



Análisis del tiempo de ascenso de pasajeros en buses de la ciudad de Bogotá usando un diseño factorial

Analyze passengers' boarding time in buses in the city of Bogota using a factorial design

Espejo-Fandiño Fredy Leandro

Ministerio de Transporte, Bogotá, Colombia

Correo: flespejof@unal.edu.co

<https://orcid.org/0000-0002-1159-7602>

Resumen

Mediante un estudio comparativo por observación se analizan en un diseño factorial los tiempos de ascenso por pasajero en rutas urbanas de la ciudad de Bogotá. A través de mediciones en terreno de los tiempos de detención de los buses en distintos paraderos de la ciudad y dando prioridad a los casos donde suceden ascensos de usuarios exclusivamente, fueron involucrados dos tratamientos en el experimento, la tipología vehicular con tres niveles (bus pequeño, bus mediano y bus grande) y la ocupación del bus con dos niveles (pocos y muchos pasajeros de pie). Se encontró que los datos originales no siguen una distribución normal y por tanto la variable respuesta fue objeto de transformación para estabilizar la varianza. Sobre los datos transformados, mediante un análisis de varianza de dos factores, se probó que tanto las diferentes tipologías vehiculares que operan en el transporte público de la ciudad como la ocupación del bus, tienen efectos en el tiempo de ascenso por pasajero. Con comparaciones múltiples se encontró que los buses pequeños, para bajos y altos niveles de ocupación, presentan tiempos de ascenso por pasajero significativamente diferentes en comparación con los buses medianos y grandes, y que la ocupación del bus genera diferencias significativas en la variable respuesta solo para el bus mediano y grande. El modelo resultante fue un diseño factorial sin interacción, el cual produjo una explicación de 46 % de variabilidad de los datos transformados.

Descriptor: Transporte público, ascenso de pasajeros, tiempo de parada, ocupación vehicular, diseño factorial.

Abstract

Using an observational comparative study, boarding times per passenger in urban routes in the city of Bogota are analyzed through a factorial design. By measuring the dwell time at different bus stops in the city, prioritizing exclusively the cases where the users get on the buses, two treatment were incorporated into the experiment, the vehicular typology in three levels (small, medium middle and large size bus) and bus occupancy in two levels (few and many passengers standing). It was found that the original data do not follow a normal distribution and therefore the response variable was transformed to stabilize the variance. On the transformed data, by means of a two-factor analysis of variance, it was proved that the different vehicle typologies that operate in the public transport of the city, such as bus occupation, have effects on the boarding time per passenger. With multiple comparisons, it was found that small buses, for low and high levels of occupancy, show significantly different boarding times per passenger when compared against the medium and large size buses, and that the occupation of the bus generates significant differences in the response variable only for the medium bus and the large bus. The resulting model was a factorial design without interaction which produced an explanation of 46 % variability of the transformed data.

Keywords: Public transportation, passengers boarding, dwell time, vehicle occupancy, factorial design.

INTRODUCCIÓN

Una apropiada caracterización de la operación del transporte público como función no solo de los tiempos viaje sino también de los tiempos de parada, es una herramienta de interés para programadores, coordinadores de la operación y gestores de planeación y control. La operación del transporte público tiene tres aspectos que determinan el nivel de servicio: capacidad, velocidad y confiabilidad. Estas variables se ven afectadas por el tiempo de recorrido que puede entenderse como la suma de dos componentes: el tiempo de viaje y el tiempo de parada. El tiempo viaje depende de varios factores como longitud de la ruta, actividad de pasajeros, número de intersecciones señalizadas, número de paradas, características de los vehículos y restricciones en la vía (Abkowitz y Engelstein, 1983). Por su parte, el tiempo de parada se define como el tiempo necesario para servir pasajeros en la puerta más ocupada, más el tiempo requerido para abrir y cerrar las puertas, además de cualquier tiempo de abordaje muerto. Cualquier otro tiempo como el que ocurre por esperar debido a una señal de tráfico, o esperar que otro vehículo se mueva, o esperar un pasajero que llega tarde a la parada, no forma parte de esta definición (Transit Cooperative Research Program, TCRP, 2013).

El análisis del tiempo de parada de un bus es importante para estudiar la confiabilidad de los sistemas de transporte público, sirve para la predicción de los tiempos de recorrido, para estimar la capacidad de una parada e incluso la capacidad de una línea de transporte (Peña y Moreno, 2014). De igual manera, puede ser sustancial para disminuir la frecuencia óptima del bus, incrementar la capacidad del bus y el subsidio óptimo requerido para operar el sistema (Jara y Tirachini, 2013; Jiang *et al.*, 2015).

El tiempo de parada, dependiendo la ciudad y el esquema de operación, puede representar porcentajes considerables del tiempo total de recorrido del bus. Por ejemplo, Tirachini (2013) determinó en Sidney que el tiempo gastado en las paradas es de 22.6 % respecto al tiempo total de viaje, mientras que Bertini y El-Geneidy (2004) lo estimaron en 16 % en la ciudad de Portland.

La estimación y el análisis del tiempo de parada en el transporte público tiene un importante acervo de estudios. En esquemas operacionales con sistema de pago al interior del bus, la mayoría de estudios coinciden que el factor preponderante en el tiempo gastado en una parada, corresponde al ascenso y descenso de pasajeros, pues es la actividad que mayor duración involucra en estos sitios (Strathman y Hopper, 1992; Zhang y Then; 2013). Adicionalmente, se ha encontrado en diversos

estudios que el tiempo de ascenso es el que determina la demora en las paradas, pues el sistema de pago y la plataforma de acceso (a veces a desnivel), hacen que los tiempos individuales de abordaje sean casi siempre superiores a los tiempos de bajada por pasajero (Dueker *et al.*, 2004; El-Geneidy y Vijayakumar, 2011).

Sin embargo, existen diferentes aspectos como la rotación de la demanda y la congestión que no son sencillos de analizar debido a que presentan una alta variabilidad, y que inciden en la poca confiabilidad de los tiempos de parada. En la mayoría de ciudades, los horarios pico representan verdaderos desafíos para el transporte, porque en este período coincide el momento de mayor congestión vehicular en las vías con el de mayor demanda de pasajeros. El resultado principal de esta combinación para el transporte público es que la regularidad, confiabilidad y nivel de servicio disminuyen notablemente en estos períodos.

En concordancia con lo anterior, la presente investigación se enfoca en analizar los tiempos de ascenso de pasajeros en rutas urbanas del sistema de transporte público colectivo de la ciudad de Bogotá, mediante un estudio comparativo por observación en un diseño factorial que incluye dos tratamientos, la tipología vehicular y los niveles de ocupación del bus. Estas rutas se caracterizan por servir largas distancias sobre corredores de demanda media, transitan en tráfico mixto, poseen sistema de pago al interior del bus y tienen un alto número de paradas.

