

## CUANDO SUBIR LA FRECUENCIA EMPEORA EL SERVICIO: COMPRENDIENDO LA “DANGER ZONE” DE LOS SERVICIOS EXPRESOS.

Carlos Olivos, Pontificia Universidad Católica, [cuolivos@uc.cl](mailto:cuolivos@uc.cl)  
Homero Larrain, Pontificia Universidad Católica, [homero@ing.puc.cl](mailto:homero@ing.puc.cl)  
Juan Carlos Muñoz, Pontificia Universidad Católica, [jcm@ing.puc.cl](mailto:jcm@ing.puc.cl)

### RESUMEN

Larrain y Muñoz (2019) introdujeron y estudiaron por primera vez el fenómeno de la *danger zone* de los servicios expresos. Este fenómeno se manifiesta en una aparente paradoja: es posible aumentar la frecuencia de un servicio expreso, manteniendo todos los demás servicios inalterados, generando un empeoramiento global del nivel de servicio recibido por los usuarios del sistema. El modelo en que se basa este resultado considera algunos supuestos propios de un enfoque analítico. En este trabajo se robustecen estos resultados mediante un análisis basado en una simulación que permite validar y extender el análisis a situaciones más realistas. Nuestros experimentos permiten concluir que la heterogeneidad de los usuarios atenúa (pero no invalida) el efecto de la *danger zone* en un paradero. También mostramos que proveer información perfecta a los usuarios no contrarresta este efecto.

*Transporte público, servicios expresos, optimización de frecuencias, micro simulación*

### ABSTRACT

Larrain y Muñoz (2019) introduced and studied the phenomenon called the danger zone of express services. This danger zone seemingly translates into a paradox: it is possible to raise the frequency of an express service without benefitting its users, or even leaving them worse off. In this work, we implement a simulation that let us validate some previous theoretical results, and extend the analysis to settings that are more realistic. Our experiments show that user heterogeneity mitigates (but does not nullify) the danger zone effect in a station. We also show that availability of perfect information for the users does not prevent this effect.

*Public transport, express services, frequency optimization, micro simulation*

## 1. INTRODUCCIÓN

Varios sistemas de transporte público, sobre todo en ciudades altamente pobladas, presentan corredores que enfrentan niveles muy altos de demanda, para lo cual ofrecen una muy alta capacidad de transporte. En estos ejes se enfrenta el riesgo de que cualquier eventualidad que reduzca su oferta pueda llevarlos al colapso. Este es el caso, por ejemplo, de Transmilenio en Bogotá, uno de los sistemas BRT más complejos y de mayor demanda de mundo (Merket et al., 2017). En algunos casos, este deterioro en el nivel de servicio sucede incluso en condiciones donde las capacidades del sistema parecieran ser suficientes en forma agregada para satisfacer la demanda que enfrentan, lo que insinúa que existen fenómenos a escala microscópica que podrían no estar considerados adecuadamente en la planificación de servicios. En Larrain y Muñoz (2019) se ofrece un ejemplo de una situación como esta, donde un aumento en la capacidad agregada del sistema gatilla un colapso en el nivel de servicio a los usuarios, como consecuencia de elecciones individuales de los usuarios en la estación.

Los autores plantean esta paradoja, a la que llaman la “*danger zone*” (que abreviaremos como DZ) de los servicios expresos, sobre un escenario basado en dos servicios que operan sobre un corredor BRT: un servicio “*all-stop*”, que se detiene en todos los paraderos en su camino, y un servicio expreso, que se detiene solo en los dos terminales (extremos) del corredor. Sobre este escenario, y usando un modelo sencillo de comportamiento de usuarios, los autores muestran que existen ciertas combinaciones de frecuencias para ambos servicios en que aumentar la frecuencia del servicio expreso podría empeorar fuertemente el servicio a los usuarios en el primer terminal, en vez de mejorarlo. Esta aparente paradoja se explica porque al aumentar la frecuencia del servicio expreso, los usuarios que pueden escoger entre ambos servicios se vuelcan al expreso. Este aumento brusco de demanda por el servicio expreso puede sobrepasar la capacidad de sus buses, generando largas colas en las paradas. Este problema persiste mientras la frecuencia del bus expreso no sea suficiente como para llevarse toda la demanda que espera por él.

Mediante un análisis teórico del fenómeno los autores determinan un rango de frecuencias en las que las esperas por este servicio expreso se vuelven extraordinariamente largas. Este rango de frecuencias y de condiciones operacionales que detonan la formación de filas y esperas en el paradero que contrarrestan los beneficios de un aumento de frecuencias es lo que los autores definen más precisamente como la *danger zone* (DZ) de los servicios expresos. Este resultado ciertamente levanta alarmas, pues indica que autoridades de un sistema podrían involuntariamente empeorar de manera sustancial las condiciones de los usuarios, mediante un esfuerzo que significa mayores costos para el sistema. En palabras sencillas, cuando un sistema entra en la DZ, todos los actores pierden.

Larrain y Muñoz (2019) identifican este fenómeno y proponen una caracterización matemática de la DZ, identificando sus límites teóricos y proponiendo expresiones para diferentes indicadores de desempeño. Los análisis de dicho trabajo se estructuran en torno a tres supuestos simplificadores

que podrían ser discutibles. En primer lugar, se estudia un sistema gobernado por procesos determinísticos, donde tanto buses como usuarios llegan a la estación a intervalos perfectamente regulares. En segundo lugar, se asume que los usuarios son homogéneos en su proceso de toma de decisión. Por último, el trabajo se enfoca en la modelación de la cola por el servicio expreso, y asume por simplicidad que la cola por el servicio all-stop es despreciable (o, equivalentemente, que su frecuencia de servicio es relativamente alta). Este supuesto facilita la modelación del comportamiento de los usuarios flexibles, ya que se puede asumir que al llegar a la estación siempre se ponen en la cola del expreso, y que recién cuando pasa un servicio all-stop deciden si quedarse en esta fila y esperar por el expreso, o si tomar el servicio all-stop. Un modelo más realista de la operación de una estación debiese capturar los efectos de la variabilidad en las tasas de llegada y en las preferencias de los usuarios, y debiese permitir la formación de colas en cualquiera de los servicios de la estación.

