

DISEÑO DE SERVICIOS SOBRE UN CORREDOR BRT CONSIDERANDO LA CAPACIDAD LIMITADA DE BUSES Y ESTACIONES

Iván Daniel Moreno, Pontificia Universidad Católica de Chile, idmoreno@uc.cl
Daniel Soto, Pontificia Universidad Católica de Chile, dusoto@uc.cl
Homero Larrain, Pontificia Universidad Católica de Chile, homero@ing.puc.cl
Juan Carlos Muñoz, Pontificia Universidad Católica de Chile, jcm@ing.puc.cl
Sara Larraín, Pontificia Universidad Católica de Chile, larrain.sarita@gmail.com

RESUMEN

Algunos modelos de optimización de frecuencias consideran la capacidad de los buses para transportar pasajeros, pero no se ha considerado aún la capacidad de las estaciones para atender buses simultáneamente. En sistemas que operan a capacidad, esta simplificación resta credibilidad al modelo, y puede conducir a soluciones distorsionadas. No es difícil imaginar instancias en las cuales no existen soluciones de transporte para movilizar toda la demanda de transporte. Este trabajo propone un método para encontrar un vector de frecuencias que minimice el costo social del corredor respetando dichas capacidades. El proceso de optimización consiste en un algoritmo de búsqueda local equipado con tres metaheurísticas que mejoran su rendimiento. Se observó que las soluciones con las restricciones de capacidad activas eran en promedio un 41,7% más costosas que aquellas donde dichas restricciones estaban relajadas. El algoritmo logró la factibilidad en el 66% de las instancias probadas, y se logró una mejora promedio de 9,67% en los costos sociales respecto del vector de frecuencias inicial.

Palabras clave: Optimización de Frecuencias, Algoritmos de Búsqueda Local, Problema de Diseño de Redes de Transporte.

ABSTRACT

Some frequency optimization models consider the capacity of buses to carry passengers, but the capacity of stations to serve buses simultaneously has not yet been considered. In systems operating at capacity, this simplification detracts credibility from the model, and can lead to distorted solutions. It is not difficult to imagine instances in which there are no transportation solutions to mobilize the entire demand. This paper proposes a method to find a vector of frequencies that minimizes the social cost of the corridor while respecting these capacities. The optimization process consists of a local search algorithm equipped with three metaheuristics that improve its performance. It was observed that solutions with active capacity constraints were on average 41.7% more expensive than those with relaxed capacity constraints. The algorithm achieved feasibility in 66% of the tested instances, and an average improvement of 9.67% in social costs with respect to the initial frequency vector was achieved.

Keywords: Frequency Optimization, Local Search Algorithms, Transport Network Design Problem.

1 INTRODUCCIÓN

La movilidad es una actividad que impacta en la calidad de vida de los ciudadanos, no solo por el tiempo de viaje que gastamos en ella, sino porque afecta la seguridad vial, deteriora la calidad ambiental de las ciudades, influye en su desarrollo económico, y consume una gran cantidad de energía. A medida que las ciudades crecen, ofrecer transporte público de alta calidad se vuelve crucial para proveer accesibilidad y movilidad y mantener los niveles de vida y prosperidad económica de los ciudadanos (Institute for Transportation & Development Policy, 2017). Dentro de los modos de transporte disponibles, el BRT se destaca por tener bajos costos de inversión y por ofrecer transporte público de alta calidad para todas las personas.

La determinación de frecuencias de cada servicio de transporte público es una de las etapas clave en el Problema de Diseño de Servicios (Ceder y Wilson, 1986). Asignar frecuencias eficientes a cada uno permite minimizar los costos sociales, considerando simultáneamente los tiempos de los usuarios y los costos de operación. Este problema es especialmente desafiante durante las horas de máxima demanda. En estos periodos, identificar una solución inicial factible para el problema puede no ser trivial, pues simplemente aumentar la frecuencia para poder transportar toda la demanda puede exceder la capacidad de las estaciones para recibir buses. Cuando la capacidad de las estaciones se agota y las bahías de detención de una estación están llenas, la frecuencia del corredor se restringe a la frecuencia que es capaz de atender dicha parada, y si se excede se formarán colas de buses para entrar (Institute for Transportation & Development Policy, 2017). En otras palabras, la capacidad de los buses supone que debe ofrecerse una frecuencia mínima que garantice que dicha capacidad no se excederá, mientras que la capacidad de las paradas impone una cota máxima para la frecuencia. En algunas instancias, ambas cotas podrían tocarse y no existir una solución para la demanda requerida de transporte.

En este contexto, los servicios que solo visitan un subconjunto de paradas sobre el corredor pueden resultar cruciales para alcanzar una solución que satisfaga las restricciones del problema. Así, los modelos deben considerar la capacidad de los buses a la hora de asignar a los pasajeros a una ruta. Sin embargo, añadir esta restricción directamente en una formulación que minimice los costos sociales no permite representar adecuadamente el comportamiento espontáneo de los usuarios, quienes debieran escoger una combinación de servicios para alcanzar sus destinos que minimice su costo individual y no necesariamente lo que beneficia al sistema global. Este problema ha mostrado ser difícil de resolver, por lo que se han propuesto heurísticas para añadir indirectamente estas restricciones (Leiva et al., 2010).

Entonces, la frecuencia y el comportamiento espontáneo de los usuarios deben ser abordados simultáneamente para alcanzar una solución a la vez factible, confiable, y eficiente. En este trabajo proponemos un algoritmo de búsqueda local que optimiza las frecuencias de los servicios, considerando el comportamiento racional de los usuarios y sin exceder la capacidad de los buses de transportar pasajeros, ni la capacidad de las paradas de atender buses. El algoritmo recibe como

entrada un conjunto de servicios con su esquema de paradas fijo; estos pueden ser servicios que se detengan en todas las estaciones (*all-stop services*), servicios zonales, o servicios expresos (*limited-stop services*). La metodología está estructurada en tres momentos: en primer lugar, se construye una solución inicial a partir de algunos criterios de demanda; en la Fase I se busca hacer factible dicha solución respecto a las capacidades, y en la Fase II se busca mejorar los costos sociales sin perder factibilidad.