ANTECEDENTES

En la literatura se encuentran numerosos estudios para la modelación de los tiempos de parada, los cuales generalmente invierten una parte importante del análisis a los tiempos de ascenso y descenso de pasajeros. Algunos autores, han centrado su investigación en determinar cómo interactúa el tiempo de parada con el diseño de la infraestructura y los aspectos del tránsito vehicular, otros con aspectos propios de los usuarios, y otros con la configuración del bus y el esquema de operación.

En el primer grupo de estudio, que se enfoca en la infraestructura, se encuentra que el tiempo de parada es función de variables como el carril adyacente sobre la vía (Valencia y Fernández, 2012), la diferencia entre una estación tipo BRT y otro tipo de paraderos (Li *et al.*, 2012; Wang *et al.*, 2016), la diferencia entre una parada sobre bahía y una en acera (Meng y Qu, 2013), el fenómeno de colas y el número de estacionamientos de la parada (Bian *et al.*, 2015), y la ubicación de paraderos a mitad de acera y cerca de la intersección (Arhin *et al.*, 2016).

En el segundo grupo, que se enfoca en los usuarios, el TCRP (2013) indica que la variabilidad del tiempo de parada depende también de la presencia de personas con sillas de ruedas, bicicletas y de las interacciones que se dan entre el conductor y pasajeros. Daamen *et al.* (2008), realizan un experimento en donde encuentran que los pasajeros que cargan maletines decrecen la capacidad de la puerta.

El tercer grupo, que corresponde al de mayor foco de investigación y que se enfoca en la configuración del bus y el esquema de operación, ha encontrado que el tiempo de parada tiene relación con el número, ancho y altura de las puertas, método de pago, presencia de escalones (Sun *et al.*, 2014; Fernández, 2015; Tirachini, 2015), así como con los períodos de operación durante el día y la tipología vehicular (El-Geneidy y Vijayakumar, 2011).

Cuestión interesante es que la ocupación del bus en el momento de la parada, no tiene tanta investigación como los aspectos anteriores, tal vez porque es más difícil de registrar y agrega mayor variabilidad al resultado. Sin embargo, se ha determinado que esta variable influye el tiempo de parada a partir de una ocupación de 60 % (Fletcher, 2012), que depende de la distribución interna del vehículo (Katz y Garrow, 2012) y que incluir la ocupación mejora la precisión de la estimación del tiempo de parada (Zhang y Then, 2013).

METODOLOGÍA

La hipótesis de la investigación, consiste en que las distintas tipologías vehiculares existentes en el transporte colectivo de la ciudad de Bogotá, involucran diferentes dimensiones de la puerta de acceso y de la configuración del bus, que pueden tener efecto sobre el tiempo de ascenso por pasajero. La ocupación del bus, sobre todo cuando se encuentra muy lleno, también puede tener efecto sobre esta respuesta.

Para probar esta hipótesis se realiza una aproximación a partir de un estudio comparativo por observación, que corresponde a una técnica estadística relacionada con el diseño de experimentos, el cual nos ayuda a identificar qué factores o variables afectan el comportamiento de un proceso productivo o de un servicio (Gutiérrez *et al.*, 2012).

Los estudios comparativos por observación son aquellos para los que se desea realizar un experimento, pero no es posible por razones prácticas o éticas, por lo que se enfocan en relacionar las respuestas con las condiciones de los tratamientos. Los factores o tratamientos son el conjunto de circunstancias creadas para el experimento, en respuesta a las hipótesis de investiga-

ción y son el centro de la misma; la unidad experimental o de observación es la entidad física expuesta al tratamiento. Las réplicas son básicamente la repetición del experimento o estudio básico (Montgomery, 2013).

De esta manera, lo que se pretende analizar es el efecto de la ocupación del bus y la tipología vehicular de rutas de transporte urbano sobre el tiempo de ascenso por pasajero y si existe alguna posible interacción entre estos factores. El enfoque más adecuado para trabajar este problema es un diseño factorial que consiste en una estrategia experimental o de observación en la que existen dos o más factores que se hacen variar en conjunto, en lugar de uno a la vez (Kuehl, 2001; Montgomery, 2013).

DISEÑO DEL TRATAMIENTO

Se realizaron observaciones en terreno del tiempo de detención de los buses colectivos, en distintos paraderos de la ciudad de Bogotá. En el período pico de la mañana de las 06:00 a 08:00 horas se tomaron mediciones en zonas residenciales de la ciudad, mientras que en el período pico de la tarde de 16:00 a 19:00 horas se tomaron en zonas de trabajo, con el fin de priorizar situaciones donde se dan exclusivamente ascensos de pasajeros. Se completaron unos 260 datos en los que se registró la tipología vehicular, el nivel de ocupación del bus antes de abrir puertas, número de pasajeros ascendiendo y tiempo de parada.

El arreglo factorial para el diseño del tratamiento consistió en dos factores cualitativos, **factor A** "tipología vehicular" con tres niveles $a = 3$ (bus pequeño, bus mediano y bus grande) y **factor B** "ocupación del bus" antes de abrir puertas con dos niveles $b = 2$ (pocos o ningún pasajero de pie y muchos pasajeros de pie). Los niveles de cada factor fueron seleccionados de acuerdo a lo mostrado en la Figura 1.

DISEÑO DEL EXPERIMENTO

Se seleccionaron de manera aleatoria $n = 15$ réplicas por unidad experimental para un total de $a \times b \times n = 90$ observaciones a trabajar en el experimento. De cada una, se calculó el tiempo individual de ascenso por pasajero como el cociente entre el tiempo de parada y el número de usuarios que subieron al bus. Los datos a analizar se muestran en la Tabla 1. Los totales de los renglones y las columnas se indican en los márgenes de la tabla y a la derecha de cada celda los totales de cada una.

