, por lo que es posible de explicar por medio de la red de la Figura 2-4 que aísla una atapa de viaje (en este caso entre el nodo 1 y el nodo 2). En esta red las personas que deben viajar desde el nodo 1 al nodo 2, tienen disponible n líneas para realizar su viaje.

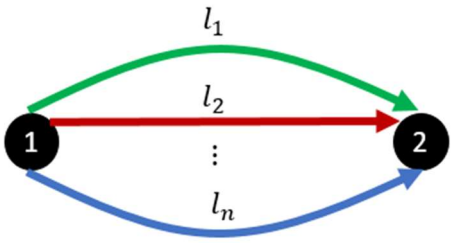


Figura 2-4: Red ejemplo 3
Fuente: Elaboración Propia

Las personas pueden considerar atractivas una, dos, hasta n líneas para realizar su viaje. Dependiendo de cuántas y cuáles líneas se decide esperar, el tiempo de viaje total esperado cambia. Esto se puede expresar a través de la ecuación (2.18), en donde el primer término representa el tiempo de espera esperado y el segundo término el tiempo de viaje esperado, este último corresponde al tiempo de viaje en vehículo de cada línea por la probabilidad de que esa línea sea la utilizada un usuario. En esta expresión el conjunto s corresponde

al conjunto de líneas consideradas atractivas por el usuario para la etapa de viaje en cuestión.

$$TV = \frac{\alpha}{\sum_{i \in s} f_i} + \sum_{i \in s} t_i \left(\frac{f_i}{\sum_{j \in s} f_j} \right) \tag{2.18}$$

Debido a que los pasajeros buscan minimizar su tiempo de viaje total esperado, la definición del conjunto óptimo de líneas atractivas consiste en identificar el conjunto s que minimiza *TV*. Este problema equivale a resolver el siguiente problema de optimización:

$$\underset{\{x_i\}}{Min} \quad \frac{\alpha + \sum_1^n t_i f_i x_i}{\sum_1^n f_i x_i} \tag{2.19}$$

$$s.a. \quad x_i = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} \ \forall\, i = 1..n \tag{2.20}$$

En que x_i es 1 si la línea *i* es considerada atractiva para esta etapa de viaje y 0 si no lo es. Esta formulación no considera los efectos de la congestión en los tiempos de espera o en los tiempos de viaje, por lo tanto, los valores óptimos de x_i no dependen de los flujos asociados a la asignación de viajes. En relación con la resolución del problema se utiliza el siguiente algoritmo (Notación de Cominetti y Correa, 2001):

- 0: $s^* = \{\emptyset\}$
- 1: $s^* \leftarrow s^* \cup \left\{ i \notin s^* : t_i = \min_{j \notin s^*} t_j \right\}$
- 2: $si \begin{cases} \min_{j \notin s^*} t_j \ \geq \ TV_{s*} & termina\ el\ algoritmo \\ \min_{j \notin s^*} t_j \ < \ TV_{s*} & vuelve\ al\ paso\ 1 \end{cases}$

Este algoritmo está basado en la idea de que si alguien considera la línea *j* como una línea atractiva para una etapa de viaje necesariamente tiene que haber considerado atractivas también todas las líneas i más rápidas que j ($t_i < t_j$).

Este algoritmo conocido como el problema de líneas comunes se representa mediante la red G (N, S*) definida por De Cea y Fernández (1993), en donde cada arco del conjunto S*, se llama sección de ruta y representa el costo para el usuario de considerar las líneas

atractivas para viajar sin transbordos entre un par de nodos cualquiera. Las líneas disponibles para viajar entre cualquier par de nodos se obtienen a partir de la red G (N, L) descrita en la sección 2.2.1, Figura 2-5 . Si la línea es considerada en cada sección de ruta de la red G (N, S*) corresponde a una línea que minimiza el costo agregado de viaje. Para el ejemplo 2, en el que se construyó una red G (N, L) si se considera que todos los tiempos de viaje y espera de los segmentos de línea son iguales, se puede obtener la siguiente red G (N, S*)

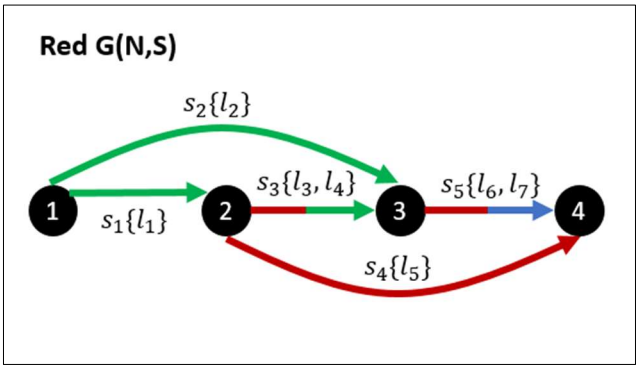


Figura 2-6: Red G (N, S) del ejemplo 2
Fuente: Elaboración propia

Si los costos de todos los segmentos de líneas no son iguales, solo las secciones de ruta s_3 y s_5 podrían cambiar y tener una línea atractiva en vez de dos, según los resultados del problema de líneas comunes. Generada la red con sus secciones de rutas (s_i), las rutas se definen al resolver un algoritmo de rutas mínimas sobre esta red G (N, S*). Como resultado se obtiene una secuencia de secciones de rutas para cada viaje. Por ejemplo, si se quiere viajar del nodo 1 al nodo 4 con solo un transbordo las opciones son viajar con s_1 y s_4 o con s_2 y s_5.

2.2.3 Estrategias mínimas

La tercera red que se puede modelar considera que los usuarios al elegir sus rutas deciden de la siguiente forma:

"*Antes de empezar cualquier viaje, un pasajero ha elegido un subconjunto de líneas para cada parada que puede encontrar en su viaje, y para cada línea un paradero de bajada*" (Nguyen & Pallottino, 1988, pág. 179).

Esta mayor flexibilidad se obtiene producto de que dependiendo de la línea que el usuario toma en un mismo punto podrá bajarse en una parada distinta. Esto da origen a un conjunto de rutas asociadas a un par origen-destino w, llamado híper-ruta, la red asociada a esta estrategia de viaje se conoce como híper-grafo. En esta red cada híper-ruta es un subgrafo, $H_P = \{N_p, E_p, \pi_p\}$, donde N_p es el conjunto de nodos que se podría visitar en la híper-ruta, E_p el conjunto de arcos que representa los segmentos de líneas descritos en 2.2.1. y arcos asociados a las esperas, y π_p vector de valores reales con dimensión $|E_p|$ (Nguyen & Pallottino, 1988). La híper-ruta H_P cumple con:

- H_P es acíclico.
- Nodo origen o_p del par w no tiene predecesor y el nodo destino d_p del par w no tiene sucesor.
- Para cada nodo $i \in N_p - \{o_p, d_p\}$ existe al menos una ruta que pasa por el nodo i.
- Las características del vector π_p son:
    - $\sum_{j/(i,j)\in E_p} \pi_{ijp} = 1 \quad \forall\, i \in N_p$
    - $\pi_{ijp} \geq 0 \qquad \forall\, i \in E_p$

A continuación, se muestra un ejemplo de híper-ruta, red ejemplo 3, con origen en r, destino s, 5 líneas, 7 posibles rutas y una demanda de 120 pax/h. En la Figura 2-7 se muestra las probabilidades, π_{ijp}, para cada arco, estas permiten determinar la probabilidad de cada ruta, tal y como se muestra en la Tabla 2-2.

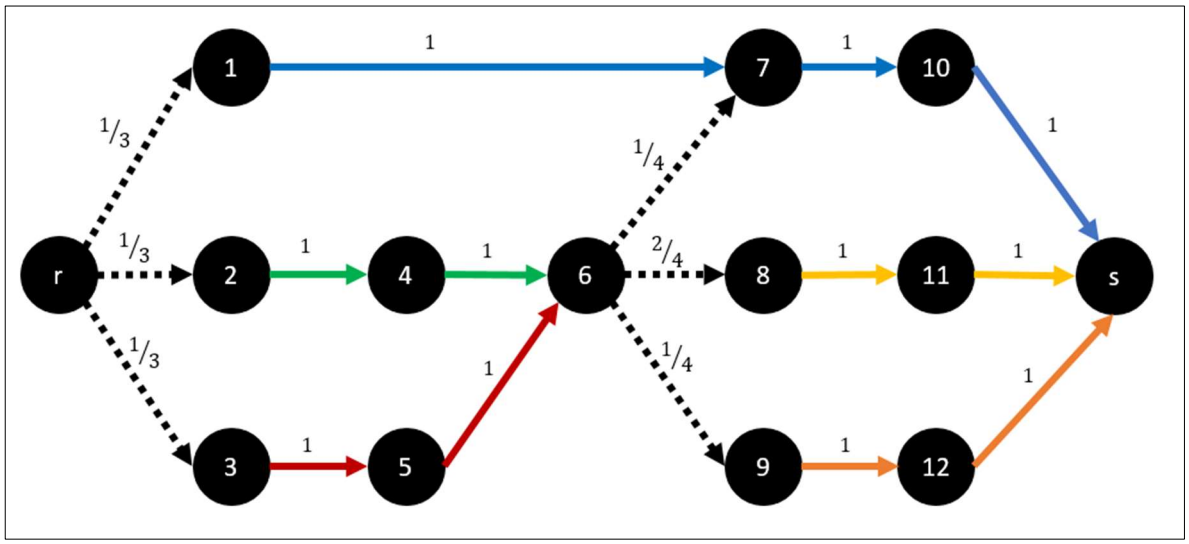


Figura 2-7: Híper-ruta del ejemplo 3
Fuente: *(Nguyen & Pallottino, 1988, pág. 180)*

Tabla 2-2: Resultados en las rutas del ejemplo 3

Rutas	Secuencia de nodos	Probabilidad π_{h_r}	Flujo
$\mathbf{h_1}$	(r, 1, 7, 10, s)	1/3	40
$\mathbf{h_2}$	(r, 2, 4, 6, 7, 10, s)	1/12	10
$\mathbf{h_3}$	(r, 2, 4, 6, 8, 11, s)	1/6	20
$\mathbf{h_4}$	(r, 2, 4, 6, 9, 12, s)	1/12	10
$\mathbf{h_5}$	(r, 3, 5, 6, 7, 10, s)	1/12	10
$\mathbf{h_6}$	(r, 3, 5, 6, 8, 11, s)	1/6	20
$\mathbf{h_7}$	(r, 3, 5, 6, 9, 12, s)	1/12	10

Fuente: (Nguyen & Pallottino, 1988, pág. 180)

Definidas las tres estructuras de viajes más usadas en redes de transporte público, interesa estudiar cómo se ha incluido la restricción de capacidad de los vehículos en los Modelos de Asignación para cada uno de estos comportamientos y sus grafos asociados.

2.3 Restricción de Capacidad

La restricción de capacidad de los vehículos en redes de transporte público altamente congestionadas tiene como resultado que algunos usuarios no puedan abordar el primer vehículo que están esperando. Esto a su vez, tiene como consecuencia que los tiempos de espera aumenten y dependiendo de su incremento las rutas que consideran los viajeros para alcanzar sus destinos podrían verse también alteradas.

El aumento del tiempo de espera asociado a la restricción de capacidad de los vehículos ha sido estudiado bajo dos enfoques (Fu et al., 2012). El primer enfoque trata el tiempo de espera por un conjunto de servicios como una función convexa, creciente con el flujo de pasajeros que espera por esos servicios y definida sobre todo el dominio para ese flujo, lo que permite exceder la capacidad inicial de un arco (De Cea & Fernández, 1993). Las funciones se asocian a cada etapa de viaje y dependen tanto de los flujos de dicha etapa como de los flujos que abordan las líneas atractivas de las etapas en paraderos aguas arribas y que tienen sus destinos aguas debajo de la estación de inicio de la etapa. Esta dependencia entre el costo de una etapa y el flujo de otra etapa tiene como consecuencias que el vector de costos de los arcos de la red presente interacciones asimétricas en los flujos que los determinan. Al tener interacciones asimétricas ya no existe un problema de optimización equivalente para resolver el problema de asignación. Los métodos más usados para abordar este problema de asignación representan la restricción de capacidad a través de una función convexa estrictamente creciente en el flujo definida sobre todo el dominio de flujos. Esta formulación de la capacidad permite que en casos de alta demanda algunos arcos presenten un flujo que excede a su capacidad.

El segundo enfoque corresponde a considerar una restricción de capacidad estricta en el uso de los vehículos como lo hace Lam et al. (1999) y Lam et al. (2002). La ventaja de

este enfoque es que los tiempos de espera adicionales provocados por los problemas de congestión son obtenidos a partir de las condiciones de equilibrio del problema y no tienen que ser determinados previamente por medio de la calibración de una función de costos, como sería el caso del modelo de De Cea y Fernández (1993). Sin embargo, considerar esta formulación implica que el tiempo de espera ya no sea una función continua y creciente con los flujos (ver imagen a de la Figura 2-8), si no que al activarse la restricción de capacidad se gatille un tiempo de espera adicional variable (ver imagen b de la Figura 2-8).

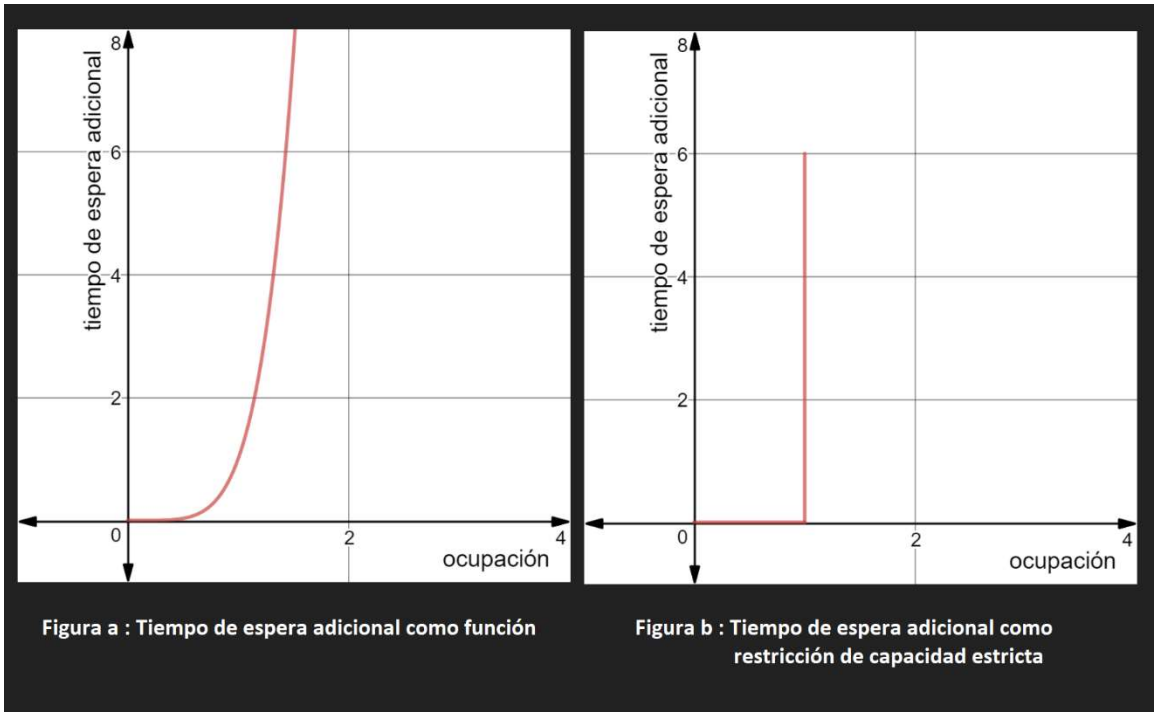


Figura 2-8: Ejemplos de representaciones de tiempos de espera adicional en función de la proporción de ocupación del vehículo.
Fuente: Elaboración propia

Ambos enfoques además proponen una forma de incluir la congestión en la generación de rutas alternativas de viaje cuando el flujo de una etapa se aproxima o alcanza su capacidad. A continuación, se detalla cómo impacta el efecto de no poder abordar el primer vehículo

en la consideración de rutas alternativas por parte del usuario. En la sección 2.3.1 se presenta un modelo de asignación que permite flujos que superan la capacidad de los vehículos, mientras la sección 2.3.2 presenta un modelo que no lo permite.

2.3.1 De Cea y Fernández (1993)

En este estudio se propone un Modelo de Asignación mediante un algoritmo de Equilibrio de Usuarios Determinístico sobre una red de rutas mínimas G (N, S*) como la descrita en 2.2.2. La modelación del tiempo de espera adicional por no poder abordar al primer vehículo se realiza a través de una función de costo que representa el tiempo de espera asociado al inicio de cada etapa de viaje. Esta función asociada a cada nodo *i* de inicio de cada sección de ruta *s* depende de:

i. El flujo que sube en la sección de ruta en estudio, V^s.
ii. El flujo que sube en el mismo nodo a otras secciones de ruta que comparten alguna línea común con las secciones de ruta *s*, V_{is}^{+}.
iii. El flujo que sube antes de *i* a líneas que pertenecen a la sección de ruta *s* y que baja después del inicio de la sección de ruta s, $\bar{V}_{is}$.

El flujo que compite por la capacidad residual con V^s (V_{is}^{+}) y el flujo que reduce capacidad ($\bar{V}_{is}$), se denomina flujo compitente, $\bar{V}_s$. De Cea y Fernández (1993), definen una función genérica de costo φ_s, asociada al tiempo de espera adicional como la ecuación (2.21).

$$\varphi_s\left(\frac{V^s + \bar{V}_s}{K_s}\right) \tag{2.21}$$

Esta función $\varphi_s(x)$ es estrictamente creciente reflejando que a mayor V^s y/o mayor $\bar{V}_s$ mayor será la probabilidad de no poder subir al primer vehículo que pase de las líneas atractivas por el nodo *i* y por lo tanto el tiempo de espera de los usuarios será mayor. Una posibilidad para representar como esta función afecta el tiempo de espera es usar una función BPR como la de la ecuación (2.22).

$$\varphi_s(V) = \beta \cdot \left(\frac{V^s + \bar{V}_s}{K_s} \right)^n \tag{2.22}$$

Esta formulación evita tener que modelar un sistema de cola en cada paradero donde las llegadas de los usuarios son aleatorias y las tasas de servicio dependen de las frecuencias de las líneas y las capacidades residuales mayores a cero de cada servicio. Enfrentar el problema con este último enfoque implicaría lidiar con problemas de factibilidad e inestabilidad numérica, producto de que los costos solo estarían definidos para flujos que no exceden a la capacidad K_s.

Así, las principales ventajas de esta formulación propuesta por De Cea y Fernández (1993) son: (1) la eficiencia del algoritmo de solución, ya que se le agrega una función convexa y (2) la escalabilidad del problema de asignación a un problema a de Equilibrio Simultáneo entre oferta y demanda, como el usado en ESTRAUS.

La principal desventaja de esta formulación en el caso en que las restricciones de capacidad sean estrictas es que permite asignaciones con arcos sobrecargados lo que implica la sobreestimación de flujos en estos arcos y subestimación en otros arcos. Además, es importante mencionar que, para cada red que se esté modelando es necesario calibrar los parámetros de la función de costo descrita en (2.22).

Esta representación del tiempo de espera modela aisladamente el impacto en una sección de ruta dada (con un conjunto de líneas atractivas fijo). Sin embargo, puede ocurrir que si el costo de una sección de ruta aumenta una línea que fue originalmente descartada al definir la sección de ruta sin incluir los efectos previstos de la demanda ahora se vuelva atractiva. Así, para modelar cómo influye la restricción de capacidad en la construcción de rutas y selección de servicios atractivos, De Cea y Fernández (1993) proponen una red G (N, S**) que incluye secciones de rutas paralelas para cada par de nodos con el objetivo de incluir las posibles líneas que podrían volverse atractivas si la demanda aumenta lo suficiente.

Por ejemplo, se tiene una red con un par de nodos y 4 secciones de líneas que los conectan (ver Figura 2-9). Suponga que al resolver el problema de líneas comunes solo la línea 1 y 2 son atractivas, sin embargo, cuando existe presencia de congestión los tiempos de espera aumentan y nuevas líneas se vuelven atractivas. Estrictamente se debería considerar una nueva sección de ruta que incluya no solo estas dos líneas sino también otras que pudieran volverse atractivas. Sin embargo, por simplicidad De Cea y Fernández (1993) lo abordan resolviendo un nuevo problema de líneas comunes solo considerando las secciones de línea restantes, es decir, como si las secciones de línea que fueron escogidas como atractivas en el problema original ya no estuvieran disponibles. Esto permite generar una nueva sección de ruta paralela atractiva con líneas distintas. El problema se puede volver a resolver, pero ahora eliminando las nuevas secciones de líneas recientemente seleccionadas hasta que todas las secciones de líneas queden asociadas a una sección de ruta, ver Figura 2-10.

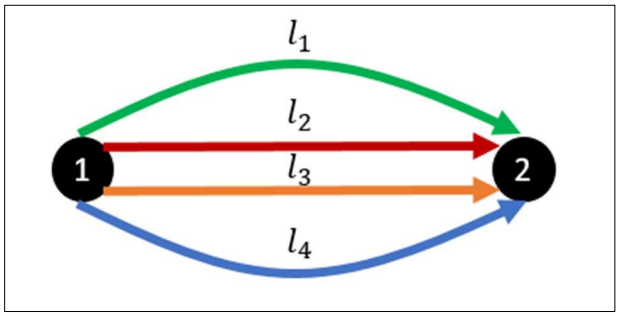


Figura 2-9: Red ejemplo 4
Fuente: Elaboración propia

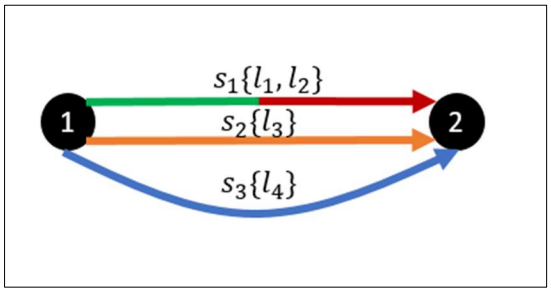


Figura 2-10: Red G(N, S*) del ejemplo 4
Fuente: Elaboración Propia

Un algoritmo de asignación sobre la red G (N, S**) permitirá que los usuarios elijan solo la primera sección de ruta hasta que el tiempo de viaje total esperado sea menor o igual que la segunda sección de ruta. Pero, cuando la congestión aumenta significativamente y el tiempo esperado de viaje por la primera sección de ruta supere cierto umbral, la segunda sección de ruta se volverá también atractiva y los usuarios se repartirán entre la primera y segunda sección de ruta. Eventualmente si la demanda es muy alta podrían elegir también la tercera sección de ruta. Este tipo de red es solo válida para los casos en que los demás costos de viaje, sin considerar el tiempo de espera, son constantes.

