



## PROYECTO DE GRADO

Presentado ante la ilustre UNIVERSIDAD DE LOS ANDES como requisito final para  
obtener el Título de INGENIERO DE SISTEMAS

# MODELO DE DEMANDA DE LA RED DE TRANSPORTE CENTRO - LA HECHICERA PARA LA LÍNEA 2 DEL TROLEBÚS MÉRIDA

Por

Br. Ulises Silva Moreno

Tutor: Prof. Magdiel Ablan Bortone

Tutor: Prof. Douglas Eduing Rivas Olivo

Junio 2017

©2017 Universidad de Los Andes Mérida, Venezuela

Atribución - No Comercial - Compartir igual 3.0 Venezuela  
(CC BY - NC - SA 3.0 VE)

# Modelo de Demanda de la Red de Transporte Centro - La Hechicera para la Línea 2 del Trolebús Mérida

Br. Ulises Silva Moreno

Proyecto de Grado — Investigación de Operaciones, 100 páginas

**Resumen:** El trabajo realizado presenta un modelo de demanda de transporte de la zona Centro - La Hechicera en la ciudad de Mérida, Venezuela, donde se representaron las estaciones de transporte público por medio de un grafo apoyado en sistemas de información geográfica; utilizando análisis de regresión lineal y optimización de la entropía se estimó el modelo de generación y distribución de viajes. Se recolectaron las variables urbanas que caracterizan a cada nodo de la red con el objetivo de explicar la dinámica de viajes en función de las características de uso de suelos de los mismos; adicionalmente, se desarrolló una aplicación web para generar diferentes matrices origen-destino que permitieron evaluar el patrón de viajes de la red contribuyendo al estudio de escenarios de la nueva ruta del Trolebus Mérida: Línea 2 (Centro - La Hechicera)

**Palabras clave:** modelo, demanda, transporte, cuatro etapas, Trolebús Mérida

# Índice

Índice de Tablas	vi
Índice de Figuras	viii
Agradecimientos	x
<b>1 Introducción</b>	<b>1</b>
1.1 Antecedentes	1
1.2 Justificación	2
1.3 Planteamiento y delimitación del problema	3
1.4 Objetivos	3
1.4.1 Objetivo general	3
1.4.2 Objetivos específicos	3
1.5 Metodología de trabajo	4
1.6 Marco teórico y conceptual	5
1.6.1 Demanda de transporte	5
1.6.2 Planificación urbana	5
1.6.3 Planificación de transporte	6
1.6.4 Modelo de transporte	6
1.6.5 Estructura de un modelo de transporte	6
1.6.6 Modelo de cuatro etapas	8
1.6.7 Modelo de regresión lineal	8
1.6.8 Modelo de maximización de la entropía	15
1.6.9 Sistemas de Información Geográfica	17

1.7	Estructura del documento . . . . .	18
<b>2</b>	<b>Descripción y Análisis de la Red de Transporte</b>	<b>19</b>
2.1	Sistema de Transporte Masivo Trolebús Mérida . . . . .	19
2.1.1	Gestión de la Línea 1: Ejido - Centro . . . . .	21
2.1.2	Nodos de transporte . . . . .	21
2.1.3	Costo del transporte . . . . .	23
2.1.4	Unidades de transporte . . . . .	24
2.1.5	Planificación de la Línea 2: Centro - La Hechicera . . . . .	24
2.2	Modelo de demanda de transporte Centro - La Hechicera . . . . .	28
2.3	Zonificación de la red de transporte . . . . .	34
2.4	Modelo de datos . . . . .	38
<b>3</b>	<b>Modelo de Generación</b>	<b>40</b>
3.1	Sistemas de información de transporte . . . . .	40
3.2	Formulación del modelo de origen y atracción de viajes . . . . .	41
3.3	Recolección de la información . . . . .	42
3.4	Metodología del análisis geográfico . . . . .	45
3.5	Ajustes de los modelos de regresión lineal múltiple . . . . .	47
3.5.1	Análisis de datos . . . . .	47
3.5.2	Pruebas de hipótesis . . . . .	51
<b>4</b>	<b>Modelo de Distribución</b>	<b>59</b>
4.1	Balance entre orígenes y atracciones . . . . .	59
4.2	Formulación del modelo de optimización . . . . .	60
4.3	Algoritmo de estimación de la matriz O-D . . . . .	65
<b>5</b>	<b>Estudio de Escenarios</b>	<b>69</b>
5.1	Escenario 1 . . . . .	71
5.2	Escenario 2 . . . . .	74
5.3	Escenario 3 . . . . .	76



<b>6 Conclusiones</b>	<b>79</b>
6.1 Recomendaciones . . . . .	80
<b>Bibliografía</b>	<b>82</b>
<b>A Bibliotecas de R</b>	<b>87</b>
A.1 Paquete: shiny v.1.0.3 . . . . .	87
A.2 Paquete: igraph v.1.0.1 . . . . .	88
A.3 Paquete: lmttest v.0.9-35 . . . . .	88
<b>B Código Fuente: Matriz O-D</b>	<b>89</b>
B.1 Archivos de código fuente . . . . .	89
<b>C Estadísticas y Datos para Simulación</b>	<b>93</b>
<b>D Pruebas Estadísticas en la Validación de Modelos</b>	<b>98</b>
D.1 Shapiro - Wilks . . . . .	98
D.2 Durbin - Watson . . . . .	99
D.3 Criterio de información de Akaike . . . . .	99
D.4 Breusch - Pagan . . . . .	100