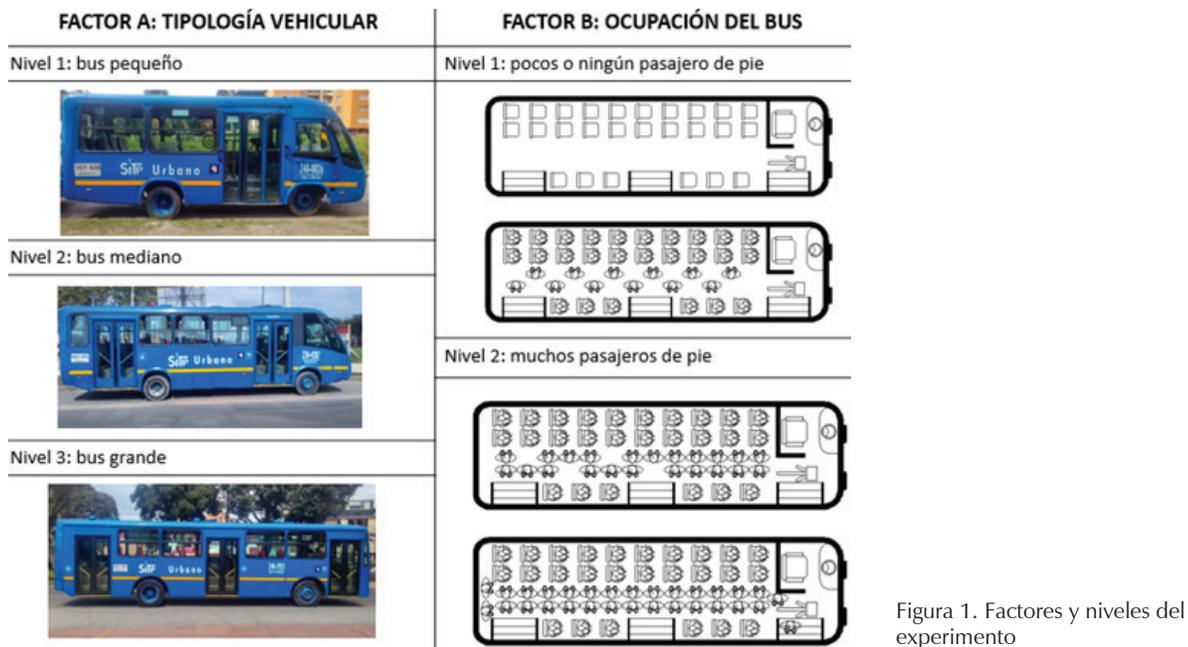


Figura 1. Factores y niveles del experimento

Tabla 1. Datos tiempo de ascenso recolectados por pasajero

Tipología vehicular	Ocupación del bus										y_i	
	Sin muchos pasajeros de pie					Con muchos pasajeros de pie						
Bus pequeño	5.66	4.74	4.93	5.19	4.15	8.41	8.06	5.13	9.21	5.98	97.6	180.2
	3.03	4.43	3.95	8.65	6.68	82.7	7.29	3.79	5.26	8.31		
	7.17	4.82	5.56	7.52	6.19	6.04	3.49	7.31	4.20	7.13		
Bus mediano	3.59	3.23	3.99	3.45	3.81	4.07	4.21	4.29	5.26	7.98	77.3	128.6
	3.77	3.96	3.53	2.44	4.13	51.4	5.30	3.54	5.94	4.65		
	2.75	4.00	3.10	3.46	2.19	5.52	3.73	5.24	4.95	8.52		
Bus grande	3.32	6.00	2.69	3.18	3.90	4.35	4.20	4.28	3.56	6.60	73.2	125.3
	4.28	3.77	2.06	3.91	2.49	52.2	4.90	3.76	7.74	3.47		
	2.73	2.70	3.19	4.11	3.85	6.58	6.49	4.39	4.88	4.50		
y_j	186					248.0					434 = y	

El análisis estadístico de los datos se realizó con base en el modelo de los efectos que se describe en (1) donde:

- μ = efecto del promedio global
- τ_i = efecto del nivel i -ésimo del factor A de los renglones
- β_j = efecto del nivel j -ésimo del factor B de las columnas
- $(\tau\beta)_{ij}$ = efecto de la interacción entre τ_i
- β_j y ε_{ijk} = componente del error aleatorio (Montgomery, 2013)

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk} \begin{cases} i = 1, 2, 3, \dots, a \\ j = 1, 2, 3, \dots, b \\ k = 1, 2, 3, \dots, n \end{cases} \quad (1)$$

El interés particular de este experimento conocidos los factores, es probar la hipótesis acerca de la igualdad de los efectos de los tratamientos de los renglones (tipología vehicular) descrita en (2), y probar la igualdad de los efectos de los tratamientos en las columnas (ocupación del bus) descrita en (3).

$$\begin{aligned} H_0 &: \tau_1 = \tau_2 = \dots = \tau_a = 0 \\ H_1 &: \text{al menos una } \tau_i \neq 0 \end{aligned} \quad (2)$$

$$\begin{aligned} H_0 &: \beta_1 = \beta_2 = \dots = \beta_b = 0 \\ H_1 &: \text{al menos una } \beta_j \neq 0 \end{aligned} \quad (3)$$

También existe interés en determinar si los tratamientos de los renglones y las columnas interactúan, por lo que se probará la hipótesis descrita en (4).

$$\begin{aligned}
 H_0 : (\tau\beta)_{ij} &= 0 \text{ para todas las } i, j \\
 H_1 : \text{al menos una } (\tau\beta)_{ij} &\neq 0
 \end{aligned}
 \tag{4}$$

Las hipótesis se prueban a partir de un análisis de varianza (ANOVA) de dos factores. En este procedimiento se calcula la suma de cuadrados para el total corregido según la ecuación (5), para los renglones o factor A ($SS_{Tipología}$) según la ecuación (6), para las columnas o factor B ($SS_{Ocupación}$) según la ecuación (7), para la interacción entre A y B ($SS_{Interacción}$) según la ecuación (8) y para el error (SS_E) en la ecuación (9).

$$SS_T = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n y_{ijk}^2 - \frac{y^2}{abn}
 \tag{5}$$

$$SS_{Tipología} = \frac{1}{bn} \sum_{i=1}^a y_i^2 - \frac{y^2}{abn}
 \tag{6}$$

$$SS_{Ocupación} = \frac{1}{an} \sum_{j=1}^b y_j^2 - \frac{y^2}{abn}
 \tag{7}$$

$$SS_{Interacción} = \frac{1}{n} \sum_{i=1}^a \sum_{j=1}^b y_{ij}^2 - \frac{y^2}{abn} - SS_{Tipología} - SS_{Ocupación}
 \tag{8}$$

$$SS_E = SS_T - SS_{Tipología} - SS_{Ocupación} - SS_{Interacción}
 \tag{9}$$

RESULTADOS

En la Tabla 2 se muestran los estadísticos descriptivos para cada unidad experimental definida. Una evaluación preliminar de las medias indica un aumento de los tiempos de ascenso por pasajero, a medida que disminuye el tamaño del bus y a medida que aumenta la ocupación del mismo. La desviación estándar parece ser más alta cuando aumenta el número de pasajeros al interior del bus.

En la Tabla 3 se muestra el ANOVA resultante de los datos originales recolectados en el estudio. No obstante, para confiar en el procedimiento es necesario que se satisfagan ciertos supuestos, específicamente, que el modelo sugerido en (1) describe de manera adecuada las observaciones, y que los errores siguen una distribución normal e independiente con media cero y varianza σ^2 constante pero desconocida (Montgomery, 2013).