Para validar los resultados de Larrain y Muñoz (2019), en este trabajo desarrollamos e implementamos una simulación cuyos elementos son los usuarios individuales, los buses, y la fila de cada bus. Para desarrollar esta simulación proponemos un modelo de comportamiento de los usuarios que les permite cambiarse de fila en función de los buses que van llegando a la parada. Los resultados de la simulación logran validar las conclusiones sobre la DZ, y abren las puertas a hacer nuevos análisis en mayor profundidad de este fenómeno. Por ejemplo, en este trabajo incluimos un experimento para determinar si proveer información a los usuarios respecto de los instantes de llegada de los próximos buses a la estación podría influir positivamente en su conducta, y de alguna forma contrarrestar la ocurrencia de la DZ en algunos escenarios.

Lo que sigue de este artículo está organizado de la siguiente forma. A continuación, presentamos una revisión de literatura que resume el desarrollo histórico de los modelos de comportamiento, y explora propuestas de distintos autores para optimizar el trazado y la frecuencia de servicios sobre corredores BRT. Luego describimos el problema a abordar y la metodología a utilizar, comenzando con un resumen de la formulación de la DZ propuesta en Larrain y Muñoz (2019), para luego especificar el funcionamiento de la simulación y su modelo de comportamiento de usuarios. Luego reportamos los resultados de tres experimentos, que nos permiten validar la simulación y corroborar la existencia de la DZ, ofreciendo nuevas perspectivas sobre este fenómeno. Finalmente, presentamos las principales conclusiones del trabajo.

## 2. REVISIÓN DE LITERATURA

En esta sección presentamos una revisión de la literatura relevante para este trabajo. Comenzamos abordando la literatura referente a servicios expresos y cómo diseñarlos. Luego, dado que la DZ está íntimamente conectada con el comportamiento de los usuarios, detallamos diferentes modelos de asignación de pasajeros en transporte público, para finalmente discutir cómo modelar las elecciones que ocurren dentro de una estación o paradero, problema conocido como el “*bus stop problem*” (BSP).

El problema de diseño de servicios expresos se enmarca en el problema más general de diseño de redes de transporte público: para una detallada revisión de literatura sobre este tema el lector puede consultar el trabajo de Desaulniers y Hickmann (2007). El problema de diseño de servicios

expresos ha sido estudiado ampliamente en la última década. En Leiva et al. (2010), se propone un modelo de optimización de servicios expresos que minimiza los tiempos de viaje y los costos operacionales. En Larrain et al. (2010) se aplica el modelo anterior con diferentes tipos de perfil de demanda, para concluir que la distancia promedio de viaje es un buen predictor del desempeño de los servicios expresos. El trabajo de Larrain y Muñoz (2016) concluye que los servicios expresos son particularmente beneficiosos en presencia de altos tiempos de detención, alta concentración de viajes en pocos pares origen destino y buses con bajas capacidades. Wang et al. (2018) presentan una metodología para determinar la estrategia óptima de operación de un servicio expreso, considerando el tamaño de la flota, las frecuencias y atributos específicos de la línea. Finalmente, un modelo para la capacidad de paradas de bus se construye en Shen et al. (2019), resultando ser más preciso que otros en literatura anterior.

Todo modelo de diseño de servicios de transporte público necesita hacer algún tipo de supuesto sobre la forma en que los usuarios del sistema hacen uso de los servicios disponibles cuando pueden escoger entre varios de ellos para realizar su viaje, es decir, de cómo se asignan los usuarios a una red dada. La forma en la que los planificadores de transporte modelan esta asignación ha evolucionado a lo largo de los años. Ford y Fulkerson (1962) sientan concisamente las bases teóricas para resolver un problema de asignación. Usando estas nociones, Dial (1967), presenta un algoritmo de asignación construido bajo el supuesto de que el tiempo de espera por un servicio se puede aproximar como la mitad del inverso de su frecuencia. Más tarde, en Chriqui y Robillard (1975) el problema adquiere nuevas dimensiones con la incorporación de líneas comunes en el análisis (i.e: servicios alternativos que permiten realizar un mismo viaje). Bajo este marco, los usuarios que esperan en una estación determinan un subconjunto de los servicios disponibles como atractivos para realizar el siguiente tramo de su viaje, y abordan el primero de ellos que visita la estación. Años después Nguyen y Pallotino (1989) generalizaron este concepto introduciendo las hiperrutas, que permiten capturar estrategias más complejas para realizar un viaje, en las que el usuario va adaptando la secuencia de servicios a tomar para alcanzar el destino según qué bus llega primero a la estación.

Los modelos antes mencionados resuelven el problema de asignación modelando las estaciones y paradas desde una perspectiva agregada, sin poner atención de cómo las configuraciones de las estaciones (puertas de embarque y filas) afectan en las elecciones de los usuarios. El problema de cómo estimar la distribución de pasajeros entre las líneas atractivas y el tiempo esperado de espera de pasajeros en el paradero, que se conoce simplemente como el “*bus stop problem*”, ha sido estudiado por Bouzaïene-Ayari et al. (2001). En este trabajo los autores realizan una revisión exhaustiva de literatura respecto a este problema, categorizando a los modelos existentes como incompletos, y por lo tanto proponiendo un nuevo modelo capaz de tratar efectivamente problemas del paradero de bus con y sin congestión, denominado GSM (*General bus Stop Model*).

Aun cuando el GSM y el modelo de comportamiento de nuestra simulación tienen focos distintos, alcanzan conclusiones similares. Ambos determinan que las estrategias de viaje de los usuarios dependen de la frecuencia de los servicios, la capacidad disponible al interior de los buses, la distribución de los intervalos entre buses y del nivel de demanda. Más aún, ambos son consistentes con la teoría de colas en lotes (i.e: colas donde se sirve a los usuarios en grupos, en vez de usuarios de a uno). Si bien esta teoría de cola en lotes existe al menos desde que fue descrita en Bailey (1954), rara vez se ha utilizado como componente de un problema de paradero de buses, lo que es

relevante dado que, como mostraremos después, asumir una tasa de servicio continua en vez de por lotes puede distorsionar en gran medida el desempeño estimado del sistema.

### 3. METODOLOGÍA

En lo que sigue, ofrecemos una descripción general de la *danger zone* “teórica” modelada en Larrain y Muñoz (2019), para luego detallar la simulación de una estación y el modelo de comportamiento de usuarios que utilizaremos para estudiar este fenómeno. En este modelo de comportamiento, que detallaremos más adelante en este capítulo, los usuarios están constantemente reevaluando el tiempo esperado de viaje de las distintas filas disponibles, y se cambian si perciben un beneficio.

#### 3.1. La *danger zone* de los servicios expresos.

La DZ se puede explicar como el fenómeno que ocurre cuando un aumento en la frecuencia de un servicio expreso (o eventualmente otro tipo) acarrea un deterioro en la calidad de servicio que perciben sus usuarios.

