

## MODELACIÓN DEL FENÓMENO DE APELOTONAMIENTO DE BUSES USANDO DATOS MASIVOS DE POSICIONAMIENTO Y TRANSACCIONES DE PAGO

Jacqueline Arriagada, Universidad de Chile, [afjg.7440@gmail.com](mailto:afjg.7440@gmail.com)

Antonio Gschwender, DTPM, [Antonio.Gschwender@dtpm.gob.cl](mailto:Antonio.Gschwender@dtpm.gob.cl)

Marcela A. Munizaga, Universidad de Chile, [mamuniza@ing.uchile.cl](mailto:mamuniza@ing.uchile.cl)

### RESUMEN

Apelotonamiento de Buses (o irregularidad de intervalo entre buses consecutivos) es un fenómeno conocido y relevante para operadores, usuarios y reguladores de servicios de buses de alta frecuencia. El efecto de este fenómeno es tal que los buses tienden a agruparse en dos o más, lo cual afecta la calidad del servicio y la eficiencia operacional. Con el objetivo de contribuir al entendimiento del fenómeno, desarrollamos modelos econométricos que indican que las variables que afectan el problema son aquellas relacionadas a: la operación (distancia, frecuencia, composición de flota, entre otras), la estructura de demanda y la infraestructura del camino.

*Palabras claves: Apelotonamiento de buses, operación, modelación*

### ABSTRACT

Bus bunching is a well-known phenomenon for operators, users and regulators of high-frequency bus services. The effect of this variability is that buses tend to group into bunches of two or more, which severely affects the quality of service and the operational efficiency. With the aim of contribute to the understanding of the phenomenon, we developed econometrics models which indicate that the variables affect the problem, are related to: operation (distance, frequency, fleet composition, and others), demand structure and infrastructure of road.

*Keywords: Bus bunching, operation, modeling*

## 1. INTRODUCCIÓN

El Apelotonamiento de buses es un fenómeno entendido como la inestabilidad (variabilidad) de los intervalos de tiempo entre buses consecutivos. Este problema es causado por desviaciones del intervalo esperado (promedio). Cuando un bus está retrasado o adelantado en sus horarios, incrementa/reduce el intervalo con el bus que le antecede y lo opuesto ocurre con el bus que le sigue.

Tanto para buses que operan por itinerario como por frecuencia, el apelotonamiento de buses incrementa el número de pasajeros que suben al bus que llega después de un intervalo largo. Esto implica uso ineficiente de la flota de buses, reducción en la capacidad del sistema de buses y la consecuencia más grave el aumento del tiempo de espera de los pasajeros.

El modelo de tiempo de espera usado por Welding (1957), Holroyd y Scraggs (1966) y Osuna y Newell (1972), el cual asume que la llegada de los pasajeros es aleatoria (supuesto para servicios que operan por frecuencia), muestra una relación directa entre el tiempo de espera promedio por persona y la variabilidad de los intervalos. En el caso de servicios que operan por itinerario, la variabilidad de los intervalos contribuye a incrementar la aleatoriedad del tiempo de llegada de pasajeros. Esto es explicado por el modelo propuesto por Turnquist (1978), pues el apelotonamiento de buses genera que los pasajeros sean menos capaces de planear los tiempos de llegada al paradero para coincidir con la llegada del bus.

Dada la importancia del tiempo de espera en la calidad de servicio de transporte público, varios estudios se han enfocado en entender el fenómeno de irregularidad de buses, usando datos de diferentes sistemas de transporte, usualmente incluyendo datos de Automatic Vehicle Location (AVL) y Automatic Passenger Counting (APC) (Moreira-Matias et al, 2012; Albright, y Figliozzi, 2012). Modelos de simulación estocástica (Bellei y Gkoumas, 2010; Man y Wenquan, 2013) y de regresión lineal (Albright y Figliozzi, 2012) han sido construidos para explicar el fenómeno apelotonamiento de buses. De acuerdo a estos estudios algunos de los factores más relevantes son: desviación del intervalo programado al inicio de la ruta (Moreira-Matias et al, 2012), número de pasajeros (que suben y bajan) (Kittelson et al, 2003; Bellei y Gkoumas, 2010; Moreira-Matias et al, 2012; Albright y Figliozzi, 2012, Man y Wenquan, 2013), largo de la ruta y número de paraderos (Kittelson et al, 2003), frecuencia programada de los buses (Moreira-Matias et al, 2012; Man y Wenquan, 2013), distancia desde el inicio de la ruta (Bellei y Gkoumas, 2010), condiciones del tráfico (incluyendo incidentes y congestión) (Kittelson et al, 2003; Man y Wenquan, 2013), trato preferencial del tránsito (pistas para buses o sistemas de Transit Signal Priority (TSP)) (Kittelson et al, 2003; Albright y Figliozzi, 2012), uso de elevador de silla de ruedas (Kittelson et al, 2003; Albright y Figliozzi, 2012), comportamiento o habilidades del conductor (Kittelson et al, 2003), y mantención de los buses (Kittelson et al, 2003).

Para mitigar el problema de irregularidad de buses, varias estrategias de control han sido propuestas, como la retención de buses en paraderos, ajuste de velocidad, salto de paradero, tiempos limitados de subida de pasajeros y una combinación de estas estrategias. Algunos autores han estudiado y/o desarrollado programas para conocer cuánto tiempo los buses deben ser retenidos en los paraderos (Daganzo, 2009; Bartholdi y Eisenstein, 2012; Chen, Q., Adida, E. and Lin, J, 2013; Lizana et al, 2014).

El objetivo de este estudio es contribuir al entendimiento del fenómeno apilamiento de buses por medio de la identificación de factores que incrementan y disminuyen la intensidad del problema. En la próxima sección se describen los casos de estudios, en la sección tres se describe el modelo y finalmente en la sección cuatro y cinco se entregan los resultados y las conclusiones obtenidas en este trabajo.