# Índice de Tablas

1.1	Tabla de parametros del modelo de transporte (Lieberman y Hillier, 2010)	7
1.2	Estructura de Datos en Regresión Lineal Multiple . . . . .	11
2.1	Tecnologías de transporte STMTM (Tromerca, 2017) . . . . .	24
2.2	Características de los nodos de la Línea 2 Centro - La Hechicera . . . .	26
2.3	Distancia entre nodos de la Línea 2 . . . . .	26
2.4	Los beneficios de las iniciativas de proyectos de transporte público . . .	28
2.5	Usos urbanos (Hernández y Molina, 2011) . . . . .	33
2.6	Identificación de los nodos y sus características geográficas y urbanas. Sistema de referenciación: WGS 84/UTM zona 18S . . . . .	37
3.1	Total de viajes diarios originados y atraídos por nodo de transporte . .	43
3.2	Datos de matricula estudiantil (ULA, 2016) y camas de hospitalización (INE, 2011) . . . . .	45
3.3	Matriz de correlación <i>pearson</i> ( $x_i, O_i$ ) . . . . .	48
3.4	Matriz de correlación <i>pearson</i> ( $x_i, D_j$ ) . . . . .	48
3.5	Estadístico $t_0$ vs <i>valor-p</i> de cada coeficiente (modelo de orígenes) . . . .	54
3.6	Estadístico $t_0$ vs <i>valor-p</i> de cada coeficiente (modelo de atracciones) . .	54
3.7	Valores observados vs predichos y residuos $\varepsilon_i$ para cada modelo $O_i$ y $D_j$	55
4.1	Estructura de matriz de viajes O-D (Wright et al., 2010) . . . . .	61
4.2	Matriz de viajes con máxima entropía de la red de transporte Centro - La Hechicera . . . . .	68
5.1	Flujo de viajes entre cada arco de la red Centro - La Hechicera. Hora: 7:00 AM - 8:30 AM . . . . .	72

5.2	Intervalos de confianza para la media de los viajes. Hora: 7:00 AM - 8:30 AM . . . . .	73
5.3	Flujo de viajes entre cada arco de la red Centro - La Hechicera. Hora: 12:00M - 2:00 PM . . . . .	75
5.4	Intervalos de confianza para la media de los viajes. Hora: 12:00 M - 2:00 PM . . . . .	75
5.5	Flujo de viajes entre cada arco de la red Centro - La Hechicera. Hora: 4:00PM - 6:00 PM . . . . .	77
5.6	Intervalos de confianza para la media de los viajes. Hora: 4:00PM - 6:00 PM . . . . .	78
C.1	Ajustes de distribuciones sentido Centro - La Hechicera. Periodo mañana: 7:00 AM - 12:00 M . . . . .	94
C.2	Ajustes de distribuciones sentido La Hechicera - Centro. Periodo mañana: 7:00 AM - 12:00 M . . . . .	94
C.3	Ajustes de distribuciones sentido Centro - La Hechicera. Periodo tarde: 12:00 M - 06:00 PM . . . . .	95
C.4	Ajustes de distribuciones sentido La Hechicera - Centro. Periodo tarde: 12:00 M - 06:00 PM . . . . .	95
C.5	Principales estadísticas de los viajes observados . . . . .	96

# Índice de Figuras

1.1	Representación del modelo de transporte en nodos y arcos . . . . .	7
1.2	El clásico modelo de cuatro etapas para la planeación del transporte (de Dios Ortúzar y Willumsen, 2008) . . . . .	9
2.1	Red de rutas y estaciones del STMTM (Tromerca, 2017) . . . . .	20
2.2	Funcionamiento de la Línea 1 (Tromerca, 2017) . . . . .	21
2.3	Estaciones Línea 1 (Tromerca, 2017) . . . . .	22
2.4	Infraestructura y acceso a nivel de estación (Tromerca, 2017) . . . . .	23
2.5	Tipos de tarjeta magnética inteligente (Tromerca, 2017) . . . . .	23
2.6	Red y nodos de transporte de la Línea 2 Centro - La Hechicera (Tromerca, 2017) . . . . .	25
2.7	Características de la red vial de un sistema de transporte masivo Trolebús (Wright et al., 2010) . . . . .	27
2.8	Propuesta de modelo de demanda de transporte para la red Centro - La Hechicera . . . . .	29
2.9	Situación relativa del área de estudio Centro - La Hechicera . . . . .	30
2.10	Plano de Uso Urbano de la ciudad de Mérida 2010 (Hernández y Molina, 2011) . . . . .	31
2.11	Leyenda empleada en la cartografía temática de uso de suelos . . . . .	32
2.12	División político territorial de la red de transporte Centro - La Hechicera . . . . .	34
2.13	Area de influencia de transporte público (Gutiérrez et al., 2013) . . . . .	35
2.14	División en nodos de la red Centro - La Hechicera . . . . .	36
2.15	Modelo de datos de la red Centro- La Hechicera . . . . .	38
3.1	Visualización geográfica del conteo de viajes . . . . .	44