En el trabajo de Larrain y Muñoz (2019) se introduce y estudia este efecto en un corredor con solo dos servicios, un *all-stop* y un expreso que opera entre los terminales del *all-stop*. Los autores asumen en su modelación que los usuarios son homogéneos, y que los buses operan a intervalos regulares. En este contexto, se define como usuario flexible a aquel que puede usar cualquiera de los dos servicios para realizar su viaje. Asumiendo que el servicio expreso es  $\tau$  unidades de tiempo más rápido que el *all-stop*, un usuario flexible posee dos estrategias razonables para realizar su viaje: esperar por el expreso, o tomar el primer bus que pase por el paradero. Comparando los tiempos esperados de viaje de cada opción, y asumiendo que todos los usuarios son idénticos en sus percepciones y preferencias, este cambio de conducta ocurre en el momento en que la frecuencia del servicio expreso,  $f_E$ , supera un umbral de  $f_E \geq \hat{f}_E = 0.5/\tau$ . Sobre este umbral, todos los usuarios flexibles preferirán esperar por el servicio expreso. Es en el momento que la frecuencia del servicio expreso cruza este umbral que puede ocurrir que una masa de pasajeros cambie de golpe su conducta, congestionando el sistema.

Por otra parte, la frecuencia mínima  $P_E$  necesaria para del servicio expreso para que su capacidad permita acarrear a todos los pasajeros flexibles es de  $P_E = L_E/b_E$ , donde  $L_E$  es la demanda (en pax/h) por viajes de usuarios flexibles, y  $b_E$  es la capacidad de los buses. Entonces, si en una instancia del problema se tiene que la frecuencia del servicio expreso es suficientemente alta para atraer todos los viajes, pero insuficiente para acoger esos viajes al interior de sus vehículos (esto es:  $\hat{f}_E \leq f_E \leq P_E$ ), tendremos que la demanda por el servicio expreso superará su capacidad deteriorando el nivel de servicio de estos usuarios. Larrain y Muñoz (2019) analizan el comportamiento del sistema dentro de este rango de frecuencias asumiendo que los usuarios forman una fila por el servicio expreso, y que esta fila crece hasta el punto donde el tiempo en esta cola logra que los usuarios sean indiferentes entre hacer la fila y esperar directamente el servicio *all-stop*. Este análisis permite estudiar las filas y tiempos de espera esperados para los usuarios del sistema para diferentes frecuencias del servicio expreso. La figura 1 muestra las curvas construidas en Larrain y Muñoz (2019).

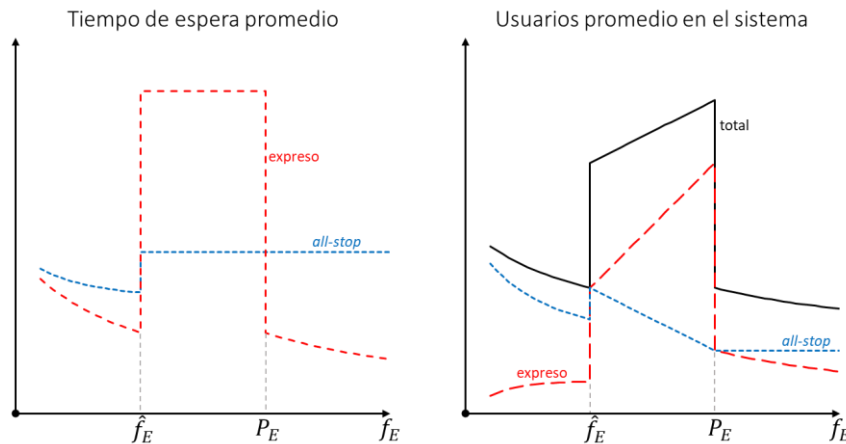


Figura 1: Esperas y número de usuarios en el sistema para la DZ teórica.

En lo que sigue, nos referiremos al modelo de Larrain y Muñoz (2019) como la DZ teórica (en contraposición a la DZ simulada que construiremos a continuación). La DZ teórica está sujeta a algunos supuestos que será necesario considerar al comparar sus resultados con los de la simulación. En primer lugar, la DZ teórica asume que los buses descargan pasajeros de la fila en forma continua, y no por lotes como sucede en la realidad (y en la simulación). Segundo, los procesos de llegada al sistema se asumen todos determinísticos y regulares, como se mencionó antes. Finalmente, esta DZ solo permite estudiar escenarios donde el servicio *all-stop* posee capacidad holgada. Este último supuesto permitió simplificar el modelo de comportamiento en el análisis, ya que permite asumir que por defecto a los usuarios flexibles les conviene ponerse en la fila del expreso, y decidir cuando pasa el *all-stop* si les conviene tomarlo o quedarse en la fila. Esto claramente no es válido si existe la posibilidad de que haya fila por el *all-stop*.

### 3.2. Estructura de la simulación.

La simulación pretende representar el mismo escenario abordado en la DZ y contempla tres elementos principales: una estación, personas y buses. La estación (que se ubica en el extremo aguas arriba del corredor, y posee capacidad ilimitada para albergar pasajeros) corresponde a un sistema con una entrada y dos salidas: una para abordar al servicio expreso y la otra para el *all-stop*. Las personas en la estación escogen formarse en una de las dos colas, en frente a la puerta del servicio que desean abordar.

Tanto buses como personas llegan a la estación siguiendo un proceso de llegada discreto a nivel individual. Cada vez que un bus sale de su puerta, se lleva a un grupo de personas de su fila, sin exceder la capacidad del bus ( $b_i$ ). Asumimos para estos experimentos que el tiempo de descarga de una cola es despreciable. La estructura general de la simulación se resume en la figura 2.

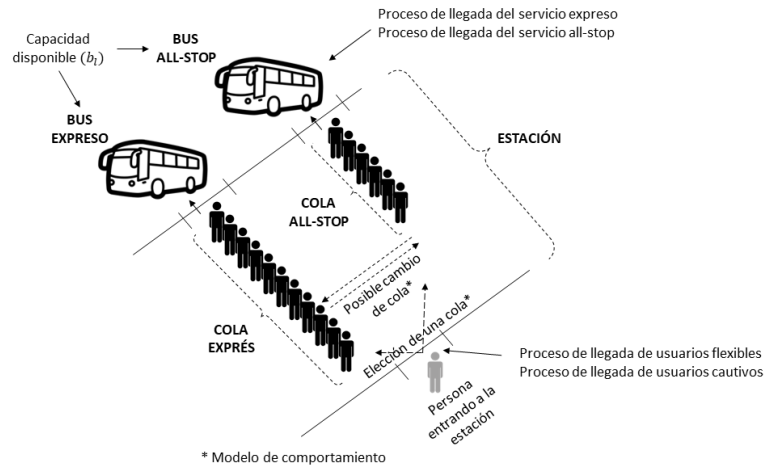


Figura 2: Esquema general del paradero a simular.