2 TRABAJOS PREVIOS

Asignar frecuencias es una de las cinco etapas del Problema de Diseño de Redes de Transporte (TNDP, o *Transport Network Design Problem*) formulado por Ceder y Wilson (1986). El objetivo de este problema suele ser encontrar un conjunto de servicios que considere simultáneamente el tiempo de viaje de los usuarios y el costo de los operadores. En el problema de optimización de frecuencias, el énfasis ha estado en atacar la complejidad computacional usando heurísticas que permitan encontrar soluciones eficientes en tiempos razonables. En ese sentido, se destacan los trabajos de Gao et al. (2004) y Park (2005).

La optimización de frecuencias está ligada a otras etapas del TNDP, según el nivel de frecuencia a ofrecer: en sistemas de baja frecuencia, está relacionado con la determinación de los horarios de pasada mientras que, en sistemas con altas frecuencias, está relacionado con la elección de las paradas del servicio. Este trabajo se centra en sistemas de altas frecuencias, pues se abordan casos en que la capacidad de los vehículos o de las paradas se vuelven una restricción activa producto de la elevada demanda. Si bien la elección de paradas para un bus es un problema de nivel táctico y la determinación de frecuencias es más bien de nivel operativo (Guihaire & Hao, 2008), la elección de ruta de los pasajeros no se puede predecir sin conocer las paradas y la frecuencia del servicio; este es el motivo por el que ambos problemas están estrechamente relacionados. Algunos trabajos que han abordado simultáneamente ambos problemas son Carrese y Gori (2002), Lee y Vuchic (2009), Szeto y Wu (2011), o Larrain (2013).

Ninguno de los trabajos nombrados anteriormente considera la capacidad de buses y/o estaciones. El primer trabajo en hacerlo es Leiva et al. (2010) que aborda la restricción de capacidad de los buses. A partir de un conjunto de servicios candidatos, se asignan frecuencias no-nulas a aquellos servicios que minimizan los costos sociales del corredor. Para incluir la capacidad de los vehículos sin impactar la asignación espontánea de los usuarios, se incluyen cotas inferiores crecientes a las frecuencias de los servicios con déficit hasta que la solución al modelo de optimización entrega una solución factible al problema global sin que las restricciones de capacidad de los buses estén activas. Similarmente, Bagloee y Ceder (2011) añaden una cota de la frecuencia mínima requerida por cada servicio para satisfacer la demanda espontánea de los usuarios.

La capacidad de los buses también puede considerarse en el modelo de elección de ruta. Se han formulado modelos que consideran la congestión asociada a que los usuarios no siempre pueden tomar el primer bus que pasa por estar lleno, como el propuesto en de Cea y Fernandez (1993) o más recientemente en Larrain et al. (2021). Estos modelos de asignación suelen ser el *nivel inferior* de metodologías modulares donde se llama iterativamente al modelo de elección de ruta para conocer el costo de la solución, como es el caso de Soto et al. (2017) o Szeto & Jiang (2014), entre

otros. La capacidad puede añadirse directamente como restricción, o puede sumarse una función de costos por congestión que considera esperas adicionales cuando la capacidad se rebasa.

No sabemos de trabajos que hayan considerado la capacidad de las estaciones en el TNDP. En Larrain y Muñoz (2008) se propone un modelo de asignación en el que el tiempo de embarque y desembarque depende del flujo del corredor. Un estudio más completo sobre la capacidad de las estaciones y su efecto en la operación del transporte público puede encontrarse en Fernandez y Planzer (2002), Fernández (2010), o Tirachini et al. (2014). Dos revisiones más completas alrededor del TNDP pueden encontrarse en Desaulniers y Hickman (2007) y Guihaire y Hao (2008).

3 FORMULACIÓN DEL PROBLEMA

En esta sección planteamos un modelo matemático cuyo objetivo es determinar las frecuencias que conviene asignar a un conjunto de distintos servicios que operan en un corredor, de forma de reducir los costos sociales del sistema, respetando las dos familias de restricciones de capacidad discutidas en las secciones anteriores.

Con este fin, definiremos un corredor como una secuencia de estaciones N . Sobre este corredor existe una demanda por viajes dada por T_w , donde $w \in W$ es un par origen/destino consistente en dos paradas $i, j \in N$. Además, sobre este corredor se define un conjunto de servicios candidatos $l \in L$, donde l es una lista de las estaciones en las que se detiene el servicio. Estos servicios difieren entre sí por el conjunto de paradas que visitan, e incluye un servicio *all-stop* en cada sentido. Para cada uno de estos servicios se conoce tt_l^w , que corresponde al tiempo de viaje el servicio l ofrece para el par w . Es importante destacar que este tiempo de viaje no depende de la asignación de viajes. Si el servicio l no conecta ambos extremos del par w (por no visitar alguno de estos dos) definiremos $tt_l^w = \infty$.

Hemos definido un servicio como una secuencia de paradas contenidas en el corredor, y que no necesariamente termina su servicio en el mismo paradero donde comienza. Definiremos L_i^+ y L_i^- como los conjuntos de servicios que comienzan y que terminan su recorrido en la parada i . Definiremos además L_i como el conjunto de servicios que se detienen en la parada i .

El objetivo del problema entonces es encontrar un vector de frecuencias $F = \{f_1, \dots, f_{|L|} \mid f_i \geq 0\}$ que minimice los costos del sistema, Z . Los costos del sistema se definen como la suma de los costos de los usuarios y los costos del operador:

$$\begin{aligned}
 Z &= \sum_{l \in L} c_l f_l + \sum_{w \in W} T_w (\theta_{WT} \cdot WT_w(F) + \theta_{TT} \cdot TT_w(F)) \\
 WT_w &= \sum_{w \in W} T_w \frac{\lambda}{\sum_{l \in L} f_l^w} \\
 TT_w &= T_w \frac{\sum_{l \in L} tt_l^w f_l^w}{\sum_{l \in L} f_l^w}
 \end{aligned} \tag{1}$$

El primer término de esta función de costos corresponde a los costos del operador, donde c_l representa el costo por unidad de frecuencia de un servicio l . El segundo término corresponde a los costos del usuario, y se calculan como la suma de los tiempos de espera más los tiempos de viaje de los usuarios. En este término es posible incorporar también un costo por transbordo, pero por simplicidad este análisis no se considerará transbordos en el modelo de elección de ruta de los pasajeros. En esta expresión los parámetros θ_{WT} y θ_{TT} son los valores sociales del tiempo de espera y de viaje de los usuarios. Las funciones $WT_w(F)$ y $TT_w(F)$ entregan el tiempo de espera y de viaje para un par w dado un vector de frecuencias F .