Antes de avanzar en la explicación de los resultados del ANOVA, se realiza la verificación del modelo fundamental a través del análisis residual. Los residuales del modelo factorial de dos factores son función del promedio de las observaciones de la celda ij -ésima y se calculan a partir de (10).

$$e_{ijk} = y_{ijk} - \bar{y}_{ij}
 \tag{10}$$

En la Tabla 4 se muestran los residuales de datos de tiempos de ascenso por pasajero. Con el fin de determinar la existencia de datos atípicos, se examinan los residuales estandarizados a partir de (11) donde MS_E corresponde al cuadrado medio del error. Suponiendo que los errores e_{ij} siguen una distribución $N(0, \sigma^2)$ se espera que 100 % de los d_{ij} se encuentren entre ± 3 . Realizando el cálculo del residuo estandarizado para los

Tabla 2. Estadísticos descriptivos datos originales

Tipología	Ocupación	Media	Desviación estándar	N
Bus pequeño	Sin muchos pasajeros de pie	5.51078	1.50237	15
	Con muchos pasajeros de pie	6.50339	1.81537	15
	Total	6.00708	1.71330	30
Bus mediano	Sin muchos pasajeros de pie	3.42507	0.58771	15
	Con muchos pasajeros de pie	5.15112	1.44028	15
	Total	4.28810	1.39236	30
Bus grande	Sin muchos pasajeros de pie	3.47806	0.96253	15
	Con muchos pasajeros de pie	4.87708	1.33984	15
	Total	4.17757	1.34911	30

Tabla 3. Análisis de varianza datos originales

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_{α}	Valor P
Tipos de tipología	63.142	2	31.571	17.713	0.000
Tipos de ocupación	42.388	1	42.388	23.782	0.000
Interacción	2.025	2	1.013	0.568	0.569
Error	149.717	84	1.782		
Total	257.273	89			

Tabla 4. Residuales datos originales

Tipología vehicular	Ocupación del bus									
	Sin muchos pasajeros de pie					Con muchos pasajeros de pie				
Bus pequeño	0.15	-0.77	-0.58	-0.32	-1.36	1.91	1.56	-1.38	2.71	-0.52
	-2.48	-1.09	-1.56	3.14	1.16	0.79	-2.71	-1.24	1.81	1.45
	1.66	-0.69	0.04	2.01	0.68	-0.47	-3.02	0.81	-2.31	0.63
Bus mediano	0.17	-0.20	0.56	0.03	0.39	-1.09	-0.95	-0.86	0.11	2.83
	0.35	0.54	0.11	-0.99	0.70	0.14	-1.61	0.79	-0.50	-1.07
	-0.67	0.58	-0.33	0.03	-1.24	0.37	-1.42	0.09	-0.20	3.37
Bus grande	-0.16	2.52	-0.79	-0.30	0.42	-0.53	-0.68	-0.60	-1.32	1.72
	0.80	0.29	-1.42	0.43	-0.99	0.02	-1.12	2.87	-1.41	-1.42
	-0.75	-0.78	-0.29	0.63	0.37	1.70	1.62	-0.48	0.00	-0.38

valores obtenidos, se obtiene que ninguno de ellos es superior a 3 desviaciones, por lo que no se observa ningún dato inusual. El residual más grande es 3.37 (correspondiente a un registro de 8.52 s por pasajero en bus mediano con muchos pasajeros de pie) y su residual estandarizado es $3.37/\sqrt{1.782} = 2.523$.

$$d_{ij} = \frac{e_{ij}}{\sqrt{MS_E}} = \frac{e_{ij}}{\sqrt{1.782}} \quad (11)$$

En contraste con el análisis de los residuales estandarizados, la gráfica de probabilidad normal de los residuales mostrada en la Figura 2, revela que los datos no siguen una distribución normal, en primer lugar porque todos los puntos no están cubiertos por la banda formada en la gráfica, y en segundo, el valor de p está por debajo de 0.05.

En la Figura 3 se muestra la gráfica de los residuales contra los valores ajustados \hat{y}_{ij} , en donde se aprecia una tendencia de la varianza de los residuales a incrementarse cuando el tiempo de ascenso por pasajero aumenta. Los residuales se encuentran con una estructura obvia y una varianza no constante, que crece en forma de embudo con la boca hacia afuera, lo cual confirma que los datos no siguen una distribución normal.

Dado que no se cumplen los supuestos del modelo, el enfoque más recomendado para abordar el problema cuando la varianza no es constante es aplicar una transformación de datos para estabilizar la varianza (Kuehl, 2001; Montgomery, 2013).

RESULTADOS DE LOS DATOS TRANSFORMADOS

Para encontrar la transformación de datos más adecuada, cuando se tiene un diseño factorial se debe seleccionar una que minimice el cuadrado medio de las interacciones. Se puede determinar la familia de potencias de las transformaciones de los datos $y^* = y^\lambda$ donde λ es el parámetro de la transformación que habrá de determinarse. Box y Cox (1964) desarrollaron un método útil que consiste en efectuar el análisis de varianza y^λ de mostrado en (12) para varios valores de λ .

$$y^\lambda = \begin{cases} \frac{y^\lambda - 1}{\lambda y^{\lambda-1}} & \lambda \neq 0 \\ \ln y & \lambda = 0 \end{cases} \quad (12)$$

Donde $\bar{y} = \ln^{-1}[(1/n) \sum \ln y]$ es la media geométrica de las observaciones y la estimación de máxima verosimilitud de λ es el valor para el cual la suma de cuadrados del error $SS_E(\lambda)$ es un mínimo que puede verificarse construyendo la gráfica de λ en función de $SS_E(\lambda)$ como se muestra en la Figura 4.

El valor recomendado para la transformación de los datos, obtenido a partir del método de Box-Cox es $\lambda = -0.05$ (el recíproco de la raíz cuadrada). Con este valor se muestran los datos de la Tabla 5, que suponen una varianza más homogénea. Los totales de los renglones y las columnas se indican en los márgenes de la tabla y a la derecha de cada celda los totales de cada una.