En relación con la repartición de flujo en las secciones de ruta, se indica que debe usarse las frecuencias efectivas en vez de las nominales. Las frecuencias efectivas se definen para cada sección de línea como la capacidad efectiva disponible de espacios por hora en los buses que visita la parada, la ecuación (2.23) representa esta frecuencia efectiva.

$$f_l'^i = \frac{\alpha_l}{w_l^i} \tag{2.23}$$

Donde el tiempo de espera de cada sección de línea, w_l^i, depende del tiempo de espera nominal más el tiempo de espera adicional por no poder abordar el primer vehículo, ambos términos se muestran en la ecuación (2.24).

$$w_l^i = \frac{\alpha_l}{f_l} + \varphi_l\left(\frac{\tilde{v}_{il}}{k_l}\right) \tag{2.24}$$

2.3.2 Lam et al. (2002)

En este estudio se presenta un Modelo de Asignación que considera un Equilibrio de Usuarios Estocástico con restricción estricta de la capacidad de los vehículos a través de un modelo Logit con la siguiente utilidad representativa.

$$V_r^w = \theta(t_r^w + u_r^w + d_r^w) \tag{2.25}$$

Donde

- t_r^w: corresponde al tiempo de viaje en el vehículo de la ruta *r* del par *w*.
- u_r^w: corresponde al tiempo espera de la ruta *r* del par *w*.
- d_r^w: corresponde al tiempo de sobrecarga de la ruta *r* del par *w*.
- θ: parámetro de calibración

Este modelo simplificado de elección de rutas en transporte público que considera la restricción de capacidad estricta tiene un problema de optimización equivalente. El problema de optimización para modelos Logit de elección de ruta fue propuesto por Fisk (1980) y modificado por Lam et al. (2002) agregando la restricción de capacidad estricta:

$$Min \sum_{s\epsilon S}(t_s + u_s)v_s + \frac{1}{\theta}\sum_{w\epsilon W}\sum_{r\epsilon R_w} h_r^w(ln(h_r^w) - 1) \tag{2.26}$$

$$\text{s.a.} \sum_{r\,\epsilon\, R_w} h_r^w = T_w \qquad \forall\, w \in W \tag{2.27}$$

$$\sum_{w\epsilon W}\sum_{r\epsilon R_w} \alpha_{sr} h_r^w = v_s \qquad \forall\, s\ \in S \tag{2.28}$$

$$\sum_{w\epsilon W}\sum_{r\epsilon R_w}\sum_{s\epsilon S} \bar{\gamma}_{as}\alpha_{sr} h_r^w \leq k_a \quad \forall\, a \epsilon\ A \tag{2.29}$$

$$h_r^w \geq 0 \qquad \forall\, r\ \in R_w, w \epsilon W \tag{2.30}$$

La notación utilizada se describe a continuación:

- h_r^w: corresponde al flujo en la ruta *r* del par *w*.
- t_s: corresponde al tiempo de viaje de la sección de ruta *s*.
- u_s: corresponde al tiempo espera de la sección de ruta *s*.
- v_s: corresponde al flujo de la sección de ruta *s*.
- T_w: corresponde a la demanda del par *w*.
- α_{sr}: corresponde al elemento de la matriz de incidencia sección de ruta- ruta, e indica si la sección de ruta *s* pertenece a la ruta *r*.

$$a_{sr}: \begin{cases} 1 & si\ s \in r. \\ 0 & e.o.c. \end{cases}$$

- γ_{as}: indica si el segmento de línea *a* pertenece a la sección de ruta *s*.

$$\gamma_{as}: \begin{cases} 1 & si\ a \in s. \\ 0 & e.o.c. \end{cases}$$

- x_s^l: proporción de viajes que de la línea *l* en la sección de línea *s*.
- k_a: capacidad del segmento de línea *a*.

Lam et al. (2002) prueban que el problema de optimización equivalente recién descrito corresponde al modelo de elección de ruta en transporte público que considera la restricción de capacidad estricta si y solo si los multiplicadores de Lagrange asociados a la restricción de capacidad son iguales al inverso aditivo de los tiempos de sobrecarga de los segmentos de línea a. Esta proposición tiene como consecuencia que el tiempo de sobrecarga de la ruta r del par w, $\mathrm{d}_\mathrm{r}^\mathrm{w}$, se obtenga directamente de las condiciones de equilibrio del problema (Lam et al., 1999).

A continuación, se presenta la demostración hecha por Lam et al. (2002) de que el problema de optimización propuesto es equivalente al modelo de elección de ruta en transporte público que considera la restricción de capacidad estricta.

Primero se plantea el Lagrangeano en la ecuación (2.31)

$$L = \sum_{s\epsilon S}(t_s + u_s)v_s \sum_{w\epsilon W}\sum_{r\epsilon R_w} \alpha_{s-r}h_r + \frac{1}{\theta}\sum_{w\epsilon W}\sum_{r\epsilon R_w} h_r^w(lnh_r^w - 1) + \sum_{w\epsilon W} l_w\left(T_w - \sum_{r\epsilon R_w} h_r^w\right) + \sum_{a\epsilon A} m_a\left(c_a - \sum_{w\epsilon W}\sum_{r\epsilon R_w}\sum_{s\epsilon S}\bar{\gamma}_{a-s}\alpha_{s-r}h_r^w\right) \tag{2.31}$$

Reemplazando la ecuación (2.28) en el Lagrangeano, se obtienen las siguientes condiciones de Kuhn-Tucker (KKT).

$$h_r^w\left(\sum_{s\epsilon S}(t_s + u_s)\alpha_s + \frac{1}{\theta}lnh_r^w - l_w - \sum_{s\epsilon S}\left(\sum_{a\epsilon A} m_a\,\bar{\gamma}_{a-s}\right)\alpha_{s-r}\right) = 0, \quad \forall\, r \in R_w, w \in W \tag{2.32}$$

$$\left(\sum_{s\epsilon S}(t_s + u_s)\alpha_s + \frac{1}{\theta}lnh_r^w - l_w - \sum_{s\epsilon S}\left(\sum_{a\epsilon A} m_a\,\bar{\gamma}_{a-s}\right)\alpha_{s-r}\right) \geq 0, \quad \forall\, r \in R_w, w \in W \tag{2.33}$$

$$m_a\left(c_a - \sum_{w\epsilon W}\sum_{r\epsilon R_w}\sum_{s\epsilon S}\bar{\gamma}_{a-s}\alpha_{s-r}h_r^w\right) = 0, \qquad \forall\, a \in A \tag{2.34}$$

$$m_a \leq c_a, \qquad \forall\, a \in A \tag{2.35}$$

$$\sum_{w\epsilon W}\sum_{r\epsilon R_w}\sum_{s\epsilon S}\bar{\gamma}_{a-s}\alpha_{s-r}h_r^w \leq c_a, \qquad \forall\, a \in A \tag{2.36}$$

$$\sum_{r\,\epsilon\, R_w} h_r^w = T_w \qquad \forall\, w \in W \tag{2.37}$$

$$h_r^w \geq 0 \qquad \forall\, r \in R_w, w \in W \tag{2.38}$$

Si h_r^w >0, entonces

$$\sum_{s\epsilon S}(t_s + u_s)\alpha_s + \frac{1}{\theta}lnh_r^w - l_w - \sum_{s\epsilon S}\left(\sum_{a\epsilon A} m_a\,\bar{\gamma}_{a-s}\right)\alpha_{s-r} = 0, \quad \forall\, r \in R_w, w \in W \tag{2.39}$$

$$lnh_r^w = -\theta \sum_{s\epsilon S}(t_s + u_s)\alpha_s + \theta\, l_w + \theta \sum_{s\epsilon S}\left(\sum_{a\epsilon A} m_a\, \bar{\gamma}_{a-s}\right)\alpha_{s-r} \qquad \forall\, r \in R_w, w \in W \tag{2.40}$$

Si se considera:

$$m_r^w = \sum_{s\epsilon S}\left(\sum_{a\epsilon A} m_a\, \bar{\gamma}_{a-s}\right)\alpha_{s-r} \quad \forall\, r \in R_w, w \in W \tag{2.41}$$

Reescribiendo ecuación (2.40) con la ecuación (2.41), se tiene la ecuación (2.42)

$$lnh_r^w = -\theta(t_r^w + u_r^w) + \theta\, l_w + \theta m_r^w \quad \forall\, r \in R_w, w \in W \tag{2.42}$$

Despejando h_r^w y usando la restricción de demanda (2.27) se tiene:

$$h_r^w = \frac{e^{-\theta(t_r^w + u_r^w) + \theta m_r^w}}{\sum_{r \in R_w} e^{-\theta(t_r^w + u_r^w) + \theta m_r^w}} \quad \forall\, r \in R_w, w \in W \tag{2.43}$$

Lo que es equivalente a un Logit en donde m_r^w=-d_r^w se interpreta como el inverso aditivo del tiempo de espera adicional de la ruta *r* del par *w*, el que se obtiene directamente de las condiciones de equilibrio del problema.

En cuanto al algoritmo de solución del problema de optimización planteado en las ecuaciones (2.26) a (2.30), se resuelve mediante un algoritmo de balanceo propuesto por Bell (1995), donde los factores a balancear son M_a y L_w, que corresponden a funciones de los multiplicadores de Lagrange del modelo, m_a y l_w.

A continuación, se muestra la equivalencia entre los factores y los multiplicadores de Lagrange, en base a la expresión para el flujo en las rutas obtenida al derivar el Lagrangeano.

$$h_r^w = e^{-\theta(t_r^w + u_r^w)} \cdot e^{\theta \sum_{a\epsilon r} m_a} \cdot e^{\theta l_w} \qquad \forall\, r \in R_w, w \in W \tag{2.44}$$

$$h_r^w = e^{-\theta(t_r^w + u_r^w)} \cdot \prod_{a\,\epsilon\, r} M_a \cdot L_w \qquad \forall\, r \in R_w, w \in W \tag{2.45}$$

Bell (1995) propone iterar con el siguiente algoritmo.

Inicialización:

$$M_a^{n=1} = 1 \qquad \forall\, a \in A \tag{2.46}$$

$$L_w^{n=1} = 1 \qquad \forall\, w \in W \tag{2.47}$$

Iteración:

Repetir estos pasos hasta que la convergencia en los M_a y L_w es alcanzada.

Para todo i:

$$\beta = \frac{k_a}{\sum_{w \epsilon W} \sum_{r \epsilon R_w} \sum_{sr \epsilon S} \bar{\gamma}_{a-sr}\, \alpha_{sr-r} h_r^w(\boldsymbol{L}_w^n, \boldsymbol{M}_a^n)} \tag{2.48}$$

$$M_a^{n+1} = min\ \{1, \beta \cdot M_a^n\} \tag{2.49}$$

Para todo w:

$$\beta = \frac{g_w}{\sum_{r \epsilon R_w} b_{wr} h_r^w(\boldsymbol{L}_w^n, \boldsymbol{M}_a^{n+1})} \tag{2.50}$$

$$L_w^{n+1} = \beta \cdot L_w^n \tag{2.51}$$

Resultados:

Para todo $r\ \in R_w, w \in W$:

$$h_r^w = e^{-\theta(t_r^w + u_r^w)} \cdot \prod_{a\ \epsilon\ r} M_a \cdot L_w \tag{2.52}$$

Para todo a:

$$f_a = \sum_{s \epsilon S} \bar{\gamma}_{as}\, \alpha_{sr} h_r^w \tag{2.53}$$

$$d_a = -\frac{(ln\, M_a)}{\theta} \tag{2.54}$$

La convergencia del algoritmo se prueba usando el Teorema del Punto de Inflexión que dice que el óptimo corresponde a la solución que minimiza la función de Lagrange respecto a las variables primales (h_r^w) y que la maximiza respecto de las variables duales (m_a y l_w). Al aplicar este teorema se buscan funciones para los multiplicadores de Lagrange asociados a la restricción de capacidad (m_a) y demanda (l_w) que maximicen el valor del Lagrangeano *L*, ecuación (2.31).

Se comienza analizando el impacto de m_a en el L, se observa que m_a está multiplicado por la diferencia entre la capacidad y el flujo en el segmento de línea, por lo tanto, dependiendo del valor que tome esta diferencia, m_a tiene que crecer o decrecer para maximizar el L. A continuación, se describen los valores que tiene que tomar m_a para los tres posibles valores de la diferencia entre la capacidad y el flujo del segmento de línea a.

1. Si $f_a > c_a$, la diferencia queda negativa y como m_a también es negativo este tiene que disminuir para que el Langrangeano crezca. Al disminuir m_a, f_a también se reduce hasta que se cumple con $f_a = c_a$.
2. Si $f_a = c_a$, la diferencia queda en 0 y el valor de m_a no influye en el Lagrangeano.
3. Si $f_a < c_a$, la diferencia queda positiva y m_a tiene que reducirse para que el Langreano crezca.

Debido a que el factor de balanceo es M_a y no m_a, hay que determinar que pasa con M_a cuando m_a crece o disminuye, para esto se puede graficar la función de $M_a = e^{\theta m_a}$, con $m_a \in [-\infty, 0]$, como se muestra en la Figura 2-11, se observa que a medida que m_a decrece M_a también decrece.

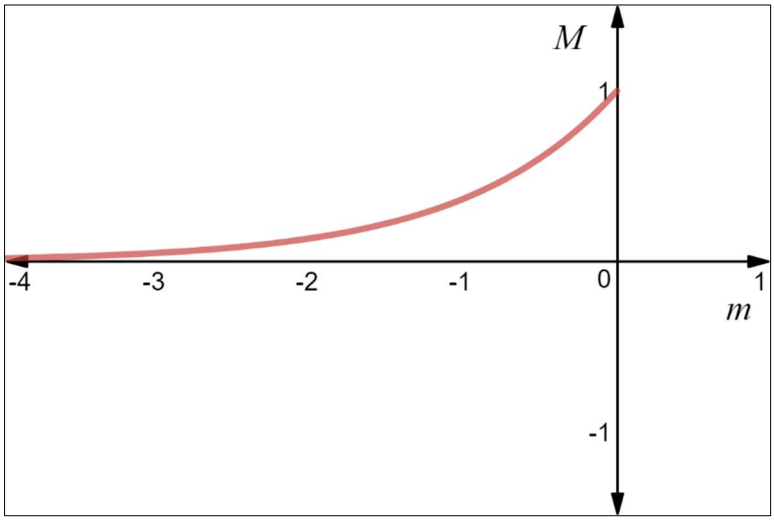


Figura 2-11: Gráfico de $M_a = e^{\theta m_a}$
Fuente: Elaboración propia.

Descrita la relación entre m_a y M_a, se pueden resumir las condiciones para que el Lagrangeano crezca mediante la ecuación (2.21).

$$\frac{c_a}{f_a}\begin{cases} < 1 & M_a \; disminuye \\ = 1 & M_a \; no \; influye \\ > 1 & M_a \; crece \end{cases} \tag{2.55}$$

Estas condiciones las cumple la función de la ecuación (2.55) descrita por Bell (1995).

Con los factores L_w, se realiza el mismo análisis y se evalúa como el L crece con L_w, para esto primero se determina como l_w influye en el Langrangeano, como este aparece multiplicando la diferencia entre la demanda y los flujos en las rutas que conectan un par w, se obtiene las siguientes conclusiones:

1. Si $\sum_{r\epsilon R_w} h_r^w > T_w$, la diferencia queda negativa, entonces l_w tiene que disminuir para que el Langrangeano crezca hasta que $\sum_{r\epsilon R_w} h_r^w = g_w$, porque la función de demanda es estricta.
2. Si $\sum_{r\epsilon R_w} h_r^w = T_w$, la diferencia queda en 0 y el valor de l_w no influye en el Lagrangeano.
3. Si $\sum_{r\epsilon R_w} h_r^w < T_w$, la diferencia queda positiva y l_w tiene que reducirse hasta que $\sum_{r\epsilon R_w} h_r^w = T_w$, porque la restricción de demanda es estricta

La relación entre l_w y L_w se muestra en la Figura 2-12, se observa que a medida que crece l_w, L_w también crece, por lo tanto la función de L_w tiene que crecer cuando la demanda es mayor que el flujo en las rutas y tiene que decrecer cuando el flujo en las rutas es mayor a la demanda, esto se cumple con la ecuación (2.18) definida por Bell (1995).

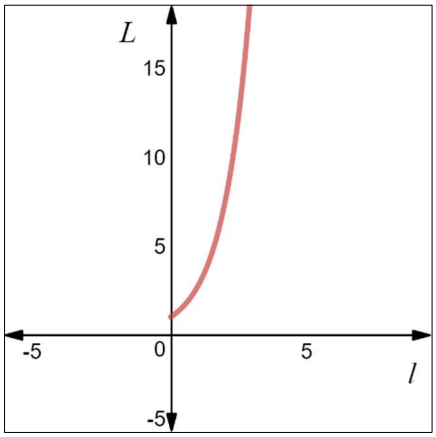


Figura 2-12: Gráfico de $M_a = e^{\theta m_a}$
Fuente: Elaboración Propia

En relación con cómo se considera el cambio del conjunto de líneas atractivas por los aumentos en los tiempos de espera, Lam et al. (2002) propone un problema de punto fijo en donde el tiempo de espera depende de las frecuencias efectivas. Estas frecuencias las define en función de los tiempos de ciclos del recorrido que a su vez dependen de los tiempos de las detenciones en cada paradero y como estos Lam et al. (2002) los modela según los usuarios que suben y bajan, dependen de la asignación.

Lo mismo ocurre con la asignación que depende de los costos asociados al viaje, y como ya se dijo los tiempos de espera dependen de la frecuencia efectiva que a su vez depende de la asignación, por lo tanto, se tiene un problema de punto fijo en donde las frecuencias efectivas se determinan de manera endógena. A continuación, se muestra dónde incide la frecuencia efectiva en el problema ($\boldsymbol{f}$).

$$Min \sum_{s \epsilon S} (t_s + u_s(\boldsymbol{f}))v_s + \frac{1}{\theta} \sum_{w \epsilon W} \sum_{r \epsilon R_w} h_r^w (lnh_r^w - 1) \tag{2.56}$$

$$\text{s.a.} \sum_{r \,\epsilon\, P_w} h_r = T_w \qquad \forall\, w \in W \tag{2.57}$$

$$\sum_{w \epsilon W} \sum_{r \epsilon R_w} \alpha_{sr} h_r = v_s \qquad \forall\, s \in S \tag{2.58}$$

$$\sum_{w \epsilon W} \sum_{r \epsilon R_w} \sum_{s \epsilon S} \bar{\gamma}_{as}(\boldsymbol{f}) \alpha_{sr} h_r^w \leq k_a(\boldsymbol{f}) \quad \forall\, a \,\epsilon\, A \tag{2.59}$$

$$h_r \geq 0 \qquad \forall\, p \in R_w, w \in W \tag{2.60}$$

En el Capítulo 3 de esta Tesis se realiza un análisis crítico de esta metodología, identificando ejemplos en los cuales la asignación que predice no se ajusta a comportamientos esperables en los usuarios de transporte público. Esta Tesis se hace cargo de esta crítica, proponiendo una metodología alternativa para este problema de asignación. Esta metodología predice flujos más coherentes con un comportamiento espontáneo de los usuarios

3 DEFINICIÓN DEL MODELO DE ASIGNACIÓN

En este capítulo se propone un Modelo de Asignación que respeta el concepto de Equilibrio de Usuarios Estocástico (S.U.E.) y responde a la restricción de capacidad de los vehículos en forma estricta. En la sección 3.1. se presenta un Modelo de Asignación Heurístico de Viajes, luego en la sección 3.2. se analiza la asignación que el modelo genera para una serie de redes de tamaño pequeño con el objetivo de realizar pruebas en su capacidad predictiva y desempeño. Se cierra el capítulo con la sección 3.3. donde se resumen el Modelo de Asignación Heurístico de Viajes con los cambios finales.

La implementación computacional del modelo se realizó en C# y la estructura de datos de la red de transporte público utilizada fue la de la red ARTP1 de la herramienta de planificación ESTRAUS. Se decidió utilizar esta estructura, porque genera una red de transporte público de rutas mínimas, G (N, S*) que es la que se considera en el modelo propuesto de esta tesis, permitiendo probar redes de distinto tamaño de manera eficiente sin ninguna modificación adicional a la estructura de datos.

3.1 Modelo de Asignación Heurístico de Viajes

A continuación, se describe el problema que se busca resolver y los supuestos con los que se trabaja. Posteriormente, se formula el Modelo de Asignación Heurístico de Viajes en base a la revisión bibliográfica realizada.

3.1.1 Definición del problema y supuestos

En este trabajo, interesa estudiar redes de transporte público sometidas a altas demandas en donde la oferta disponible del sistema es utilizada al límite, provocando problemas de congestión que determinan las decisiones de los viajeros. En particular, interesa estudiar el caso en que los usuarios se rehúsan a abordar un vehículo porque su capacidad ha sido

alcanzada. Esto podría ocurrir por ejemplo si los vehículos pueden llevar solo pasajeros sentados.