Dos consideraciones para conocer los tiempos $WT_w(F)$ y $TT_w(F)$. Primero, es necesario modelar la forma en que los pasajeros eligen sus rutas a partir de F , es decir, el proceso de asignación de pasajeros a la red. Dado que en este contexto existen servicios alternativos para realizar un viaje, supondremos que los pasajeros que viajan en un par w escogen un conjunto de líneas atractivas para realizar su viaje sin transbordos, tomando la primera de ellas que pase. Este subconjunto de líneas atractivas, $L_w \subseteq L$, será el que minimice el tiempo esperado de viaje en el par w , dado por la suma del tiempo esperado de viaje en vehículo más la espera en estación. La ecuación (2) permite conocer el tiempo esperado de viaje del par w (ETT_w) utilizando el primer bus que pase del conjunto de líneas atractivas.

$$ETT_w = \frac{\sum_{l \in L_w} f_l \cdot tt_l^w + \lambda}{\sum_{l \in L_w} f_l} \quad (2)$$

En esta expresión, el parámetro λ corresponde a una constante de proporcionalidad entre el intervalo promedio entre los buses del conjunto de servicios, y la espera resultante, y refleja la distribución de probabilidades detrás del proceso de llegada de buses a la estación (para Poisson se usa $\lambda = 1$). La solución de este problema se puede obtener en forma eficiente aplicando el algoritmo propuesto por Chriqui y Robillard (1975).

Segundo, es necesario exigir que la asignación de viajes sea coherente con una decisión racional a nivel individual. Sin ello, el modelo podría asignar viajeros a ciertos servicios solo con el fin de minimizar costos sociales pero sin que estas decisiones sean las que los individuos hubieran tomado.

Para que la asignación de frecuencias resulte en una operación factible en estado de régimen, debemos exigir además que en cada nodo la frecuencia agregada de los servicios que terminan su recorrido en ese nodo iguale a la de los servicios que comienzan su servicio allí:

$$\sum_{l \in \mathcal{L}_i^+} f_l = \sum_{l \in \mathcal{L}_i^-} f_l, \quad \forall i \in \mathcal{N} \quad (3)$$

Conocida la elección de ruta de los pasajeros, podemos verificar si la oferta entregada de cada servicio es suficiente para transportar la demanda que captura. Si definimos $T_i^l(F)$ como el flujo de pasajeros que utiliza el servicio l al salir de la estación i cuando las frecuencias son F , esta verificación se expresa con la siguiente desigualdad:

$$T_i^l(F) \leq cap_l \cdot f_l, \quad \forall l \in L, \forall i \in N \quad (4)$$

En esta ecuación, cap_l representa la capacidad que tienen los buses de la línea l . Para calcular $T_i^l(F)$, se asumirá que los T_w usuarios que viajan en un par w y que poseen un conjunto L_w de servicios atractivos se reparten entre estos servicios en forma proporcional a sus frecuencias. Para simplificar la notación, definimos W_i^l como el conjunto de pares w que comienzan o pasan por el nodo i , donde el servicio l es atractivo. Con esto, podemos expresar $T_i^l(F)$ como:

$$T_i^l(F) = \sum_{w \in W_i^l} T_w \cdot \frac{f_l}{\sum_{l' \in L_w} f_{l'}} \quad (5)$$

Además, la restricción de capacidad en estaciones exige que la frecuencia agregada del conjunto \mathcal{L}_i formado por los buses de las líneas que se detienen en la estación i no supere un umbral dado δ_i :

$$\sum_{l \in \mathcal{L}_i} f_l \leq \delta_i, \quad \forall i \in N \quad (6)$$

Con esto, el problema puede ser formulado como la minimización de los costos sociales (1), sujeto a las restricciones de conservación de frecuencias (3), capacidad en los buses y estaciones (4 y 6), y la no negatividad del vector F . En conclusión, el modelo asignará la flota disponible a un subconjunto de los servicios disponibles con el objetivo de minimizar los costos sociales, pero permitiendo a cada usuario viajar de la forma que minimice su tiempo de viaje individual.

4 METODOLOGÍA

La metodología propuesta para abordar este problema se estructura en tres fases: la Fase 0 determina una serie de puntos iniciales bajo distintos criterios que son luego forzados a satisfacer el balance de flujo en estaciones (3); luego, la Fase I busca soluciones que satisfagan las restricciones de capacidad (4) y (6). Finalmente, el objetivo de la Fase II es reducir lo más posible Z sin salir del dominio factible.

Cada etapa tiene una función objetivo y unos criterios de evaluación distintos, que se enuncian en su sección correspondiente. Las Fases I y II tienen un algoritmo de solución común, que se describe en la sección 4.4. En las páginas siguientes detallaremos los componentes de nuestro algoritmo de solución.

4.1 Fase 0 (Punto de partida)

En esta etapa se define una serie de posibles soluciones iniciales, que son factibles respecto de la conservación de frecuencias (3). Para garantizar que la evolución del algoritmo posterior de búsqueda conserve el balance de flujo en cada estación, el vector de frecuencias solo se alterará sumando y/o restando un mismo nivel de frecuencia en cada uno de los arcos que deberán componer un ciclo. Un ciclo está formado por secuencias de servicios, en el que la primera parada de cada servicio es igual a la última parada del servicio que le precede en el ciclo.

Como primer paso se crea un vector de frecuencias inicial, a partir del cual se comenzarán a buscar soluciones que se acerquen a la factibilidad. En dominios complejos como los de este problema es posible que existan zonas de infactibilidad de las que sea difícil salir, es decir en que un algoritmo de búsqueda local converja a una solución óptima en que la factibilidad no se haya alcanzado. Asimismo, es muy probable que el dominio de soluciones factibles no sea conexo. Es por esto que este algoritmo diversifica la búsqueda a través de múltiples puntos de partida que permitan converger hacia un mayor número de soluciones distintas. Así, se definieron seis puntos de partida:

- 1 *Frecuencia mínima para todos:* se asigna a cada servicio la frecuencia mínima requerida de modo que la frecuencia agregada en cada arco del corredor cubra la demanda de viajes. Esta frecuencia se conoce como f_{min} . Es decir, cada servicio comienza con una frecuencia suficiente para transportar por sí mismo toda la demanda de su recorrido.
- 2 *Frecuencia mínima distribuida en cada sentido:* En este caso, se calcula f_{min} para cada sentido, y se distribuye equitativamente entre el número de servicios que compone cada sentido del corredor.
- 3 *Frecuencia a partir de la demanda:* en este caso se reparte f_{min} entre cada servicio de forma proporcional a los viajes que puede atender. Es decir, para cada línea l se calcula la demanda de todos los pares w que comunica, y se distribuye f_{min} proporcional a ella.
- 4 *Frecuencia de acuerdo a cantidad de viajes que inician o terminan en cada estación:* en este punto de partida se asigna a cada línea la capacidad suficiente para atender la demanda de su origen o destino más cargado. Para cada estación se calcula el número de viajes que comienzan o terminan en ella, y se divide entre el número de líneas que la visitan; así se conoce la demanda promedio a transportar por cada línea.
- 5 *Frecuencia de acuerdo a cantidad de viajes que pasan:* esta solución inicial es similar a (4), pero para cada estación se calcula también el número de viajes que pasan. De nuevo, la frecuencia de cada línea se fija dividiendo la demanda de la estación más cargada entre la capacidad del bus.
- 6 *Frecuencia equitativa:* en este punto, se asigna la misma frecuencia a todos los ciclos de estaciones que existan en la red (por ejemplo, en la figura previa, los ciclos formados por las estaciones 1-2-1, o 4-5-6-4). Luego se distribuye este flujo entre las líneas que usan los ciclos de estaciones. Por defecto, $f = \frac{f_{min}}{|K|}$ siendo K el conjunto de ciclos únicos de estaciones.

Una vez definido el punto de partida, resta hacer factible en (3) la solución. Para lograrlo, se crea un grafo $G(N, C)$ con un arco que une la primera y última estación de cada servicio. Si existen arcos paralelos, se agrupan en uno solo y se suma la frecuencia de cada uno. Este grafo permite identificar ciclos de servicios. Para lograr el balance en flujo, se igualan las frecuencias de todos los arcos que componen cada uno de los ciclos posibles.

Existe además una rutina previa a la Fase I que se encarga de hacer cumplir la capacidad de las estaciones para atender buses (6), llamada “Descongestionar”. Para ello, se revisa cada estación y

se reduce equitativamente la frecuencia de todos los servicios que desembarcan en ella hasta que la parada quede sin exceso de frecuencia. Esta rutina preliminar permite aliviar rápidamente la congestión en estaciones que causan algunos puntos de partida con frecuencias demasiado grandes.

4.2 Fase I (o de factibilidad)

El objetivo de esta etapa es ajustar la solución inicial de modo que satisfaga las restricciones de capacidad (4) y (6). Para lograrlo, se implementó un algoritmo de búsqueda local, que permite transitar de una solución a otra vecina en pasos que no alteren el balance de flujo de buses (3). Los criterios de parada de esta fase son: cumplir el tiempo máximo de ejecución, o alcanzar una solución factible.

Para elegir la dirección de movimiento, se emplea una Función Ponderada de Infactibilidad (FPI). Este indicador se puede calcular para cualquier solución, y refleja qué tanto se están violando ambas familias de restricciones de capacidad (tanto de capacidad de los vehículos como de las paradas). La formulación es la siguiente:

$$FPI = \alpha * DB + \overline{\beta * cap} * DPFPI = \alpha * DB + \underline{\beta * cap} * DE \quad (7)$$

Los parámetros α y β corresponden a constantes de calibración que permitan ponderar el déficit de capacidad en buses y estaciones. El primer término (DB) está asociado al déficit de capacidad de los buses para transportar pasajeros, relacionado con (4). Corresponde a la suma del déficit de capacidad (en pasajeros por hora) de todas las líneas en cada una de sus secciones de línea. Se calcula así:

$$DB = \sum_{i \in N} \sum_{l \in L_i} \sum_{\substack{w \in W \\ 0_w < i \\ D_w > i+1}} \max \left(T_w * \frac{f_l^w}{\sum_{h \in L} f_h^w} - cap_l * f_l ; 0 \right) \quad (8)$$

El segundo término (DE) está asociado al déficit de las estaciones para atender buses, relacionado con (6). Corresponde a la suma, para cada estación del corredor, del exceso en frecuencia de buses que se debieran detener en una estación por sobre la capacidad de dicha estación. La fórmula para calcularlo es:

$$DE = \sum_{i \in N} \max \left(\sum_{l \in L} f_l * x_l^i - \delta_i ; 0 \right) \quad (9)$$

Una solución será factible para el problema si tanto DB como DE son cero. Considerando que ambos están calculados en distintas unidades (pasajeros/hora y buses/hora respectivamente), se opta por homogeneizar dimensionalmente FPI en (7), multiplicando DE por la capacidad promedio de los buses.

4.3 Fase II (o de mejora)

El objetivo de esta etapa es encontrar la mejor solución posible dentro del dominio factible. Para ello se recibe como entrada la solución entregada en la Fase I, y se usa el mismo algoritmo de búsqueda local para encontrar mejores soluciones sin salir del dominio factible. Sin embargo, la función empleada para evaluar la conveniencia de cada nueva solución es (1), que estima los costos sociales del corredor. La variable de decisión en esta expresión es la frecuencia de cada línea; aumentar la frecuencia encarece los costos del operador, pero disminuye los tiempos de espera y de viaje.

Con el fin de descartar soluciones infactibles, se añade a la función objetivo el FPI de la Fase I multiplicado por un ponderador γ suficientemente alto para que el algoritmo nunca abandone el dominio factible. Por lo tanto, la función objetivo de la Fase II es la siguiente:

$$FPI' = CS + \gamma \cdot FPI \quad (10)$$

Como criterio de término de esta etapa se contempla carecer de movimientos de mejora o alcanzar el tiempo máximo de ejecución.