3.2	Buffer de influencia (300 m) para cada nodo de transporte . . . . .	46
3.3	Diagramas de dispersión ( $y$ vs $x_i$ ) . . . . .	49
3.4	Verificación de supuestos del modelo de orígenes de viajes ( $O_i$ ) . . . . .	56
3.5	Verificación de supuestos del modelo de atracciones de viajes ( $D_j$ ) . . . . .	57
3.6	Visualización geográfica de los valores ajustados $\hat{O}_i$ y $\hat{D}_j$ . . . . .	58
4.1	Representación de los 3 de estados de información del sistema . . . . .	62
5.1	Parámetros de la matriz de viajes en la aplicación web . . . . .	70
5.2	Matriz de viajes estimada. Hora: 7:00 AM - 8:30 AM . . . . .	71
5.3	Flujo entre cada arco ( $i,j$ ). Hora: 7:00 AM - 8:30 AM . . . . .	73
5.4	Matriz de viajes estimada. Hora: 12:00 M - 2:00 PM . . . . .	74
5.5	Flujo entre cada arco ( $i,j$ ). Hora: 12:00 M - 2:00 PM . . . . .	76
5.6	Flujo entre cada arco ( $i,j$ ). Hora: 4:00 PM - 6:00 PM . . . . .	78
B.1	Pantalla principal de la aplicación Matriz O-D . . . . .	90
C.1	Captura de viajes 7:00 AM - 6:00 PM . . . . .	96
C.2	Captura de viajes 7:00 AM - 6:00 PM . . . . .	97

# Agradecimientos

A Dios, por ser mi guía a lo largo de mi carrera, por ser mi fortaleza y brindarme una vida llena de aprendizajes.

A mi madre, quien ha sido mi apoyo incondicional, por estar siempre a mi lado brindándome consejos y la oportunidad de tener una excelente educación.

A mis hermanos y al resto de mi familia, por llenar mi vida de alegría y siempre estar presentes cuando los necesito.

A la profesora Magdiel Ablan, tutor de este trabajo, por su esfuerzo, dedicación y sabiduría que me han orientado en la elaboración de este proyecto de grado.

Al profesor Jacinto Davila y Carlos Echeverria, por su revisión del trabajo de grado y los aportes al mismo.

A todos mis amigos, quienes durante mi camino me han brindado el aliento para concluir este trabajo de grado, y me han enseñado a ser una mejor persona cada día.

A la ilustre Universidad de Los Andes, por ser la casa de estudios que me acogió durante mi carrera.

A la Empresa Tromerca, por la oportunidad brindada en la elaboración de este proyecto de grado.

A la profesora Magdalena Parra y al profesor Douglas Rivas, por ser participe en la realización de este proyecto.

A todas aquellas personas que de alguna manera u otra han contribuido para el logro de mis objetivos

# Capítulo 1

## Introducción

### 1.1 Antecedentes

En la década de los años 1960, surgió formalmente el algoritmo de las cuatro etapas: generación, distribución, partición modal y elección de ruta; un enfoque integral de estudio del transporte reconocido ampliamente por planificadores para modelar la demanda de viajes de personas. A grandes rasgos, este esquema maneja el fenómeno espacial del transporte entre áreas geográficas al estudiar a partir de encuestas, cuántos viajes se originan y atraen, en dónde inician y terminan estos viajes, qué modo de transporte utilizan, y cuáles rutas emplean en su movimiento ([Moreno, 2006](#)). A continuación se mencionan diversos estudios de modelos encontrados en la literatura, donde se ha logrado estimar la demanda de transporte basados en el enfoque anterior y en otras diferentes metodologías, entre estas se encuentran:

[Vozzi y Acquaviva \(2011\)](#), donde aplicaron el modelo clásico de las cuatro etapas para realizar un pronóstico plausible de la demanda mediante el software privativo TransCAD, calibrado para un modelo del sistema de transporte del área metropolitana del Rosario, Argentina, con el fin de evaluar su desempeño actual y el impacto en las distintas alternativas de gestión que se tengan para dicho sistema. En [Puebla et al. \(2008\)](#), se propuso una alternativa a la clásica metodología de las cuatro etapas, basada en sistemas de información geográfica, con la que se estima la demanda en estaciones de la red del metro de Madrid; con el propósito de explicar y predecir los

viajes (entradas de usuarios al sistema) se plantean dos modelos: uno a partir de las tasas de penetración del uso del metro en función de la distancia, y otro generado mediante un modelo de regresión lineal múltiple que incorpora las variables relativas a la estación y su entorno. Por otra parte, en [Gutiérrez et al. \(2013\)](#) se formula un modelo alternativo al estudio anterior que consiste en la aplicación de regresión geográficamente ponderada para estimar la demanda del metro de Madrid; los resultados obtenidos para este modelo representan un mejor ajuste y su significancia estadística permite alcanzar resultados más realistas que los encontrados en [Puebla et al. \(2008\)](#). [Quintero et al. \(2011\)](#) desarrollaron un modelo de regresión lineal que cuantifica el número de viajes vehiculares generados por cuatro conjuntos residenciales ubicados en la Av. Alberto Carnevali del Municipio Libertador en el Estado Mérida, en base a características urbanas de la zona. En otro sentido, [Galán \(2005\)](#) estima los determinantes en la demanda de transporte público y privado en la ciudad de Monterrey México: por medio de un modelo de posesión de coches y otro de elección modal, se crean posibles escenarios que permitan ver cuáles serían las políticas de transporte que pudieran incrementar en diez puntos porcentuales el uso del transporte público.