Es importante destacar que en este estudio la capacidad máxima de un vehículo no puede sobrepasarse, a diferencia de otros modelos en que la capacidad puede ser superada con el consiguiente deterioro en el nivel de servicio, tal como el tiempo de espera y comodidad experimentados por los viajeros. En términos prácticos, el Modelo de Asignación propuesto solo genera asignaciones donde el flujo por cada arco es menor o igual a su capacidad. Para estos casos, el fenómeno de congestión se presenta cuando la asignación de flujos en algún segmento de línea alcanza su capacidad máxima y en que existe una demanda que no puede ser satisfecha por estos segmentos. Para estos casos, la asignación de equilibrio forzará a estos usuarios a detectar rutas alternativas atractivas. Para que estas rutas alternativas se vuelvan atractivas se modelará una penalización en el tiempo de espera de las rutas que han alcanzado su capacidad. Esta penalización en el tiempo de espera las vuelve menos atractivas para los usuarios.

La capacidad máxima de un vehículo está definida por la densidad máxima de pasajeros al interior del vehículo y corresponde a un dato exógeno del problema que debe ser definida por el modelador. La capacidad máxima de una línea está determinada por la capacidad del vehículo multiplicada por la frecuencia de la línea.

Otro supuesto del problema es que la oferta de la red modelada es suficiente para transportar la demanda de viajes en la red, en otras palabras, los viajes que se están asignando son viajes que efectivamente podrían viajar en la red de transporte público y, por lo tanto, si es que algunas de las rutas atractivas para algún viaje no tienen capacidad disponible, existen otras rutas con un nivel de servicio menor, que esos usuarios podrían usar para completar su viaje.

En relación con la complejidad de los usuarios para estructurar sus viajes en la red, tal como se vio en el capítulo 2 de esta tesis, la literatura contempla: itinerarios mínimos, rutas y estrategias mínimas. Se descartó los itinerarios mínimos, porque este enfoque no es realista en situaciones con un alto número de ejes de transporte público y alta congestión

en que no todos los usuarios pueden viajar por su ruta de menor costo. Entre los enfoques de rutas y estrategias mínimas, si bien ambos se pueden modificar para considerar los casos en donde existen problemas de congestión se opta por el enfoque de rutas mínimas debido a su mayor simplicidad para implementar, ya que existen programas de planificación que lo utilizan (STEP Y ESTRAUS) y de donde se puede obtener la estructura de datos y codificación para redes reales. Por otro lado, las estrategias mínimas corresponden a un proceso de decisión mucho más complejo que no todos los usuarios de transporte público consideran (Raveau et al., 2014).

En relación con cómo toman las decisiones los viajeros, se consideró que todos son personas racionales que buscan minimizar su costo de viaje esperado y que poseen información perfecta de la red. Este supuesto, no implica que el modelador posea información perfecta de la forma cómo los viajeros toman sus decisiones. De hecho, al tratarse de una asignación estocástica, se asume que el modelador reconoce no contar con información perfecta.

3.1.2 Definición del Modelo

Definido el problema y sus supuestos, se inició el proceso de construcción del Modelo Heurístico de Asignación de Viajes. Para esto se determinó el modelo base con el que se iba a trabajar y luego se fue modificando de manera de incluir los efectos de la restricción de capacidad en los costos de viaje.

a) Modelo de Asignación base

El Modelo de Asignación base elegido es del tipo Equilibrio de Usuarios Estocástico S.U.E. Se descartó un modelo que considerara un Equilibrio de Usuarios Determinístico, D.U.E., porque si bien es posible incluir la congestión en la modelación, su conceptualización es más restrictiva que el S.U.E., ya que asume que todos los usuarios

se comportan de igual manera a ojos del modelador. Por el contrario, en el modelo del tipo S.U.E. se asocia un error a la utilidad de cada alternativa permitiendo capturar los efectos que no pueden ser detectados por el planificador, pero que sí influyen en la decisión del usuario.

Los Modelos de Asignación basados en el concepto de S.U.E. se sustentan en la teoría de Utilidad Aleatoria; en esta, los individuos eligen dentro de un conjunto de alternativas aquella que le genera mayor utilidad. Esto se puede formular matemáticamente mediante probabilidades tal como se presenta en la ecuación (3.1).

$$P_i = \Pr\{ U_i > U_j \, , \forall \, j \neq i\} \tag{3.1}$$

Donde P_i es la probabilidad de elegir la alternativa *i*, y se define como la probabilidad de que la utilidad percibida por el usuario para la alternativa *i* (U_i), sea mayor a las utilidades percibidas para las otras alternativas *j* disponibles ($U_j \, , \forall \, j \neq i$). Esta utilidad el modelador no puede observarla, pero sí puede definir y observar la utilidad representativa, V_i, y asociar a esta un error ε_i por elegir una alternativa *i* que explica los otros factores que considera el individuo en su decisión y que no ve el modelador. De esta forma la ecuación (3.1) queda como la ecuación (3.2).

$$P_i = \Pr\{ V_i + \varepsilon_i > V_j + \varepsilon_j \, , \forall \, j \neq i\} \tag{3.2}$$

Como el error no puede ser observado se modela como una variable aleatoria según los supuestos de comportamiento que defina el modelador. Al asociar una función de densidad de probabilidades al error, la ecuación (3.2) se puede reescribir de manera tal de obtener la probabilidad de elegir la alternativa *i* mediante la distribución de probabilidades acumulada de la diferencia entre los errores de dos alternativas, ecuación (3.3).

$$P_i = \Pr\{ \varepsilon_j - \varepsilon_i < V_i - V_j \, , \forall \, j \neq i\} \tag{3.3}$$

En la literatura, existen diversas especificaciones para la función de densidad del error que dan origen a distintos tipos de modelos, Logit, GEV, Probit y Logit Mixto (Train, 2009,

pág. 26). En esta tesis se utilizará un modelo Logit para caracterizar la elección de rutas de transporte público. Este modelo considera que los errores son independientes e idénticamente distribuidos Gumbel. La ventaja de utilizar esta distribución es que: la media de los errores es 0, los errores están definidos para todos los reales y las probabilidades de elección del modelo Logit son fáciles de manejar (para más detalle ver Anexo 6.1). La desventaja de utilizar un modelo tipo Logit es que el conjunto de alternativas no debe presentar correlación, esto se puede manejar mediante una buena formulación de la utilidad representativa, V_i, de manera tal que explique bien la decisión del usuario y el error modelado sea solo el asociado a la percepción personal del usuario (Train, 2009, pág. 41). En esta tesis se abordará la correlación considerando solo rutas razonables, es decir, rutas que no impliquen transbordos en la misma línea, esto debería disminuir la correlación entre el conjunto de rutas alternativas, pero no lo elimina completamente.

Definidas las características más importantes del Modelo de Asignación S.U.E. base con la formulación de un modelo Logit, se comienza a caracterizar cómo afecta en las decisiones del viajero la restricción activa de capacidad de los vehículos.

b) Tiempos de esperas y restricción de capacidad

El primer efecto que se identifica es el aumento del tiempo de espera por no poder abordar el primer vehículo, este aumento ha sido estudiado en la literatura bajo dos enfoques tal y como se vio en el capítulo 2.

El enfoque utilizado en esta tesis es el que considera una restricción estricta en el modelo (Lam et al., 2002). Se selecciona este enfoque, ya que solo interesa modelar el efecto de no poder abordar el primer vehículo debido a que no existe capacidad. Si se considerara el modelo de De Cea y Fernández (1993) se modelarían tiempos de espera adicionales para situaciones en que existe capacidad disponible. Por ejemplo, si en un paradero todos los buses pasan al 65% de capacidad y el número de pasajeros esperando en este paradero

equivale a un 20% de la capacidad del bus, el modelo asignaría a los usuarios un mayor tiempo de espera por subir a un bus más congestionado, pero no por no poder abordar al bus.

Es importante recordar que estos modelos están pensados para valores promedios en un intervalo de tiempo y si bien no todos los vehículos tienen la misma distribución de llegadas y tampoco la misma cantidad de pasajeros, lo que importa es modelar cuando el no poder abordar un vehículo solo depende de la restricción de capacidad.

La formulación de Lam et al. (2002) modela una asignación estocástica, Logit, en transporte público basado en el problema de optimización equivalente propuesto por primera vez por Fisk (1980) y al que se le agrega la restricción de capacidad estricta según la notación de la red G (N, S). En ese estudio se muestra que el modelo cumple con el S.U.E. y restricción de capacidad si y solo si los multiplicadores de Lagrange asociados a la restricción de capacidad son iguales al inverso aditivo de los tiempos de sobrecarga asociados a cada arco.

El algoritmo de solución utilizado para resolver el problema de optimización es el propuesto por Bell (1995) y utilizado por Lam et al. (2002), que basa su convergencia en un algoritmo de balanceo. Los factores de balanceo de este problema corresponden a los multiplicadores de Lagrange asociados a las restricciones de capacidad y demanda (para más detalle ver sección 2.3.2).

La resolución de este problema se define para una red de asignación fija G (N, S*) en la que a cada sección de ruta se le ha asociado un tiempo de espera adicional mayor o igual a 0. El modelo ajusta este tiempo para así elevar los costos de las rutas más congestionadas de tal forma que ninguna ruta se cargue con más flujo que el que pueden llevar sus vehículos. La ventaja de esta formulación es que los tiempos de espera asociados a no poder abordar un vehículo se obtienen directamente de las condiciones de equilibrio del problema.

La Figura 3-1 presenta el algoritmo que permite determinar los tiempos de espera. Este algoritmo de resolución es un algoritmo iterativo en el que, para una red dada, con sus rutas, costos y demandas se realiza en cada iteración una asignación Logit y si producto de esa asignación la carga de algún segmento de línea excede su capacidad, se aumenta su tiempo de espera, por consecuencia aumenta también el costo de las etapas que utilizan esas líneas y de las rutas asociadas a estas etapas. Con los nuevos costos resultantes se vuelve asignar y revisar la restricción de capacidad, este proceso itera hasta que los flujos en los segmentos de línea satisfacen la restricción de capacidad de los vehículos y algún criterio de convergencia definido por el modelador.

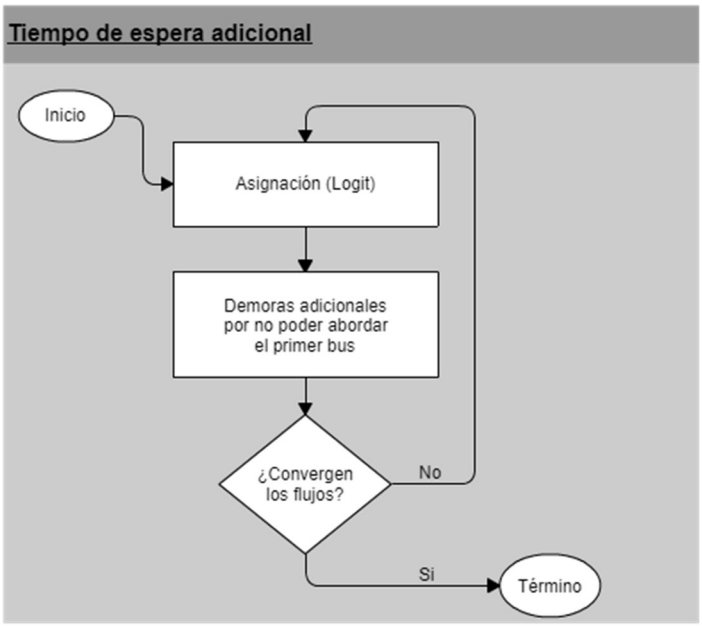


Figura 3-1: Obtención de tiempo de espera adicional
Fuente: Elaboración propia

Si bien este problema genera una asignación final que no sobrepasa la capacidad de los vehículos existe una inconsistencia en este planteamiento y es que al fijar la red G (N, S*), no se están considerando los aumentos en los costos de viajes y sus efectos en la definición del conjunto de líneas comunes. Esto se abordará en la siguiente sección.

c) Conjunto de líneas comunes y restricción de capacidad.

En cada etapa los viajeros esperan un conjunto de líneas que ellos determinan como atractivas o comunes y que les permite minimizar sus tiempos de viajes totales esperados, tal y como lo plantea Chriqui y Robillard (1975). Un supuesto de este enfoque es que no existen problemas de congestión y que las personas pueden subir al primer bus que pasa dentro del conjunto de servicios atractivos. Esto genera un problema para esta tesis, pues en este caso interesa analizar casos en los que los buses circulan a capacidad por ciertas paradas impidiendo ser abordados.

En la literatura se ha estudiado cómo incluir los efectos de la congestión en la definición del conjunto de líneas atractivas mediante el uso de frecuencias efectivas. Las frecuencias efectivas buscan representar el efecto de que no todos los buses que pasan en una red congestionada tienen toda su capacidad disponible para los usuarios que están esperando (De Cea & Fernández, 1993). Matemáticamente la frecuencia efectiva de una línea l, al pasar por un paradero i, se define como:

$$f_l'^i = \frac{\alpha_l}{w_l^i} \tag{3.4}$$

Donde α_l, depende de la distribución de los intervalos de llegada de los buses de la línea l y de los tiempos de llegada de las personas al paradero i, y w_l^i es el tiempo de espera de la línea l en el paradero i, que depende de cuán llenos llegan los buses de la línea l al paradero i y de cuántas personas suben y bajan allí. Es decir, a medida que aumenta el nivel de utilización de la línea, aumentan sus tiempos de espera y como resultado disminuye su frecuencia efectiva. Esto tiene como consecuencia que los tiempos de viaje totales esperados crezcan y aparezcan nuevas líneas atractivas para el usuario.

Cominetti y Correa (2001) también generan un conjunto de líneas atractivas considerando frecuencias efectivas. La diferencia con el modelo anterior es que reescriben el problema de líneas comunes, considerando teoría de colas, y muestran que para el caso particular de

llegadas exponenciales de buses con capacidades variables y llegadas de los usuarios mediante un proceso Poisson, se puede aplicar el concepto de frecuencias efectivas directamente en el problema planteado por Chriqui y Robillard (1975). Esto implica, que la red de asignación se obtiene a partir de las condiciones de equilibrio del problema, en este estudio se demuestra matemáticamente -usando argumentos de punto fijo- la existencia de una red de asignación de equilibrio.

Finalmente, Lam et al. (2002), también incluye los efectos de la capacidad de los vehículos en la construcción del conjunto de líneas atractivas por medio de las frecuencias efectivas. La diferencia de este trabajo con los dos modelos anteriores radica en que la frecuencia efectiva depende del número de buses y tiempo de ciclo del bus, y este último depende del tiempo de detención de bus, el que se obtiene como una función de los pasajeros que suben y bajan en cada paradero. La obtención de la frecuencia efectiva se realiza mediante un problema de punto fijo.

El enfoque utilizado en el modelo propuesto de asignación en esta tesis, es una combinación entre lo que propone Lam et al. (2002) y De Cea y Fernández (1993). Se propone determinar una red de asignación mediante un problema de punto y fijo tal y como lo hace Lam et al. (2002) y en vez de definir las frecuencias en función del número de buses y tiempo de ciclo se utiliza la formulación y de frecuencia efectiva de De Cea y Fernández (1993).

Este es un problema de punto fijo porque la construcción de la red G (N, S*) necesita considerar las frecuencias realmente experimentadas por los usuarios para la definición de sección de rutas. A su vez, las frecuencias experimentadas por los usuarios, que en esta tesis se consideran como frecuencias efectivas, dependen de la asignación que entrega el modelo de elección elegido, por tanto, es necesario iterar hasta que la construcción de la red y la asignación de flujos sean coherentes entre ellas.

Dicho esto, se muestra el proceso iterativo al cual se recurre para construir la red G (N, S*). En la Figura 3-2, se muestran dos etapas necesarias para construir una red de asignación. La primera etapa corresponde a la determinación de frecuencias efectivas de

cada sección de línea *l* y la segunda etapa consiste en determinar la red G (N, S*) en base al problema de líneas comunes de Chriqui y Robillard (1975) a partir de las frecuencias efectivas de la primera etapa.

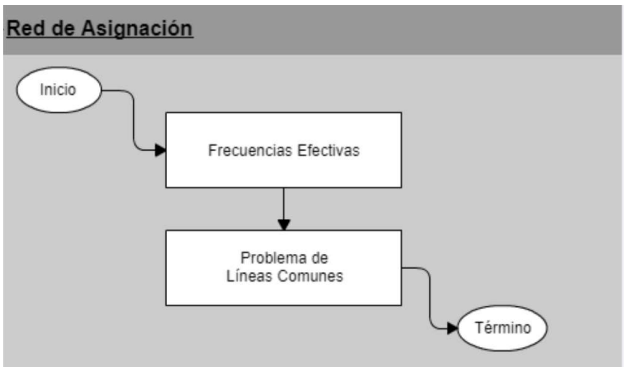


Figura 3-2: Obtención de Red de Asignación
Fuente: Elaboración Propia.

Es importante mencionar que debido a que las frecuencias efectivas de cada sección de línea l dependen de los flujos resultantes de la asignación vigente, en la primera iteración las frecuencias efectivas serán iguales a las frecuencias nominales.

Para ejemplificar este proceso iterativo, se puede considerar una red de tres nodos y dos líneas, como la de la Figura 3-3.

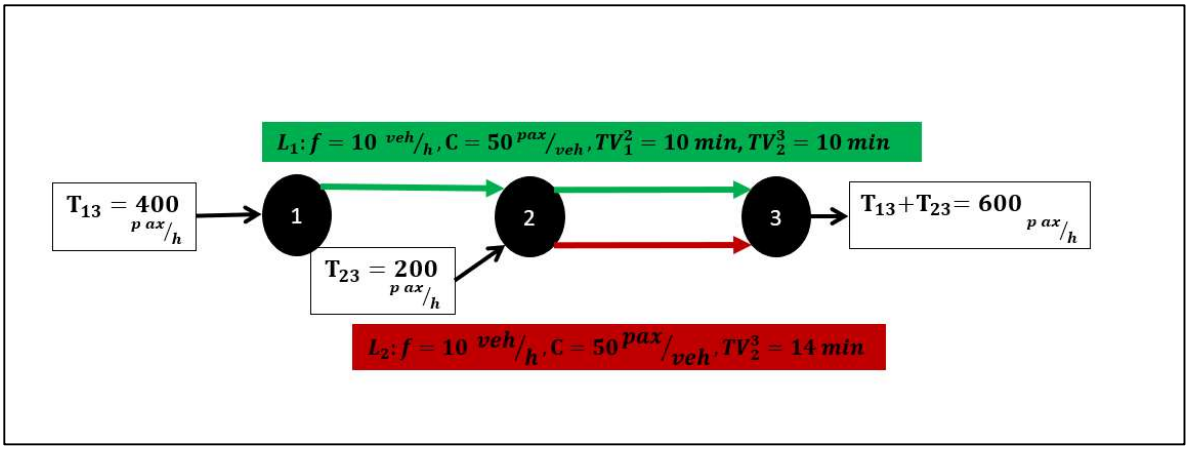


Figura 3-3 Red ejemplo de Líneas Comunes
Fuente: Elaboración propia

En este caso en la primera iteración de la red de asignación, considerando las frecuencias nominales de la red, se tendría la siguiente red G (N, S*) (Ver Figura 3-4)

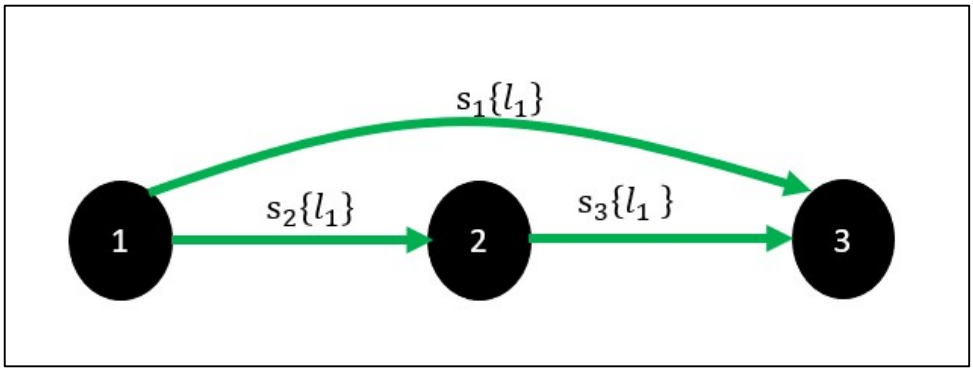


Figura 3-4 Red G (N, S*) de ejemplo de Líneas Comunes - Iteración 1
Fuente: Elaboración propia

Al aplicar un modelo de asignación se tendrían 400 pax/h viajando en la sección de ruta s_1 y 200 pax/h viajando en la sección de ruta s_3. El problema de esta asignación es que al revisar la carga del segmento de línea asociado a la línea verde entre 2 y 3, sería de 600 pax/h lo que sobrepasa la capacidad de este mismo (500 pax/h).

En este caso el comportamiento esperado es que los viajeros de 2 a 3, vean varios vehículos a capacidad de la línea verde y que no puedan abordar al primer vehículo que estén esperando de esta. Por esto, es importante que la construcción de la red G (N, S*) considere la disminución de la frecuencia debido a los resultados de la asignación del mismo problema.

Este problema de capacidad de los vehículos en la definición de las líneas atractivas se pretende abordar mediante el uso de frecuencias efectivas que representan la frecuencia de los vehículos con capacidad en una parada. En términos matemáticos lo que se está haciendo es disminuir las frecuencias nominales debido a posibles problemas de capacidad en algunos nodos y según el modelo de asignación que se considere se obtendrá en la segunda iteración una frecuencia efectiva de 2 a 3 de la línea 1 menor a la nominal, que podría o no ser suficientemente baja para que el tiempo total de viaje esperado de la línea 1 entre 2 y 3 aumente de manera tal que la línea 2 comience a ser atractiva para los viajeros que están esperando en 2. Este proceso podría tardar 2 o más iteraciones, se calculará la red de asignación hasta que la red sea coherente con la asignación, en otras palabras, se calculará la red hasta que las frecuencias efectivas converjan.

En este ejemplo la red de asignación esperada G (N, S*) corresponde a la Figura 3-5, en esta red se consideran como atractivas la línea verde y roja para los viajeros de 2 a 3, lo que permite obtener resultados de asignación más realistas y que respeten la restricción de capacidad de los vehículos.