4.4 Algoritmo de búsqueda local

El algoritmo de búsqueda local es la herramienta común utilizada las fases I y II. La única diferencia entre ellas está en la función objetivo, que fue explicada en cada sección correspondiente. Es un algoritmo del tipo *First Improvement* en el que se revisan las soluciones secuencialmente y se mueve en dirección de la primera solución que cumple un criterio de aceptabilidad. El funcionamiento general de esta búsqueda local se muestra a continuación:

1. Hacer $i = 1$. Definir como solución actual v_i al vector inicial de frecuencias entregado como entrada. Calcular $Z(v_i)$.
2. Hacer una lista R que contiene las líneas de v_i ordenadas de mayor a menor peso en la función objetivo. Elegir la primera línea l .
3. Identificar un ciclo al que pertenezca l .
4. Alterar f_l en una unidad de paso. El vector que contiene la frecuencia resultante es el vecino v_{i+1} .
5. Calcular $Z(v_{i+1})$.
6. Si $Z(v_{i+1}) \leq Z(v_i)$, se acepta la nueva solución. Hacer $v_i = v_{i+1}$. En caso contrario, ir a 4. Hacer $R = R \setminus l$.
7. Si existen líneas por revisar, ir a 2.
8. Si se agotaron todas las líneas y no se cumple la condición de término, ir a 2. En caso contrario se termina el algoritmo y se entrega la solución.

Considerando que solo se desea explorar soluciones vecinas que no alteren el balance de flujo de las estaciones, los movimientos del algoritmo de búsqueda se hacen en ciclos de servicios. Para hacer más claros estos movimientos, en la figura siguiente se muestra un grafo $G(N, C)$ con 4 estaciones y 6 servicios. Para alterar la frecuencia de algunos servicios, los movimientos permitidos exigen identificar un ciclo no dirigido de servicios y aumentar o quitar frecuencia a los servicios

del ciclo; si en un sentido se agregan o quitan buses, en el sentido contrario debe efectuarse la acción opuesta. A modo de ejemplo:

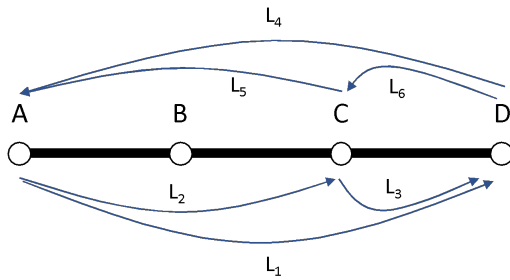


Figura 1. Red de ejemplo

- Añadir o quitar frecuencia a $\{L1, L4\}$.
- Añadir o quitar frecuencia a $\{L5, L6\}$, y hacer lo opuesto con $\{L4\}$.

Una vez definida la dirección de movimiento, resta definir el tamaño del paso (es decir, el aumento o reducción en el flujo de los servicios que componen el ciclo). Para ello, se estiman los costos sociales para un tamaño de paso de 10 bus/h y se reduce a la mitad n_r veces a lo largo de la ejecución. Se elige aquel

tamaño que mayor ahorro traiga. El parámetro n_r es dado por el usuario, y entre más grande sea, más se acercará al ahorro óptimo, a costo de un mayor tiempo de ejecución.

En el caso de la Fase I esta etapa concluye una vez que el algoritmo identifica un vector de frecuencias factible en continuidad de flujos, que respeta las capacidades de buses y estaciones. En el caso de la Fase II en que se busca minimizar los costos sociales del corredor, se concluye cuando ya no es posible encontrar soluciones que mejoren la solución vigente o se alcanza el tiempo máximo para la búsqueda.

4.4.1 Diversificación de la búsqueda

El dominio factible de este problema se caracteriza por tener muchos óptimos locales. Así, se hace necesario incluir en el algoritmo mecanismos que le permitan escapar de ellos y así explorar un espacio amplio del dominio. Por eso es relevante considerar varios puntos iniciales factibles para comenzar la búsqueda. Además, se implementaron tres metaheurísticas que pretenden romper la lógica de moverse solo hacia soluciones mejores. Cada una de ellas puede usarse o no para complementar el algoritmo de base. A continuación, se describe cada una de ellas, mientras que en la sección de resultados se muestra el desempeño de cada una de estas metaheurísticas para un experimento estructurado en torno a un escenario real.

- 1 *Simulated Annealing*: Esta metaheurística permite moverse hacia soluciones de peor desempeño siguiendo una distribución de probabilidad que se concentra progresivamente hacia las soluciones mejores a medida que se avanza en la búsqueda.
- 2 *Destroy & Repair*: Cuando esta metaheurística está activa, cada cierto número de iteraciones definido por el usuario, el algoritmo acepta el movimiento hacia el primer vecino que se explore, independientemente de si trae una mejora o no. Este movimiento alterará arbitrariamente la frecuencia de uno de los servicios, lo que representa un salto en el dominio.
- 3 *Lista tabú*: Para no repetir soluciones y caer en un bucle de empeorar y mejorar entre las mismas soluciones, se implementó una lista tabú para no repetir las últimas n_{FO} soluciones que están en la lista.

5 EXPERIMENTOS

En esta sección se muestran los experimentos formulados para conocer el rendimiento del algoritmo programado. Los escenarios se basan en tres corredores creados a partir de corredores reales de BRT: la avenida Caracas (Bogotá, Colombia), y las avenidas Grecia y Pajaritos (Santiago, Chile). Para cada uno de estos corredores se produjeron cuatro instancias, en las que varía el número de estaciones y el número de líneas que las atienden; el resto de las características son comunes a cada corredor. La capacidad de los buses es de 150 pax/h, y la capacidad de las estaciones es de 180 bus/h. La Tabla 1 resume las características de las doce instancias.

Corredor	Demanda	Número de estaciones	Número de líneas
Caracas	43550 pax/h	10, 20, 40, 40	17, 22, 29, 41
Grecia	37810 pax/h	10, 20, 40, 40	26, 31, 31, 60
Pajaritos	20546 pax/h	10, 20, 40, 40	11, 27, 28, 56

Tabla 1. Descripción de las Instancias

Se formularon dos experimentos que analizan el desempeño del algoritmo respecto de dos distintas dimensiones: en primer lugar, desde el punto de vista de las soluciones de partida para el algoritmo; en segundo lugar, analiza el rendimiento con las distintas mejoras de diversificación que se implementaron. A menos que se diga lo contrario, se ejecutó cada experimento en las doce instancias de prueba.