## 2. DESCRIPCIÓN DE LOS CASOS DE ESTUDIO

Los datos utilizados en este estudio provienen del sistema de transporte público de Santiago, Transantiago, donde todos los buses están equipados con dispositivos GPS que registran el tiempo y la posición del vehículo cada 30 segundo (Cortés et al, 2011). Adicionalmente, la información detallada de demanda está disponible gracias al sistema AFC (Munizaga y Palma, 2012). Se utilizaron datos de los servicios 506 y su variante expreso 506e, entre 15 de abril de 2013 y 19 de abril de 2013 y del servicio 206 los días 17 y 19 de abril de 2013 (los demás días fueron omitidos por presencia de feria en segmentos de calle por los que transita el servicio normalmente).

El servicio 506 opera durante todo el día y toda la noche todos los días de la semana, mientras que el servicio 506e opera durante horarios de alta demanda de lunes a viernes. Ambos servicios comparten el mismo camino y los datos utilizados corresponden a los del sentido poniente a oriente. El largo de la ruta de estos servicios es de 32,6 km, en donde el servicio 506 tiene 82 paraderos y el 506e 43 paraderos. La distancia promedio entre paraderos es de 393 metros para el servicio regular y 758 metros el servicio expreso.

Por otro lado el servicio 206 opera todos los días de la semana entre 5:30AM y 0:40AM, con un largo de ruta de 20,7 km, sirviendo a 58 paraderos que presentan una separación promedio de 354 metros. Los datos utilizados de este servicio corresponden a los del sentido sur a norte.

Tres de los paraderos de los servicios 506-506e están equipados con dispositivos extra de pago (zonas paga), mientras que dos paraderos del servicio 206 son zonas paga. Las zonas paga funcionan como estaciones de buses durante periodos y paraderos de alta demanda. Los semáforos que están localizados a lo largo de la ruta de los servicios estudiados están coordinados en grupos dentro de diferentes redes. La programación de estos semáforos en red es optimizada con el objetivo de minimizar la demora total de los vehículos dentro de una red.

Con respecto a la infraestructura, se diferencian cuatro tipos diferentes de prioridad de buses: tráfico mixto, pistas solo bus, corredor segregado y vía exclusiva. Los segmentos de tráfico mixto no consideran ningún tipo de prioridad al transporte público sobre vehículos particulares, luego la calle es compartida. Las pistas solo bus están localizadas a la derecha de la calle, y son de exclusividad para vehículos de transporte público. Estas no presentan segregación física, y está permitido que vehículos privados entren a estas vías cuando necesitan doblar a la derecha, por lo que no siempre son respetadas. Los segmentos de corredor segregado corresponden a vías centrales segregadas físicamente de dos pistas o una pista con adelantamiento en los paraderos, en donde pueden circular solamente buses del transporte público. Las vías exclusivas son pistas

restringidas solo para la circulación de vehículos del transporte público, estas operan de lunes a viernes (excepto festivos) entre las 7:30 a las 10:00 hrs y entre las 17:00 y 21:00 hrs. En particular los servicios 506-506e utilizan tipos de prioridad correspondientes a tráfico mixto, pistas solo bus y corredor segregado, mientras que el servicio 206 utiliza los tipo de prioridad tráfico mixto, corredor segregado y vía exclusiva. Se hace notar que el segmento de dos pistas solo bus que utiliza el servicio 506 y su variante 506e, tiene la particularidad de encontrarse recurrentemente con autos estacionados al costado derecho (ver Imagen 1), lo cual impide que los buses tengan disposición total de ambas pistas para ellos.



**Imagen 1: Avenida Manuel Antonio Matta, Santiago, Chile**

El intervalo de tiempo promedio entre buses consecutivos del servicio 506 es de 7-8 minutos, lo cual disminuye a 6 minutos durante los horarios de alta demanda en la mañana y en la tarde. El servicio 506e presenta un intervalo promedio entre 10-11 minutos, mientras que el intervalo promedio del servicio 206 está entre 7-8 minutos durante el día, no mostrando mayor variación durante los horarios de alta demanda. El servicio 506e opera únicamente con buses de tipo articulado (capacidad 160 personas), mientras que los otros dos servicio operan tanto con buses articulados como no articulados (capacidad 100 personas).

Con respecto a la demanda el servicio 506 tiene en promedio 89 subidas de pasajeros por expedición a lo largo del día, mientras que durante los periodos de alta demanda en la mañana y en la tarde este número asciende a 118 pasajeros por expedición para el servicio 506 y 162 para el servicio 506e. Por otro lado el servicio 206 presenta en promedio 83 subidas de pasajeros por expedición, número que asciende a 123 pasajeros durante periodos de alta demanda en la mañana y en la tarde. Se hace notar que esta demanda considera únicamente aquellos pasajeros que pagaron su viaje, pues los viajes con evasión de tarifa no son registradas en el sistema AFC y representan cerca del 25% de los viajes. Finalmente el servicio 506 y 506e son operados por la empresa Metbus, mientras que el servicio 206 es operado por la empresa Subus.