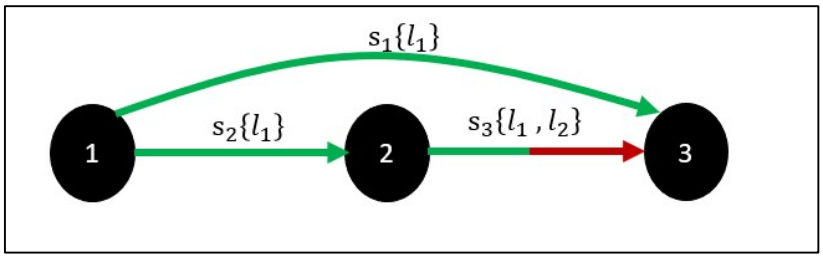


Figura 3-5: Red G (N, S*) final -Ejemplo Líneas Comunes
Fuente: Elaboración propia

Definido los modelos que se van a utilizar para la asignación, la obtención del tiempo de espera de adicional y la definición de la red de asignación, se propone el siguiente Modelo Heurístico de Asignación de Viajes (ver Figura 3-6).

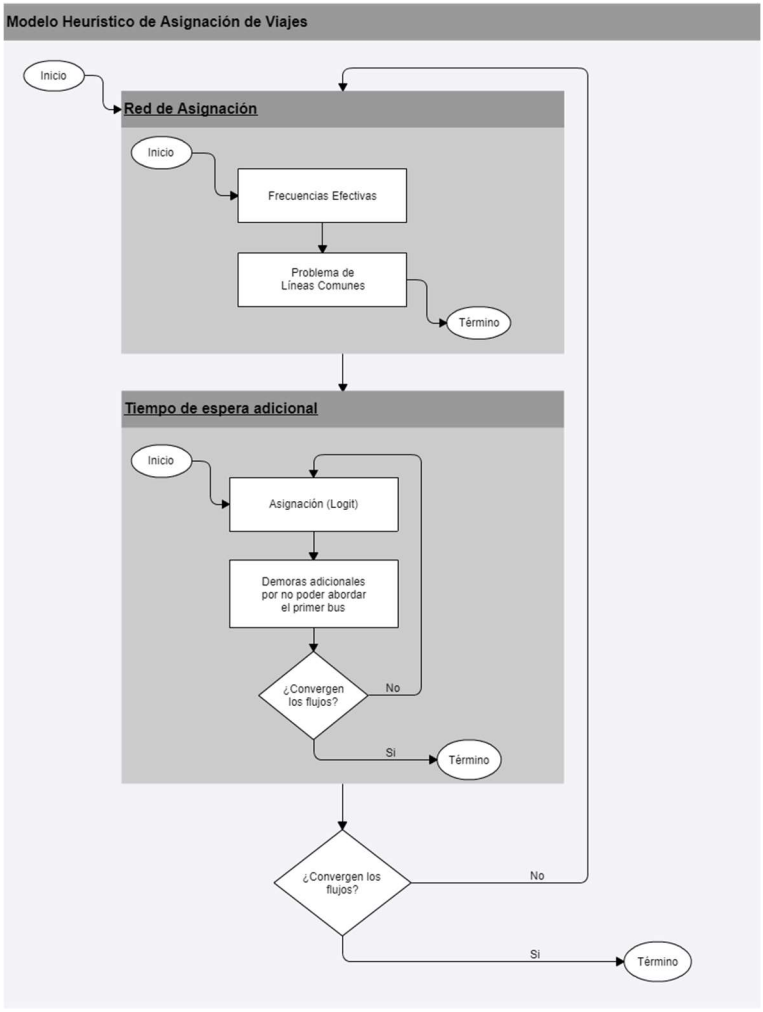


Figura 3-6: Modelo Heurístico de Asignación de Viajes
Fuente: Elaboración Propia

La primera etapa del modelo, Red de Asignación, tiene por objetivo definir la red G (N, S*) con la que se va a iterar, para esto se tienen dos subetapas: la primera define las frecuencias efectivas y la segunda utiliza este resultado para determinar las líneas comunes de cada sección de ruta.

Como ya se mencionó anteriormente en la primera iteración, las frecuencias efectivas se consideran iguales a las frecuencias nominales. A continuación, se describirá cómo se definen las frecuencias efectivas para la iteración 2 en adelante, a partir de los resultados de la asignación realizada en la segunda etapa del Modelo Heurístico de Asignación de Viajes, Tiempo de espera adicional.

En primera instancia, se reescribe la frecuencia efectiva enunciada en la ecuación (3.4) en función de las secciones de línea sl, la iteración $n+1$ del modelo, con $n \geq 1$ y el tiempo de espera para esta sección de línea (te_{sl}^{n+1}) obtenido en la segunda etapa del modelo propuesto.

$$f_{sl}^{n+1} = \frac{\alpha_l}{te_{sl}^{n+1}} \tag{3.5}$$

El tiempo de espera de la iteración $n+1$, obtenido de la segunda etapa del modelo, queda definido por la ecuación (3.6).

$$te_{sl}^{n+1} = te_{sl}^{n} + d_{sl}^{n} \tag{3.6}$$

Donde:

- te_{sl}^{n+1}: Es el tiempo de espera de la sección de línea sl de la iteración $n+1$.
- te_{sl}^{n}: Es el tiempo de espera de la sección de línea sl de la iteración n.
- d_{sl}^{n}: Es el tiempo de espera adicional por no poder abordar el primer vehículo de la sección de línea sl en la iteración n. Este tiempo es un resultado de la segunda etapa del modelo propuesto.

A partir de las ecuaciones (3.5) y (3.6), se obtiene una formulación para las frecuencias efectivas de la iteración $n+1$, con $n \geq 1$.

$$f_{sl}^{n+1} = \frac{\alpha_l}{\left(te_{sl}^n + d_{sl}^n\right)} = \frac{\alpha_l}{\left(\frac{\alpha_l}{f_{sl}^n} + d_{sl}^n\right)} = \frac{1}{\left(\frac{1}{f_{sl}^n} + \frac{d_{sl}^n}{\alpha_l}\right)} \quad (3.7)$$

Estas frecuencias efectivas representan las frecuencias con que las personas ven un vehículo con capacidad disponible. Debido a que esta formulación solo representa vehículos que no tienen problemas de capacidad no se viola el supuesto de que los viajeros esperan un conjunto de líneas y suben al primer vehículo que pasa dentro del conjunto, por lo tanto, se puede utilizar directamente la formulación de Chriqui y Robillard (1975) cambiando solo las frecuencias nominales por las frecuencias efectivas recién definidas.

Una forma de interpretar estas dos etapas iterativas del Modelo Heurístico de Asignación de Viajes es considerar la siguiente situación. En una ciudad como Santiago, cuando comienzan las clases de colegios, institutos y universidades, las personas viajan en la mañana eligiendo la ruta de menor costo según su conocimiento previo de la red, luego de realizar su viaje, los usuarios pueden conservar su elección de ruta o pueden cambiar su decisión producto de los problemas de congestión experimentados en los primeros días.

Es esperable que los usuarios vayan probando distintas configuraciones hasta encontrar la combinación que les genera menor costo de viaje. Además, como los costos experimentados dependen de las decisiones que tomen los otros usuarios, encontrar la mejor ruta puede ser un proceso que tome varios días incluso semanas. Sin embargo, es razonable suponer que los usuarios encontrarán la ruta que les genera mayor utilidad y que por lo tanto la elección de rutas del sistema se estabilice en una solución de equilibrio.

La situación recién descrita es similar a cómo se resuelve el problema de punto fijo que se propone, ya que en la primera iteración del problema genera una asignación que considera los costos de una red sin congestión, la segunda asignación se realiza en base a los resultados de los costos de la primera iteración y este proceso se repite hasta que los flujos de asignación convergen, o como pasa en la realidad hasta que los usuarios seleccionan una ruta que crean que les genera mayor utilidad.

3.2 Pruebas en redes ficticias

Luego de definir el Modelo Heurístico de Asignación de Viajes, se efectuó la implementación computacional del modelo en C#, con el objetivo de realizar pruebas en la capacidad predictiva y desempeño del modelo, en redes ficticias.

3.2.1 Capacidad predictiva del modelo

La capacidad predictiva del modelo se evaluó en dos redes ficticias, la primera es una red simple de 4 nodos y dos líneas, mientras la segunda red tiene 6 nodos, 6 líneas y mayor demanda que la primera red. Para ambos ejemplos la valorización del tiempo de espera y tiempo de viaje de los usuarios se consideró igual a 1.

a) Red ficticia 1

La primera red ficticia y más simple, se define como una calle unidireccional en donde existen cuatro paraderos y el tiempo de viaje experimentado en cada arco es de 10 minutos. Los servicios ofrecidos en la red son dos y sus atributos (frecuencia, tiempos de viaje y capacidad de los vehículos) asociados se muestran en la Figura 3-7. La demanda de la red es de 500 pasajeros/ hora que viajan desde el nodo 1 al nodo 4.

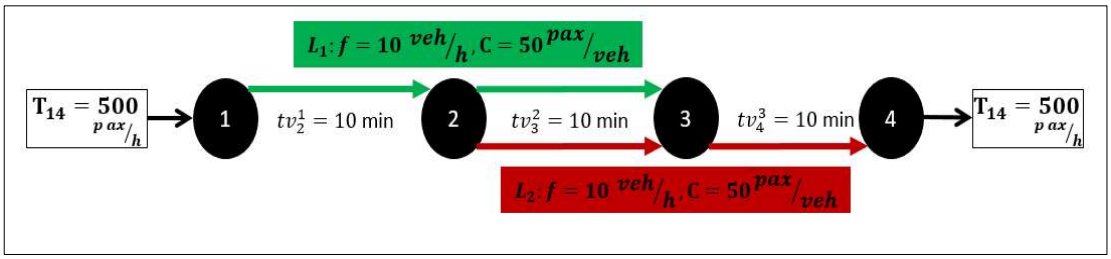


Figura 3-7: Red ficticia 1
Fuente: Elaboración propia

Lo primero que se hace es construir la red G(N, S*) asociada a la red (ver Figura 3-8) y definir las rutas alternativas para los viajeros. Ruta 1 (h_1) y ruta 2 (h_2) y sus niveles de servicio asociados (ver Tabla 3-1), se descarta la ruta que considera las secciones de ruta, s_2 - s_3 - s_5 ya que no es razonable que un viajero se baje en el nodo 2 para esperar un conjunto de líneas comunes que considera como atractiva una línea que ya estaba usando. Además, por construcción el costo total de viaje siempre será mayor que el costo de las otras dos rutas, ya que los tiempos de viaje son iguales, pero siempre tendrá un tiempo de espera adicional por una parada añadida.

En este caso, ambas rutas exigen tomar primero la línea verde y luego la línea roja, pero los usuarios pueden elegir entre transbordar en 2 o en 3, lo que define las rutas 1 y 2 respectivamente.

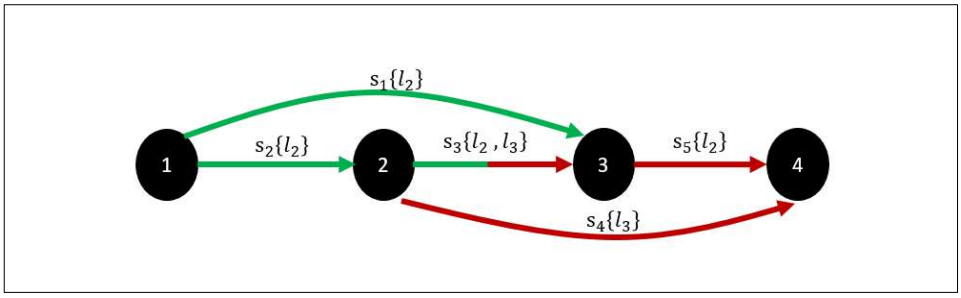


Figura 3-8: Red G(N,S) de red ficticia 1
Fuente: Elaboración Propia

Tabla 3-1: Rutas asociadas a la red ficticia 1

Demanda	Rutas	Secciones de ruta	Tiempo de viaje	Tiempo de espera
T_{14}=500	h_1	s_2 - s_4	30	6
	h_2	s_1 - s_5	30	6

Fuente: Elaboración propia

El comportamiento esperado de los viajeros al escoger rutas de acuerdo a un S.U.E. sometido a una restricción de capacidad de los vehículos depende de cómo se considera el tiempo de espera al inicio de cada etapa de viaje. En esta tesis, se considera que el tiempo de espera adicional solo se activa cuando una de las secciones de línea de las secciones de ruta ha alcanzado su capacidad. En la red ficticia 1 los 500 pasajeros/hora pueden viajar sin generar problemas de congestión en ningún paradero, es decir, no existe ningún tiempo de espera adicional en las rutas ofrecidas por no poder abordar un vehículo. Por lo tanto, la solución de equilibrio responderá a una elección Logit entre ambas rutas y será de 250 viajeros eligiendo la primera ruta y transbordando en el nodo 2 y 250 viajeros eligiendo la segunda ruta y transbordando en el nodo 3, ya que los costos totales de viajes asociados a cada ruta son iguales y por consecuencia las probabilidades de elegir cada ruta también son iguales.

Es interesante contrastar esta solución con la que arrojaría el modelo de De Cea y Fernández (1993), en que se modela una función BPR creciente con los flujos. Para ese caso como los pasajeros que suben aguas arriba aumentan el tiempo de espera de los viajeros aguas abajo aunque exista capacidad disponible, la solución de equilibrio esperada sería que todos eligieran la ruta 1 y transbordaran en 2, pues de este modo se subirían a vehículos vacíos en la segunda etapa de su viaje. En cambio, si transbordaran en 3, el modelo con BPR asociaría un tiempo de espera mayor que los que se transbordan en 2, pues abordarían los buses con pasajeros. Así, podrían cambiar unilateralmente su elección de ruta hacia la primera y percibir un costo menor.

Los resultados obtenidos mediante el modelo iterativo propuesto fueron consistentes con la solución de equilibrio esperada (ver Tabla 3-2). Si bien, en este caso la capacidad predictiva del modelo es correcta, no es necesario abordar la etapa del Modelo de Asignación asociada a la restricción de capacidad, ya que al no existir problemas de congestión el algoritmo realiza una iteración de la etapa de tiempo adicional sin obtener tiempos de esperas asociados a no poder abordar el primer bus que pasa.

Tabla 3-2: Resultados de modelo propuesto en red ficticia 1

Demanda	Rutas	Secciones de ruta	Tiempos de viaje	Tiempos de espera	Tiempo de espera adicional	Flujo
T_{14}=500	h_1	s_2 - s_4	30	6	0	250
	h_2	s_1 - s_5	30	6	0	250

Fuente: Elaboración propia

La opción más simple de congestionar más esta red, respetando la factibilidad del problema (que la red sea capaz de transportar a todos los viajeros) es generar una demanda adicional de 500 pasajeros/hora del nodo 2 al nodo 3 (ver Figura 3-9). Sin embargo, esta configuración no provoca problemas de congestión solo lleva al sistema al límite. La nueva demanda entre 2 y 3 usaría la sección de ruta 3, repartiendo 250 viajan por la línea 1 y 250 por la línea 2 (ver Tabla 3-3 y Tabla 3-4) dado que ambas líneas cuentan con su capacidad disponible.

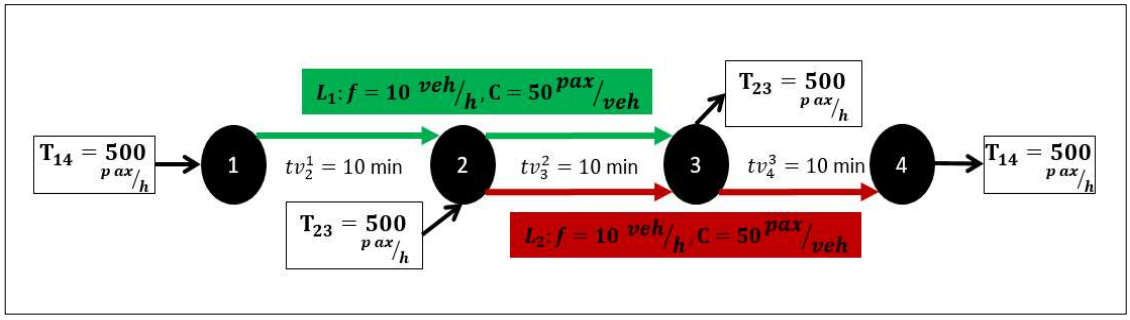


Figura 3-9: Red ficticia 1 modificada para incluir demanda entre 2 y 3
Fuente: Elaboración propia

Tabla 3-3: Resultados de modelo propuesto en red ficticia 1 modificada

Demanda	Rutas	Secciones de ruta	Tiempos de viaje	Tiempos de espera	Tiempo de espera adicional	Flujo
T_{14}=500	h_1	s_2 - s_4	30	6	0	250
	h_2	s_1 - s_5	30	6	0	250
T_{23}=500	h_3	s_3	10	1,5	0	500

Fuente: Elaboración propia

Tabla 3-4: Resultados del modelo propuesto en los segmentos de línea para red ficticia 1 modificada

Rutas	Flujo	Tiempos de viaje	Tiempos de espera	x_1^{12}	x_1^{23}	x_2^{23}	x_2^{34}
h_1	250	30	6	250	-	250	250
h_2	250	30	6	250	250	-	250
h_3	500	10	1,5	-	250	250	-
Total de flujo en segmentos de línea				**500**	**500**	**500**	**500**

Fuente: Elaboración propia

En vista de esta situación, para poder probar el algoritmo activando la restricción de capacidad se opta por considerar una demanda de un par origen destino mayor que la capacidad de la o las rutas disponibles más rápidas. Por lo tanto, la red ficticia 2 a probar debiese presentar una mayor demanda y para no violar el supuesto de que la oferta del sistema tiene que ser capaz de satisfacer esa demanda se agregaran más líneas, pero con menores niveles de servicios.

b) Red ficticia 2

La segunda red se construye a partir de la primera, agregando dos nuevas rutas alternativas, una para los viajeros que van desde el nodo 1 a 4 y otra para los que van desde el nodo 2 a 3. La primera ruta corresponde a un servicio que tiene un mayor tiempo de viaje que las rutas actuales, pero igual tiempo de espera total del viaje de 1 a 4 usando las líneas 1 y 2. La segunda ruta que conecta los nodos 2 y 3, es una ruta con mayor tiempo de viaje, mayor tiempo de espera y con un transbordo. Además, se aumenta la demanda de 2 a 3 a 600 pasajeros/hora (ver Figura 3-10), con el objetivo de generar problemas de congestión en el arco más demandado por los usuarios del sistema.

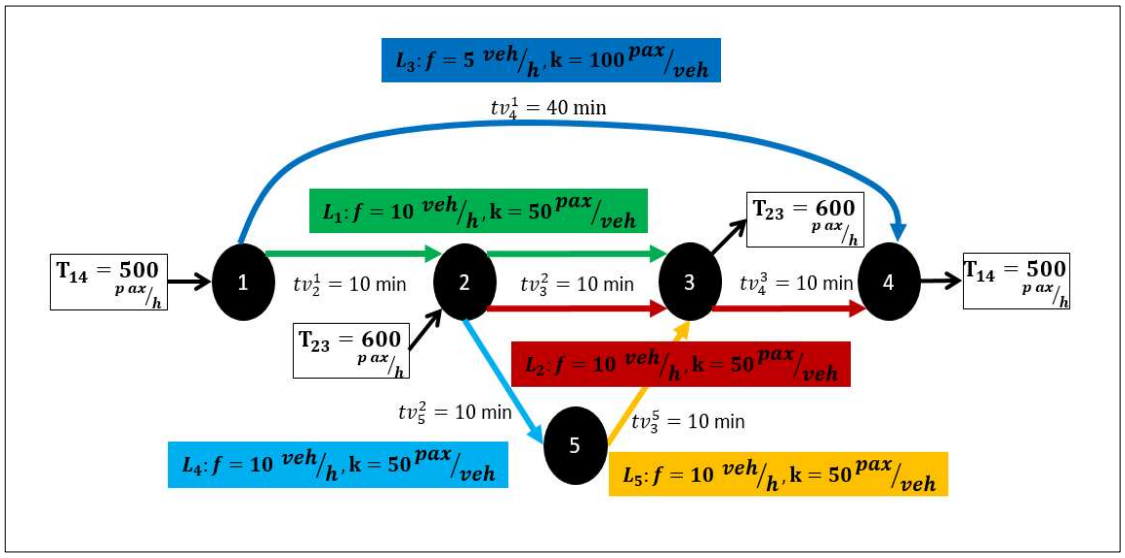


Figura 3-10: Red ficticia 2
Fuente: Elaboración propia

La nueva red G(N,S*) a esta red se muestra en la Figura 3-11 y los niveles de servicio asociados a cada ruta se muestran en la Tabla 3-5.