5.1 Puntos de partida y Fase I

Punto de partida	Eficacia
1	58,75%
2	83,30%
3	66,66%
4	53,80%
5	91,60%
6	27,86%

Tabla 2. Eficacia de los distintos puntos de partida para encontrar factibilidad

En primer lugar, se estudia la calidad de cada punto de partida. Un buen punto de partida es aquel que encuentra una solución factible en ambas familias de restricciones de capacidad y por tanto tiene un FPI = 0. La Tabla 2 muestra la eficacia de cada punto de partida para encontrar soluciones factibles. Destacan los puntos 2 y 5, con una eficacia mayor al 80% de las instancias estudiadas. Cuando una solución parte muy lejos del dominio factible, el tiempo de ejecución puede no ser suficiente. Ejemplo de ello son los puntos de partida 1 y 6 que obtienen los peores desempeños, pues asignan frecuencias iniciales muy desproporcionadas.

Metaheurística		Al terminar Fase I	
		Infactible	Factible
Simulated Annealing	No	19%	81%
	Sí	25%	75%
Destroy & Repair	No	19%	81%
	Sí	17%	83%
Lista tabú	No	19%	81%

Transcurrido el tiempo máximo de ejecución (15 minutos), el 66% de todas las instancias convergieron a una solución factible. En el 11% de las instancias, la infactibilidad se concentra en la restricción de algunas estaciones cuya capacidad es violada en un nivel menor (menos de 5%). En estos casos la solución alcanzada resulta útil para poder encontrar una buena solución para el problema. La solución infactible alcanzada sugiere que, si toleramos un nivel de hacinamiento máximo al interior de los buses ligeramente superior al considerado inicialmente, habremos encontrado una solución factible al diseño operacional para el corredor. Así, parece razonable en estos casos relajar dicha capacidad en los servicios afectados y seguir adelante con la fase II del algoritmo.

	Sí	19%	81%
--	----	-----	-----

La Tabla 2 confirma que comenzar con un buen punto de partida es fundamental. Esto confirma que el desempeño de los algoritmos de búsqueda en dominios complejos está condicionado fuertemente por el punto de partida. Por otra parte, la Tabla 3 muestra que las metaheurísticas no mejoran significativamente el desempeño del algoritmo en Fase I, caso contrario a lo que se muestra en los resultados de la Fase II, en la que ellas resultan esenciales para escapar de óptimos locales.

Tabla 3. Rendimiento de las diferentes metaheurísticas en Fase I

5.2 Eficiencia del algoritmo y Fase II

El objetivo de este segundo experimento es analizar el desempeño del algoritmo para encontrar las soluciones factibles con menores costos sociales. Solo se experimentó con los tres puntos de partida de mayor eficacia en la Fase I (puntos 2, 3, y 5). En primer lugar, y contrastando con los resultados en la Fase I, la Tabla 4 muestra que el desempeño del algoritmo mejora significativamente con metaheurísticas que exploren mejor el dominio. En esta tabla se consideran solamente los experimentos para los que se encontró una solución factible en Fase I.

Metaheurística adicional	Promedio de FO		Mejora
	Inicio Fase II	Fin Fase II	
Simulated Annealing	4766013	4304980	9,67%
Destroy & Repair		4366982	8,37%
Lista Tabú		4603578	3,41%
Ninguna		4604420	3,39%

Tabla 4. Rendimiento de las metaheurísticas en Fase II

El algoritmo de búsqueda local mejorado con *Simulated Annealing* reduce en hasta un 9,67% los costos sociales (tiempo de viaje de los usuarios más costo de los operadores).

Resulta interesante analizar el efecto de nuestro algoritmo en el vector de frecuencias. La Figura 2 muestra la frecuencia de cada línea en tres momentos del algoritmo: las correspondientes al punto de partida, a la solución tras la Fase I (factibilidad), y luego de la Fase II (mejora). Se muestran los resultados para el Escenario Caracas con 20 estaciones comenzando en el punto de partida 5 (que asigna frecuencias mayores a las requeridas) ya que muestra mejor el funcionamiento de las dos fases. Las líneas 1 y 2 corresponden a los servicios all-stop; las líneas 3 al 10 son servicios expresos en el sentido de ida, y las líneas 11 al 22 son servicios expresos en el sentido de regreso.

Resulta interesante analizar el efecto de nuestro algoritmo en el

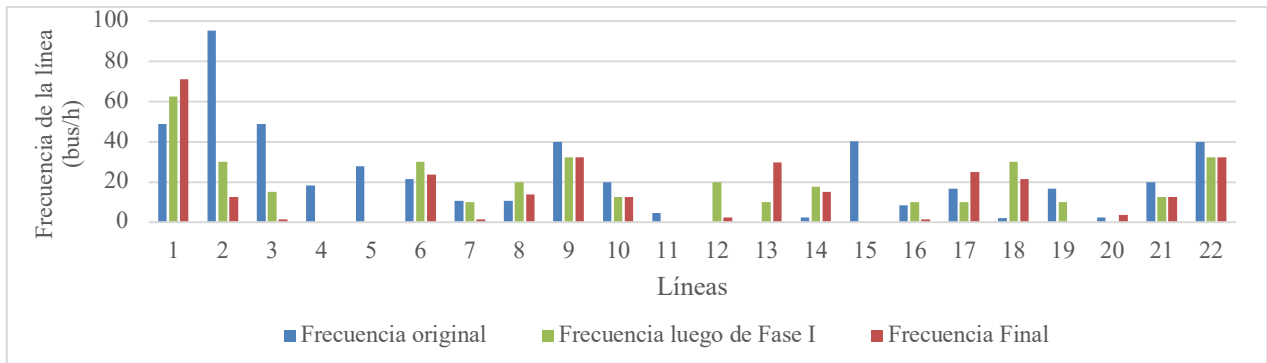


Figura 2. Comparación de los vectores de frecuencia a lo largo de la ejecución

El punto de partida acierta en el orden de magnitud de las frecuencias de algunas líneas que serán mantenidas luego de la Fase II (ver líneas 1, 6, 9, o 22). Este acierto se explica por el criterio de demanda que utiliza el punto de partida 5. La Fase I tiene la misión de aliviar la congestión en buses y estaciones, por lo que respectivamente aumenta la frecuencia de algunas líneas y reduce la de otras. La Fase II en cambio balancea las frecuencias dentro del dominio factible para minimizar los costos sociales. Para ello saca de operación algunas líneas que resultan ineficientes (ver líneas 3, 16, o 19), y asigna esta capacidad liberada a líneas eficientes (ver líneas 1 o 13). Es atractivo que la solución final exige menos líneas, lo que es además más fácil de comprender para el usuario.