### 3.3. Modelo de comportamiento de usuarios.

El modelo de comportamiento utilizado en la simulación es sencillo: cada vez que un usuario debe escoger una fila (ya sea al llegar a la estación, o al evaluar un cambio de fila) el usuario escogerá la fila cuyo tiempo estimado de viaje (TEV) completo (espera más viaje en vehículo) sea la menor posible. Es decir, el usuario escogerá la fila del servicio  $l$  cuyo TEV esperado  $TEV_l = w_l + t_l$  sea menor.

El tiempo de viaje en vehículo,  $t_l$ , lo asumimos como conocido por cada individuo y constante para cada servicio. Sin embargo, el tiempo de espera  $w_l$  estimado por el pasajero dependerá de diversos factores: la frecuencia del servicio, el largo de la fila, la capacidad de los buses, la información sobre el tiempo de pasada del último bus, y la disponibilidad de información externa. Si los usuarios (o algunos de ellos) cuentan con información externa que predice en tiempo real la llegada de los próximos buses, entonces el tiempo de espera  $w_l$  se vuelve un dato exógeno y preciso.

En el caso en que el usuario no cuente con información externa sobre el tiempo de espera, nuestro modelo asume que lo estima como:

$$w_l = \underbrace{(1 - \kappa_l) \cdot \frac{0,5}{f_l} + \kappa_l \cdot \left(\frac{1}{f_l} - \pi_l\right)}_{\text{tiempo hasta el primer bus}} + \underbrace{\frac{\left\lceil \frac{s_l}{b_l} \right\rceil - 1}{f_l}}_{\text{tiempo restante}} \quad (1)$$

En la expresión (1) el tiempo de espera en la fila del servicio  $l$  para un usuario en particular es estimado en dos partes: el tiempo de espera hasta que pase el primer bus asociado a esa fila, y el tiempo restante entre ese primer bus y el bus donde el usuario estima que podrá subirse. El tiempo de espera para el primer bus considera si el usuario ha visto pasar el último bus del servicio  $l$  o no, representado por el parámetro  $\kappa_l$ . Si el usuario no ha visto pasar el último bus ( $\kappa_l = 0$ ) el tiempo de espera para el primer bus es estimado como la mitad del intervalo promedio del servicio. Si el usuario vio pasar un bus ( $\kappa_l = 1$ ) hace  $\pi_l$  unidades de tiempo, el tiempo esperado hasta el primer

bus lo estima como  $1/f_l - \pi_l$ . El tiempo restante de espera es simplemente el intervalo promedio,  $1/f_l$ , multiplicado por el número de buses estimados para lograr abordar, que dependen de  $s_l$ , el número de personas delante del usuario en la fila, y de  $b_l$ , la capacidad de los buses tipo  $l$ .

El modelo hasta acá descrito es determinístico: dos usuarios, en igualdad de condiciones, siempre evaluarán las filas con los mismos costos, y por ende tomarán la misma decisión. Para considerar heterogeneidad en los usuarios incorporamos una componente de aleatoriedad en sus decisiones, es decir, podemos decir que el tiempo total estimado por un usuario para una fila  $l$  tendrá la forma:

$$TEV_l = \varepsilon \cdot w_l + t_l \quad (2)$$

En esta expresión  $\varepsilon$  corresponde a un error multiplicativo que se aplica a la espera. En nuestra simulación asumimos que este error distribuye normal con media 1 y desviación estándar  $\sigma$  dependiente del escenario estudiado.

### 3.4. Modelo de cambios de fila.

En nuestra simulación, un usuario evaluará y escogerá una fila en dos posibles eventos: cuando el usuario llega a la estación, o cuando un bus llega a una puerta. El proceso de cambio de fila se ilustra en la figura 3.

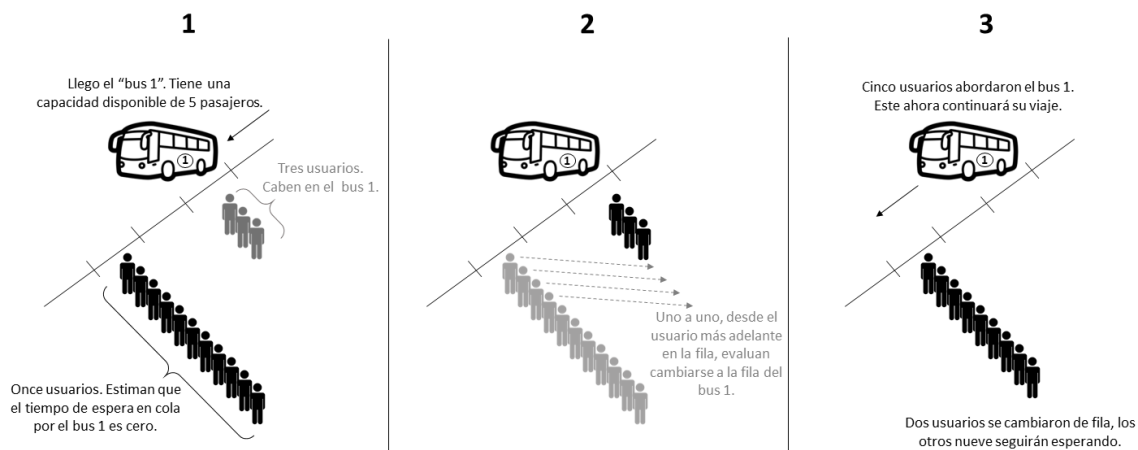


Figura 3: Proceso de cambio de fila.

Para modelar el proceso de cambio de fila es relevante definir un orden en el que los usuarios toman esta decisión, de forma de evitar movimientos en bloque donde los usuarios puedan quedar peor de lo que estaban con el cambio. Para evitar esta situación asumiremos que los usuarios evalúan cambiarse comenzando por quien está primero en la fila y terminando por quien está último, y que este proceso ocurre una vez que los usuarios de la fila respectiva que se pueden subir al bus ya lo han hecho.