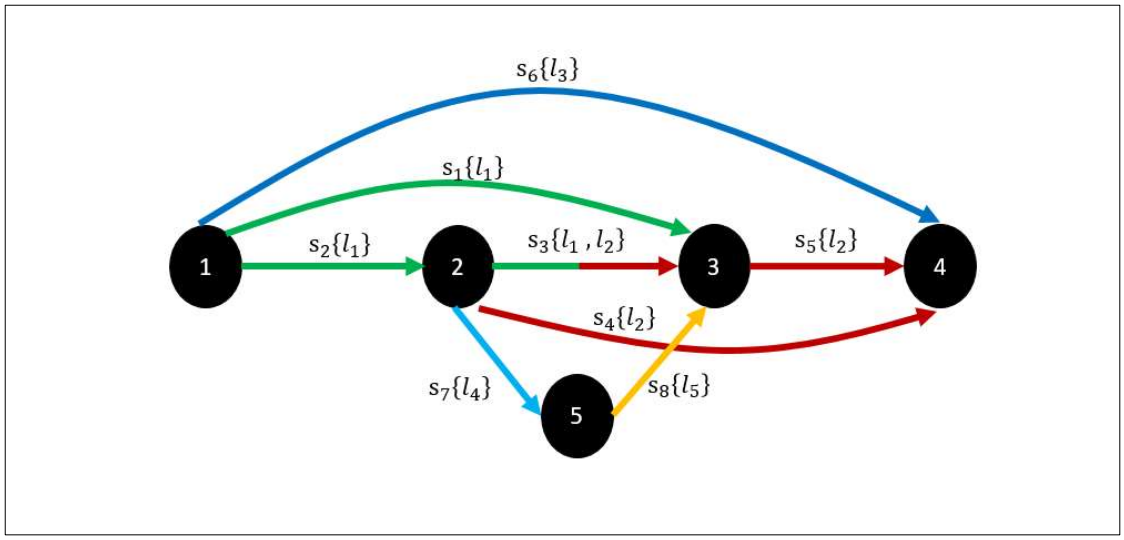


Figura 3-11: Red G(N,S) de la red ficticia 2
Fuente: Elaboración propia

Tabla 3-5: Rutas asociadas a la red ficticia 2

Demanda	Rutas	Secciones de ruta	Tiempos de viaje	Tiempos de espera
T_{14}=500	h_1	s_2 - s_4	30	6
	h_2	s_1 - s_5	30	6
	h_4	s_6	40	6
T_{23}=600	h_3	s_3	10	1,5
	h_5	s_7 - s_8	20	6

Fuente: Elaboración propia

Se descartan las rutas que por construcción siempre tendrán un costo total de viajes mayor que el costo de las otras rutas (s_2 - s_3 - s_5 y s_2 - s_7 - s_8 - s_5) y que por lo tanto no son razonables como rutas alternativas al usuario.

El comportamiento esperado para la red ficticia 2, es que todos quieran viajar por las rutas de menor costo, para los que van de 1 a 4, las rutas de menor costo son la primera y segunda ruta, y para los viajeros de 2 a 3, la tercera ruta. Sin embargo, la carga asociada a

esta asignación generaría resultados que no respetan la restricción de capacidad en los segmentos de línea que comienzan en el nodo 2 y terminan en el nodo 3. En efecto, para la carga de la línea 1 (x_1^{23}) y de la línea (x_2^{23}), se tendrían 550 pasajeros/hora y las capacidades asociadas a estos segmentos de línea serían de 500 pasajeros/hora (ver Tabla 3-6).

Tabla 3-6: Resultados esperados sin restricción de capacidad para red ficticia 2

Rutas	Flujo	Tiempos de viaje	Tiempos de espera	x_1^{12}	x_1^{23}	x_2^{23}	x_2^{34}	x_3^{14}	x_4^{25}	x_5^{53}
$\mathbf{h_1}$	250	30	6	250	-	250	250	-	-	-
$\mathbf{h_2}$	250	30	6	250	250	-	250	-	-	-
$\mathbf{h_4}$	0	40	6	-	-	-	-	0	-	-
$\mathbf{h_3}$	600	10	1,5	-	300	300	-	-	-	-
$\mathbf{h_5}$	0	20	6	-	-	-	-	-	0	0
Total de flujo en segmentos de líneas				**500**	**550**	**550**	**500**	**0**	**0**	**0**
Capacidades en segmentos de línea				**500**	**500**	**500**	**500**	**500**	**500**	**500**

Fuente: Elaboración propia

En este caso es razonable suponer que la restricción de capacidad se activa en los segmentos de línea entre 2 y 3 y no debería afectar a quienes viajan de 1 a 4 en términos de tiempos de espera, ya que tanto la ruta 1 y ruta 2, que poseen los menores tiempos de viajes poseen capacidad suficiente para llevar a los 500 pax/hora. Sin embargo, los usuarios seguramente modificaran su comportamiento, privilegiando transbordar en 3 (ruta 2) así se evitan acceder a la línea 2 en el nodo 2 junto a una demanda que excede la capacidad. Es justamente en el nodo 2 donde se debiese observar el mayor cambio de comportamiento, pues quiénes acceden allí a la red exceden la capacidad de la ruta predilecta para esos viajes.

Este problema, el modelo lo resuelve asociando un tiempo de espera adicional a los segmentos con problemas de capacidad. En este caso, las rutas en que sus etapas parten

usando la línea 1 o línea 2 en el nodo 2, debiesen tener asociado un tiempo de espera adicional por no poder abordar al primer vehículo que están esperando. Los usuarios que pasan por estos segmentos de líneas, pero que se suben al bus antes o después no deberían percibir un tiempo de espera adicional, ya que ni aguas arriba ni aguas abajo se generan problemas de congestión.

Teniendo en cuenta la situación descrita, el comportamiento esperado es que los viajeros de 1 a 4 opten por la segunda ruta, así el transbordo se realiza en el nodo 3, donde no existen problemas de congestión y por lo tanto no perciben un tiempo de espera adicional al transbordar. En cuanto, a las personas que van de 2 a 3, lo que se esperaría es que 500 opten por viajar por la línea 2 que es la más rápida, pues la línea 1 ya no se encontraría disponible porque llega a capacidad al nodo 2 y nadie baja de ella. Los 100 pasajeros restantes elegirían la ruta alternativa que si bien es más larga, ofrece capacidad disponible para poder completar su viaje.

La asignación esperada se muestra en la Tabla 3-7. En la tabla se presenta también el resultado esperado por el modelo de asignación para el tiempo de espera total de cada ruta. En este caso, se observa que solo en las rutas congestionadas, primera y tercera, los usuarios tienen un tiempo de espera percibido mayor que en el caso sin congestión, lo que explica la asignación de viajes a rutas alternativas en cada caso (ver Tabla 3-6).

Tabla 3-7: Resultados esperados para red ficticia 2

Rutas	Flujo	Tiempos de viaje	Tiempos de espera	x_1^{12}	x_1^{23}	x_2^{23}	x_2^{34}	x_3^{14}	x_4^{25}	x_5^{53}
$\mathbf{h_1}$	0	30	16.6	0	-	0	0	-	-	-
$\mathbf{h_2}$	500	30	6	500	500	-	500	-	-	-
$\mathbf{h_4}$	0	40	6	-	-	-	-	0	-	-
$\mathbf{h_3}$	500	10	14.4	-	0	500	-	-	-	-
$\mathbf{h_5}$	100	20	6	-	-	-	-	-	100	100
Total de flujo en segmentos de líneas				**500**	**500**	**500**	**500**	**0**	**100**	**100**
Capacidades en segmentos de líneas				**500**	**500**	**500**	**500**	**500**	**500**	**500**

Fuente: Elaboración propia

Sin embargo, esta asignación no es la que arroja el modelo propuesto por Lam et al. (2002). Al buscar una asignación factible este modelo asume que usuarios aguas arriba cambiarán su comportamiento para permitir un mejor uso de la red por parte de los usuarios aguas abajo. En particular, la asignación predicha por Lam et al. (2002) se presenta en la Tabla 3-8, en que usuarios optan por la ruta 1 para evitar que usuarios que van desde 2 a 3 usen la ruta 5. Este comportamiento parece poco plausible.

Así ese modelo arroja tiempos de espera adicionales para la primera y segunda ruta iguales (ver Tabla 3-8). Esto ocurre porque el tiempo de espera adicional de cada ruta, que se obtiene a partir de las condiciones de equilibrio del problema, es igual a la suma de los tiempos de espera asociados a cada segmento de línea de la ruta y no a la suma de los tiempos de espera asociados solo al inicio de cada etapa.

Tabla 3-8: Resultados obtenidos con modelo propuesto para red ficticia 2

Demanda	Rutas	Secciones de ruta	Tiempos de viaje	Tiempos de espera fijo	Tiempo de espera adicional	Flujo
T_{14}=500	h_1	s_2 - s_4	30	6	9.3	200.4
	h_2	s_1 - s_5	30	6	9.3	200.4
	h_4	s_6	40	6	0	99.2
T_{23}=600	h_3	s_3	10	1.5	4.6	600.0
	h_5	s_7 - s_8	20	6	0	0.0

Fuente: Elaboración propia

Esta formulación de Lam et al. (2002) para el tiempo de espera adicional, se basa en un Modelo de Asignación Estocástica propuesto para transporte privado por Bell (1995) en el que uno de los supuestos del problema es que la red se encuentra en estado estacionario y los arcos se comportan *store-and-forward*. Este comportamiento de envío, como dice el nombre, implica que los viajeros que utilizan un arco son retenidos en cada arco antes de poder ser enviados tal y como ocurre con los vehículos en cada calle. Cada vehículo solo avanza al siguiente arco cuando existe capacidad disponible.

Este comportamiento dista mucho del que se obtiene en transporte público en que cada viajero sube a un servicio cuando existe capacidad disponible en el paradero de subida, poder o no abordar un vehículo no depende de los usuarios que esperan en paraderos posteriores del recorrido. Por ejemplo, si un viajero toma un servicio al principio del recorrido, su tiempo de espera no se verá afectado por cuántas personas están esperando en el segundo paradero, aunque su destino sea el tercer paradero o el último paradero del recorrido, cuando la persona subió al recorrido su tiempo de espera terminó y la congestión aguas abajo, podría implicar mayores tiempos de detención, aumentando el tiempo de viaje percibido, pero no el tiempo de espera.

Al obtener el resultado de la Tabla 3-8, se volvió a revisar la formulación del problema y se detectó que la proposición que dice que el problema es equivalente a un Modelo de Asignación S.U.E. en transporte público con restricción de capacidad de los vehículos, si y solo si los multiplicadores de Lagrange asociados a la restricción de capacidad, m_a, son iguales al inverso aditivo de las demoras asociadas a la solución de equilibrio, $-d_a$, no es válida para transporte público. Esto se debe a que la proposición define el tiempo de espera adicional asociado a cada ruta en transporte público como la suma de los tiempos adicionales de todos los segmentos de línea por los que pasa la ruta (ver ecuación (3.8), la notación se encuentra explicada en el capítulo 2) y no como la suma de los tiempos de espera adicionales de cada etapa o sección de ruta.

$$d_r^w = -m_r^w = -\sum_{sr \in r} m_{sr} = -\sum_{sr \in r}\sum_{a \epsilon A} \bar{\gamma}_{a-sr}\, m_a = -\sum_{a \in A} m_a \sum_{sr \epsilon r} \bar{\gamma}_{a-sr} \qquad (3.8)$$

En redes de transporte público las esperas se perciben solo al principio de cada etapa, es decir el tiempo de espera adicional debiese ser como el que se define en la ecuación (3.9).

$$d_r^w = -m_r^w = -\sum_{sr \in r} m_{sr} = -\sum_{s \in r}\sum_{a \epsilon A\ /o(a)=o(sr)} \theta \bar{\gamma}_{as}\, m_a \qquad (3.9)$$

Donde el tiempo de espera percibido en cada etapa depende de la congestión percibida al subir al primer segmento de línea ($o(a) = o(sr)$) y no depende de la congestión en segmentos posteriores. Al hacer este cambio en el algoritmo de Lam et al. (2002) se obtuvieron los resultados de la Tabla 3-9.

Tabla 3-9: Resultados del modelo propuesto con nueva formulación del tiempo de espera adicional

Demanda	Rutas	Secciones de ruta	Tiempos de viaje	Tiempos de espera fijo	Tiempo de espera adicional	Flujo
T_{14}=500	h_1	s_2 - s_4	30	6	10.6	0.0
	h_2	s_1 - s_5	30	6	0	500.0
	h_4	s_6	40	6	0	0.0
T_{23}=600	h_3	s_3	10	1.5	12.7	513.6
	h_5	s_7 - s_8	20	6	0	86.4

Fuente: Elaboración propia

Al analizar los resultados con la nueva formulación del tiempo de espera adicional, tampoco se obtienen los resultados esperados presentados en la Tabla 3-7, si bien ahora el tiempo de espera adicional considerado en cada iteración de la segunda etapa del Modelo Heurístico de Asignación de Viajes es el esperado, las frecuencias efectivas asociadas a este tiempo y que influyen en la primera etapa del modelo no son correctas, generando resultados que no respetan la restricción de capacidad, ver Tabla 3-10.

Tabla 3-10: Resultados en los segmentos de líneas del modelo propuesto con nueva formulación del tiempo de espera adicional

Rutas	Flujo	Tiempos de viaje	Tiempos de espera	x_1^{12}	x_1^{23}	x_2^{23}	x_2^{34}	x_3^{14}	x_4^{25}	x_5^{53}
$\mathbf{h_1}$	0.0	30	16.8	0.0	-	0.0	0.0	-	-	-
$\mathbf{h_2}$	500.0	30	6	500.0	500.0	-	500.0	-	-	-
$\mathbf{h_4}$	0	40	6	-	-	-	-	0	-	-
$\mathbf{h_3}$	513.6	10	14.21	-	0	500.0	13.6	-	-	-
$\mathbf{h_5}$	86.4	20	6	-	-	-	-	-	86.4	86.4
Total de flujo en segmentos de líneas				**500.0**	**500.0**	**500.0**	**513.6**	**0**	**86.4**	**86.4**
Capacidades en segmentos de línea				**500**	**500**	**500**	**500**	**500**	**500**	**500**

Fuente: Elaboración propia

Esto ocurre, porque la formulación de la función de las frecuencias efectivas presentada en la ecuación (3.7) es asintótica a 0, esto implica que, aunque la línea asociada a una sección de línea en particular no tenga capacidad disponible en el nodo i donde se inicia la sección de línea l, si el costo de viaje de la sección de línea le permite al usuario minimizar el tiempo de viaje esperado de la sección de ruta, esta sección de línea se seguirá considerando atractiva. En el caso del ejemplo en cuestión, la ruta 3 sigue incluyendo dentro de segunda etapa la línea 1 como atractiva y si bien su frecuencia efectiva es pequeña, 0,03 veh/hora, esto no impide que se le asignen viajeros al segmento de 2 a 3 de la línea1, generando problemas de capacidad (ver Tabla 3-10).

Además del inconveniente de tener una función asintótica a 0, la función de frecuencias efectivas tiene el problema de ser decreciente con el número de iteraciones. Esto implica que independientemente del conjunto de líneas atractivas que se considere en cada iteración, las frecuencias efectivas se mantendrán o disminuirán, es decir, nunca volverán a mejorar. El problema de esta situación es cuando la asignación de flujo a una sección de ruta pasa de una iteración con problemas de congestión a una situación sin problemas

de congestión, es en este caso en donde las frecuencias efectivas mayores a las frecuencias nominales pierden sentido.

Dicho estos dos inconvenientes en la formulación se propone redefinir la frecuencia efectiva como se presenta en la ecuación (3.10).

$$f_{sl}^{n+1} = \begin{cases} \dfrac{1}{\left(\dfrac{1}{f_{sl}^{n}} + \dfrac{d_{sl}^{n}}{\alpha_l}\right)}, & \dfrac{v_a - \sum_{\forall l \,/\, d(a)=d(sl)} v_{sl}\beta_{asl}}{C_a * f_l} < 1 \\ & con\ a/\ l(a) = l(sl)\ \wedge\ d(a) = o(sl)\ \wedge\ o(l) = o(sl) \\ 0, & en\ otro\ caso \end{cases} \tag{3.10}$$

Donde:

β_{asl}: Indica si el segmento de línea a se relaciona con la sección de línea sl.

C_a: Es la capacidad del segmento de línea a.

L: Es el conjunto de líneas de la red (l).

O: Es el conjunto de nodos orígenes de la red (o).

D: Es el conjunto de nodos destinos de la red (d).

Para el primer caso, en donde existe capacidad disponible para los pasajeros que esperan en el inicio de la sección de línea sl, se define la frecuencia efectiva como una función dependiente de las demoras adicionales de cada sección de línea resultantes de la iteración anterior. La frecuencia nominal f_{sl}^{n} y la constante de regularidad de intervalos α_l, son considerados parámetros de la función. De esta forma, si el resultado de una iteración de la asignación de flujos no posee problemas de congestión, la frecuencia efectiva considerada en la siguiente iteración para construir la red será igual a la frecuencia nominal.

Además de este cambio, cada vez que al conjunto de líneas atractivas se agrega una nueva línea atractiva o más, las frecuencias efectivas se definen iguales a las frecuencias nominales, ya que no realizar este cambio puede distorsionar la repartición de flujos y/o el tiempo de espera. Por ejemplo, si una sección de ruta tiene disponible tres secciones de línea y al resolver la primera iteración del modelo solo son atractivas dos líneas, el modelo asignará a estas dos líneas el flujo de las rutas que utilicen esta sección de ruta. Si esta

configuración genera problemas de congestión, la siguiente iteración incluirá la tercera línea y la repartición de los flujos de las dos primeras líneas será proporcional a su frecuencia efectiva y la de la tercera a la de su frecuencia nominal. Pero, qué pasaría si al agregar la tercera línea los problemas de congestión desaparecen, ¿sería correcto castigar las líneas más frecuentes y más rápidas? La respuesta es no, porque si las personas esperan tres líneas con capacidad disponible, lo razonable es que las líneas más frecuentes tengan mayor flujo que las menos frecuentes. Es por esto, que cada vez que cambia el conjunto de líneas atractivas de una iteración a otra, las frecuencias efectivas se redefinen iguales a las frecuencias nominales. Estos cambios, generan los resultados de la Tabla 3-11.|

Para el segundo caso, en el que la línea al pasar por el nodo de inicio de una etapa no tiene capacidad disponible la frecuencia efectiva será igual a 0, es decir, no se podrá considerar en el conjunto de líneas atractivas.

Tabla 3-11: Resultados del Modelo Heurístico de Viajes final

Demanda	Rutas	Secciones de ruta	Tiempos de viaje	Tiempos de espera fijo	Tiempo de espera adicional	Flujo
T_{14}=500	h_1	s_2 - s_4	30	6	20.6	0.0
	h_2	s_1 - s_5	30	6	0	500.0
	h_4	s_6	40	6	0	0.0
T_{23}=600	h_3	s_3	10	3	11.4	500.0
	h_5	s_7 - s_8	20	6	0	100.0

Fuente: Elaboración propia.

Se observa, que los resultados obtenidos son iguales a los esperados y con esta prueba se cierra el Modelo Heurístico de Asignación de Viajes, que respeta el S.U.E. y la restricción de capacidad de los vehículos.

3.2.2 Desempeño del modelo para la red ficticia 2

Finalizadas las pruebas de la capacidad predictiva se analiza el desempeño del algoritmo de solución descrito en el capítulo 2 para la red ficticia 2. Este algoritmo si bien resuelve el problema de manera correcta, realiza muchas iteraciones para obtener los resultados de la asignación, en la Tabla 3-12 se muestra la cantidad de iteraciones del Modelo de Heurístico de Asignación de Viajes y las iteraciones realizadas para obtener el tiempo de espera adicional, el máximo de iteraciones se fijó en 1000. Se observa que en las iteraciones 2 y 3 del modelo, alcanza el máximo de iteraciones de la etapa que calcula el tiempo de espera adicional, lo que significa que en estas dos iteraciones el tiempo de espera adicional no es suficiente para que la asignación respete la restricción de capacidad, sin embargo esto no perjudica la convergencia total del modelo, porque solo hace que la siguiente iteración parta con un tiempo de espera adicional un poco menor que el esperado y este se corrige en las iteraciones posteriores del modelo general.

Tabla 3-12: Iteraciones del Modelo Heurístico de Asignación de Viajes para la red ficticia 2

Iteración del modelo	Iteraciones de Etapa: Tiempo de espera adicional
1	205
2	1000
3	1000
4	93
5	2

Fuente: Elaboración propia

A pesar de que el tiempo de ejecución para alcanzar la convergencia del modelo es bastante pequeño, 2.7 segundos, la cantidad de iteraciones que realiza puede ser un número que aumente significativamente los tiempos de ejecución en redes de gran tamaño, es por esto que se decide reevaluar las funciones de M_a y L_w.

Lo primero que se identifica es que L_w queda únicamente definido por M_a. En efecto, mediante la restricción de demanda y la ecuación (2.45) del problema se obtiene una expresión para $e^{\theta l_w}$, tal y como se muestra a continuación.

$$\sum_{r \epsilon R_w} h_r^w = T_w = \sum_{r \epsilon R_w} e^{-\theta(t_r^w + u_r^w)} \cdot e^{\theta m_r^w} \cdot e^{\theta l_w} \qquad \forall\, r \in R_w, w \epsilon W \tag{3.11}$$

$$T_w = e^{\theta l_w} \sum_{r \epsilon R_w} e^{-\theta(t_r^w + u_r^w)} \cdot e^{\theta m_r^w} \cdot \qquad \forall\, r \in R_w, w \epsilon W \tag{3.12}$$

$$e^{\theta l_w} = \frac{T_w}{\sum_{r \epsilon R_w} e^{-\theta(t_r^w + u_r^w)} \cdot e^{\theta m_r^w}} \qquad \forall\, r \in R_w, w \epsilon W \tag{3.13}$$

Reemplazando la ecuación (3.13) en (2.44), se obtiene:

$$h_r^w = e^{-\theta(t_r^w + u_r^w)} \cdot e^{\theta m_r^w} \cdot \frac{T_w}{\sum_{r \epsilon R_w} e^{-\theta(t_r^w + u_r^w)} \cdot e^{\theta m_r^w}} \qquad \forall\, r \in R_w, w \epsilon W \tag{3.14}$$

Por lo tanto, L_w no es el factor que permita mejorar el algoritmo de solución ya que cualquiera sea el valor de M_a, L_w queda totalmente determinado por este, incluso si reemplaza esta formulación en el algoritmo de solución, las iteraciones obtenidas son las mismas que presentadas en la Tabla 3-12.