En la Tabla 6 se muestra el ANOVA resultante de los datos transformados del estudio en cuestión. Dado que

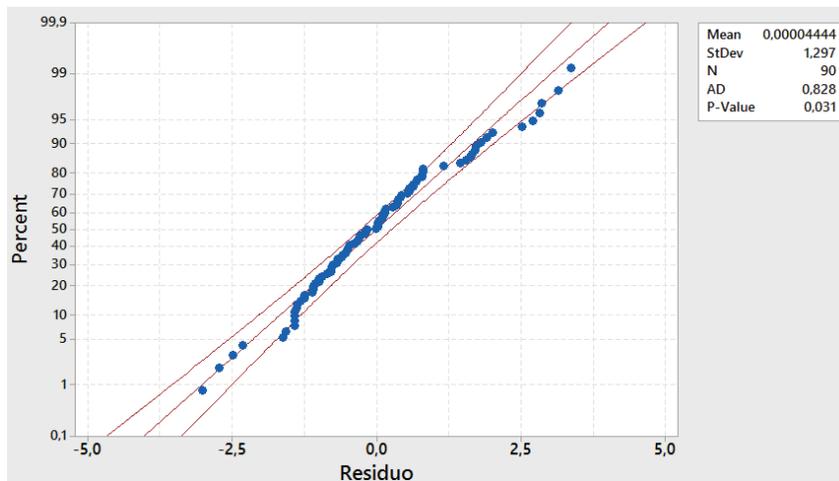


Figura 2. Gráfica de probabilidad normal residuales datos originales

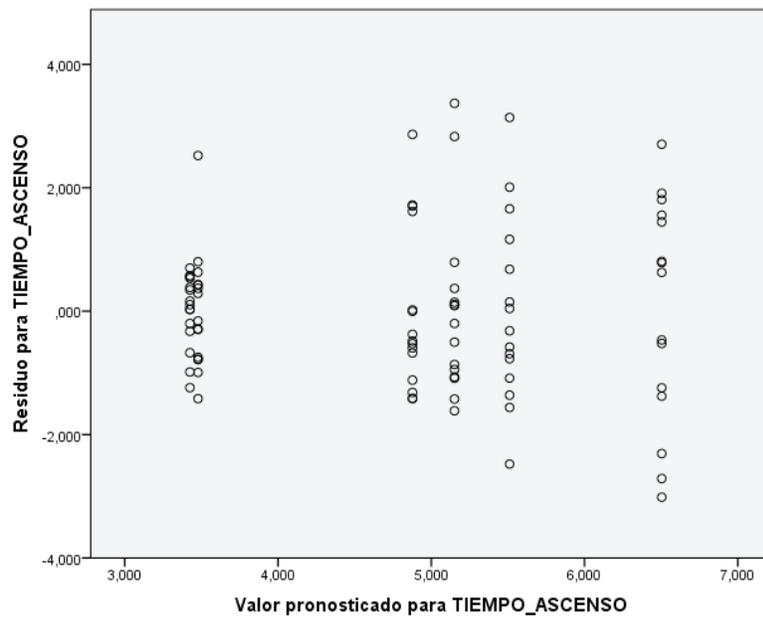


Figura 3. Gráfica de los residuales contra \hat{y}_{ij}

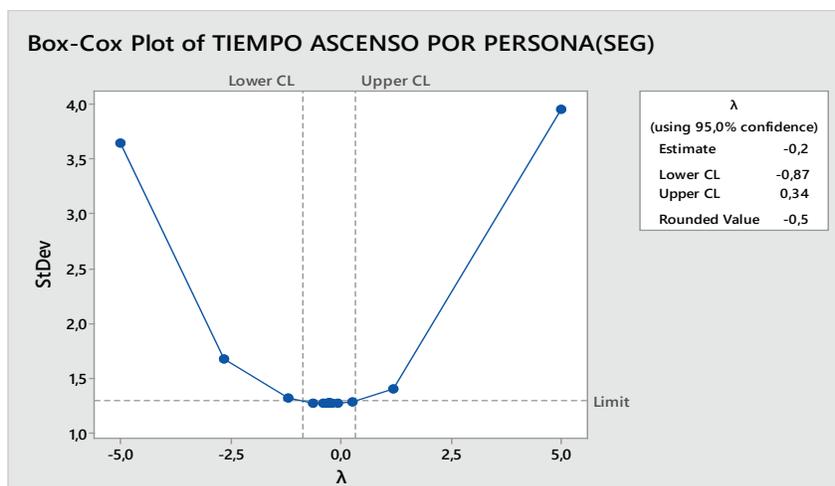


Figura 4. Gráfica Box-Cox de λ contra $SS_e(\lambda)$

$F_{0.05, 2,84} = 3.12 > 2.351$ se concluye que no existe interacción significativa entre la tipología y la ocupación, es decir, se rechaza la hipótesis alternativa planteada en (4). Por su parte, como $F_{0.05, 2,84} = 3.12 < 17.995$ y $F_{0.05, 1,84} = 3.97 < 30.755$ se concluye que los efectos principales de la tipología y la ocupación en sus valores transformados son significativos, esto es, se rechaza la hipótesis nula planteada en (2) y (3).

Para entender mejor que los efectos de cada factor transformado son significativos, en la Figura 5 se muestran las respuestas promedio para cada combinación de los tratamientos. Aunque la interacción no es significativa,

y por tanto las rectas no se cruzan entre sí, se observa que para ocupaciones bajas y altas del bus, el valor de la respuesta transformada es más baja para el bus pequeño y que los buses más grandes tienen un valor de respuesta transformado más alto.

Como el ANOVA indicó que las medias de los renglones y las filas difieren entre sí, se procede a realizar la comparación entre las medias individuales de cada una para identificar las diferencias significativas. Dentro de los métodos de comparaciones múltiples se usará la prueba de Tukey (Montgomery, 2013).

Tabla 5. Datos transformados tiempos de ascenso por pasajero

Tipología vehicular	Ocupación del bus											y_i	
	Sin muchos pasajeros de pie					Sin muchos pasajeros de pie					y_i		
	0.42	0.46	0.45	0.44	0.49	0.34	0.35	0.44	0.33	0.41			
Bus pequeño	0.57	0.48	0.50	0.34	0.39	6.6	0.37	0.51	0.44	0.35	0.35	6.1	12.6
	0.37	0.46	0.42	0.36	0.40		0.41	0.54	0.37	0.49	0.37		
	0.53	0.56	0.50	0.54	0.51		0.50	0.49	0.48	0.44	0.35		
Bus mediano	0.52	0.50	0.53	0.64	0.49	8.2	0.43	0.53	0.41	0.46	0.49	6.8	15.0
	0.60	0.50	0.57	0.54	0.68		0.43	0.52	0.44	0.45	0.34		
	0.55	0.41	0.61	0.56	0.51		0.48	0.49	0.48	0.53	0.39		
Bus grande	0.48	0.52	0.70	0.51	0.63	8.2	0.45	0.52	0.36	0.54	0.54	7.0	15.2
	0.61	0.61	0.56	0.49	0.51		0.39	0.39	0.48	0.45	0.47		
y_j						23						19.8	42.8 = y

Tabla 6. Análisis de varianza datos transformados

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Tipos de tipología	0.134	2	0.067	17.995	0.000
Tipos de ocupación	0.115	1	0.115	30.755	0.000
Interacción	0.018	2	0.009	2.351	0.102
Error	0.314	84	0.004		
Total	0.581	89			