## 1.2 Justificación

El aumento de la demanda por uso del transporte público en la ciudad de Mérida ha sido consecuencia del constante crecimiento poblacional y la forma en que se distribuyen las actividades como el trabajo, centros educativos y recreacionales. Un servicio de transporte sin los atributos necesarios que permitan gestionar adecuadamente la demanda existente puede ser completamente inútil; por este motivo, para enfocar e invertir los recursos existentes del proyecto de infraestructura de la Línea 2 del Trolebús que cubrirá la ruta Centro - La Hechicera, es necesario hacer uso de las herramientas dentro de la planificación urbana, de entre las cuales destacan los modelos de demanda, pues permiten estimar el número de viajes que confluyen en una zona de estudio; además contribuyen en la descripción del comportamiento de viajes futuros que presentará un nuevo sistema de transporte a implementar, logrando así identificar sus problemas y establecer una base más sólida para la solución de los mismos.



## 1.3 Planteamiento y delimitación del problema

Los problemas asociados al transporte público en la ciudad de Mérida son cada vez más serios; éstos se ven reflejados en el congestionamiento masivo que se sostiene en la ciudad a lo largo de día, producto del crecimiento de su población año tras año. Adicionalmente se ha generado una demanda tal que sobrepasa las capacidades de los sistemas de transporte público que ofertan sus servicios en las principales rutas. Por este motivo, la empresa Tromerca ha decidido expandir el sistema de rutas del Trolebús Mérida, a través del desarrollo de la Línea 2 que abarcará la ruta Centro - La Hechicera. Ciertamente, los proyectos de infraestructura de transporte poseen largos plazos y costos elevados y, dado que los recursos son limitados, para la planificación del proyecto de la nueva Línea del Trolebús se requiere la estimación de la demanda de transporte de la zona a través de un modelo, para su correcta evaluación, un modelo que sirva como base para la toma de decisiones y generación de políticas por parte de los planificadores de transporte, orientadas a maximizar los beneficios de la comunidad merideña.

## 1.4 Objetivos

### 1.4.1 Objetivo general

Desarrollar un modelo de la demanda de transporte de la red que cubre la ruta Centro - La Hechicera bajo la aplicación de las herramientas estadísticas y optimización.

### 1.4.2 Objetivos específicos

- Dividir, por medio de un grafo de transporte, el área de estudio en zonas o nodos codificados.
- Estudiar la tipología de uso de suelos de la red y sus características geográficas.
- Muestrear los datos de viajes y las variables urbanísticas que intervienen en la dinámica de la demanda.

- Almacenar las variables involucradas en una base de datos por medio de un sistema de información geográfica.
- Explicar y estimar la cantidad de viajes totales diarios originados y atraídos en cada una de los nodos.
- Obtener y estudiar la matriz origen-destino para los viajes diarios de la red de transporte.
- Estudiar, mediante el modelo de demanda, escenarios futuros sobre la Línea 2 del Trolebús Mérida.

## 1.5 Metodología de trabajo

El enfoque de modelización que se implementó en este trabajo es una adaptación de la estructura clásica del modelo de cuatro etapas. Los resultados aquí presentados fueron obtenidos mediante implementaciones en el software R.

Se realizó una descripción geográfica detallada de la zona Centro - La Hechicera y se representó por medio de una red de transporte.

Se determinaron las variables relevantes a la demanda y se procesaron por medio de un sistema de información geográfica utilizando el software QGIS.

Se desarrolló el modelo de generación para obtener la cantidad de viajes originados ( $O_i$ ) y atraídos ( $D_j$ ) en cada nodo, aplicando la técnica de regresión lineal múltiple (Montgomery, 2005) donde, a partir de la información capturada, se explica además las relación entre las características urbanas de la red con la demanda de transporte.

Se contruyó la matriz origen-destino: mediante la estimación de un modelo de distribución se obtuvo la cantidad de viajes totales ( $V_{ij}$ ) entre un nodo  $i$  y un nodo  $j$ , utilizando el concepto de máxima entropía (de Dios Ortúzar y Willumsen (2008)).

A través del framework *Shiny* para R, se desarrolló una aplicación web interactiva que permite generar, manipular y presentar diferentes matrices de viajes y así analizar distintos escenarios de planeación de la Línea 2 Centro - La Hechicera.

## 1.6 Marco teórico y conceptual

En el marco del estudio de la demanda de transporte como un fenómeno urbano, dinámico y complejo, es apropiado manejar conceptos útiles, tales como planificación urbana, planificación y modelos de transporte, modelos de regresión lineal, modelos de maximización de la entropía y sistemas de información geográfica, lo cual permite facilitar el mejor aprovechamiento de la información plasmada en este trabajo.

### 1.6.1 Demanda de transporte

Los servicios de transporte surgen como consecuencia de la necesidad que tienen los individuos de realizar actividades que implican desplazamientos; por ejemplo, ir al trabajo, llevar los niños al colegio, ir al médico, entre otros. De acuerdo con esta definición, se dice que la demanda de transporte es derivada. En general, se demanda viajar con el objeto de realizar alguna actividad localizada en el espacio (destino) y en el tiempo. Estamos, por tanto, ante una demanda cualitativa y diferenciada en el sentido de que existen viajes por múltiples motivos, a distintas horas del día y que pueden realizarse en distintos modos de transporte ([Salas y García, 2003](#)).

### 1.6.2 Planificación urbana

La complejidad que conlleva el modelado de la demanda del transporte, es uno de los principales retos dentro de la planificación urbana. Sánchez de Madariaga como se cita en [Ornés \(2009\)](#), la define como el proceso de descripción, análisis y evaluación de las condiciones de las ciudades y su funcionamiento, para poder generar propuestas de diseño y formular proyectos que permitan regular la dinámica urbana y ambiental de toda la ciudad.