Dicho esto, se analizó la expresión de M_a (ecuación 2.49) de manera de obtener una función que en cada iteración permita al algoritmo avanzar más rápido respetando las condiciones de la ecuación (2.55). A continuación, se muestra un ejemplo para un segmento de línea a cualquiera en que se muestra cómo se comporta la expresión de la ecuación (2.54) en cada iteración:

- $n = 1 \rightarrow\ M_a^1 = 1$
- $n = 2 \rightarrow\ M_a^2 = \min\left[1, M_a^1 \cdot \left(\frac{c_a}{f_a}\right)^{n=1}\right]$
- $n = 3 \rightarrow\ M_a^3 = \min\left[1, M_a^2 \cdot \left(\frac{c_a}{f_a}\right)^{n=2}\right]$
- $\vdots$

- $n = k \rightarrow M_a^k = \min\left[1, M_a^{k-1} \cdot \left(\frac{c_a}{f_a}\right)^{n=k-1}\right]$

Es decir, cuando $\left(\frac{c_a}{f_a}\right)^n$ decrece, M_a^n decrece y si $\left(\frac{c_a}{f_a}\right)^n$ crece M_a^n crece, siempre y cuando no sea mayor a 1, ya que como M_a^n representa $e^{\theta m_a} = e^{-\theta d_a}$ y los tiempos de espera adicionales asociados a los segmentos de línea a solo pueden ser mayores o iguales a 0, M_a^n solo puede tomar valores entre $(0,1]$. Teniendo esto en cuenta, una forma de hacer que M_a^n aumente o disminuya más rápido es considerar $\frac{c_a}{f_a}$ como una potencia, tal y como se muestra en la ecuación (3.15).

$$M_a^{n+1} = min\ \{1, \left\{\left(\frac{c_a}{f_a}\right)^q\right\}^n \cdot M_a^n\} \tag{3.15}$$

Para obtener el valor de q, se probaron los siguientes valores enteros $q = \{1,2,3,4,5,6\}$, tal y como se muestra en la Tabla 3-13, el objetivo de esta prueba era determinar el exponente q de la ecuación (3.15) que permitiera disminuir el número de iteraciones del sub-problema del tiempo de espera, sin perjudicar el número de iteraciones del modelo general. Esta prueba arrojó que $q = 4$, era el exponente más adecuado que permitía reducir en más iteraciones el sub-problema sin aumentar las iteraciones del modelo.

Tabla 3-13: Iteraciones del modelo para distintos M_a

Iteración del modelo	**Iteraciones de Etapa: Tiempo de espera adicional**					
	x=1	x=2	x=3	x=4	x=5	x=6
1	205	106	72	55	44	37
2	1000	924	638	490	399	337
3	1000	1000	1000	1000	1000	1000
4	93	48	32	24	1000	1000
5	2	2	2	2	19	15
6	-	-	-	-	2	2

Fuente: Elaboración propia

3.3 Modelo de Asignación Heurístico de Viajes

Con la evaluación de la capacidad predictiva y desempeño en las redes ficticias, finalmente se proponen los siguientes ajustes del Modelo Heurístico de Asignación de Viajes.

Tabla 3-14 Resumen de ajustes en la formulación del Modelo de Asignación de Viajes Heurístico

Etapas		**Ajustes**
1.	**Red de asignación**	
1.1.	Frecuencias efectivas	Formulación (3.10) con actualización de frecuencias efectivas a nominales cuando aumenta el conjunto de líneas atractivas a considerar.
1.2.	Problema de líneas comunes	-
2.	**Tiempo de espera adicional**	
2.1.	Asignación Logit	-
2.2.	Demoras adicionales por no poder abordar el primer vehículo.	Formulación (3.9) con el objetivo de considerar las demoras adicionales solo cuando los pasajeros abordan el vehículo.
2.3	Convergencia	Aumento del exponente q a 4 del factor $\left(\frac{c_a}{f_a}\right)^q$ con el objetivo de aumentar la convergencia del algoritmo de solución.

Fuente: Elaboración Propia

4 ANÁLISIS DE RESULTADOS

En este capítulo se evaluará la capacidad predictiva del Modelo Heurístico de Asignación de Viajes, con los ajustes realizados en el capítulo 3, en una red real altamente congestionada como la red de Metro de Santiago, Año 2013. Se eligió este corte temporal, porque se cuenta con datos para la oferta, la demanda y un modelo de asignación calibrado para año 2012, proporcionado por Metro de Santiago (2012). La caracterización de la red se presentara en la sección 4.1.

La evaluación de la capacidad predictiva presentada en la sección 4.2. se realizará en dos etapas. Primero se evaluará la capacidad predictiva de la primera etapa del modelo propuesto en la sección 4.2.1, Red de Asignación, comparando la formulación de conjunto de líneas propuesta en el capítulo 3 (ecuación 3.10 y la actualización a frecuencias nominales cuando se agrega una línea al conjunto de líneas atractivas) y el caso en que no se considera una frecuencia efectiva para la creación de líneas comunes, se muestra la importancia de considerar la restricción de capacidad en la creación de la red.

Luego, en la sección 4.2.2 se evaluará la formulación del tiempo de espera adicional obtenido en la segunda etapa del modelo (descrito en el capítulo 3, ecuación 3.9), para esto se contrastarán los resultados con tres modelos:

- Asignación tipo Logit, calibrada específicamente para metro (Raveau et al., 2011) en el que no se considera explícitamente la restricción de capacidad, pero sí se incluye una penalidad en la función de utilidad asociada a no poder subir a un vehículo, y que depende de la probabilidad de no subir.
- El segundo modelo se basa en el modelo de asignación de rutas de transporte público estocástica mencionado en el párrafo anterior (Raveau et al., 2011), pero reemplazando la penalidad asociada a la probabilidad de no subir por la formulación de De Cea y Fernández (1993) que incluye un tiempo de espera adicional percibido por los usuarios producto de que no pueden subir al primer vehículo que pasa.

- El tercer modelo corresponde al modelo de esta tesis, donde se obtienen tiempos de espera adicionales a partir de los resultados de asignación que permiten que algunos usuarios busquen alternativas a las rutas que ellos tomarían en ausencia de estas esperas adicionales, y que permiten que ningún flujo en la red exceda a su capacidad, tal y como se describió en el capítulo 3, con las modificaciones descritas en la formulación (3.9).

A continuación, se describe la oferta y demanda considerada en el escenario propuesto para analizar los resultados del modelo desarrollado en esta tesis.

4.1 Caracterización de Red Metro Santiago

4.1.1 Red Metro 2013

La red considerada corresponde a la red de operación de METRO del año 2013 desde 06:00 horas hasta 09:00 horas. En esta red se consideran cinco líneas de metro (líneas 1, 2, 4, 4A y 5) , en la que existen dos tipos de operación:

- Servicios expresos tipo *skip-stop*: Estos corresponden a recorridos identificados con los colores rojo y verde, en que cada tipo de tren se detienen solo en algunas estaciones (ver
-
- Figura *4-1*). Este tipo de operación se presenta solo en las líneas 2, 4 y 5.
- Servicios normales: Estos corresponden a la operación estándar de una línea en que cada tren recorre la línea de metro completa deteniéndose en todas las estaciones para que los pasajeros puedan abordar o descender del tren. Las líneas que presentan este tipo de operación corresponden a Línea 1 y Línea 4A.

Figura 4-1 Red de metro 2013
Fuente: www.metro.cl

Además de la operación basada en las detenciones de los servicios, existen bucles de operación, que corresponde a servicios que en vez de recorrer la línea completa recorren un subconjunto consecutivo de estaciones con el objetivo de reducir el tiempo

de ciclo de cada recorrido y así aumentar la capacidad de transporte que ofrece metro. En la Red de Metro 2013, se opera con dos bucles para la línea 1 y dos bucles para la línea 5.

La combinación entre la operación de bucles y servicios expresos o normales dan origen a 26 líneas de operación, tal y como se muestra en la Tabla 4-1. Esta tabla incluye también el intervalo de operación que es el inverso multiplicativo de la frecuencia de trenes en cada servicio, y la capacidad de cada tren. La multiplicación de la capacidad de los trenes y la frecuencia de operación arroja la capacidad de transporte ofrecida por cada servicio.

Tabla 4-1: Características de líneas de la Red de Metro (2013)

Código de línea	Sentido	Línea de metro	Servicio	Bucle	Intervalo de operación [min]	Capacidad trenes [Pax/tren]
1	Ida	Línea 1	Normal	San Pablo - Manquehue	3.3	1469
2	Regreso	Línea 1	Normal	Manquehue - San Pablo	3.3	1469
3	Ida	Línea 1	Normal	Pajaritos - Los Dominicos	3.2	1469
4	Regreso	Línea 1	Normal	Los Dominicos- Pajaritos	3.2	1469
5	Ida	Línea 2	Rojo	-	5.7	1431
6	Regreso	Línea 2	Rojo	-	5.7	1431
7	Ida	Línea 2	Verde	-	5.9	1431
8	Regreso	Línea 2	Verde	-	5.9	1431
9	Ida	Línea 4	Rojo	-	4.9	1610
10	Regreso	Línea 4	Rojo	-	4.9	1610
11	Ida	Línea 4	Verde	-	4.9	1610
12	Regreso	Línea 4	Verde	-	4.9	1610
13	Ida	Línea 4A	-	-	4.3	805
14	Regreso	Línea 4A	-	-	4.3	805
15	Ida	Línea 5	Rojo	Quinta Normal - Vicente Valdés	19	1255
16	Regreso	Línea 5	Rojo	Vicente Valdés - Quinta Normal	19	1255
17	Ida	Línea 5	Verde	Quinta Normal - Vicente Valdés	19	1255
18	Regreso	Línea 5	Verde	Vicente Valdés - Quinta Normal	19	1255
19	Ida	Línea 5	Rojo	Pudahuel -Vicente Valdés	36.5	1255
20	Regreso	Línea 5	Rojo	Vicente Valdés - Quinta Normal	36.5	1255
21	Ida	Línea 5	Verde	Pudahuel -Vicente Valdés	36.5	1255
22	Regreso	Línea 5	Verde	Vicente Valdés - Quinta Normal	36.5	1255
23	Ida	Línea 5	Rojo	-	6.7	1255
24	Regreso	Línea 5	Rojo	-	6.7	1255
25	Ida	Línea 5	Verde	-	6.7	1255
26	Regreso	Línea 5	Verde	-	6.7	1255

Fuente: Elaboración propia con datos entregados por SECTRA (Año 2012).

La codificación de la red de metro se hizo a partir de los datos enviados por SECTRA para el año 2012. Para esto se solicitó la red de transporte público de Santiago codificada para ESTRAUS y se seleccionó solo los datos de la codificación de metro que identificaban las

líneas de metro con sus respectivas características operacionales. Luego se actualizó la oferta según datos enviados por Metro, principalmente se agregaron bucles de operación que no estaban operativos el año 2013 y se cambiaron las frecuencias de los recorridos acordes a los presentados en la Tabla 4-1.

4.1.2 Red G (N, S*) y rutas de metro

Para el conjunto de rutas alternativas entre cada par origen-destino y requerido por el Modelo de Asignación tipo Logit se usaron las rutas del estudio de Metro de Santiago (2012), tal y como se mencionó en la sección 1.2.3 la generación de rutas para este modelo es un dato exógeno.

La información de las rutas contemplaba líneas utilizadas y transbordos realizados, esta información estaba en un archivo Excel y se procesó para que fuera compatible con la estructura de datos utilizada.

La estructura de datos utilizada para construir la red G(N,S*) fue la misma que genera el utilitario ARTP1 de ESTRAUS, se selecciona esta estructura de datos porque tiene un manejo de información de la red eficiente y entrega la red G(N,S*) que se usa en esta tesis para modelar las rutas que eligen los usuarios.

Para que el modelo entendiera las rutas de viajes entre cada par origen-destino era necesario generar una secuencia de secciones de ruta según los resultados obtenidos del utilitario ARTP1. Para esta tarea se hizo una macro que hizo equivalente la información del estudio de Metro de Santiago (2012) con la estructura de datos asociada a la red G(N,S*) del metro.

Algunas de las consideraciones más importantes de la macro fueron:

1. Recodificación de cada ruta en etapas de metro, generando un nombre para la estación de subida y bajada de cada etapa. Además, era importante identificar el sentido de la línea porque en la estructura de datos seleccionada cada estación se

codifica con dos nodos de esta forma los tiempos asociados a los cambios de andén se pueden incluir en la modelación.

2. Definidas las etapas se hicieron diccionarios para transformar los nombres de estaciones a los nodos de input de la codificación de ARTP1.
3. Con las rutas definidas según los nodos equivalentes del input de la codificación se obtuvieron las secciones de ruta asociadas al output del utilitario ARTP1.
4. Definida la equivalencia de nombres se agregaron rutas entre pares orígenes-destinos que no se consideraron en el estudio de Metro de Santiago (2012), pero que si tenían viajes asociados según los datos recolectados de las validaciones Bips.
5. Además, se agregaron rutas que consideraban la operación de expresos del metro, ya que el estudio original no tenía esta consideración.

4.1.3 Demanda Metro Santiago 2013

La demanda utilizada para la modelación se obtuvo a partir de los datos Bips (validaciones de los usuarios para pagar su tarifa en el sistema integrado de transporte público de Santiago, Transantiago) de abril 2013. Estos datos son procesados según la metodología de Munizaga y Palma (2012) de manera tal de reconstruir las distintas etapas de viaje de los usuarios de la red Transantiago; la construcción de etapas tiene tres supuestos fuertes:

1. Solo se consideran las etapas del viaje que tienen asociada una validación de la tarjeta Bips en el inicio de la etapa. Este supuesto deja fuera a los evasores en los cálculos, pero para el caso de metro no debería ser determinante.
2. Al no contar con validaciones en las bajadas, el término de cada etapa se define como la parada de mayor beneficio para el usuario según su siguiente transacción. Para el caso del metro la estación de bajada considerada es la más cercana a su siguiente validación y para el término de etapa de los buses se considera el paradero que minimiza el costo generalizado de viaje.

3. Las etapas consecutivas de un viaje tienen una restricción tiempo máximo de 30 minutos entre el fin de una etapa y el inicio de la siguiente.

Los datos fueron procesados en PostgreSQL y se obtuvieron solo las etapas asociadas a metro. Para el caso de las etapas en metro no se tiene información si el viajero realizó un transbordo o no, solo se tiene información de la estación de subida y la estación de bajada, lo que es suficiente para generar la matriz origen-destino asociada a la red de metro.

Para cada par origen destino se hizo un proceso similar a la construcción de rutas, donde a cada estación se le asignó un nodo de la codificación. En este caso no fue necesario identificar el sentido del nodo, porque se definieron centroides por cada estación que conectaban los nodos asociados a cada sentido de la línea o en otras palabras el centroide conecta los viajes a los dos andenes de cada estación.

Para definir qué viajes se iban a considerar en la matriz de modelación se evaluaron las etapas de metro de las Bips dentro de una hora en el período punta de la mañana. Para esto, se evaluaron tres intervalos de tiempo: 7:00-8:00, 7:30-8:30 y 8:00-9:00, en cada intervalo se contaron los viajes que iniciaron dentro del periodo, viajes que su hora media de viaje estaba comprendida dentro del intervalo y viajes en que el término de viaje estaba dentro del intervalo (Ver Tabla 4-2).

Tabla 4-2: Demanda de Metro, Año 2013

Intervalo	Hora de inicio de viaje	Hora media de viaje	Hora de término de viaje
7:00 - 8:00	245.106	219.576	186.144
7:30 - 8:30	273.448	274.729	256.910
8:00 - 9:00	236.281	266.641	273.065

Fuente: Elaboración propia con datos de Metro, 2013

Se observa en la Tabla 4-2 que la mayor demanda se obtiene al considerar los viajes con hora media de viaje entre las 07:30 y 08:30. Se utiliza esta matriz de viajes porque interesa ver los resultados en una red altamente congestionada.

Definida la red de Metro de Santiago, su oferta de transporte y su demanda por viajes obtenidos a partir de las Bips 2013, se comienza con la implementación de los modelos comparativos para posteriormente realizar el análisis de resultados.

Es importante mencionar que se eligió el año 2013, porque se contaba con un estudio de Metro de Santiago (2012) en dónde se calibró un modelo de elección de rutas de transporte público tipo Logit, basado en la formulación de Raveau et al. (2011), pero la demanda correspondía al año 2013, por lo que se optó por dejar la oferta acorde a ese año y utilizar como Modelo de Asignación el presentado en el estudio de Metro de Santiago en diciembre de 2012.

4.2 Análisis de capacidad predictiva del modelo

En esta sección se evaluará la capacidad predictiva del modelo en las dos sub-etapas del Modelo Heurístico de Asignación de Viajes

4.2.1 Análisis de capacidad predictiva de etapa: Red de Asignación

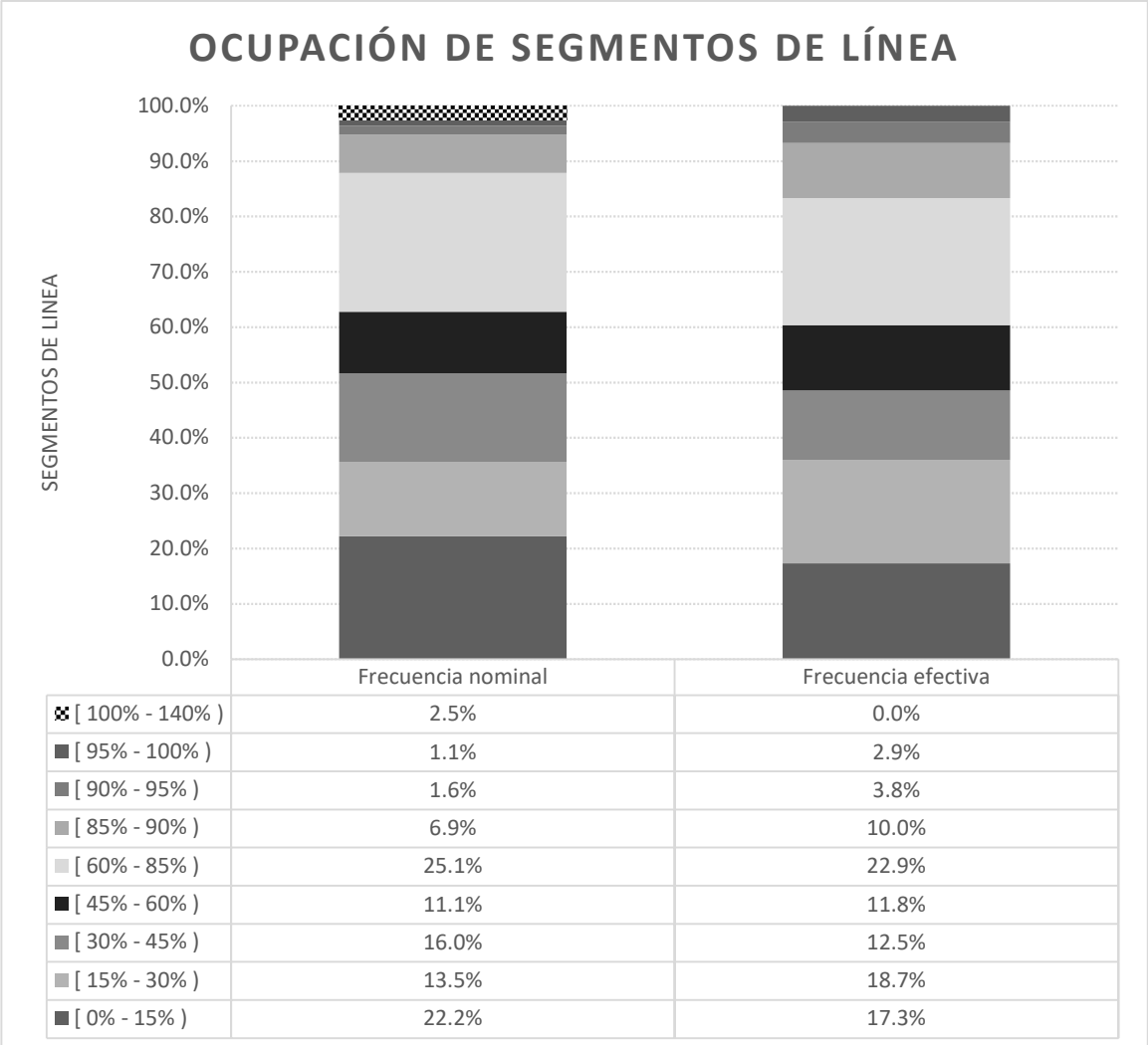
Para evaluar la capacidad predictiva de la primera etapa del modelo propuesto se muestran los resultados que se obtendrían sin considerar las frecuencias efectivas (frecuencia nominal) en la construcción de la red y se contrastarán con los resultados obtenidos con el modelo propuesto que considera frecuencias efectivas descrito en el capítulo 3 en la ecuación 3.10.

Cabe recordar que al generar el conjunto de líneas comunes para cada par de estaciones mediante las frecuencias nominales, se omite la congestión en las líneas que enfrentan los usuarios. Así, ante un alto nivel de demanda (como es el caso en cuestión) los flujos que se obtienen exceden la capacidad en algunos segmentos de línea. Esto se debe a que al

cargar la red, una línea congestionada seguirá recibiendo prioridad si su frecuencia nominal lo determina, aun cuando la frecuencia que vean los usuarios sea muy baja o nula.

En la Tabla 4-3 se presentan los resultados obtenidos de las dos formulaciones.

Tabla 4-3: % de segmentos de línea cuyo flujo presenta distintos niveles de saturación, medido como el porcentaje que representan el flujo de la capacidad del segmento.



	Frecuencia nominal	Frecuencia efectiva
[100% - 140%)	2.5%	0.0%
[95% - 100%)	1.1%	2.9%
[90% - 95%)	1.6%	3.8%
[85% - 90%)	6.9%	10.0%
[60% - 85%)	25.1%	22.9%
[45% - 60%)	11.1%	11.8%
[30% - 45%)	16.0%	12.5%
[15% - 30%)	13.5%	18.7%
[0% - 15%)	22.2%	17.3%

Fuente: Elaboración propia

Se observa que:

- Como era de esperar al considerar solo frecuencias nominales en la construcción de la red G (N, S*) no se respeta la restricción de capacidad
- La formulación de frecuencias efectivas representa de mejor manera los problemas de congestión, ya que se obtienen más arcos con ocupaciones mayores o iguales a 85% y menos con ocupaciones menores al 15%.