### 3. MODELACIÓN

Con el objetivo de explicar el fenómeno apilamiento de buses se utilizan modelos de regresión lineal, de la forma en que se muestra en la ecuación (1), donde  $Z$  representa la variable dependiente o estudiada,  $x_i$  representan las variables explicativas o independientes y  $\beta_i$  son los coeficientes de cada variable.

$$Z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n \quad (1)$$

### 3.2 Variable Dependiente: Indicadores de Apelotonamiento de Buses

Como medición del apelotonamiento de buses se utilizan dos indicadores, un indicador continuo basado en la variabilidad de los intervalos observados, propuesto por Byon et al. (2011) y un indicador discreto que toma valor uno cuando dos buses están demasiado cerca uno de otro, y cero en caso contrario, este último indicador fue propuesto por Bellei and Gkoumas (2010). En la ecuación (2) se muestra la fórmula de indicador continuo y en la ecuación (3) se muestra la fórmula de indicador discreto, en donde el indicador continuo se calcula por paradero ( $n$ ) y periodo de tiempo específico ( $p$ ) y el indicador discreto se calcula por expedición ( $m$ ) y paradero ( $n$ ).

$$IPO_{np} = \frac{\sum_m \left( \frac{H_{mn}}{H_{np}^*} \right)^2}{N_{np}} \quad (2)$$

$$I_{mn} = \begin{cases} 1 & \text{si } H_{mn} \leq \frac{H_{np}^*}{4} \\ 0 & \text{si } H_{mn} > \frac{H_{np}^*}{4} \end{cases} \quad (3)$$

En la ecuación (2) y (3)  $H_{mn}$  representa el intervalo de tiempo entre la expedición  $m - 1$  y la expedición  $m$  en el paradero  $n$ ;  $H_{np}^*$  es el intervalo de tiempo promedio en el paradero  $n$  y periodo de tiempo  $p$  y  $N_{np}$  el número total de intervalos observados en el paradero  $n$  y periodo de tiempo  $p$ . El indicador continuo es utilizado para formar la variable dependiente tal cual se muestra en la ecuación 1 y el indicador discreto se utiliza por medio del promedio de su valor en cada paradero-periodo evaluado (ver ecuación 4).

$$Y_{np} = \frac{\sum_m I_{mn}}{N_{np}} \quad (4)$$

Los valores promedio del indicador discreto ( $Y$ ) y el indicador continuo ( $IPO$ ) fueron calculados para cada paradero a lo largo de la ruta, usando periodos de tiempo de 30-minutos para el servicio 206 y 506, excepto entre las 2AM y 5AM en que los datos del servicio 506 fueron agregados en un periodo de tres horas, y periodos de una hora para el servicio 506e (los periodos de tiempo fueron escogidos en base a la consideración del número de observaciones en cada caso). En resumen cada medida de paradero-periodo de tiempo es tomado como un punto de observación. Los intervalos de tiempo entre buses fueron medidos como la diferencia entre el instante en que cada bus pasó por el paradero evaluado. Esta diferencia es obtenida con una interpolación lineal de pulsos GPS si es que no se observa ningún pulso en el paradero. Si es que dos pulsos o más son localizados en el mismo paradero, se utiliza un promedio de tiempo. Los datos de buses fuera de servicio o de paraderos-periodos donde uno o más pulsos GPS están ausentes, fueron removidos de la base de datos en estudio.

A modo de ejemplo los valores obtenidos para el servicio 506 están representados en la figura 1, donde cada celda está pintada según el valor correspondiente de acuerdo a la escala mostrada en

la parte derecha de cada diagrama, donde el color rojo representa apelotonamiento de buses severo y el color azul representa regularidad entre los buses.

Ambos indicadores muestran valores razonables, indicando que el apelotonamiento de buses ocurre durante los horarios de mayor actividad durante el día, después de las 6AM, y que es más intenso hacia el final de la ruta. Comparando los resultados de ambos indicadores, se puede observar que el indicador discreto presenta mayor sensibilidad al apelotonamiento más severo, capturando algunos casos “rojos” que no son detectados por el indicador continuo, mientras que el indicador continuo es opuesto, capturando algunos casos de apelotonamiento que no son capturados por el indicador discreto. Esto sugiere que el indicador continuo puede ser más apropiado para la detección temprana del problema de irregularidad de buses.

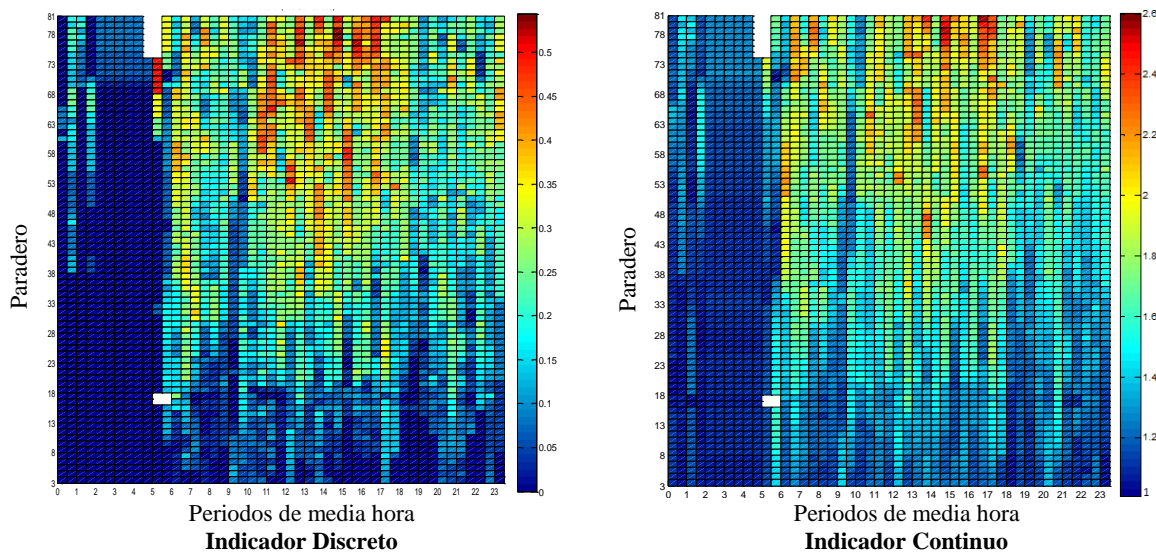


Figura 1: Indicadores de apelotonamiento de buses del servicio 506

La identificación de estos dos indicadores permite desarrollar dos modelos de regresión lineal, una para cada indicador como variables dependiente. Una transformación de Box-Cox fue aplicada a los indicadores para obtener variables dependientes con una distribución de probabilidad que se asemeje en mayor medida a la distribución de probabilidad normal. Los coeficientes obtenidos del procedimiento de transformación Box-Cox son -1,58 y -4,56 para el indicador continuo y el indicador discreto respectivamente. Por simplicidad, estos factores fueron aproximados a -1 y a -4, resultando las transformaciones que se observan en la ecuación (5) y (6).