### 1.6.3 Planificación de transporte

La planificación de transporte es una de las tareas principales de los Ingenieros de Transporte. De acuerdo con [Monge \(2011\)](#), un planificador de transporte estudia la demanda presente y futura de movilidad de personas y material. Estos proyectos están precedidos por estudios de movimientos e involucran a los diferentes medios de transporte. Por lo tanto, es la fase fundamental del proceso de desarrollo de proyectos de infraestructura, pues es la que permite conocer los problemas, diseñar o crear soluciones y, en definitiva, optimizar y organizar los recursos para enfocarlos a atender la demanda de movilidad existente.

### 1.6.4 Modelo de transporte

Los modelos de transporte son utilizados como herramientas de análisis, predicción y apoyo general a la toma de decisiones en la planificación del transporte tanto público como privado. Los problemas que se atacan son básicamente de minimización de tiempos de viajes y aseguramiento de la calidad del servicio del transporte público. En la construcción de modelos matemáticos de transporte se trabaja a partir de un enfoque tomado de la economía donde la demanda es la necesidad de movimiento para un cierto período de tiempo desde un origen a un destino y la oferta es la red vial y el sistema de transporte público entre dicho origen y destino. La metodología empleada usualmente para modelar el transporte es el denominado algoritmo de las cuatro etapas ([Mauttone Vidales, 2002](#)).

Los modelos de transporte proporcionan una herramienta objetiva para evaluar las principales inversiones en infraestructura y preparar planes regionales de transporte a largo plazo ([Martin y McGuckin, 1998](#)).

### 1.6.5 Estructura de un modelo de transporte

La estructura del modelo general de transporte se representa por medio de la red de la figura 1.1, donde hay  $m$  orígenes o fuentes y  $n$  destinos, cada uno representado por un nodo. Los arcos representan las rutas que enlazan los orígenes y los destinos. El arco  $(i, j)$ , conduce dos clases de información: el costo de unidad de transporte  $c_{ij}$  y la

cantidad transportada  $X_{ij}$  desde el origen  $i$  hacia el destino  $j$ . La cantidad de oferta en el origen  $i$  es  $a_i$  y la cantidad de demanda en el destino  $j$  es  $b_j$  (Taha, 2004),.

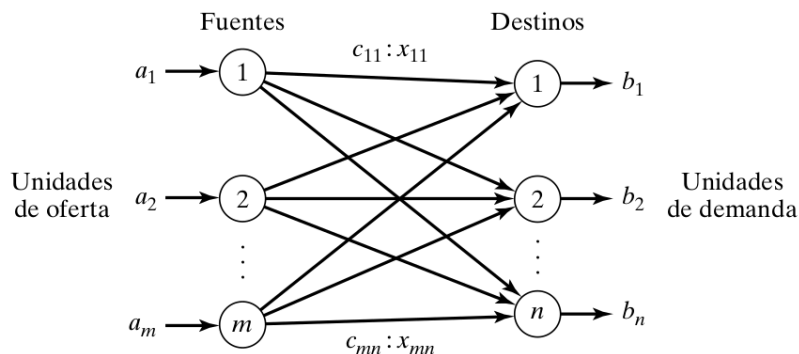


Figura 1.1: Representación del modelo de transporte en nodos y arcos

La misma información representada por la red de la figura 1.1, se puede formular por medio de una tabla de parámetros en el formato de la tabla 1.1, donde se observan los datos correspondientes a la oferta y demanda, siendo estos los parámetros del modelo (Lieberman y Hillier, 2010).

Tabla 1.1: Tabla de parametros del modelo de transporte (Lieberman y Hillier, 2010)

Origen	Destino				Oferta
	1	2	...	n	
1	$X_{11}$	$X_{12}$	...	$X_{1n}$	$a_1$
2	$X_{21}$	$X_{22}$	...	$X_{2n}$	$a_2$
:	:	:	...	:	:
m	$X_{m1}$	$X_{m2}$	...	$X_{mn}$	$a_m$
<b>Demanda</b>	$b_1$	$b_2$	...	$b_n$	

#### 1.6.5.1 Soluciones factibles

En Lieberman y Hillier (2010), la hipótesis en que se basa un problema de transporte es que éste tiene una solución factible si existe un balance entre los parámetros del modelo, es decir, si y sólo si:

$$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j$$

En nuestro caso, para el modelo de transporte de pasajeros, es necesario emplear términos mucho más específicos, empleando los términos: viajes originados ( $O_i$ ) a la cantidad de oferta en el nodo  $i$  y viajes atraídos ( $D_j$ ) a la cantidad demandada por el nodo  $j$ .

### 1.6.6 Modelo de cuatro etapas

En la literatura existente sobre modelos de predicción de la demanda del transporte, queda en evidencia que la mayoría se basa en el clásico modelo de cuatro etapas, surgido en los años 50 (figura 1.2). En teoría, la robustez de este modelo lo convierte en la mejor herramienta para el análisis y predicción de la demanda, pero en la práctica, muestra varios inconvenientes (Puebla et al., 2008). Uno de los principales problemas encontrados en la aplicación de la metodología de cuatro etapas en este trabajo, es que el mismo concierne al modelado de la demanda para todos los modos de transporte (público, privado, entre otros), además estuvo presente el largo plazo de tiempo que este— requiere, los recursos y múltiples encuestas que lo alimentan, por lo que se optó en este trabajo, por construir un modelo más simple como una alternativa al modelo de cuatro etapas, que produzca una estimación aceptable (de manera rápida y con bajos recursos) de la demanda para el modo de transporte público. El mismo es presentado en la sección 2.2.