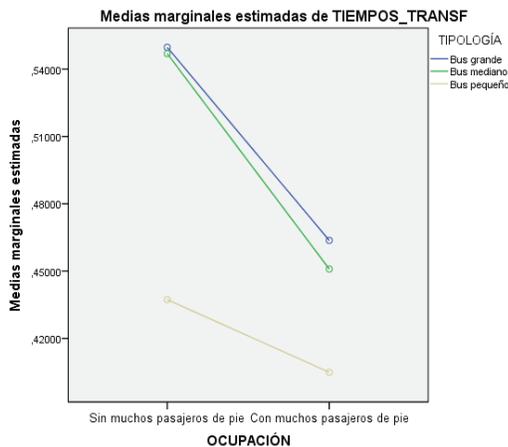


Figura 5. Gráfica de tipología-ocupación datos transformados

Para tamaños de muestras iguales como es este caso, la prueba de Tukey declara que dos medias son significativamente diferentes si el valor absoluto de sus diferencias muestrales excede el cálculo mostrado en (13). El valor de $q_{0.05}(3,84) = 3.384$ para los tipos de tipología se obtiene por interpolación de la tabla de puntos porcentuales del rango studentizado.

$$T_{0.05} = q_{0.05}(a, f) \sqrt{\frac{MS_E}{n}} = q_{0.05}(3,84) \sqrt{\frac{MS_E}{n}} = (3,384) \sqrt{\frac{0,0037}{15}} = 0,053 \quad (13)$$

Siendo, f número de grados de libertad asociados al MS_E .

Las comparaciones por pares dan como resultado los valores mostrados en la Tabla 7. Aquellos que son superiores a $T_{0.05}$ son los que presentan diferencias significativas. Para bajos niveles de ocupación, el tiempo de ascenso por persona (transformado) no es significativamente diferente para bus grande y mediano, mientras que es significativamente inferior cuando se realiza en un bus pequeño en comparación con el bus grande y mediano. Para niveles de ocupación altos, es solo significativamente inferior cuando ocurre en bus pequeño comparado con bus grande.

Donde $T_{0.05} = 0.044$ para los tipos de ocupación se obtienen las comparaciones por pares mostradas en la Tabla 8. Para este caso, la variable respuesta difiere significativamente en función de los niveles de ocupación para el bus mediano y bus grande, y no así para el bus pequeño.

Ahora bien, se calcula la proporción de variabilidad en los datos explicada por el modelo de varianza, a través del parámetro "R cuadrada" que se define en (14).

$$R^2 = \frac{SS_{MODELO}}{SS_{TOTAL}} = \frac{SS_{TIPOLOGÍA} + SS_{OCUPACIÓN} + SS_{INTERACCIÓN}}{SS_{TOTAL}} \quad (14)$$

$$R^2 = \frac{0.134 + 0.115 + 0.018}{0.314} = 0.4596$$

El resultado de R^2 indica que cerca de 46 % de la variabilidad de los tiempos transformados de ascenso por persona se explica por la tipología, la ocupación y su interacción.

Como se mencionó anteriormente, es necesaria la verificación del modelo elaborado a través del análisis residual. En la Tabla 9 se muestran los residuales de los datos transformados. El residual más grande es 0.147 y su residual estandarizado es $0.147/\sqrt{0.004} = 2.405$ lo que no representa un valor atípico potencial.

La gráfica de probabilidad normal mostrada en la Figura 6, para esta ocasión con los residuales obtenidos a partir de los datos transformados, revela ahora un único punto por fuera de la banda formada en la gráfica (antes eran seis), y el valor de p es superior a 0.05 lo que indica una distribución normal de los datos.

En la Figura 7 se muestra la gráfica de los residuales contra los valores ajustados de los datos transformados, que en esta oportunidad no presentan una estructura obvia, con una varianza más constante sin la forma de embudo que tenía antes de la transformación.

La estimación de los parámetros del modelo de los efectos para el diseño factorial mostrado en (1) puede ser resuelto según Khuel (2001) como $\hat{\mu} = \bar{y}$ (media global como el gran promedio), $\hat{\tau}_i = \bar{y}_i - \bar{y}$ (efectos de los tratamientos de los renglones como el promedio del renglón menos el gran promedio), $\hat{\beta}_j = \bar{y}_j - \bar{y}$ (efectos de los tratamientos de las columnas como el promedio de la columna menos el gran promedio), $\hat{\tau\beta}_{ij} = \bar{y}_{ij} - \bar{y}_i - \bar{y}_j + \bar{y}$ (interacción ij -ésima como el promedio de la celda ij -ésima más el gran promedio, menos el efecto del renglón i -ésimo y menos el efecto de la columna j -ésima).

Tabla 7. Comparaciones por pares tipología-datos transformados

Comparación	Sin muchos pasajeros de pie	Con muchos pasajeros de pie
Bus grande vs bus pequeño	$ -0.11 > T_{0.05} = 0.053$	$ -0.06 > T_{0.05} = 0.053$
Bus grande vs bus mediano	$ 0.00 < T_{0.05} = 0.053$	$ 0.01 < T_{0.05} = 0.053$
Bus mediano vs bus pequeño	$ -0.11 > T_{0.05} = 0.053$	$ -0.046 < T_{0.05} = 0.053$

Tabla 8. Comparaciones por pares ocupación-datos transformados

Comparación	Bus pequeño	Bus mediano	Bus grande
Ocupación baja vs ocupación alta	$ 0.03 < T_{0.05} = 0.044$	$ 0.10 > T_{0.05} = 0.044$	$ 0.09 > T_{0.05} = 0.044$

Tabla 9. Residuales datos transformados

Tipología vehicular	Ocupación del bus									
	Sin muchos pasajeros de pie					Con muchos pasajeros de pie				
Bus pequeño	-0.017	0.022	0.013	0.002	0.054	-0.06	-0.053	0.037	-0.075	0.004
	0.137	0.038	0.066	-0.097	-0.05	-0.035	0.109	0.031	-0.058	-0.05
	-0.064	0.018	-0.013	-0.073	-0.035	0.002	0.13	-0.035	0.083	-0.03
Bus mediano	-0.019	0.01	-0.046	-0.009	-0.035	0.045	0.037	0.032	-0.015	-0.097
	-0.032	-0.044	-0.015	0.093	-0.055	-0.016	0.081	-0.041	0.013	0.044
	0.056	-0.047	0.021	-0.009	0.13	-0.025	0.067	-0.014	-0.002	-0.108
Bus grande	-0.001	-0.142	0.06	0.011	-0.043	0.016	0.024	0.02	0.066	-0.074
	-0.066	-0.034	0.147	-0.044	0.085	-0.012	0.052	-0.104	0.073	0.074
	0.055	0.059	0.01	-0.056	-0.04	-0.074	-0.071	0.013	-0.011	0.008