4.2.2 Análisis de capacidad predictiva de etapa: Tiempo de espera adicional

Como el objetivo de esta sección es comparar la capacidad predictiva de la formulación propuesta para el tiempo de espera (que determina los flujos predichos para cada sección de ruta), es necesario primero definir el modelo base de asignación y luego especificar las variaciones en la implementación del tiempo de espera contra las que se van a comparar los resultados del Modelo Heurístico de Asignación de Viajes. Es importante mencionar que para toda la comparación se considera las frecuencias efectivas corregidas, lo que permite aislar solo los efectos de cambiar la consideración de tiempo espera adicional de manera distinta en el modelo

4.2.2.1 Especificación de función de utilidad de modelo base

En el estudio de Metro de Santiago (2012) la formulación del Modelo de Asignación se sustenta en el modelo de Raveau et al. (2011), en el que las variables consideradas en la función de utilidad corresponden a: tiempo de viaje, tiempo de espera, tiempo de caminata en transbordos, transbordo en subida, transbordo sin escalera mecánica, porcentaje medio de ocupación, probabilidad de sentarse, probabilidad de subir, costo angular, percepción de alejarse y percepción de devolverse.

En el estudio de Metro de Santiago (2012), las probabilidades de no poder subir y la de sentarse (del modelo original), se modelan como variables binarias que se activan cuando la ocupación es mayor al 85% o menor al 15% respectivamente. En esta tesis, estas variables binarias fueron reemplazadas por una sola variable continua que representa la ocupación del tren al subir (ecuación 4.1). Esta modificación se realizó, porque las variables binarias producían que las funciones de utilidad no fueran continuas y que el algoritmo de solución no convergiera.

$$Ocupación\_subir_{sr} = \left(\frac{V_{sr} + \tilde{V}_{sr}}{K_{sr}} \right) \tag{4.1}$$

Donde:

- V_{sr} : Corresponde al flujo de la sección de ruta sr.
- $\tilde{V}_{sr}$: Corresponde al flujo compitente de la sección de ruta sr.
- K_{sr} : Corresponde a la capacidad de la sección de ruta sr.

Además, el Estudio de Metro de Santiago (2012) agrega a la función de utilidad del modelo de elección de rutas de transporte público, variables especificas asociadas a estaciones de transbordos y a las líneas de metro 4 y 4A, para mejorar el ajuste del modelo. El modelo resultante del Estudio de Metro de Santiago (2012) con sus ponderadores y atributos se muestran en la Tabla 4-4.

Tabla 4-4: Atributos función utilidad

Atributo		Ponderador
Componentes temporales	Tiempo de Viaje	-0,123
	Tiempo de Espera	-0,171
	Tiempo de Caminata	-0,279
Experiencia al transbordar	Transbordos en Subida	-0,206
	Transbordos sin Escalera Mecánica	-0,366
	Santa Ana	-0,999
	Los Héroes	-0,724
	Baquedano	-0,635
	Vicente Valdés	-0,131
	Tobalaba	-0,113
	Vicuña Mackenna	-0,073
	La Cisterna	-0,107
	San Pablo	-0,524
Conocimiento y uso	Línea Nueva	-0,220
	Porcentaje Medio de Ocupación	-0,795
	Ocupación al Subir	-0,216
Topología del mapa	Costo Angular	-0,043
	Percepción de Alejarse	-0,400
	Percepción de Devolverse	-0,489

Fuente: Estudio Metro de Santiago (2012)

Es importante mencionar que la variable de ocupación captura tanto a los usuarios que no abordan el primer tren por falta de capacidad como a los que no abordan por que la ocupación es muy alta. Debido a que de alguna forma está capturando el efecto de no poder abordar al primer vehículo que se está esperando se decide excluir del modelo base para que todos los modelos a contrastar tengan en su formulación variables que consideren los costos de viajes pero no la restricción de capacidad.

4.2.2.2 Metodología de comparación

En esta sección se presenta una comparación del desempeño del modelo propuesto utilizando como Modelo de Asignación base el descrito en la sección anterior. Además, se presentan dos formulaciones alternativas para el tiempo de espera adicional con el

objetivo de evaluar los resultados de la predicción considerando formulaciones alternativas.

4.2.2.2.1 Formulación Raveau et al. (2011)

Como ya se mencionó, a pesar de que este modelo no considera explícitamente la restricción de capacidad de los vehículos y su impacto en el tiempo de espera, sí contempla un costo asociado a la ocupación al interior del vehículo en el nodo en que los usuarios suben a él. Esta penalización combina el efecto de no poder subir si los vehículos vienen demasiado llenos con el efecto de incomodidad al interior del vehículo o de no poder viajar sentado. De hecho, si se analiza el detalle de la formulación que se muestra en la ecuación (4.1), se observa que la penalización asociada al nivel de ocupación al subir al vehículo es un caso particular del componente de tiempo de espera adicional que propone De Cea y Fernández (1993), considerando el caso en que $\beta = 1,\ m = 1$ y un ponderador asociado al tiempo de espera como $\frac{\theta_{ocupación\ al\ subir}}{\theta_{espera}}$.

4.2.2.2.2 Formulación De Cea y Fernández (1993)

La segunda formulación del tiempo de espera adicional a comparar con el modelo propuesto es la descrita por De Cea y Fernández (1993) y que en la actualidad es la que se ocupa en los programas de planificación de ESTRAUS.

Los parámetros de esta función utilizados en este análisis corresponden a los calibrados por SECTRA (Año 2012). Estos parámetros indican que el tiempo de espera adicional modelado tiene los mismos parámetros de calibración, β y m, para todas las líneas de metro. Por lo tanto, para cada sección de ruta sr se obtiene la siguiente expresión para el tiempo de espera adicional:

$$d_{sr}^{n} = \beta \left(\frac{V_{sr} + \tilde{V}_{sr}}{K_{sr}} \right)^{m} = 1.5 \left(\frac{V_{sr} + \tilde{V}_{sr}}{K_{sr}} \right)^{6} \quad (4.2)$$

Definidos los dos modelos de comparación se puede resumir los tres modelos cuyo desempeño se compara como:

- Modelo Raveau et al. (2011): Modelo base + una penalización asociada a la ocupación al subir al vehículo definida por Raveau et al. (2011).
- Modelo De Cea y Férnandez (1993): Modelo base + un tiempo de espera adicional definido tal y como se presenta en el estudio de De Cea y Férnandez (1993).
- Modelo propuesto por esta tesis, con los ajustes finales del capítulo 3..

4.2.2.3 Resultados de la comparación

La comparación se hizo en base a la información de perfiles de carga entregados por metro para el año 2013 y se evalúa como ajusta la forma de la carga (correlación) y el error absoluto medio porcentual (MAPE).

Los perfiles de carga muestran cuántos pasajeros están a bordo de los trenes al salir en una estación, para esto suma los pasajeros que suben y abordaron aguas arribas el tren con destino aguas abajo menos los pasajeros que se bajan en la estación. La primera estación de la línea de metro siempre tendrá una carga igual a todos los pasajeros que suben a ésta y la última estación tendrá flujo 0, ya que todos los pasajeros se habrán bajado. Por simplicidad en los perfiles de carga no se mostrará la última estación con flujo 0.

Como referencia de perfil de carga se usan datos proporcionados por SECTRA (2012), los datos entregados corresponden al intervalo entre 7:30 y 8:30 y son obtenidos por Metro para el año 2013, mediante el pesaje de los pasajeros y considerando un peso promedio de 80 kg por persona.

A continuación se presentan los resultados obtenidos al cargar directamente la matriz de viajes descritas en la sección 4.1.3 y la oferta en las secciones 4.1.1 y 4.1.2. y los datos proporcionados por SECTRA (2012).

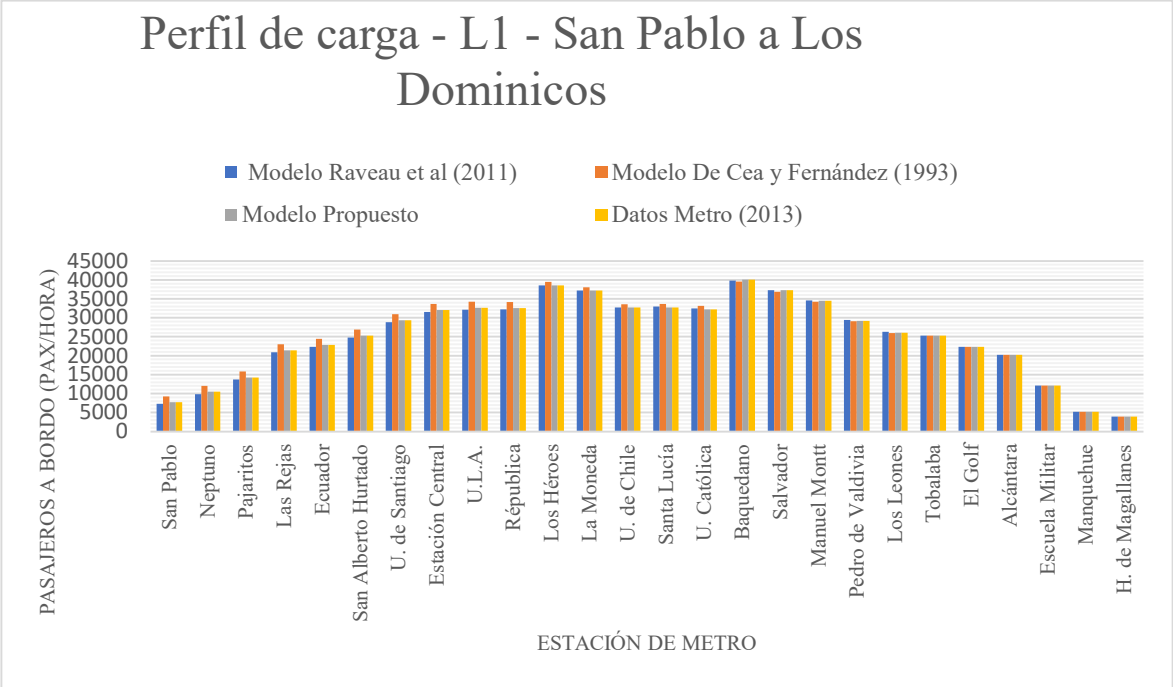


Figura 4-2: Perfil de carga L1
Fuente: Elaboración propia a partir de datos de Metro (2013)

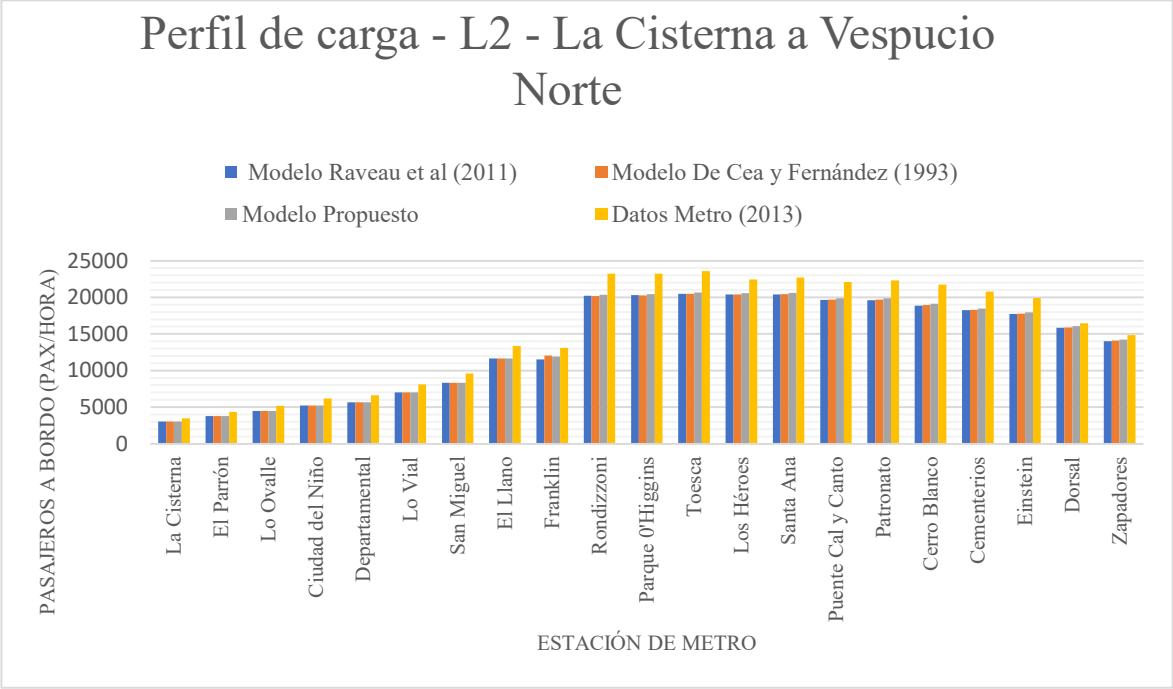


Figura 4-3: Perfil de carga L2
Fuente: Elaboración propia a partir de datos de Metro (2013)

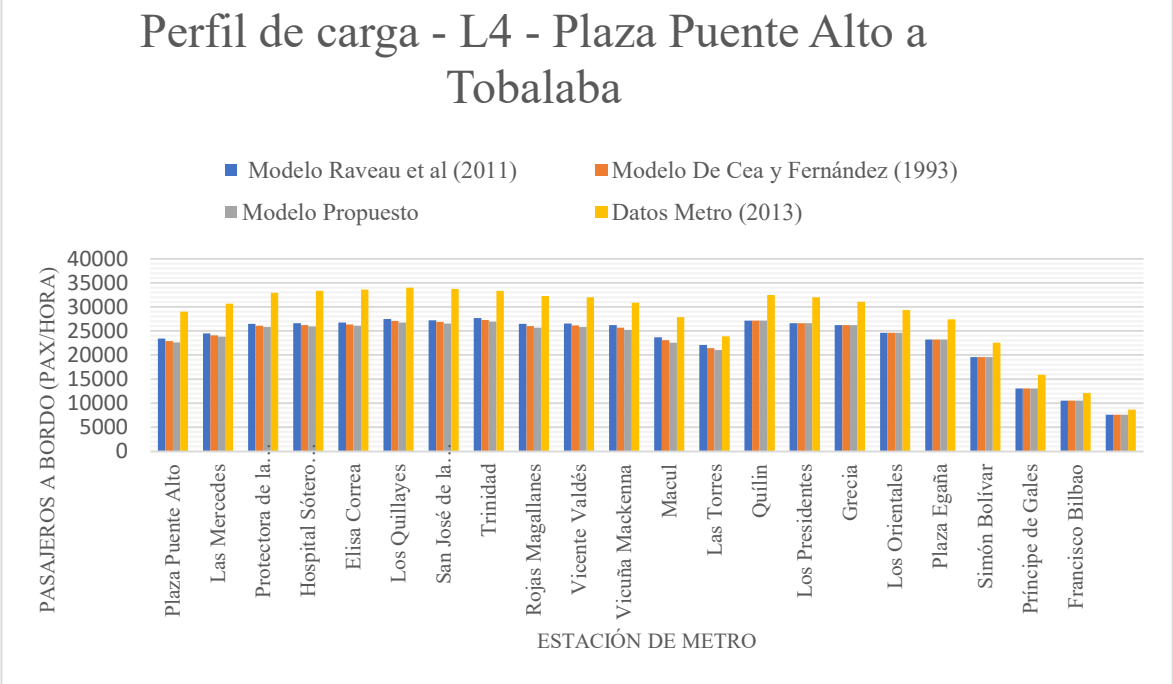


Figura 4-4: Perfil de carga L4
Fuente: Elaboración propia a partir de datos de Metro (2013)

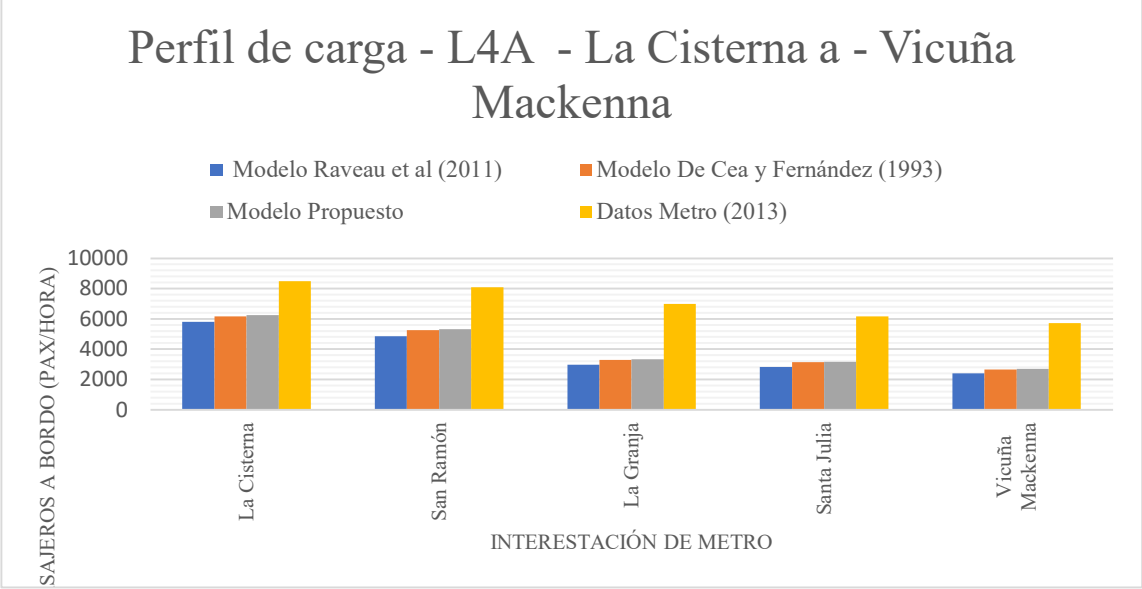


Figura 4-5: Perfil de carga L5
Fuente: Elaboración propia a partir de datos de Metro (2013)

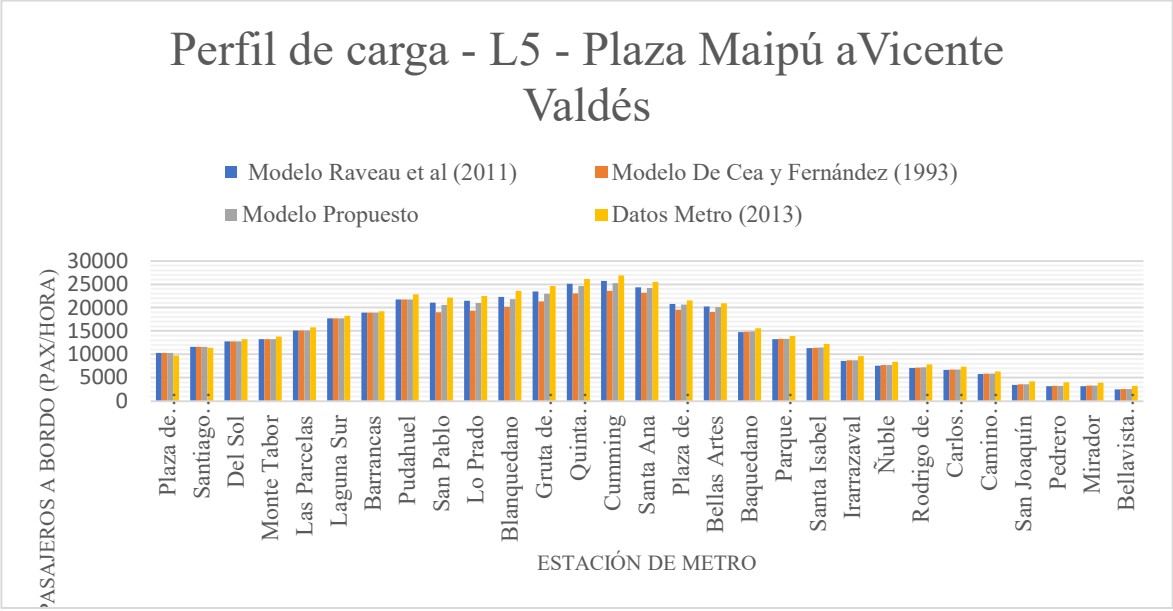


Tabla 4-5 Comparación de correlación y error porcentual absoluto ponderado de líneas de Metro en sentido más cargado

	Modelo Propuesto		Modelo Raveau et al. (2011)		Modelo De Cea y Fernández (1993)	
	Correlación	Error porcentual absoluto ponderado	Correlación	Error porcentual absoluto ponderado	Correlación	Error porcentual absoluto ponderado
L1	99,7%	10,9%	99,7%	11,4%	99,3%	9,1%
L2	99,8%	10,5%	99,8%	11,4%	99,8%	11,2%
L4	99,0%	18,7%	99,2%	17,0%	99,2%	17,0%
L4A	96,2%	54,6%	95,8%	61,6%	96,2%	55,5%
L5	99,8%	5,8%	99,4%	5,3%	99,9%	8,8%

Fuente: Elaboración propia a partir de datos de Metro (2013).

Al observar el detalle de la asignación de flujos de estos modelos se encuentra con que todos los modelos ajustan bien en la forma con que se cargan los viajes en las estaciones,

correlaciones superiores a 99% excepto en L4A, pero con errores de viajes muy variables en las líneas, entre 5% y 60% respecto a lo observado, siendo la L4A con los peores resultados.

Es importante recalcar que esto se debe a que el modelo de asignación base del estudio de Metro de Santiago (2012) no logra capturar bien el comportamiento de la línea 4 al igual que los modelos calibrados para ESTRAUS, obtenidos de SECTRA(2012), A esto se le suma la discrepancia que existe en la forma que se miden los perfiles de carga y cómo se obtienen los datos de demanda. Los perfiles de carga se obtienen directamente del pesaje de los trenes de metro, considerando un peso promedio de 80 kg. En cambio la demanda se obtiene a partir de las validaciones Bips, pero el destino del viaje es una predicción como se explica en Munizaga y Palma (2012). Dicho esto no se puede concluir que uno de los modelos lo hace mejor, porque todos ajustan bien en la forma pero hay diferencias importante en las cantidades de pasajeros.

Debido a estos 2 problemas: datos y capacidad se decide modificar la demanda con el objetivo de llevar al sistema al límite, respetando los supuestos del problema y poder probar mejor el desempeño predictivo de modelos sometido a altos niveles de congestión.

Los resultados obtenidos de acuerdo con la carga de cada segmento de línea según los tres modelos de tiempo de espera adicional que se están contrastando se muestran en la Tabla 4-6.