### 1.6.7 Modelo de regresión lineal

En muchas situaciones prácticas existe una variable dependiente o de respuesta  $y$ , que se cree está relacionada o depende de  $k$  variables independientes o regresores, por ejemplo  $x_1, x_2, x_3, \dots, x_k$ . La relación que existe entre estas variables se caracteriza por un modelo matemático llamado modelo de regresión (Montgomery, 2005).

En Pulido y De la Vara Salazar (2008), se enuncia el modelo de la siguiente manera:

Sean  $x_1, x_2, \dots, x_k$  variables independientes o regresoras, y sea  $y$  una variable de respuesta, entonces el modelo de regresión lineal múltiple con  $k$  variables independientes

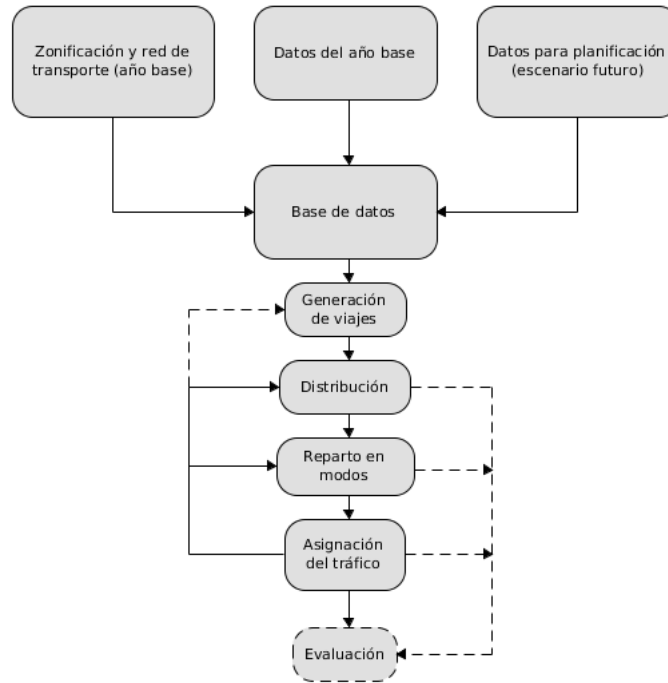


Figura 1.2: El clásico modelo de cuatro etapas para la planeación del transporte (de Dios Ortúzar y Willumsen, 2008)

es el polinomio de primer orden:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad (1.1)$$

donde los  $\beta_j$  son los parámetros del modelo que se conocen como coeficientes de regresión y  $\varepsilon$  es el error aleatorio, con media cero,  $E(\varepsilon) = 0$  y  $V(\varepsilon) = s^2$ . Si en la ecuación (1.1),  $k = 1$ , estamos ante un caso de regresión lineal simple y el modelo es una línea recta; si  $k = 2$ , tal ecuación representa un plano. En general, la ecuación (1.1) representa un hiperplano en el espacio de  $k$  dimensiones generado por las variables  $\{x_j\}$ .

El término lineal del modelo de regresión se emplea debido a que la ecuación (1.1) es función lineal de los parámetros desconocidos  $\beta_0, \beta_1, \dots, \beta_k$ .

Es frecuente que en la práctica se requieran modelos de mayor orden para explicar el comportamiento de  $y$  en función de las variables regresoras. Por ejemplo, supongamos que se tienen dos variables independientes y que se sospecha que la relación entre  $y$  y

las variables independientes es cuadrática, por ello quizá se requiera un polinomio de segundo orden como modelo de regresión:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \varepsilon \quad (1.2)$$

La ecuación anterior puede reducirse a un modelo de regresión lineal múltiple, puesto que es una función lineal de los parámetros desconocidos  $\beta_0, \beta_1, \dots, \beta_{22}$ . Además si se definen  $x_3 = x_1 x_2$ ,  $\beta_3 = \beta_{12}$ ,  $x_4 = x_1^2$ ,  $\beta_4 = \beta_{11}$ ,  $x_5 = x_2^2$  y  $\beta_5 = \beta_{22}$ , entonces la ecuación (1.2) puede escribirse así:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \varepsilon \quad (1.3)$$

La ecuación (1.3) tiene la misma forma que el modelo general de regresión lineal múltiple. Con base en lo anterior, [Montgomery \(2005\)](#) establece que existe una amplia gama de modelos de primer y segundo orden que pueden reducirse a la forma general de la expresión (1.1).

A continuación, se analizarán los principales aspectos del análisis de regresión útiles en esta investigación, como el método para estimar los parámetros y propiedades de dichas estimaciones.

### 1.6.7.1 Estimación de los parámetros del modelo

#### 1.6.7.1.1 Método de mínimos cuadrados

El método de mínimos cuadrados se utiliza de manera típica para estimar los coeficientes del modelo de regresión. Se supone que se cuenta con  $n > k$  observaciones de la variable respuesta  $y_1, y_2, \dots, y_n$ . Junto con cada respuesta observada  $y_i$  se tendrá una observación de cada uno de los regresores  $x_{ij}$ , donde  $x_{ij}$  denota la observación  $i$ -ésima de la variable  $x_j$ , tal como se muestra en la tabla 1.2 ([Montgomery, 2005](#)).