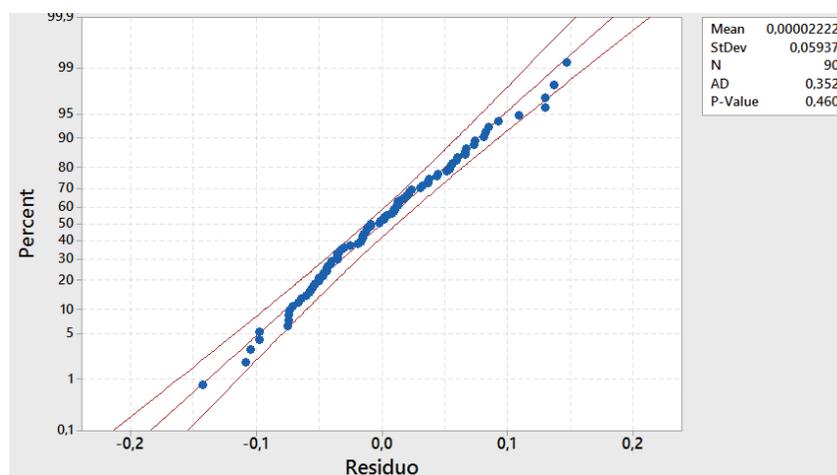


Figura 6. Gráfica de probabilidad normal residuales datos transformados

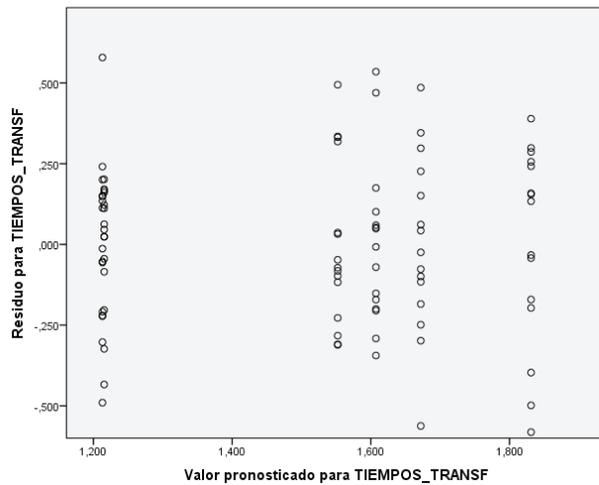


Figura 7. Gráfica residuales datos transformados contra \hat{y}_{ij}

De esta manera, la ecuación de regresión del diseño factorial de datos transformados con $\lambda = -0.5$ queda resuelta en (15).

$$(Tiempos\ de\ ascenso\ por\ pasajero)^{-0.5} = 0.47558 - 0.05448\tau_1 + 0.02335\tau_2 + 0.03113\tau_3 + 0.03573\beta_1 - 0.03573\beta_2 - 0.01955\tau\beta_{11} + 0.01955\tau\beta_{12} + 0.01226\tau\beta_{21} - 0.01226\tau\beta_{22} + 0.00728\tau\beta_{31} - 0.00728\tau\beta_{32} \tag{15}$$

Los estadísticos descriptivos se muestran en la Tabla 10, nótese que la media de cada nivel de factor es igual a la suma de los efectos correspondientes mostrados en la ecuación (15). Para este caso, los tiempos transformados tienden a ser más altos cuando aumenta el tamaño del bus y cuando disminuye la ocupación.

Tabla 10. Estadísticos descriptivos datos transformados

Tipología	Ocupación	Media	Desviación estándar	N
Bus pequeño	Sin muchos pasajeros de pie	0.43729	0.06061	15
	Con muchos pasajeros de pie	0.40492	0.06517	15
	Total	0.42110	0.06399	30
Bus mediano	Sin muchos pasajeros de pie	0.54693	0.05437	15
	Con muchos pasajeros de pie	0.45094	0.05453	15
	Total	0.49893	0.07243	30
Bus grande	Sin muchos pasajeros de pie	0.54973	0.07238	15
	Con muchos pasajeros de pie	0.46370	0.05772	15
	Total	0.50672	0.07779	30

Por último, se procede a comprobar si las 15 réplicas fueron suficientes para rechazar la hipótesis nula con una potencia de prueba de 95 % = $1 - \beta$, a un nivel de significancia $\alpha = 0.05$ y considerando la desviación estándar de los datos transformados de 0.08079. Calculando el parámetro Φ para el factor A según la ecuación (16) y suponiendo los valores de las réplicas n , es posible hacer uso de las curvas de operación característica (Montgomery, 2013) de donde se obtienen los valores de riesgo β mostrados en la Tabla 11. Nótese que con $n = 12$ réplicas se hubiera proporcionado la potencia deseada.

$$\Phi^2 = \frac{bn \sum_{i=1}^a \epsilon_i^2}{a\sigma^2} \tag{16}$$

$$= \frac{2n \sum [(-0.05448)^2 + (0.02335)^2 + (0.03113)^2]}{3 \times 0.08079^2} = 0.4579n$$

Tabla 11. Valores de riesgo β para valores de potencia

n	Φ	v_1	v_2	β
3	0.95698	2	84	0.620
4	1.17205	2	84	0.420
10	1.35337	2	84	0.081
11	1.51311	2	84	0.060
12	1.91395	2	84	0.040
13	2.13986	2	84	0.025
14	2.24431	2	84	0.018
15	2.34410	2	84	0.015

CONCLUSIONES

El diseño factorial elaborado a partir del tiempo de detención de los buses en distintos paraderos de la ciudad de Bogotá, permitió validar que existen variables asociadas a la configuración y ocupación del bus que inciden en los tiempos de ascenso por pasajero, coincidiendo con los autores descritos en el estado del arte.

Se encontró que los datos de tiempos de ascenso por pasajero recolectados no siguen una distribución normal, y en concordancia con lo descrito inicialmente, involucrar aspectos como la ocupación del bus puede hacer más complejo el análisis de la operación del transporte público. El mejor ajuste o transformación de los datos encontrado a través del método Box-Cox fue utilizar el recíproco de la raíz cuadrada con $\lambda = -0.05$. La transformación efectuada permitió estabilizar la varianza de los datos, una distribución normal de la variable respuesta y mejoró en definitiva el ajuste del modelo a los datos.

Sobre los datos transformados se validaron hipótesis como que las diferentes tipologías vehiculares que operan en la ciudad de Bogotá tienen efectos en el tiempo transformado de ascenso por pasajero, y que la ocupación del bus tiene efecto sobre la misma variable de respuesta. Específicamente se encontró, que para bajos niveles de ocupación el tiempo transformado es significativamente inferior en buses pequeños en comparación con los buses grandes y medianos, y, que para altos niveles de ocupación, el tiempo transformado es significativamente inferior para buses pequeños en comparación con buses grandes. También, que para buses grandes y medianos, la variable respuesta es significativamente diferente teniendo en cuenta los niveles de ocupación.