$$\text{Transformed IPO} = \frac{1}{\text{IPO}} \quad (5)$$

$$\text{Transformed Y} = \frac{1}{(Y + 1)^4} \quad (6)$$

Finalmente los valores promedio, máximo y mínimo de los indicadores y sus transformaciones se muestran en la tabla

**Tabla 1: Descripción estadística de indicadores y sus transformaciones**

Variables	Servicio 506			Servicio 506e			Servicio 206		
	Prom	Min	Max	Prom	Min	Max	Prom	Min	Max
IPO index	1,5	1	4,5	1,3	1	2,2	1,7	1	4,5
Transformed IPO index	0,7	0,2	1	0,8	0,5	1	0,6	0,2	1
Y index	0,2	0	0,8	0,1	0	0,5	0,2	0	0,75
Transformed Y index	0,7	0,1	1	0,8	0,2	1	0,6	0,1	1

### 3.2 Variables Independientes

Los factores usados como variables explicativas en los modelos son las siguientes: el número consecutivo de paradero, que permite capturar el largo de la ruta recorrida; la distancia entre paraderos; el tipo de prioridad de vía; el tipo de paradero (tradicional o zona paga); el efecto de los semáforos; el número de semáforos coordinados en una red; la existencia del bloqueo de tráfico; la frecuencia programada; incidentes; la variabilidad de intervalo de tiempo entre buses consecutivos al inicio de la ruta; presencia de congestión; presencia de semáforos aislados (no coordinados); la distribución del tipo de buses (articulados y no articulados); variación en la actividad de buses consecutivos (detención y no detención); subida de pasajeros y la variación entre el número total de subidas por expedición y entre el número total de bajadas por expedición; la antigüedad del bus medida en años y el operador de los buses.

Para incorporar el efecto de los semáforos, se consideró que la demora en semáforos tiene dos componentes: la demora uniforme, que corresponde a la demora promedio de un vehículo multiplicada por la probabilidad de parar en el semáforos, cuando el grado de saturación es menor a uno; y un segundo componente causado por la aleatoriedad y la sobresaturación de la cola (en este caso el grado de saturación es mayor o igual a uno). Como el valor del grado de saturación y el factor de carga, el cual es utilizado para calcular la probabilidad de detenerse, es ignorado, se consideró solamente la demora promedio de los buses, la cual está dada por la ecuación (7).

$$D = \frac{c * (1 - u)}{2} * N \quad (7)$$

Donde D es la demora promedio de buses en semáforos, c es el tiempo de ciclo de los semáforos, u es la proporción del tiempo de verde en el tiempo de ciclo, y N es el número de semáforos entre el paradero evaluado y el anterior a este. El tiempo de ciclo y de verde son calculados como el tiempo de ciclo promedio y el tiempo de verde promedio de los semáforos que se encuentran en la sección entre el paradero evaluado y el anterior a este, luego se obtiene la demora promedio de buses en semáforos por cada paradero. Cuando los semáforos están coordinados en redes, como en este caso, se puede lograr una reducción significativa en la demora uniforme (Gibson et al, 2015), para incorporar este efecto variables dummy son utilizadas para indicar si la sección entre el paradero evaluado y el anterior a este presenta semáforos coordinados. Los datos de tiempos de

semáforos se obtuvieron de la Unidad Operativa de Control de Tránsito (UOCT), unidad gubernamental encargada de la programación y coordinación de semáforos.

El bloqueo de tráfico en paraderos debido a la congestión ha sido identificado como una de las causas de variabilidad de buses. Específicamente, cuando un paradero está cerca de una intersección semaforizada se puede generar un tiempo muerto, atribuido a la cola generada por la luz roja (Salvin et al, 2013). Por esto, se incorporó el potencial efecto del bloqueo de tráfico como una variable dummy que indica si el paradero anterior al evaluado está localizada a menos de 30 metros antes de un semáforo.

Para identificar incidentes, se utilizaron datos históricos del Directorio de Transporte Público Metropolitano (DTPM), el cual indica el problema, el horario de inicio y fin de este, la fecha y los servicios de buses afectados. Para los servicios analizados, se identificaron incidentes reportados como “desviación no programada” y “bus en pana”; sin embargo, la situación registrada de falla del bus no generó un efecto de irregularidad en el siguiente, puesto que otro bus llega a reemplazar al bus fuera de servicio casi al mismo instante en que este deja de funcionar, razón por la cual este estudio ignora este incidente. Los incidentes son capturados con una variable que indica la proporción de expediciones que han sido desviadas en algún lugar anterior al paradero evaluado.

Con respecto a la congestión, se utilizó la metodología procesada por Cortés et al. (2011) para calcular la velocidad comercial de los buses. Se identificó la variable congestión como una variable dummy que toma valor uno cuando la velocidad comercial es 15km/hr o menos y su desviación estándar es 30% o menos de la velocidad comercial del correspondiente segmento de vía y periodo de tiempo.

El número de subidas y bajadas de pasajeros es obtenido del sistema AFC, usando la metodología propuesta por Munizaga and Palma (2012) para estimar los paraderos de bajadas.

Las variables dummy utilizadas para modelar son descritas en la tabla 1 con el respectivo porcentaje de veces en que estas toman valor. Las variables continuas son descritas en la tabla 2, donde se muestra el promedio de estas en las observaciones utilizadas.