A partir de la estructura de datos de la tabla 1.2, en [Pulido y De la Vara Salazar \(2008\)](#), se establece que el modelo de regresión lineal múltiple puede escribirse de la siguiente manera:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \varepsilon$$



Tabla 1.2: Estructura de Datos en Regresión Lineal Multiple

$\mathbf{y}$	$\mathbf{x}_1$	$\mathbf{x}_2$	$\dots$	$\mathbf{x}_k$
$y_1$	$x_{11}$	$x_{21}$	$\dots$	$x_{k1}$
$y_2$	$x_{12}$	$x_{22}$	$\dots$	$x_{k2}$
$\vdots$	$\vdots$	$\vdots$	$\dots$	$\vdots$
$y_n$	$x_{1n}$	$x_{2n}$	$\dots$	$x_{kn}$

$$y_i = \beta_0 + \sum_{j=1}^k \beta_j x_{ji} + \varepsilon_i, i = 1, 2, 3, \dots, n \quad (1.4)$$

El metodo de los mínimos cuadrados consiste en obtener  $\beta$  de la ecuación (1.4), de tal modo que la suma de los cuadrados de los errores  $\varepsilon_i$  se minimice, por lo tanto, al despejar los errores, elevarlos al cuadrado y sumarlos, se obtiene:

$$S = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^k \beta_j x_{ji})^2$$

Para minimizar los errores, se debe derivar la función  $S$  con respecto a cada parámetro  $\beta_j$ , es decir:

$$\frac{\partial S}{\partial \beta_j}, (\forall j = 0, 1, \dots, k)$$

Las  $p = k + 1$  ecuaciones resultantes se igualan a cero. [Montgomery \(2005\)](#) denomina estas ecuaciones como ecuaciones normales de mínimos cuadrados. La solución simultanea de las  $p$  ecuaciones normales serán los estimadores de mínimos cuadrados de los coeficientes de regresión  $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$ . Con el propósito de resolver las ecuaciones normales de una manera práctica y sencilla, se optó por utilizar en este trabajo (en término de los datos de la ecuación (1.4)), su notación matricial tal como se refiere en [Pulido y De la Vara Salazar \(2008\)](#)

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon$$

donde:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{21} & \dots & x_{k1} \\ 1 & x_{12} & x_{22} & \dots & x_{k2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1n} & x_{2n} & \dots & x_{kn} \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}$$

$$\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Nótese que las matrices y vectores están denotados en negrita y en mayúscula y minúscula, respectivamente;  $\mathbf{y}$  es un vector  $n \times 1$  de las observaciones,  $\mathbf{X}$  es una matriz  $n \times p$  de los datos de las variables independientes,  $\beta$  es un vector  $p \times 1$  de los coeficientes de regresión, y  $\varepsilon$  es un vector  $n \times 1$  de los errores aleatorios.

Se quiere encontrar el vector de los estimadores de mínimos cuadrados  $\hat{\beta}$ , que minimice:

$$\begin{aligned} S &= \sum_{i=1}^n \varepsilon_i^2 = \varepsilon' \varepsilon = (\mathbf{y} - \mathbf{X}\beta)'(\mathbf{y} - \mathbf{X}\beta) \\ &= \mathbf{y}'\mathbf{y} - \beta'\mathbf{X}'\mathbf{y} - \mathbf{y}'\mathbf{X}\beta + \beta'\mathbf{X}'\mathbf{X}\beta \\ &= \mathbf{y}'\mathbf{y} - 2\beta'\mathbf{X}'\mathbf{y} + \beta'\mathbf{X}'\mathbf{X}\beta \end{aligned}$$

La última igualdad se debe a que  $\beta' \mathbf{X}' \mathbf{y}$  es una matriz  $1 \times 1$ , o un escalar y, por lo tanto, su transpuesta  $(\beta' \mathbf{X}' \mathbf{y})' = \mathbf{y}' \mathbf{X} \beta$  es el mismo escalar. Ahora, se evalúa que los estimadores de mínimos cuadrados satisfagan:

$$\left. \frac{\partial S}{\partial \beta} \right|_{\hat{\beta}} = -2\mathbf{X}' \mathbf{y} + 2\mathbf{X}' \mathbf{X} \hat{\beta} = 0$$

Esto último implica que:

$$\mathbf{X}' \mathbf{X} \hat{\beta} = \mathbf{X}' \mathbf{y} \quad (1.5)$$

Para resolver esta ecuación en términos de  $\hat{\beta}$ , se multiplican ambos lados de la ecuación (1.5) por la matriz inversa de  $\mathbf{X}' \mathbf{X}$ , y se obtiene que el estimador de mínimos cuadrados de  $\beta$  es:

$$\hat{\beta} = (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{y} \quad (1.6)$$

Por lo tanto, el modelo ajustado está dado por:

$$\hat{y} = \mathbf{X} \hat{\beta} \quad (1.7)$$

En notación escalar, el modelo ajustado es:

$$\hat{y}_i = \hat{\beta}_0 + \sum_{j=1}^k \hat{\beta}_j x_{ji}, i = 1, 2, 3, \dots, n \quad (1.8)$$

#### 1.6.7.1.2 Estimación de la varianza $\sigma^2$

Según Pulido y De la Vara Salazar (2008) para hacer inferencias sobre  $\beta$  o, en general, sobre el modelo de regresión, on el fin de verificar su validez, es necesario encontrar estimar la varianza  $\sigma^2$ . A partir de la ecuación (1.7), se puede obtener que el vector de residuos que esta dado por:

$$\varepsilon = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{X} \hat{\beta} \quad (1.9)$$