Categoría (se detiene en...)	Frecuencia total asignada (bus/h)
Superexpreso (menos de la mitad de las paradas)	133,36
Expreso (la mitad o más de las paradas)	94,88
All-stop (todas las paradas)	83,76

Tabla 5. Distribución de la frecuencia según tipo de servicio

El número de paradas por servicio permite evaluar qué tipo de líneas prefiere el algoritmo de la Fase II. La Tabla 5 muestra la frecuencia total asignada a las líneas en tres categorías creadas a partir del número de paradas.

La preferencia por servicios que omiten más paradas puede obedecer a la concentración de la demanda en pocos pares origen-destino. Los servicios all-stop siguen cumpliendo un rol importante en la conectividad de la red, y son útiles para personas que hacen viajes cortos o entre pares con poca demanda.

Un último experimento consistió en relajar las restricciones de capacidad en estaciones. Para ello, se ejecutó de nuevo la Fase II de algunos experimentos, pero haciendo nula la penalidad por exceder esta capacidad. Esto le permite al algoritmo explorar soluciones fuera del dominio factible para reducir los costos sociales, y nos permite observar el efecto de reconocer la capacidad limitada de las estaciones en los modelos de diseño de servicios. Dado que el dominio de búsqueda es más acotado cuando las restricciones están activas, las soluciones factibles son más costosas en general. De hecho, los experimentos donde las restricciones de capacidad están activas tuvieron en promedio una función objetivo 41,67% más costosa que aquellos donde las restricciones estaban relajadas.

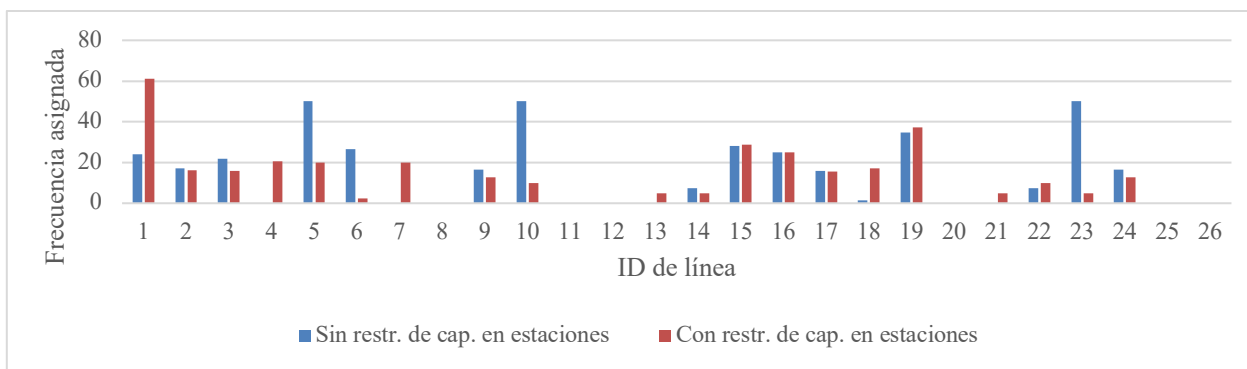


Figura 3. Efecto de las restricciones de capacidad en el vector de frecuencias

A modo de ejemplo, la Figura 3 compara los vectores de frecuencia resultantes en ambos casos, para el corredor Grecia con 20 estaciones. Las líneas 1 y 2 son servicios *all-stop*. La observación más clara es que el conjunto de líneas que forman parte de la solución no se altera al relajar la restricción; las líneas 11, 12, 20, 25, y 26 permanecieron sin buses asignados en ambas ejecuciones. El requerimiento de flota aumentó al relajar la restricción de capacidad (de 346 a 393 buses), pues ignorar la capacidad de las estaciones para atender buses le permite al modelo utilizar más vehículos. Este fenómeno se observó en la mayoría de las instancias estudiadas. Además de esto, el modelo prefirió una solución con más buses asignados a los servicios expresos. También puede observarse que la distribución de la frecuencia cambia significativamente entre ambas soluciones: esto puede deberse en parte a que el algoritmo explora regiones distintas del dominio.

Las restricciones de capacidad también reducen el tamaño del dominio factible, lo que limita la capacidad de acción de los algoritmos de búsqueda; para explorar este dominio eficientemente deben implementarse metaheurísticas, como Simulated Annealing. De hecho, el ahorro promedio en los costos sociales logrado por el algoritmo pasa de 9,67% con las restricciones impuestas a 39% con las restricciones relajadas. Naturalmente, estas soluciones carecen de utilidad en la vida real pues implicarían un mal nivel de servicio o carecen de imposibilidad física u operacional.

6 CONCLUSIONES

Este trabajo aborda el diseño de servicios de buses sobre un corredor con capacidad de adelantamiento considerando no solo la capacidad finita de los vehículos, sino también la capacidad limitada de las paradas de atender buses. Se presentó una metodología para encontrar un vector de frecuencias para un conjunto de servicios dado, que minimiza los costos sociales sobre el corredor. Hasta donde la revisión bibliográfica nos permite concluir, este trabajo es el primero que considera la restricción de capacidad de las paradas en el proceso de diseño de redes de transporte incluyendo servicios expresos, a pesar de que algunos de los sistemas BRT más importantes del mundo ven afectada su operación por estas restricciones. Cuando la capacidad de las paradas se excede en sistemas muy demandados, se crea un cuello de botella que puede limitar la capacidad de transporte del corredor. En general, los trabajos relacionados a optimización de frecuencias han pasado por alto este fenómeno, por lo que este trabajo debiera ser un aporte a la construcción de redes más robustas.

Un elemento que complejiza la solución al problema es que identificar una solución inicial factible no es simple. Frecuencias muy bajas podrían generar buses que exceden su capacidad, mientras frecuencias muy altas podrían generar paradas que exceden su capacidad de atender buses. Es posible imaginar escenarios en que el problema simplemente no admite una solución factible. Nuestro modelo procura identificar un vector de frecuencias que satisfaga la capacidad de buses y estaciones. La frecuencia de un servicio es una variable que es ajustada por los operadores para responder a cambios en la oferta o la demanda, por lo que el vector entregado por nuestro algoritmo debería ser ajustado a las necesidades operacionales del corredor.

Se realizaron experimentos para conocer el desempeño de cada una de las fases que conforman la metodología para una serie de escenarios tomados de corredores reales en que se asumió que los usuarios nunca escogerían transbordar para alcanzar sus destinos. Se concluyó que considerar las restricciones de capacidad disminuye el requerimiento de flota (o se usa más eficientemente), no altera el conjunto de líneas seleccionadas para formar parte de la solución, y se prefieren en mayor medida los servicios *all-stop* sobre los expresos.