La política descrita antes afecta a qué usuarios en definitiva son los que se benefician con un cambio de fila, pero no debería generar un gran impacto en indicadores globales asociados al desempeño global de la estación ya que si el usuario que se cambia de fila es uno y otro, no afectan los largos de las filas ni las esperas promedio.



## 4. EXPERIMENTOS Y RESULTADOS

En esta sección reportamos y analizamos los resultados de los experimentos simulados. Comenzaremos por describir los parámetros operativos y de demanda del corredor utilizado como caso base para conducir nuestros experimentos. Luego, detallamos los resultados de tres experimentos:

- **Experimento 1: Caso determinístico, usuarios homogéneos.** Este experimento tiene como fin probar la simulación frente a un escenario que se puede modelar con las fórmulas propuestas en el trabajo original de Larrain y Muñoz (2019). Los resultados de dicho trabajo serán referidos desde ahora como la “DZ teórica”. Este experimento busca validar que la simulación efectivamente logra capturar el efecto de la DZ, a pesar de trabajar con un modelo de comportamiento diferente.
- **Experimento 2: Caso estocástico con usuarios heterogéneos.** Este experimento introduce aleatoriedad en los procesos de llegada de usuarios y de buses. También considera heterogeneidad en los usuarios, por medio de la introducción de un error aleatorio en sus estimaciones de los tiempos de espera. Estos efectos podrían ser estudiados por separado, pero en pruebas preliminares se observó que el efecto de la aleatoriedad en las tasas de llegada al sistema no generaba un impacto importante en los resultados por sí solo.
- **Experimento 3: Información perfecta.** Este tercer experimento busca entender el efecto de la disponibilidad de información en la existencia de la DZ, para diferentes grados de penetración en los usuarios.

Para cada uno de los experimentos se estudia el corredor para diferentes niveles de frecuencia del servicio expreso usando tres indicadores de desempeño: los tiempos de espera, los largos de las colas, y el tiempo total de viaje de los usuarios. Todos los resultados reportados (para cada escenario estudiado) son el promedio de 300 corridas de la simulación.

### 4.1. Descripción del caso base.

La instancia base en la que realizamos nuestros experimentos considera una demanda por viajes de extremo a extremo del corredor (usuarios flexibles) de  $T_F = 900$  pax/h, y una demanda por viajes originados en la estación modelada pero con destino dentro del corredor (usuarios cautivos) de  $T_C = 3.600$  pax/h. El tiempo de viaje entre los extremos del corredor de los servicios all-stop y expreso fueron fijados en 30 y 20 minutos, lo que implica que el expreso ofrece un ahorro de  $\tau = 10$  minutos en tiempo de viaje. Asumiremos que la capacidad remanente con la que llegan los buses a la estación es fija, e igual a  $b_l = 85$  pasajeros. Finalmente, se consideró una frecuencia del servicio *all-stop* (fija) de 60 buses por hora.

### 4.2. Experimento 1: Caso determinístico, usuarios homogéneos.

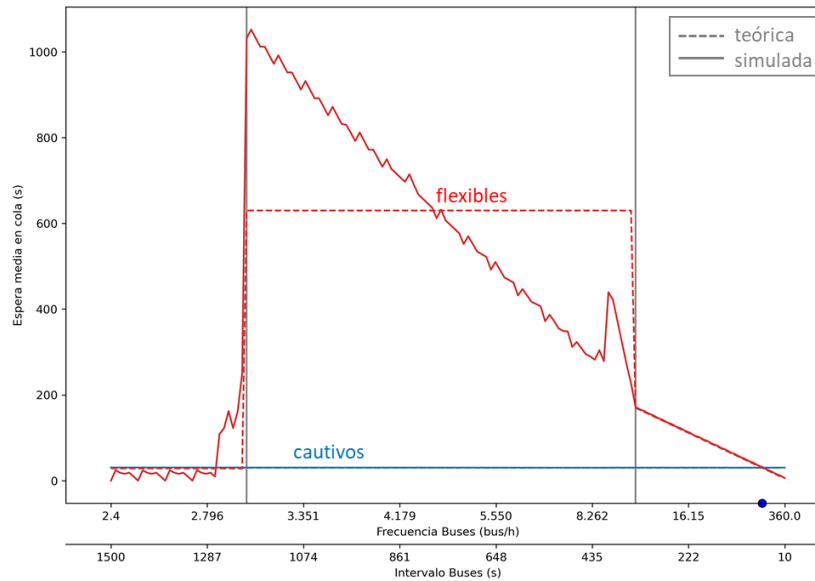


Figura 4: Esperas medias, caso determinístico, usuarios homogéneos.

En la figura 4 se ilustran, para diferentes niveles de frecuencia del servicio expreso, las esperas experimentadas en promedio por los usuarios flexibles (en rojo) y cautivos (en azul). Las líneas sólidas corresponden a los resultados de la simulación, mientras que las líneas punteadas corresponden a los resultados obtenidos según la DZ teórica para las curvas de espera. Las líneas verticales, graficadas en gris, corresponden a los límites de la DZ teórica.

Podemos concluir, en primer lugar, que los bordes de la DZ teórica predicen con gran exactitud los bordes de la DZ simulada. Esto no es una sorpresa para el borde derecho, que proviene simplemente de comparar la capacidad total con la demanda, y que no tendría por qué diferir sustancialmente entre ambos métodos. Sin embargo, la coincidencia del lado izquierdo de ambos métodos es un resultado destacable, ya que son producto de diferentes modelos de comportamiento, y sirve como validación de los modelos de comportamiento usados en ambos casos. La línea horizontal asociada a la espera de usuarios cautivos refleja que para todas las frecuencias del expreso, la cola del servicio all-stop siempre permitió subir al primer bus que pasaba.

Una diferencia que salta a la vista entre la DZ teórica y la simulada es que la primera predice una espera constante dentro de los rangos de la DZ para el servicio expreso, mientras en la segunda la espera promedio presenta una caída sustancial al aumentar la frecuencia del expreso (al interior de la DZ). Sin embargo, podemos notar que la DZ teórica es un buen estimador del promedio de la espera en este rango. Esto se debe a que la fórmula de estimación de las esperas propuesta en Larrain y Muñoz (2019) estima esta espera asumiendo que la tasa de atención de los buses es un proceso continuo. Esto no solo entrega una aproximación razonable del tiempo de espera, sino además evita sobre ajustar los efectos que se producen por patrones de sincronización de los vehículos de ambos servicios, consecuencia de su regularidad en los intervalos.