El ANOVA resultante de los datos transformados rechazó la hipótesis que suponía interacción entre ambos factores, es decir, que el efecto de uno de los tratamientos no cambia significativamente cuando existe un cambio de nivel en el otro tratamiento. El modelo entonces corresponde a un diseño factorial sin interacción, el cual produjo una explicación de 46 % de variabilidad de los tiempos transformados de ascenso por pasajero. Es posible que este valor de R^2 aumente si en futuras investigaciones se considera el efecto que tiene el número de personas totales que ascienden por parada, así como variables asociadas a la edad del usuario.

Por último, se resalta que el estudio de los tiempos de parada se convierte en un insumo de importancia para la operación del transporte público, que puede mejorar sustancialmente los itinerarios de programación de buses y el entendimiento de la regularidad y

confiabilidad de los servicios. Este aspecto debería analizarse por los concesionarios y los entes de regulación como parte de la búsqueda de mejoras del sistema.

REFERENCIAS

- Abkowitz, M.D. & Engelstein, I. (1982). Factors affecting running time on transit routes. *Transportation Research, Part A: General*, XVII(2), 107-113.
- Arhin, S. et al. (2016). Optimization of transit total bus stop time models. *Journal of traffic and transportation engineering*, 3(2), 146-153. <https://doi.org/10.1016/j.jtte.2015.07.001>
- Bertini, R.L. & El-Geneidy, A.M. (2004). Modeling transit trip time using archived bus dispatch. *Journal of Transportation Engineering*, 130(1), 56-67. [https://doi.org/10.1061/\(ASCE\)0733-947X\(2004\)130:1\(56\)](https://doi.org/10.1061/(ASCE)0733-947X(2004)130:1(56))
- Bian, B., Zhu, N., Ling, S. & Ma, S. (2015). Bus service time estimation model for a curbside bus stop. *Transportation Research, Part C*, 103-121. <http://dx.doi.org/10.1016%2Fj.trc.2015.06.011>
- Box, G. & Cox, D. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 26(2), 211-252.
- Daamen, W., Lee, Y.-c. & Wiggeraad, P. (2008). Boarding and alighting experiments. *transportation research record. Journal of the Transportation Research Board*, (2042), 71-81. <https://doi.org/10.3141%2F2042-08>
- Dueker, K.J., Kimpel, T.J. & Strathman, J.G. (2004). Determinants of bus dwell time. *Journal of Public Transportation*, 7(1), 21-40. <http://doi.org/10.5038/2375-0901.7.1.2>
- El-Geneidy, A.M., Horning, J. & Krizek, K.J. (2011). Analyzing transit service reliability using detailed data from automatic vehicular locator systems. *Journal of Advanced Transportation*, XVI(1), 66-79. <https://doi.org/10.1002/atr.134>
- El-Geneidy, A.M. & Vijayakumar, N. (2011). The effects of articulated buses on dwell and running times. *Journal of Public Transportation*, XIV(3), 63-86. <http://doi.org/10.5038/2375-0901.14.3.4>
- Fernández, R., Valencia, A. & Seriani, S. (2015). On passenger saturation flow in public transport doors. *Transportation Research, Part A*, 78, 102-112.
- Fletcher, G. & El-Geneidy, A. (2012). The effects of fare payment and crowding on dwell time: A fine grained analysis. *Transportation Research Record. Journal of the Transportation Research Board*, 124-132. <https://doi.org/10.3141%2F2351-14>
- Gutiérrez-Pulido, H. & de la Vara Salazar, R. (2012). *Análisis y diseño de experimentos*. Tercera ed. México: McGRAW-HILL/interramericana editores, S.A., de C.V.
- Jara-Díaz, S. & Tirachini, A. (2013). Urban bus transport: open all doors for boarding. *Journal of Transportation Economics and Policy*, 47(1), 91-106.
- Jiang, Z., Xie, C., Ji, T. & Zou, X. (2015). Dwell time modelling and optimized simulations for crowded rail transit lines based on

- train capacity. *PROMET-Traffic&Transportation*, XXVII(2), 125-135.
- Katz, D. & Garrow, L.A. (2012). The impact of bus door crowding on operations and safety. *Journal of Public Transportation*, XV(2), 71-93. <https://www.researchgate.net/deref/http%3A%2F%2Fdx.doi.org%2F10.5038%2F2375-0901.15.2.4>
- Kuehl, R.O. (2001). *Diseño de experimentos*. Segunda ed. México: International Thomson Editores, S.A. de C.V.
- Li, F., Duan, Z. & Yang, D. (2012). Dwell time estimation models for bus rapid transit stations. *Journal of Modern Transportation*, XX(3), 168-177.
- Meng, Q. & Qu, X. (2013). Bus dwell time estimation at bus bay: a probabilistic approach. *Transportation Research, Part C*, 61-71. <http://dx.doi.org/10.1016/j.trc.2013.08.007>
- Montgomery, D. (2013). *Design and analysis of experiments*. Octava ed. New Jersey: John Wiley & Sons, Inc.
- Peña, C. & Moreno, E. (2014). Delay at bus stops of Transmilenio transport system according to parameters measured "in situ". Case study Bogotá-Colombia. *Procedia Social and behavioral sciences*, 121-129. <https://doi.org/10.1016/j.sbspro.2014.12.123>
- Strathman, J.G. & Hopper, J.R. (1992). *Empirical analysis of bus transit on-time performance*, Portland OR: s.n.
- Sun, L. *et al.* (2014). Models of bus boarding and alighting dynamics. *Transportation Research, Part A*, Volumen 69, 447-460. <http://dx.doi.org/10.1016/j.tra.2014.09.007>
- Tirachini, A. (2011). Estimation of travel time and the benefits of upgrading the fare payment technology in urban bus services. *Transportation Research, Part C*, Volumen XXX, 239-256.
- Tirachini, A., Camus, L. & Cruz, D. (2015). Modelación de tiempos de detención de buses en paraderos de Santiago y Valparaíso-Viña del Mar. *Ingeniería de Transporte*, 19(2), 103-116.
- Transit Cooperative Research Program, Transportation Research Board. (2013). *Transit Capacity and Quality of Service Manual*. Tercera ed. Washington, DC: National Academy of Sciences.
- Valencia, A. & Fernandez, R. (2012). Influencia del espaciamiento de paraderos en la velocidad comercial en corredores de transporte público de superficie. *Ingeniería de transporte*, 16(1), 3-8.
- Wang, C. *et al.* (2016). Modeling bus dwell time and time lost serving stop in China. *Journal of Public Transportation*, XIX(3), 55-77. <http://doi.org/10.5038/2375-0901.19.3.4>
- Zhang, C. & Teng, J. (2013). *Bus dwell time estimation and prediction: A study case in Shanghai-China*. Shenzhen, s.n.