**Tabla 2: Porcentaje de casos indicados por variables independientes de tipo dummy**

<b>Variables Independientes de tipo Dummy</b>	<b>Servicio 506</b>	<b>Servicio 506e</b>	<b>Servicio 206</b>
Una pista solo bus	10,6%	10,5%	0%
Dos pistas solo bus	6,5%	2,4%	0%
Corredor segregado	29,4%	31%	42,6%
Vía exclusiva	0%	0%	7,5%
Zona paga	1%	2,76%	0,6%
Paradero anterior es zona paga	1%	3%	0,6%
Dos semáforos coordinados	17,7%	19,3%	30,3%
Tres semáforos coordinados	12,9%	15,8%	6,1%
Cuatro semáforos coordinados	2,2%	13%	0%



Variables Independientes de tipo Dummy	Servicio 506	Servicio 506e	Servicio 206
Seis semáforos coordinados	0%	6,8%	0%
Semáforos coordinados en dos redes diferentes (4 y 2)	0%	2,7%	0%
Semáforos coordinados en dos redes diferentes (3 y 8)	0%	2,4%	0%
Bloqueo de tráfico	32,3%	23,4%	30,1%
Congestion	5,2%	3,8%	3%
Presencia de semáforo no coordinado	0,8%	0,6%	0,18%
El porcentaje de buses articulados que han pasado por el paradero está entre 21% y 40%	19%	0%	19,5%
El porcentaje de buses articulados que han pasado por el paradero está entre 41% y 60%	24%	0%	28,1%
El porcentaje de buses articulados que han pasado por el paradero está entre 61% y 80%	26,2%	0%	28,5%
El porcentaje de buses articulados que han pasado por el paradero está entre 81% y 100%	14,2%	0%	17,8%
Número total de observaciones	<b>15786</b>	<b>1196</b>	<b>3325</b>

Tabla 3: Descripción estadística de variables independientes

Variables Independientes	Servicio 506			Servicio 506e			Servicio 206		
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Frecuencia programada [buses/hora]	8,5	3	14	5,2	5	5,5	7,9	5	11
Distancia entre el paradero evaluado y el anterior a este [km]	0,4	0,2	1,1	0,8	0,2	2,2	0,4	0,1	0,6
Proporción de expediciones que han sido desviadas	0,01	0	1	0,01	0	0,8	0	0	0
Indicador IPO en el despacho	1,2	1	2	1,2	1	3,3	1,5	1	2,5
Indicador IPO transformado en el despacho	0,9	0,5	1	0,9	0,3	0,9	0,7	0,4	1
Indicador Y en el despacho	0,02	0	0,4	0,01	0	0,3	0,1	0	0,5
Indicador Y transformado en el despacho	0,9	0,3	1	0,9	0,4	1	0,7	0,2	1
Edad promedio de buses [años]	3,8	0	5	5	5	5,0	6	3	7,5
Subidas de pasajeros acumuladas promedio por expedición en cada paradero	42	0	206,3	75,1	0	326	30,5	0	331
Coefficiente de variación de subidas de pasajeros acumuladas promedio por expedición en cada paradero	0,4	0	1,4	0,6	0	2,5	0,06	0	1,13
Coefficiente de variación de bajadas de pasajeros acumuladas promedio por expedición en cada paradero	0,4	0	1,2	0,9	0	2,7	0,04	0	1,14
Demora uniforme promedio [segundos]	25,9	0	120,5	55	0	272,5	24,7	0	111,5
Proporción de expediciones que son del mismo tipo de bus que la expediciones anterior a esta	0,6	0	1	1	1	1	0,5	0	1
Proporción de expediciones que realizan la misma acción que la expedición anterior a esta	0,6	0	1	0,6	0	1	0,6	0	1
Número total de observaciones	<b>15786</b>			<b>1196</b>			<b>3325</b>		

### 3. RESULTADOS DE MODELOS

Se estimaron modelos robustos para evitar correlación serial entre paraderos y heterocedasticidad de los errores, problemas que fueron detectados en las estimaciones iniciales. También se usaron

las pruebas estadísticas t de Student y F de Fisher para corroborar la eficiencia y significancia de las variables.

Los modelos finales se muestran en la tabla 4, en donde se detalla el coeficiente ( $\beta_{no\ estandarizado}$ ) junto a su t estadístico y el coeficiente estandarizado ( $\beta_{estandarizado}$ ) de cada variable independiente ( $X_i$ ). Los coeficientes tipificados o estandarizados (ver ecuación 8) ayudan a valorar la importancia relativa de cada variable independiente en la variable dependiente o estudiada (Z). Los modelos finales fueron obtenidos removiendo variables no significativas y variables con signo no esperado, además se agruparon variables que tenían coeficientes similares o que su agrupación mejoraba ya sea la significancia o el signo de la variable. Ambos modelos presentan un valor F significativo estadísticamente, además ambos modelos fueron validados con respecto a los modelos iniciales mediante la prueba F para modelos restringidos.

$$\beta_{estandarizado} = \beta_{no\ estandarizado} * \frac{Desviación\ Típica\ (X_i)}{Desviación\ Típica\ (Z)} \quad (8)$$

Los resultados de ambos modelos indican que las variables número de paradero o largo de la ruta, apelotonamiento de buses en el despacho o inicio de la ruta y frecuencia programada son los factores de mayor relevancia para el fenómeno en estudio. Estos tres factores presentan relación directa con el apelotonamiento de buses y son estadísticamente significativos.

Adicionalmente las variables relacionadas a la variación de la demanda son significativas en ambos modelos, indicando que un incremento en la diferencia de subidas y bajadas entre las expediciones contribuye a aumentar el apelotonamiento de buses. La variable que indica subida de pasajeros promedio por expedición resultó ser significativa solo en el modelo de indicador continuo, mostrando una relación directa entre esta y la irregularidad de buses. En esta línea también, la variable que indica la proporción de expediciones que realizan la misma actividad (parar o no parar) que la expedición anterior, está presente en ambos modelos y es estadísticamente significativa, indicando que una actividad más homogénea entre las expediciones permite disminuir la intensidad de la irregularidad de buses.