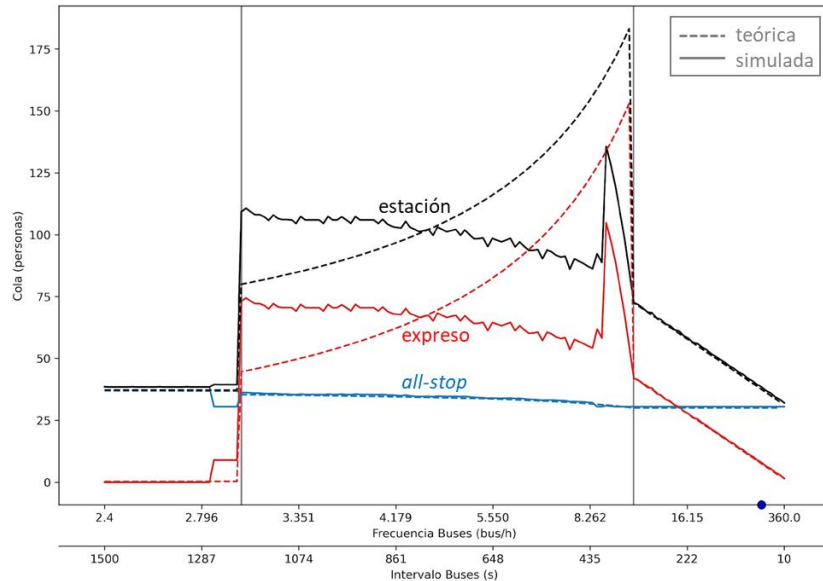


Figura 5: Largos promedio de cola, caso determinístico, usuarios homogéneos.

Los largos promedio de las filas en el sistema se entregan en la figura 5. En esta figura se entregan las curvas simuladas (línea sólida) y teóricas (líneas punteadas) para las filas de cada servicio, y del sistema completo. Acá los resultados parecen divergir, pero para comprender lo que sucede, observemos el *peak* que se forma en torno a  $f_E = 8,5$  buses por hora (cerca del borde derecho de la DZ).

Este *peak* ocurre porque dentro de la DZ los usuarios que se cambian hacia el servicio expreso lo harán si la fila de este servicio es menor que un cierto umbral. En la DZ teórica, este umbral crece en forma continua con la frecuencia del servicio; sin embargo, como el proceso de descarga de la fila es por bloques, lo que hacen los usuarios de la simulación es definir estos umbrales como múltiplos de la capacidad (remanente promedio) de los buses. Es decir, un usuario calcula que le conviene cambiarse si tiene que esperar a lo más un número entero dado de buses, y esto define los umbrales de largo de fila. Como estos umbrales dan saltos discretos conforme crece la frecuencia del expreso, el gráfico de colas (y también de esperas) muestra saltos en su comportamiento.

Estos saltos explican por qué las colas de la DZ simulada parecieran tender a la baja en pequeños tramos, a pesar de que vista en forma más agregada la DZ todavía presenta una tendencia al alza en los largos de fila.

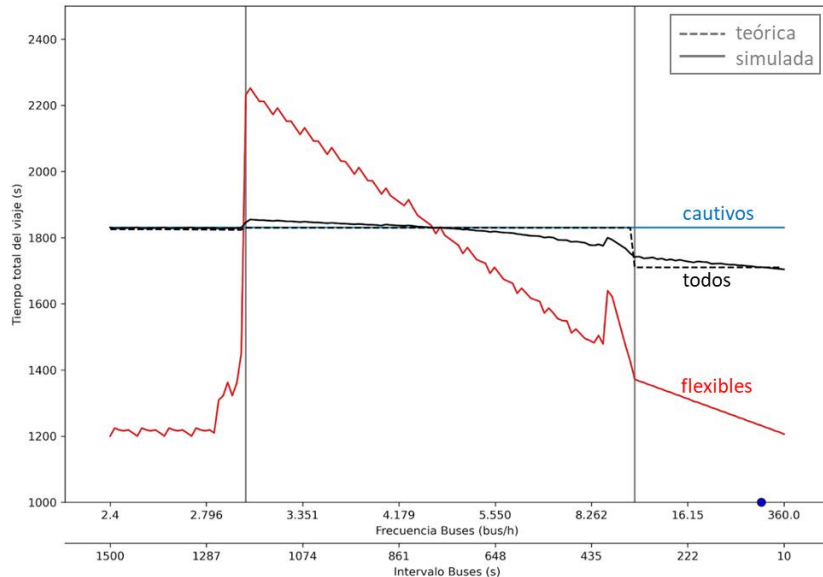


Figura 6: Tiempos totales de viaje, caso determinístico, usuarios homogéneos.

Si observamos los tiempos totales de viaje de los usuarios (figura 6) podemos notar que los tiempos esperados de viaje de los usuarios de esta parada (curva negra) presentan pequeñas divergencias entre la teoría y la simulación, pero el efecto DZ sigue presente. El modelo teórico (que considera tasas continuas de descarga) predice un tiempo total de viaje constante dentro de la DZ ligeramente superior a los 1800 segundos. La simulación presenta tiempos totales para los usuarios flexibles que dependen de la frecuencia bajando desde aproximadamente 2200 hasta 1400 segundos. Así, la simulación muestra que el aumento en tiempo de viaje para los usuarios flexibles es aún mayor que en los resultados teóricos en el borde izquierdo de la DZ. Esto refuerza la advertencia respecto de subir la frecuencia a un servicio expreso de baja frecuencia cuando los usuarios no esperan por él. Los usuarios seguramente transmitirán que si la frecuencia fuese un poco mejor ellos esperarían por un servicio que valoran cuando lo toman. Si la autoridad accede con ese aumento de frecuencia pequeño, gatillaría un impacto especialmente desastroso para el sistema.

#### 4.3. Experimento 2: Caso estocástico, usuarios heterogéneos.

En este experimento se introduce aleatoriedad en los procesos de llegada. El proceso de llegada de pasajeros a la estación es modelado como un proceso de Poisson con la tasa respectiva, y el proceso de llegada de buses a la estación es aleatorizado agregando un error normal con desviación estándar igual a un 15% del intervalo promedio entre dos buses. Además, se incorpora el error  $\sigma$  presentado en (2) en la estimación de los tiempos de espera de forma de generar heterogeneidad en el comportamiento de los usuarios, como se detalla en la metodología. La simulación fue ejecutada usando  $\sigma = 0$  (usuarios 100% homogéneos), y de 10, 20, y 30%.