Tabla 4-6 Resultados de carga de segmentos de línea

	Cantidad de segmentos de líneas			% de segmentos de líneas		
Ocupación	**Modelo Propuesto**	**Modelo Raveau et al. (2011)**	**Modelo De Cea y Fernández (1993)**	**Modelo Propuesto**	**Modelo Raveau et al. (2011)**	**Modelo De Cea y Fernández (1993)**
[0% - 25%)	173	203	205	31,5%	36,9%	37,3%
[25% - 30%)	25	18	21	4,5%	3,3%	3,8%
[30% - 35%)	26	25	24	4,7%	4,5%	4,4%
[35% - 40%)	23	26	26	4,2%	4,7%	4,7%
[40% - 45%)	20	14	17	3,6%	2,5%	3,1%
[45% - 50%)	24	24	24	4,4%	4,4%	4,4%
[50% - 55%)	20	24	20	3,6%	4,4%	3,6%
[55% - 60%)	21	14	14	3,8%	2,5%	2,5%
[60% - 65%)	18	15	15	3,3%	2,7%	2,7%
[65% - 70%)	17	12	10	3,1%	2,2%	1,8%
[70% - 75%)	17	18	14	3,1%	3,3%	2,5%
[75% - 80%)	27	24	25	4,9%	4,4%	4,5%
[80% - 85%)	47	36	51	8,5%	6,5%	9,3%
[85% - 90%)	55	29	53	10,0%	5,3%	9,6%
[90% - 95%)	21	37	24	3,8%	6,7%	4,4%
[95% - 100%)	16	13	7	2,9%	2,4%	1,3%
[100% - 104%)	0	18	0	0,0%	3,3%	0,0%

Fuente Elaboración Propia

Se observa que:

- El modelo que considera la ocupación al subir, Modelo Raveau et al. (2011) no garantiza resultados en donde se respete la restricción de capacidad en todos los segmentos de líneas. Esto implica además que hay otros segmentos cuyos flujos están subestimados.

- Tanto el modelo propuesto como el de De Cea y Fernández (1993) predicen una asignación que respeta la restricción de capacidad para esta red de Metro, 2013 con esta demanda modificada.

- El modelo propuesto predice una mayor cantidad de segmentos de líneas congestionados, que los otros modelos. En particular, se ve que predice 2,9% de segmentos de línea de la red con una ocupación igual o mayor a 95% y menor al 100%, versus el 2,4% de los segmentos de líneas predichos por la formulación de Raveau et al. (2011) y el 1,3% de la formulación de De Cea y Fernández (1993). Incluso si se compara los segmentos de línea sobre un 85% el modelo propuesto considera 16,7%, mientras que el de De Cea y Fernández (1993) considera un 15,3% y el de Raveau et al. (2011) un 14,4 % excluyendo los resultados que superan la restricción de capacidad.

- Para el caso de los segmentos de líneas con menor ocupación el modelo propuesto también es el que presenta mejores resultados, porque es el que presenta menos segmentos de líneas con una ocupación menor a 25 %, lo que tiene más sentido con las cargas observadas durante los periodos punta.

Los últimos dos puntos son relevantes en los resultados de asignación de una red altamente congestionada, porque es esperable que primero se congestionen los servicios más atractivos alcanzando o acercándose bastante a su capacidad límite, luego se esperaría que también se congestionen los servicios un poco menos atractivos. Este resultado tiene más sentido que se repartan la carga de manera más homogénea en la red, por eso también es relevante que los segmentos de líneas con baja ocupación sean menos, porque probablemente estos solo se ocupan cuando las líneas más atractivas ya no tienen capacidad.

5 CONCLUSIONES

En esta tesis se ha elaborado un Modelo de Asignación que genera una predicción de flujos consistente con un Equilibrio de Usuarios Estocástico. S.U.E. y una restricción estricta de capacidad de los vehículos. El Modelo Heurístico de Asignación de Viajes propuesto se estructuró en torno a dos etapas iterativas que se hacen cargo de dos efectos distintos en que la restricción de capacidad activa influye en la asignación: (1) se impacta la definición de conjunto de líneas atractivas, pues líneas previamente ignoradas pasan a ser atractivas y (2) genera aumentos en los tiempos de espera percibidos por los usuarios, debido a que no pueden abordar el primer vehículo, lo que influye en la asignación de los viajes por las secciones de ruta de la red.

Para la definición del modelo propuesto, se analizó la bibliografía en detalle y se mostró que el modelo más visible que aborda este problema, propuesto por Lam et al. (2002) genera asignaciones que respetan la restricción de capacidad de los vehículos en redes de transporte público, pero no el S.U.E. . Esto ocurre porque los autores abordan este equilibrio resolviendo un problema de optimización equivalente, pero para que este problema de optimización sea válido es necesario que las demoras adicionales de cada ruta sean igual a las demoras de todos los segmentos de línea que componen la ruta. Sin embargo, considerar este tipo de demora no es compatible con el comportamiento normal de los usuarios en redes de transporte público, aun cuando si lo es respecto de redes *store-and-forward*, como las redes de transporte privado.

En esta tesis se propuso una nueva formulación para el tiempo de espera adicional, considerando aumentos en los tiempos de espera solo al abordar un vehículo, que respeta los supuestos de comportamiento del S.U.E y que se puede obtener como resultado de la asignación de flujos en la red, a diferencia de la formulación de De Cea y Fernández (1993), donde es necesario calibrar una función de tiempo de espera adicional para los datos en particular. El algoritmo de solución utilizado para obtener el tiempo de espera adicional de la asignación se basó en el algoritmo desarrollado por Bell (1995). En este

proceso se redefinió la formulación base del multiplicador de Lagrange M_a generando mejoras significativas en la convergencia del algoritmo original.

En cuanto a la modelación de las frecuencias efectivas de los servicios en un paradero se obtuvieron dos conclusiones relevantes. La primera es que las frecuencias efectivas deben ser capaces de capturar cuando un servicio ya no tiene capacidad, esto implica que su frecuencia efectiva deber ser 0 y por tanto el servicio no debería aparecer dentro del conjunto de líneas comunes. A pesar de que si se evalúa el costo generalizado del viaje al agregar la línea en el conjunto de líneas comunes el costo generalizado disminuya, si se considera una frecuencia efectiva igual 0 el tiempo de espera es infinito y por tanto la línea no puede ser parte del conjunto de líneas atractivas. Eliminar la condición asintótica de la función de las frecuencias efectivas resolvió este problema.

La segunda conclusión tiene relación con que la función de frecuencias efectivas tiene que ser capaz de capturar cuando el sistema se congestiona o descongestiona de una iteración a otra, es por esto la función no puede ser decreciente con la cantidad de iteraciones sino más bien tiene que ser dependiente de los resultados de asignación, en particular de las demoras endógenas del problema, y de la frecuencia nominal de una línea. Así en casos donde en una iteración se descongestiona una línea en un paradero la frecuencia efectiva de la siguiente iteración aumenta, en cambio, cuando una línea se congestiona la frecuencia efectiva disminuye, para esto se agregó que la frecuencia considerada para la asignación sea la nominal cada vez que en una iteración se agrega un línea atractiva.

Al aplicar el algoritmo propuesto para una red real de tamaño medio, como el Metro de Santiago (año 2013), se obtienen soluciones de equilibrio que respetan el concepto de S.U.E. y la restricción estricta de capacidad de los vehículos, no así con otro tipo de formulaciones para el tiempo de espera adicional como la propuesta por Raveau et al. (2011). Esto implica que las soluciones que no respetan la restricción de capacidad de los vehículos sobreestimen los flujos en algunos arcos y se subestimen en otros, generando a su vez costos asociados al viaje sobre- o subestimados y por lo tanto obteniendo predicciones de flujo de pasajeros alejadas de la realidad.

Para el caso de la formulación de tiempo de espera adicional del modelo de De Cea y Férnandez (1993) en la red de Metro, la asignación resultante respeta la restricción de capacidad de los vehículos para el caso hipotético, pero se obtienen segmentos menos congestionados que con la formulación del modelo propuesto lo que pareciera alejar la calidad de la solución respecto de la solución buscada. Esta distorsión ocurre porque el modelo de De Cea y Fernández (1993) adiciona un tiempo de espera siempre que la línea esté en uso aun cuando ésta no alcance capacidad. Esto que es razonable en un contexto en que es posible asignar flujos que superan la capacidad, no sería razonable en el caso de una restricción de capacidad estricta como si los usuarios debieran siempre tomar asientos en los vehículos en que viajan.

Además cabe mencionar, que si bien el modelo que considera los tiempos de esperas adicionales bajo la formulación de De Cea y Fernández (1993) cumple con la restricción de capacidad estricta para el caso de Metro de Santiago (2013), tiene en consideración la creación del conjunto de líneas comunes de manera dinámica en función de los resultados de asignación de cada iteración, por lo tanto que cumpla con las restricciones de capacidad de los vehículos no es solo efecto de considerar el tiempo de espera adicional con este modelo, si no también es efecto de que no se consideran dentro del conjunto de líneas atractivas líneas que superan la capacidad en estaciones anteriores. Esta es la principal característica a destacar del Modelo Heurístico de Asignación de Viajes propuesto, frente a otros modelos, que no incluyen la restricción de capacidad estricta en la construcción de la red de asignación.

En relación a futuras líneas de investigación parece interesante estudiar el modelo con demandas dinámicas en que se relaja el supuesto de que la oferta debe ser suficiente para la demanda dada en todo momento. Este tipo de modelación tendría más sentido para microsimulaciones en donde en un intervalo más pequeño de tiempo la demanda puede sobrepasar a la oferta y esperar la oferta del siguiente intervalo. Esta aproximación al problema exigiría que la oferta durante el periodo completo de análisis fuese capaz de movilizar todos los viajes de la red. Este enfoque permitiría capturar mejor el efecto de no poder abordar el primer vehículo que el usuario está esperando y se obtendría cargas de

segmentos de líneas más acordes a la realidad, en donde los subintervalos *peak* de demanda tendrían mayores cargas que los intervalos de tiempo en donde el sistema ya está descongestionando.

Otra variación del modelo propuesto es estudiar cómo modelar casos en que la red no fuese capaz de llevar todos los viajes. Para enfrentar estos casos se podría agregar al modelo rutas residuales con capacidades lo suficientemente grandes para llevar esta demanda que no es capaz de ser movilizada por la red. Esto podría ser atractivo para enfrentar escenarios futuros en que la demanda pudiera exceder la oferta actual. Los resultados darían luces respecto de dónde podrían ocurrir problemas de capacidad insuficiente, orientando por ejemplo, a una eventual planificación de inversiones al sistema de transporte urbano.

Uno de los desafíos que se tiene al usar rutas residuales con un modelo Logit, es que al ser este un modelo basado en la diferencia de los costos de las rutas alternativas para pares orígenes-destinos con el mismo origen y con los problemas de congestión asociados a la primera etapa, las rutas más cortas se verían más perjudicadas que las rutas más largas. Esto ocurriría porque las rutas más cortas se acercarían más rápido al costo de la ruta residual que las rutas más largas y por lo tanto el modelo tendería a enviar este exceso de demanda primero a las rutas más cortas en vez de las rutas más largas. De esta forma el modelo privilegiaría descongestionar pares orígenes-destino más cercanos, lo que sería perjudicial en la planificación de una ciudad, ya que lo ideal es descongestionar la red de manera de obtener el mayor beneficio global y no de unos pocos.

Finalmente, de este tesis se puede concluir que es primordial considerar los efectos de la restricción de capacidad tanto para la construcción del conjunto de líneas atractivas considerado en la red de asignación, como la inclusión de los tiempos de esperas adicionales en los paraderos de subida de los pasajeros. Ambos efectos son los que determinan si un pasajero puede o no subir a un vehículo, si solo se consideran los tiempos de esperas adicionales y no se hace un cambio en la construcción de la red, eliminando las líneas que ya alcanzaron el máximo de capacidad, se pueden obtener resultados que no

respetan la restricción de capacidad. Este efecto lo resuelve esta tesis generando una heurística que considera solo líneas con capacidad en la creación del conjunto de líneas atractivas.

En cuanto a la principal ventaja de considerar el tiempo de espera adicional propuesto de esta tesis, es que no es necesario calibrar una función adicional como es el caso de De Cea y Fernández (1993). Los tiempos de esperas considerados en la asignación son obtenidos directamente de los resultados de asignación.

BIBLIOGRAFÍA

Metro de Santiago. (2012, Diciembre). *Análisis de Demanda para Diseño de Estaciones de Transbordo de los Proyectos Línea 3 y 6.* Santiago.

Bell, M. G. (1995). Stochastic user equilibrium assignment in networks with queues. *Transportation Research B 29(2)*, 125-137.

Chriqui, C., & Robillard, P. (1975). Common bus lines. *Transportation Sci 9*, 115-121.
Cominetti, R., & Correa, J. (2001). Common-lines and passenger assignment in congested. *Transportation science, 35*(3), 250-267.

Cortés, C., P., J.-M., E., M., & C., P. (2013). Stochastic transit equilibrium. *Transportation Research Part B*, 29-44.

Daganzo, C. F., & Sheffi, Y. (1977). On stochastic models of traffic assignment. *Transportation science, 11(3)*, 253-274.

De Cea, J. (1984). I Congreso Chileno Ingeniería de Transporte. *Análisis de distintos tipos de modelos de asignación de redes de transporte público* (pp. 269-286). Santiago de Chile: Sociedad Chilena de Ingeniería de Transporte.

De Cea, J. (2012, Agosto 1). Equilibrio en redes de transporte ICT3283. *Apuntes de Clases*. Santiago, Región Metropolitana, Chile: Pontificia Universidad Católica de Chile.

De Cea, J., & Fernández, E. (1993). Transit assignment for congested public transport systems: an equilibrium model. *Transportation science, 27(2)*, 133-147.

Fisk, C. (1980). Some developments in equilibrium traffic assignment. *Transportation Research Part B: Methodological, 14(3)*, 243-255.

Fu, Q., Liu, R., & Hess, S. (2012). A review on transit assignment modelling approaches to congested networks: a new perspective. *Procedia-Social and Behavioral Sciences, 54*, 1145-1155.

Herrera, J. C., & Sommariva, P. (2013, Agosto 1). Modelos Avanzados de Tráfico ICT3273. *Apuntes de Clases*. Santiago, Región Metropolitana, Chile: Pontificia Universidad Católica de Chile.

J. C., & N. S.-M. (2011). Wardrop Equilibria. In *Wiley encyclopedia of operations research and management science.*

Lam, W. H., Gao, Z. Y., Chan, K. S., & Yang, H. (1999). A stochastic user equilibrium assignment model for congested transit networks. *Transportation Research Part B: Methodological, 33(5)*, 351-368.

Lam, W. H., Zhou, J., & Sheng, Z. H. (2002). A capacity restraint transit assignment with elastic line frequency. *Transportation Research Part B*, 919-938.

Munizaga, A., & Palma, C. (2012). Estimation of a disaggregate multimodal public transport Origin-Destination matrix from passive smartcard data from Santiago, Chile. *Transportation Research Part C*, 9-18.

Nguyen, S., & Pallottino, S. (1988). Equilibrium traffic assignment for large scale transit networks. *European journal of operational research, 37*(2), 176-186.

Raveau, S., & Muñoz, J. (2014). Transportation Research Board 93rd Anual Meeting. *Analyzing route choice strategies on transit network.* Washington DC, EE.UU.

Raveau, S., Muñoz, J. C., & De Grange, L. (2011). A topological route choice model for metro. *Transportation Research Part A: Policy and Practice 45.2*, 138-147.

SECTRA. (2019, septiembre). *MESPE: Metodología para análisis de sistemas de transporte en grandes ciudades y ciudades de tamaño medio*. Retrieved from http://www.sectra.gob.cl/metodologias/estraus.htm

Sheffi, Y. (1985). *Urban transportation networks.* New Jersey: PRENTICE-HALL, INC.

Train, K. E. (2009). *Discrete choice methods with simulation.* Cambridge university press.

Wardrop, J. G. (1952). Road paper. Some theoretical aspects of road traffic research. *Proceedings of the institution of civil engineers, 1(3)*, 325-362.

A N E X O S

ANEXO A: FORMULACIÓN DE PROBABILIDAD DE ELECCIÓN DE LOGIT.

Las probabilidades de elección de cada alternativa i en los modelos de elección, dependen de la función de densidad de probabilidades asociadas al error que se está modelando. En el caso del modelo Logit, la función densidad de los errores distribuyen Gumbel como se muestra en la ecuación (6.1) .

$$f(\varepsilon_j) = e^{-\varepsilon_j} e^{-e^{-\varepsilon_j}} \tag{6.1}$$

La distribución acumulada asociada se formula en la ecuación (6.2)

$$F(\varepsilon_j) = e^{-e^{-\varepsilon_j}} \tag{6.2}$$

La probabilidad de elección de una alternativa i esta dada por la probabilidad de que la utilidad de la alternativa i, sea mayor que la utilidad de las otras alternativas j ($U_j \,, \forall\, j \neq i$), esto se muestra en las ecuaciones (6.3) y (6.4).

$$P_i = \Pr\{\, U_i > U_j \,, \forall\, j \neq i\} \tag{6.3}$$

$$P_i = \Pr\{\, V_i + \varepsilon_i > V_j + \varepsilon_j \,, \forall\, j \neq i\} \tag{6.4}$$

Reescribiendo la probabilidad de elegir una alternativa en función del error ε_j, se obtiene la ecuación (6.5).

$$P_i = \Pr\{\, \varepsilon_j < \varepsilon_i + V_i - V_j \,, \forall\, j \neq i\} \tag{6.5}$$

Si se considera ε_i dado, se tiene que la ecuación (6.5) es la distribución acumulada para cada error ε_j evaluado en $\varepsilon_i + V_i - V_j$. Además, como los ε_j son i.i.d en el modelo Logit, la probabilidad de elegir la alternativa i dado ε_i, se puede escribir como la ecuación (6.6).

$$P_{i|\varepsilon_i} = \prod_{\forall j \neq i} e^{-e^{-(\varepsilon_i + V_i - V_j)}} \tag{6.6}$$

Como ε_i, en realidad no esta dado, para obtener la probabilidad de elegir la alternativa i se necesita resolver la ecuación (6.7).

$$P_i = \int_{-\infty}^{+\infty} P_{i|\varepsilon_i} \cdot f(\varepsilon_i) \cdot d\varepsilon_i \tag{6.7}$$

La resolución de la ecuación (6.7) se muestra en las ecuaciones (6.8) a (6.17)

$$P_i = \int_{-\infty}^{+\infty} \prod_{\forall j \neq i} e^{-e^{-(\varepsilon_i + V_i - V_j)}} \cdot e^{-\varepsilon_i} \cdot e^{-e^{-\varepsilon_i}} \cdot d\varepsilon_i \tag{6.8}$$

$$P_i = \int_{-\infty}^{+\infty} e^{-\sum_{\forall j \neq i} e^{-\varepsilon_i - V_i + V_j}} \cdot e^{-\varepsilon_i} \cdot e^{-e^{-\varepsilon_i}} \cdot d\varepsilon_i \tag{6.9}$$

$$P_i = \int_{-\infty}^{+\infty} e^{-\left(e^{-\varepsilon_i - V_i}\right) \cdot \sum_{\forall j \neq i} e^{V_j}} \cdot e^{-e^{-\varepsilon_i}} \cdot e^{-\varepsilon_i} \cdot d\varepsilon_i \tag{6.10}$$

$$P_i = \int_{-\infty}^{+\infty} e^{-(e^{-\varepsilon_i}) \cdot \left(e^{-V_i}\right) \cdot \sum_{\forall j \neq i} e^{V_j} - e^{-\varepsilon_i}} \cdot e^{-\varepsilon_i} \cdot d\varepsilon_i \tag{6.11}$$

$$P_i = \int_{-\infty}^{+\infty} e^{\left\{-(e^{-\varepsilon_i}) \cdot \left[\left(e^{-V_i}\right) \cdot \sum_{\forall j \neq i} e^{V_j} + 1\right]\right\}} \cdot e^{-\varepsilon_i} \cdot d\varepsilon_i \tag{6.12}$$

$$P_i = \frac{e^{\left\{-(e^{-\varepsilon_i}) \cdot \left[\left(e^{-V_i}\right) \cdot \sum_{\forall j \neq i} e^{V_j} + 1\right]\right\}}}{\left[\left(e^{-V_i}\right) \cdot \sum_{\forall j \neq i} e^{V_j} + 1\right]} \Bigg|_{-\infty}^{+\infty} \tag{6.13}$$

$$P_i = \lim_{\varepsilon_i \to \infty} \frac{e^{\left\{-\left(1/e^{\varepsilon_i}\right) \cdot \left[\left(e^{-V_i}\right) \cdot \sum_{\forall j \neq i} e^{V_j} + 1\right]\right\}}}{\left[\left(e^{-V_i}\right) \cdot \sum_{\forall j \neq i} e^{V_j} + 1\right]} - \lim_{\varepsilon_i \to -\infty} \frac{1 \Big/ e^{\left\{(e^{-\varepsilon_i}) \cdot \left[\left(e^{-V_i}\right) \cdot \sum_{\forall j \neq i} e^{V_j} + 1\right]\right\}}}{\left[\left(e^{-V_i}\right) \cdot \sum_{j \forall \neq i} e^{V_j} + 1\right]} - \tag{6.14}$$

$$P_i = \frac{1}{\left[\left(e^{-V_i}\right) \cdot \sum_{j \neq i} e^{V_j} + 1\right]} - 0 \tag{6.15}$$

$$P_i = \frac{1}{\left(e^{-V_i}\right)\left[\cdot \sum_{\forall j \neq i} e^{V_j} + 1 \Big/ \left(e^{-V_i}\right)\right]} \tag{6.16}$$

$$P_i = \frac{e^{V_i}}{\left[\cdot \sum_{\forall\, j \neq i} e^{V_j} + e^{V_i}\right]} = \frac{e^{V_i}}{\sum_{\forall j} e^{V_j}} \tag{6.17}$$

Finalmente se obtiene la probabilidad de elegir la alternativa i (ecuación (6.17)).