La distancia entre paraderos se observa solo en el modelo de indicador discreto, como una variable estadísticamente significativa que se relaciona directamente al apelotonamiento de buses. Esto se debe a que los conductores tienden a acelerar los buses cuando el segmento entre paraderos lo permite, generando que el intervalo con el bus que le antecede disminuya. Con respecto al tipo de prioridad de bus, ambos modelos identifican que la presencia de un corredor segregado o vía exclusiva reduce el apelotonamiento de buses, factores que resultan estadísticamente significativos. Este resultado es consistente con la información entregada por expertos (equipo de centro de monitoreo de buses de DTPM), quienes han observado menos irregularidad de intervalo entre buses en las rutas que usan corredor segregado.

Tabla 4: Modelos Finales

Variables	Indicador Continuo		Indicador Discreto	
	Coeficientes (t-stat)	Coeficientes tipificados	Coeficientes (t-stat)	Coeficientes tipificados
Número consecutivo de paradero	-0.003 <b>(-57.6)</b>	-0.396	-0.006 <b>(-54.3)</b>	-0.383
Distancia entre el paradero evaluado y el anterior a este [km]			-0.021 <b>(-2)</b>	-0.014
Dos pistas solo bus (Variable dummy)	-0.012 <b>(-2.5)</b>	-0.015		
Corredor Segregado (Variable dummy)	0.011 <b>(4.8)</b>	0.031	0.033 <b>(6.4)</b>	0.047
Vía Exclusiva (Variable dummy)	0.094 <b>(7.7)</b>	0.06	0.191 <b>(8.3)</b>	0.064
Demora uniforme promedio	-0.0001 <b>(-1.9)</b>	-0.014		
Dos semáforos coordinados (variable dummy)			0.015 <b>(2.6)</b>	0.019
Dos o tres o cuatro semáforos coordinados (variable dummy)	0.003 <b>(1.2)</b>	0.009		
Bloqueo de tráfico (variables dummy)			-0.01 <b>(-2.2)</b>	-0.014
Frecuencia programada [buses/hora]	-0.012 <b>(-28.4)</b>	-0.177	-0.03 <b>(-36.6)</b>	-0.234
Proporción de expediciones que han sido desviadas	-0.075 <b>(-6.1)</b>	-0.034	-0.092 <b>(-3.5)</b>	-0.022
Irregularidad en despacho	0.624 <b>(67.1)</b>	-0.433	0.319 <b>(27.4)</b>	0.197
Congestión (variable dummy)	-0.005 <b>(-1.1)</b>	-0.006		
Proporción de expediciones que son del mismo tipo de bus que la expediciones anterior a esta	0.013 <b>(2.8)</b>	0.021	0.016 <b>(1.7)</b>	0.014
Proporción de expediciones que realizan la misma acción que la expedición anterior a esta	0.014 <b>(4.3)</b>	0.023	0.032 <b>(4.5)</b>	0.028
Subidas de pasajeros acumuladas promedio por expedición en cada paradero	-0.0001 <b>(-3.4)</b>	-0.025		
Coeficiente de variación de subidas de pasajeros acumuladas promedio por expedición en cada paradero	-0.068 <b>(-17.4)</b>	-0.104	-0.103 <b>(-12.5)</b>	-0.082
Coeficiente de variación de bajadas de pasajeros acumuladas promedio por expedición en cada paradero	-0.036 <b>(-12.5)</b>	-0.067	-0.086 <b>(-12.8)</b>	-0.082
El porcentaje de buses articulados que ha pasado está entre 21% y 40% (variable dummy)	-0.048 <b>(-13.4)</b>	-0.11	-0.123 <b>(-15.6)</b>	-0.146
El porcentaje de buses articulados que ha pasado está entre 41% y 80% (variable dummy)	-0.052 <b>(-15.6)</b>	-0.153	-0.109 <b>(-14.9)</b>	-0.165
El porcentaje de buses articulados que ha pasado está entre 81% y 100% (variable dummy)	-0.039 <b>(-11.5)</b>	-0.092	-0.083 <b>(-11.3)</b>	-0.101
Operador servicio 506-506e (variable dummy)	0.039 <b>(11.1)</b>	0.083	0.127 <b>(16.8)</b>	0.143
Constante	0.428 <b>(42.4)</b>		0.837 <b>(50.5)</b>	
R2	0.466		0.292	
Error	0.125		0.276	

La variable una pista solo bus no está presente en ninguno de los modelos, por lo que se puede decir que en términos de regularidad de buses, este tipo de vía no presenta diferencias con el tipo de vía tráfico mixto, lo cual se puede explicar porque este tipo de prioridad de vía no siempre es respetado por auto particulares. Por otro lado la variable dos pistas solo bus, se hace presente de manera estadísticamente significativa en el modelo de indicador continuo, revelando que en este tipo de vía se incrementa el apelotonamiento de buses, lo cual es contradictorio al resultado deseable, no obstante esta relación directa puede deberse a los vehículos estacionados permanentemente en estas vías (como se explicó en el caso de estudio).

De acuerdo a los modelos, la presencia de zona paga en un paradero o en el paradero anterior al evaluado, no presenta efecto en el apelotonamiento de buses.

Con respecto al efecto de los semáforos, el modelo de indicador continuo muestra que la demora uniforme promedio (D) intensifica el fenómeno, lo cual indica que el número de semáforos y la demora en estos, incrementa el apelotonamiento de buses. En oposición, se encuentra el efecto positivo de los semáforos coordinados, variable que es detectada por el modelo de indicador discreto para el caso de dos semáforos coordinados, y por el modelo con indicador continuo para el caso de cuatro o menos semáforos coordinados. Se hace notar que esta última variable (cuatro o menos semáforos coordinados) no es significativa y su efecto combinado fue verificado con la prueba estadística F. El efecto de bloqueo de tráfico en apelotonamiento de buses, explicado por una variable dummy que captura la proximidad de paraderos a un semáforo, está presente solo en el caso de indicador discreto con significancia estadística y mostrando un efecto de incremento en el problema estudiado.