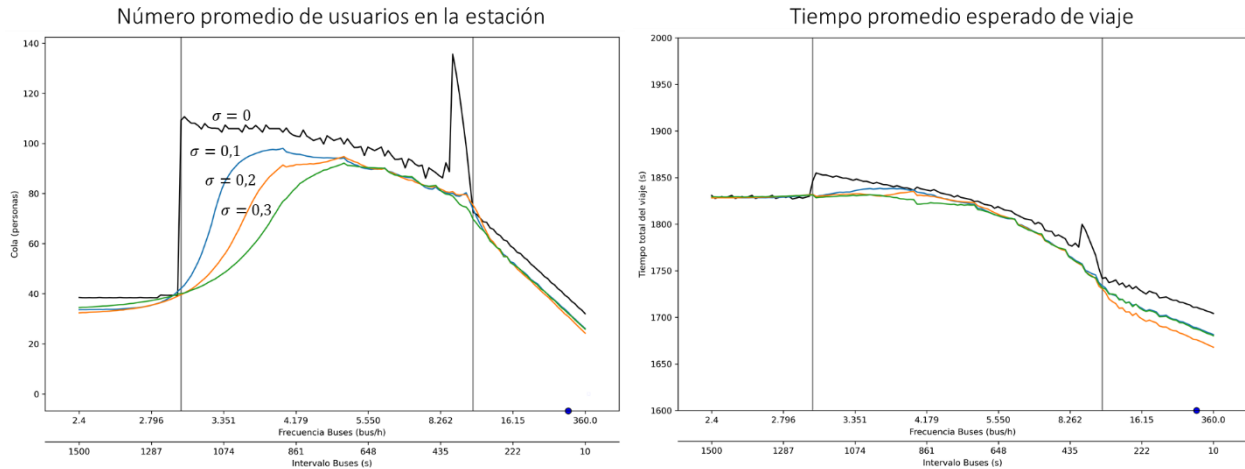


Figura 7: Resultados caso estocástico, usuarios homogéneos.

En la figura 7 se entregan los resultados de la simulación para el caso estocástico con usuarios heterogéneos. En estos gráficos se entrega en negro la curva obtenida en la simulación determinística del primer experimento. A grandes rasgos, es posible observar que la introducción de heterogeneidad en el comportamiento de usuarios tiende a suavizar las curvas, y que entre mayor la variabilidad de este comportamiento, el *peak* del largo máximo de las colas en la estación ocurre para frecuencias mayores.

Esto se puede explicar recordando que el límite izquierdo de la DZ ocurre cuando los pasajeros flexibles comienzan a ver un beneficio en comprometerse a esperar al servicio expreso, en lugar de mantener sus opciones abiertas. En un ambiente determinístico, esto sucede en el umbral dado por  $\hat{f}_E$ ; en cambio, si consideramos heterogeneidad en el comportamiento, existe una probabilidad de que los usuarios estén dispuestos a comprometerse con el servicio expreso, y esta aumenta entre mayor es la frecuencia de dicho servicio, pero el proceso de cambio de preferencias se da en forma gradual, como reflejan los gráficos.

Por otra parte, podemos observar que los tiempos totales de viaje no se ven muy afectados por la heterogeneidad de los usuarios. También es posible observar que los tiempos de espera no aumentan en este caso con el comienzo de la DZ. Esto claramente significa que el efecto en el caso estocástico (y por lo tanto más realista) es más tenue que lo predicho por la teoría determinística, pero tampoco lo desacredita: recordemos que un aumento de frecuencia implica un costo por parte del operador que no se está viendo reflejado en ningún efecto positivo en el sistema.

#### 4.4. Experimento 3: Información perfecta.

En este experimento tomamos como base el escenario del experimento 2 con  $\sigma = 10\%$ , y definimos cuatro niveles de penetración de una tecnología que entrega al usuario una predicción perfecta del tiempo de espera asociado a cada cola. Los niveles de penetración considerados fueron de 0, 33,3, 66,6, y 100% de los usuarios.

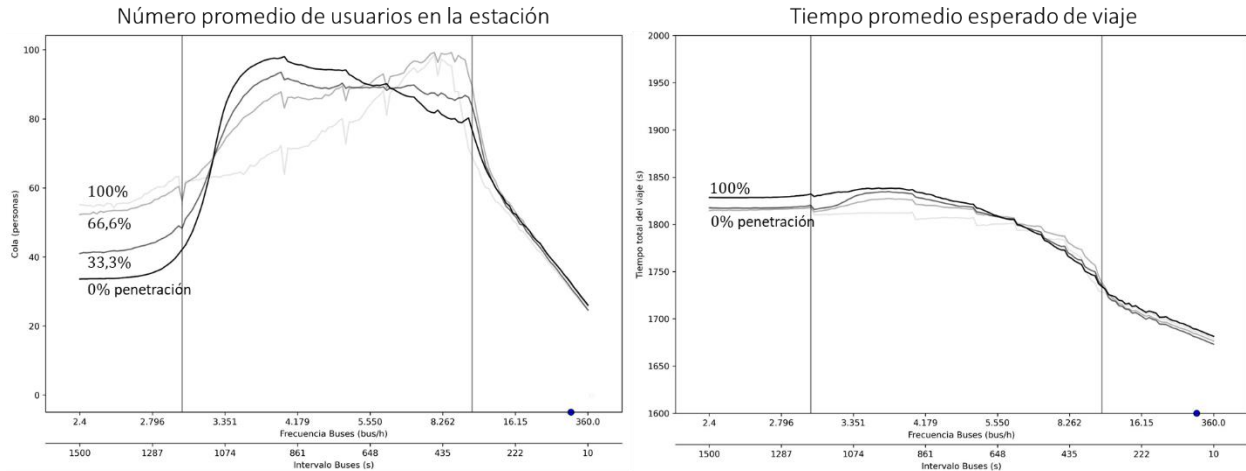


Figura 8: Resultados para diferentes niveles de penetración de información perfecta.

En la figura 8 se entrega el número promedio de personas en el sistema, y el tiempo total de viaje de los usuarios, para diferentes niveles de penetración de la tecnología que ofrece información perfecta respecto de los tiempos de espera. Una primera tendencia que llama la atención es que la disponibilidad de información pareciera adelantar el comienzo de la DZ, aumentando las colas en el sector donde las frecuencias del servicio expreso son todavía bajas. Sin embargo, lo que sucede es que cuando existe información perfecta, los usuarios que llegan justo antes de que pase un bus expreso (independiente de cuán baja sea su frecuencia) se pondrán en su fila, cosa que sin información perfecta nadie haría.