Algunas otras conclusiones significativas de estos experimentos son:

- a) El vector inicial de frecuencias condiciona significativamente el desempeño del algoritmo para encontrar soluciones factibles. En este trabajo propusimos seis estrategias distintas para determinar los vectores iniciales; los puntos de partida que fijaban una flota inicial más acertada tuvieron el mejor desempeño a la hora de encontrar factibilidad (hasta 95% de éxito).
- b) El hallazgo anterior se explica porque las restricciones de capacidad afectan el rendimiento los algoritmos de búsqueda local. En ese sentido, resulta importante mejorar los algoritmos de base con metaheurísticas que les permitan escapar de regiones complejas del dominio para encontrar mejores soluciones. De las tres metaheurísticas implementadas en este trabajo, Simulated Annealing fue la que mejoró en mayor magnitud el desempeño del algoritmo base en Fase II: el ahorro promedio respecto de la solución de la Fase I pasó de 3,39% sin la metaheurística activa a 9,67% en las instancias en las que sí lo estaba.
- c) El rendimiento del algoritmo está condicionado a la eficiencia en el esquema de paradas de los servicios entregados. Si bien nuestra metodología no incluye la etapa de elección de paradas, el algoritmo deja fuera a los servicios menos eficientes asignándoles frecuencia nula, como se muestra en la Figura 3. El diseño de servicios integral, considerando la decisión de omitir o no paradas, podría mejorar el desempeño global de la metodología.

BIBLIOGRAFÍA

- Asadi Bagloee, S., & Ceder, A. A. (2011). Transit-network design methodology for actual-size road networks. *Transportation Research Part B: Methodological*, 45(10), 1787–1804. <https://doi.org/10.1016/j.trb.2011.07.005>
- Carrese, S., & Gori, S. (2002). An Urban Bus Network Design Procedure. *Transportation Planning*, 177–195.
- Ceder, A., & Wilson, N. (1986). Transportation Research. *Bus Network Design*, 1(4), 331–344. <https://sci-hub.tw/https://www.sciencedirect.com/science/article/abs/pii/0191261586900470#aep->

abstract-id5

- Chriqui, C., & Robillard, P. (1975). Common Bus Lines. *Transportation Science*, 9(2), 115–121. <https://doi.org/10.1287/trsc.9.2.115>
- de Cea, J., & Fernandez, E. (1993). Transit assignment for congested public transport systems: An equilibrium model. *Transportation Science*, 27(2), 133–147. <https://doi.org/10.1287/trsc.27.2.133>
- Desaulniers, G., & Hickman, M. D. (2007). Chapter 2 Public Transit. *Handbooks in Operations Research and Management Science*, 14(C), 69–127. [https://doi.org/10.1016/S0927-0507\(06\)14002-5](https://doi.org/10.1016/S0927-0507(06)14002-5)
- Fernández, R. (2010). Modelling public transport stops by microscopic simulation. *Transportation Research Part C: Emerging Technologies*, 18(6), 856–868. <https://doi.org/10.1016/j.trc.2010.02.002>
- Fernandez, R., & Planzer, R. (2002). On the capacity of bus transit systems. *Transport Reviews*, 22(3), 267–293. <https://doi.org/10.1080/01441640110106328>
- Gao, Z., Sun, H., & Shan, L. L. (2004). A continuous equilibrium network design model and algorithm for transit systems. *Transportation Research Part B: Methodological*, 38(3), 235–250. [https://doi.org/10.1016/S0191-2615\(03\)00011-0](https://doi.org/10.1016/S0191-2615(03)00011-0)
- Guihaire, V., & Hao, J. K. (2008). Transit network design and scheduling: A global review. *Transportation Research Part A: Policy and Practice*, 42(10), 1251–1273. <https://doi.org/10.1016/j.tra.2008.03.011>
- Institute for Transportation & Development Policy. (2017). *BRT Planning Guide*.
- Larrain, H., & Muñoz, J. C. (2008). Public transit corridor assignment assuming congestion due to passenger boarding and alighting. *Networks and Spatial Economics*, 8(2–3), 241–256. <https://doi.org/10.1007/s11067-007-9054-3>
- Larrain, H., Suman, H., & Muñoz, J. C. (2021). Route based equilibrium assignment in congested transit networks. *Transportation Research Part C: Emerging Technologies*, 127.
- Larrain Izquierdo, H. (2013). Diseño de servicios expresos para buses. *Pontificia Universidad Católica de Chile, Escuela de Ingeniería*, 266.
- Lee, Y.-J., & Vuchic, V. R. (2009). Transit Network Design with Variable Demand. *Journal of Transportation Engineering*, 135(8), 491–505. [https://doi.org/10.1061/\(ASCE\)0733-947X\(2009\)135:8\(491\)](https://doi.org/10.1061/(ASCE)0733-947X(2009)135:8(491))
- 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: Methodological*, 44(10), 1186–1201. <https://doi.org/10.1016/j.trb.2010.01.003>
- Park, S. J. (2005). Bus Network Scheduling with Genetic Algorithms and Simulation. *Master Thesis*, 283.
- Soto, G., Larrain, H., & Muñoz, J. C. (2017). A new solution framework for the limited-stop bus service design problem. *Transportation Research Part B: Methodological*, 105, 67–85. <https://doi.org/10.1016/j.trb.2017.08.026>
- Szeto, W. Y., & Jiang, Y. (2014). Transit route and frequency design: Bi-level modeling and hybrid artificial bee colony algorithm approach. *Transportation Research Part B: Methodological*, 67, 235–263. <https://doi.org/10.1016/j.trb.2014.05.008>
- Szeto, W. Y., & Wu, Y. (2011). A simultaneous bus route design and frequency setting problem for Tin Shui Wai, Hong Kong. *European Journal of Operational Research*, 209(2), 141–155. <https://doi.org/10.1016/j.ejor.2010.08.020>
- Tirachini, A., Hensher, D. A., & Rose, J. M. (2014). Multimodal pricing and optimal design of

urban public transport: The interplay between traffic congestion and bus crowding.
Transportation Research Part B: Methodological, 61, 33–54.
<https://doi.org/10.1016/j.trb.2014.01.003>