Como se esperaba, la ocurrencia de incidentes en un expedición (bus desviado, sin programación previa, de su camino) contribuye a aumentar el apelotonamiento de buses. Esta variable es significativa en ambos modelos.

Congestión es una variable que fue identificada en la literatura y en entrevistas a expertos del monitoreo de buses de Transantiago, sin embargo esta se presenta solo en modelo con indicador continuo y sin significancia estadística al 95%. Para este caso, el modelo indica que si el paradero-periodo de tiempo está sujeto a congestión, el apelotonamiento de buses aumenta.

En la línea de distribución de flota ambos modelos capturan este efecto mediante tres grupos de porcentajes de buses articulados que han pasado por el paradero-periodo evaluado (0-20, 21-80, 81-100). Los grupos en los extremos (0-20, 81-100) son aquellos que presentan una flota más homogénea; luego, una flota no homogénea incrementa el apelotonamiento de buses. Adicionalmente, una flota homogénea de buses no articulados es preferible sobre buses articulados. Un efecto similar es capturado de manera desagregada con la variable que indica el porcentaje de buses que son del mismo tipo que el anterior (articulado o no articulado), la cual aparece en ambos modelos, aunque solo es significativa en el modelo de indicador continuo, y contribuye a disminuir el apelotonamiento de buses.

Finalmente la variable operador servicio 506-506e es una variable significativa en ambos modelos, e indica una disminución del apelotonamiento de buses para este operador en específico comparado con el operador del servicio 206.

Variables que inicialmente fueron incorporadas en los modelos y que no se presentan en los modelos finales por consideraciones estadísticas son la edad promedio de los buses, algunas combinaciones de variables dummy referidas a semáforos coordinados y a porcentaje de buses articulados, variables referidas a pistas solo bus, a paradero-periodo zona paga y a semáforo aislado.

Con respecto a la importancia relativa de cada variable independiente dentro de la ecuación, se tiene que en el modelo con indicador continuo las variables en orden descendiente de importancia son: irregularidad en despacho, número consecutivo de paradero, frecuencia programada, porcentaje de buses articulados que ha pasado está entre 41% y 80%, porcentaje de buses articulados que ha pasado está entre 21% y 40%, coeficiente de variación de subidas de pasajeros acumuladas promedio por expedición en cada paradero, porcentaje de buses articulados que ha pasado está entre 81% y 100%, operador servicio 506-506e, coeficiente de variación de bajadas de pasajeros acumuladas promedio por expedición en cada paradero, vía exclusiva, proporción de expediciones que han sido desviadas, corredor segregado, subidas de pasajeros acumuladas promedio por expedición en cada paradero, proporción de expediciones que realizan la misma acción que la expedición anterior a esta, proporción de expediciones que son del mismo tipo de bus que la expediciones anterior a esta, dos pistas solo bus, demora uniforme promedio, dos o tres o cuatro semáforos coordinados y congestión.

En el modelo con indicador discreto las variables independientes en orden descendiente de importancia son: número consecutivo de paradero, frecuencia programada, irregularidad en despacho, porcentaje de buses articulados que ha pasado está entre 41% y 80%, porcentaje de buses articulados que ha pasado está entre 21% y 40%, operador servicio 506-506e, porcentaje de buses articulados que ha pasado está entre 81% y 100%, coeficiente de variación de subidas de pasajeros acumuladas promedio por expedición en cada paradero, coeficiente de variación de bajadas de pasajeros acumuladas promedio por expedición en cada paradero, vía exclusiva, corredor segregado, dos semáforos coordinados, proporción de expediciones que han sido desviadas, proporción de expediciones que realizan la misma acción que la expedición anterior a esta, distancia entre el paradero evaluado y el anterior a este, bloqueo de tráfico y proporción de expediciones que son del mismo tipo de bus que la expediciones anterior a esta.

#### **4. CONCLUSIONES**

En este trabajo se describe el fenómeno de apelonamiento de buses mediante un indicador de tipo continuo y otro de tipo discreto. Usando datos disponibles de la operación del sistema de transporte público (AVL y AFC), se construyeron mapas de apelonamiento de buses y se calibraron modelos de regresión lineal para explicar el fenómeno. Se desarrollaron dos modelos: uno con el indicador continuo como variable dependiente y otro con el indicador discreto como variable dependiente. Basados en la revisión bibliográfica y en consultas a expertos de monitoreo de buses de Transantiago, construimos hipótesis con respecto a variables relevantes en el fenómeno estudiado, que luego fueron validadas mediante la modelación de los datos. Algunos resultados previos y opiniones de expertos fueron confirmados y otros no.

De acuerdo a los datos analizados, los factores que contribuyen a incrementar el apelotonamiento de buses son los siguientes (en orden descendente de importancia):

- Largo de la ruta
- Frecuencia programada alta
- Despacho irregular de buses
- Flota no homogénea, es decir, combinación de buses articulados y no articulados
- Alta demanda y variabilidad de esta
- Buses desviados
- Semáforos (demora uniforme y posibilidad de intersecciones con bloqueo de tráfico)
- Distancia entre paraderos alta
- Vías con vehículos estacionados

Factores que contribuyen a disminuir el apelotonamiento de buses son los siguientes (en orden descendente de importancia):

- Presencia de vía exclusiva
- Presencia de corredor segregado
- Actividad (detenerse o no detenerse) homogénea en paraderos entre buses consecutivos
- Coordinación de semáforos

También se hace notar la diferencia entre operadores a nivel de apelotonamiento de buses, puesto que ambos modelos mostraron como resultado que el operador del servicio 506-506e genera menos apelotonamiento de buses que el operador del servicio 206.