Otro efecto interesante que estos gráficos reflejan es que entre los bordes de la DZ los diferentes niveles de penetración parecieran cambiar la inclinación o la posición del *peak* de los largos de cola, pero sin cambiar el nivel promedio del indicador dentro de la DZ. Si a esto sumamos el hecho de que los tiempos totales del sistema tampoco muestran un beneficio asociado a un alza en la frecuencia de los buses expresos, podemos concluir que la disponibilidad de información perfecta no es una buena herramienta para combatir la DZ, a pesar de lo que la intuición podría haber dictado.

## 5. CONCLUSIONES

El problema de diseño de servicios para transporte público es un problema desafiante y de gran relevancia en un mundo cada vez más urbanizado, y en el que es urgente encontrar formas de promover un estilo de vida sustentable en lo medioambiental. A pesar de que muchos autores han estudiado este problema con o sin foco en los servicios expresos, hasta donde tenemos conocimiento los modelos para apoyar el diseño de servicios trabajan desde una perspectiva agregada, que no captura los efectos que pueden tener las decisiones de los usuarios en las estaciones del sistema, sobre todo en presencia de congestión.

En este trabajo hemos mostrado que el fenómeno de la *danger zone* de los servicios expresos es un riesgo concreto, y que puede surgir en condiciones realistas de una operación de transporte público. Por medio de una micro simulación, que modela a nivel individual las decisiones de los usuarios

en una estación, y que considera las filas que los pasajeros hacen por abordar a un servicio, hemos confirmado que aumentar la capacidad de un servicio no necesariamente acarrea beneficios para los usuarios de este servicio, e incluso puede deteriorar la calidad de servicio entregada.

Un resultado específico de este trabajo que también puede resultar de utilidad en otros contextos es que es posible modelar la congestión de un paradero asumiendo que los usuarios se pueden cambiar entre filas cuando existen incentivos a hacerlo, y que este modelo de comportamiento entrega resultados similares a los predichos por un modelo de líneas atractivas como el utilizado para desarrollar la teoría de la DZ original.

El trabajo acá reportado es solo el comienzo de lo que la simulación desarrollada nos puede permitir analizar sobre el fenómeno de la DZ. Nuestra simulación y nuestro nuevo modelo de comportamiento abre las puertas para estudiar la DZ en presencia de corredores con múltiples servicios y estaciones (esto era simplemente inabordable con el modelo teórico anterior), o incluso a nivel de red. Esta simulación también permitirá estudiar el problema con un foco en la optimización de frecuencias, en donde las frecuencias de todos los servicios se consideran variables. Resulta de particular interés estudiar si tiene sentido práctico, en un sistema correctamente optimizado, entregar para los servicios expresos frecuencias menores al umbral derecho de la DZ. Resulta muy atractivo entender también el impacto que pueda tener la capacidad limitada de las estaciones por albergar simultáneamente pasajeros en fila. Esto, pues cuando la capacidad está colmada, se generaría una cola en su exterior que probablemente sería FIFO afectando también a los usuarios de los demás servicios.

En función de lo reportado, creemos que es necesario seguir estudiando los fenómenos que ocurren en los sistemas de transporte público a nivel microscópico. A la luz de nuestros resultados, resulta recomendable verificar mediante simulación cualquier resultado construido con un modelo de diseño de servicios, ya que los efectos no capturados por los modelos convencionales podrían explicar en parte la mala calidad de servicio percibida en algunos sistemas BRT alrededor del mundo. Entender mejor la *danger zone* de los servicios expresos nos llevará a hacer mejores diseños de redes BRT, con un mejor entendimiento de los fenómenos locales que ocurren en las paradas, que si no son incorporados correctamente pueden llevar al deterioro de un sistema aparentemente bien diseñado.

## BIBLIOGRAFÍA

Bailey, N. (1954) On Queueing Processes with Bulk Service. **Journal of the Royal Statistical Society, Series B**, Vol 16, No. 1, 80-87.

Bouzaiene-Ayari, B. Gendreau, M. Nguyen, S. (2001) Modeling Bus Stops in Transit Networks: A Survey and New Formulations. **Transportation Science**, Vol 35, No. 3, 304-321.

Chriqui, C., Robillard, P. (1975) Common bus lines. **Transportation Science**, 9 115-121.

Dial, R. (1967) Transit Pathfinder Algorithm. **Highway Res. Record** 205 67-85.

- Ford, L. y Fulkerson, D. (1962) Flows in Networks. **Princeton University Press**.
- Hoque, M., Sultana, S., Nasrin, S. (2016) Towards Promoting Efficient BRT in Dhaka. 3rd **Conference on Sustainable Urban Mobility**, 26-27 Mayo 2016, Volos, Grecia.
- Larrain, H., Giesen, R., Muñoz, J.C. (2010) Choosing the right express services for bus corridor with capacity restrictions. **Transportation Research, Record: J. Transp. Res. Board** 2197, 63–70.
- Larrain, H., Muñoz, J.C. (2016) When and where are Limited-stop Bus Services Justified? **Transportmetrica A: Transport Science**, 12(9), 811–831.
- Larrain, H., Muñoz, J.C. (2019) The danger zone of express services: When increasing frequencies can deteriorate the level of service. **Transportation Research Part C: Emerging Technologies**, 113, 213-227.
- Leiva, C., Muñoz, J.C., Giesen, R., Larrain, H. (2010) Design of limited-stop services for an urban bus corridor with capacity constraints. **Transportation Research Part B** 44 (10), 1186–1201.
- Merkert, R., Mulley, C., Hakim, M.M. (2017) Determinants of bus rapid transit (BRT) system revenue and effectiveness—A global benchmarking exercise. **Transportation Research Part A: Policy and Practice**, 106, 75-88.
- Nguyen, S., & Pallottino, S. (1989) Hyperpaths and shortest hyperpaths. **In Combinatorial Optimization** (pp. 258-271). Springer, Berlin, Heidelberg.
- Shen, M., Gu, W., Hu, S., Cheng, H. (2019) Capacity approximations for near-and far-side bus stops in dedicated bus lanes. **Transportation Research Part B: methodological**, 125, 94-120.
- Spiess, H y Florian, M. (1989) Optimal Strategies: A New Assignment Model for Transit Network. **Transportation Research, B** 83-102.
- Wang, D.Z., Nayan, A., Szeto, W.Y. (2018) Optimal bus service design with limited stop services in a travel corridor. **Transportation Research Part E: Logistics and Transportation Review**, 111, 70-86.