Estos resultados proveen algunos lineamientos para diseñar y operar los servicios de buses, tomando en consideración la importancia del problema de apelotonamiento de buses. De acuerdo a los resultados, nosotros recomendamos a los operadores que los servicios se operen con flotas del mismo tipo de bus, ser cuidadosos con la regularidad en el despacho, y si es posible diseñar servicios que pasen por corredor segregado o vía exclusiva.

Los resultados de modelación que se obtuvieron confirman resultados previos de diferentes autores que fueron mencionados en la introducción de este trabajo. La relación entre el apelotonamiento de buses y la distancia recorrida fue identificada por Bellei y Gkoumas (2010), y el efecto que genera la demanda en la irregularidad de buses fue identificada por Bellei y Gkoumas (2010), Moreira-Matias et al. (2012) y Man y Wenquan (2013). La relación positiva entre la frecuencia y el apelotonamiento de buses fue identificada por Moreira-Matias et al. (2012) y Man y Wenquan (2013). La relación directa entre el problema estudiado y la irregularidad al inicio de la ruta (despacho de buses irregulares) fue identificada por Moreira-Matias et al. (2012). Sin embargo, hay otros efectos mencionados en la literatura que no fueron encontrados en nuestros modelos. Por ejemplo, no encontramos una relación significativa entre la congestión y el apelotonamiento de buses, la cual fue encontrada por Man y Wenquan (2013).

Los diagramas de apelotonamiento de buses presentados en este trabajo, muestran como el fenómeno cambia en el tiempo y el espacio. Con respecto a la comparación entre el indicador continuo y discreto, el indicador continuo explica de mejor manera aquellos casos de intensidad media de apelotonamiento de buses, mientras que el indicador discreto explica de mejor manera aquellos casos extremos o severos de apelotonamiento de buses.

## AGRADECIMIENTOS

El trabajo de investigación presentado en este artículo está parcialmente financiado por Fondef (Grant D10I1002) y el Instituto de Sistemas Complejos de Ingeniería, Chile (Grants ICM P-05-004-F, CONICYT FBO16).

## REFERENCIAS

Albright, E. y Figliozzi, M. (2012). Analysis of the Impacts of Transit Signal Priority on Bus Bunching and Performance. **Proceedings of the Conference on Advanced Systems for Public Transport (CASPT)**, Santiago, Chile.

Bartholdi, J.J., y Eisenstein, D.D. (2012). A self-coordinating bus route to resist bus bunching. **Transportation Research Part B: Methodological**, 46, 481-491.

Bellei, G. y Gkoumas, K. (2010). Transit Vehicles' Headway Distribution and Service Irregularity. **Public Transport**, 2, 269-289.

Byon, Y.J., Cortés, C.E., Martinez, F.J., Munizaga, M.A., y Zúñiga, M. (2011). Transit Performance Monitoring and Analysis with Massive GPS Bus Probes of Transantiago in Santiago, Chile: Emphasis on Development of Indices for Bunching and Schedule Adherence. **Transportation Research Board 90th Annual Meeting**, 11-2233.

Chen, Q., Adida, E. y Lin, J. (2013b). Implementation of an iterative headway-based bus holding strategy with real-time information. **Public Transport**, 4(3), 165-186.

Cortés, C.E., Gibson, J., Gschwender, A., Munizaga, M.A. y Zúñiga, M. (2011). Commercial bus speed diagnosis based on GPS-monitored data. **Transportation Research Part C: Emerging Technologies**, 19(4), 695-707.

Daganzo, C. (2009). A Headway-Based approach to eliminate Bus Bunching. **Transportation Research Part B**, 43, 913-921.

Gibson, J., Munizaga, M.A., Schneider, C., y Tirachini, A. (2015). Median Busways versus Mixed-Traffic: Estimation of Bus Travel Time under Different Priority Conditions with Explicit Modelling of Delay at Traffic Signals. **Transportation Research Board 94th Annual Meeting**, 15-1260.

Holroyd, E.M. y Scraggs, D.A. (1966). Waiting Times for Buses in Central London. **Traffic Engineering and Control**, 8, 158-160.

Kittelson, P. B., Quade, K., y Hunter-Zaworski, K. M. (2003). **Transit Capacity and Quality of Service Manual**. TCRP Report 100, National Research Council, Washington, D.C.

- Lizana, P., Muñoz, J.C., Giesen, R. y Delgado, F. (2014). Bus Control Strategy Application: Case Study of Santiago Transit System. **Procedia Computer Science**, 32, 397-404.
- Man, Z. y Wenquan, L. (2013). Factors affecting headway regularity on bus routes. **Journal of Southeast University**, 29, 99-102.
- Moreira-Matias, L., Ferreira, C., Gama, J., Mendes-Moreira, J. y De Sousa, J. F. (2012). Bus bunching detection by mining sequences of headway deviations. In Advances in Data Mining. **Applications and Theoretical Aspects**. Springer Berlin Heidelberg, 77-91.
- Munizaga, M.A., y Palma, C. (2012). Estimation of a disaggregate multimodal public transport Origin–Destination matrix from passive smartcard data from Santiago, Chile. **Transportation Research Part C: Emerging Technologies**, 24, 9-18.
- Osuna E.E., y Newell G.F. (1972). Control strategies for an Idealized Public Transportation System. **Transportation Science**, 6, 52-72.
- Slavin, C., Feng, W., Figliozzi, M. y Koonce, P. (2013). A Statistical Study of the Impacts of SCATS Adaptive Traffic Signal Control on Traffic and Transit Performance. **Transportation Research Record**, 2356, 117-126.
- Turnquist M. A. (1978). A model for investigating the effects of service frequency and reliability on bus passenger waiting times. **Transportation Research Record** 663, 70-73.
- Welding, P.I. (1957). The Instability of a Close-Interval Service. **Operational Research Quarterly**, 8, 133-142.