Para desarrollar un estimador de este parámetro, se considera la suma de cuadrados de los residuos:

$$SC_E = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$SC_E = \sum_{i=1}^n \varepsilon_i^2 = \varepsilon' \varepsilon$$

$$SC_E = (\mathbf{y} - \mathbf{X}\hat{\beta})'(\mathbf{y} - \mathbf{X}\hat{\beta}) = \mathbf{y}'\mathbf{y} - 2\hat{\beta}'\mathbf{X}'\mathbf{y} + \hat{\beta}'\mathbf{X}'\mathbf{X}\hat{\beta} \quad (1.10)$$

De acuerdo con la ecuación (1.5):  $\mathbf{X}'\mathbf{X}\hat{\beta} = \mathbf{X}'\mathbf{y}$ , esta última ecuación toma la siguiente forma:

$$SC_E = \mathbf{y}'\mathbf{y} - \hat{\beta}'\mathbf{X}'\mathbf{y} \quad (1.11)$$

La suma de cuadrados residual o del error dada por esta última expresión tiene  $n - k - 1$  grados de libertad, donde  $k + 1$  es igual al número de parámetros estimados en el modelo. Entonces, el cuadrado medio del error es:

$$CM_E = \frac{SC_E}{n - k - 1}$$

Se obvió para este trabajo la demostración de que el valor esperado de  $CM_E = \sigma^2$  (Pulido y De la Vara Salazar, 2008) y, por lo tanto, un estimador insesgado de  $\sigma^2$  está dado por:

$$\hat{\sigma}^2 = CM_E \quad (1.12)$$

### 1.6.7.1.3 Propiedades de los estimadores

Para Montgomery (2005), el método de mínimos cuadrados produce un estimador insesgado del parámetro  $\beta$ . Esto puede demostrarse fácilmente estudiando el valor esperado de  $\hat{\beta}$  de la siguiente manera:

$$\begin{aligned} E(\hat{\beta}) &= E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}] \\ &= E[(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}\beta + \varepsilon)] \\ &= E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\varepsilon)] \end{aligned}$$

$$= \beta$$

Esto se debe a que  $E(\varepsilon) = 0$  y  $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X} = I$ . Por lo tanto,  $\hat{\beta}$  es un estimador insesgado de  $\beta$ . Además, se puede demostrar que la matriz de covarianza de  $\hat{\beta}$  es:

$$Cov(\hat{\beta}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1} \quad (1.13)$$

### 1.6.8 Modelo de maximización de la entropía

A continuación, se resumen los principios básicos de la maximización de la entropía como medida de información en la estadística.

Las medidas de información tienen su origen en la Teoría de la Información. Las entropías miden la información contenida en una distribución, es decir, la incertidumbre acerca de los resultados de un experimento aleatorio, siendo las entropías de Shannon y de Rényi las medidas clásicas de este tipo [Cunat \(2006\)](#). Por su parte [López-Ospina \(2013\)](#), establece que el objetivo de la máxima entropía es generar las distribuciones de probabilidad que soporten el mayor nivel de incertidumbre entre todas las distribuciones de probabilidad que son factibles bajo cierto tipo de información.

Con respecto a las definiciones anteriores, de acuerdo con [Cunat \(2006\)](#), consideremos un experimento aleatorio  $A$  con dos posibles resultados, con probabilidades  $p_1, p_2 (p_i \geq 0, i = 1, 2, p_1 + p_2 = 1)$ . La incertidumbre acerca del posible resultado en caso de realizar el experimento depende de las probabilidades de los resultados, pues si se consideran los experimentos:

$$A_1 = \begin{pmatrix} a_1 & b_2 \\ 0.5 & 0.5 \end{pmatrix}, A_2 = \begin{pmatrix} a_1 & b_2 \\ 0.999 & 0.001 \end{pmatrix}$$

El primer experimento aleatorio contiene más incertidumbre sobre el resultado que el segundo. Es lógico pensar que en  $A_2$  el resultado  $a_1$  tiene mayor probabilidad de ocurrencia. Las medidas de entropía asignan un valor numérico a cada distribución de probabilidad, materializando la idea intuitiva de mayor o menor incertidumbre.

#### 1.6.8.1 Entropía de Shannon

Matemáticamente, se define la entropía de Shannon de la siguiente forma ([Cunat, 2006](#)):

Sea:

$$\Delta_n = P = (p_1, p_2, \dots, p_n) : p_i \geq 0, i = 1, \dots, n, \sum_{i=1}^n p_i = 1$$

El conjunto formado por todas las distribuciones de probabilidad asociadas a una variable aleatoria  $x$  que toma un número finito de valores  $x_1, x_2, \dots, x_n$ ; se denomina entropía de la variable aleatoria  $x$  o entropía de la distribución  $P = (p_1, p_2, \dots, p_n)$  a la expresión:

$$H(x) = H(p_1, p_2, \dots, p_n) = - \sum_{i=1}^n p_i \log p_i \quad (1.14)$$

Los logaritmos se pueden tomar con respecto a cualquier base que sea mayor que la unidad. La indeterminación  $p_i \log p_i$  con  $p_i = 0$  se resuelve definiendo  $p_i \log p_i = 0$  si  $p_i = 0$ . Es decir, la función  $f(x) = -x \log x$  definida en  $(0, 1)$  se extiende por continuidad a  $[0, 1)$ , definiendo:

$$f(x) = \begin{cases} -x \log x & \text{si } x > 0; \\ 0 & \text{si } x = 0. \end{cases}$$

La función  $H(x)$  evalúa la cantidad de información que se espera obtener al conocer el valor de  $x$ . A su vez, mide el nivel de incertidumbre asociado con próximas realizaciones de  $x$  (López-Ospina, 2013). Como ejemplo, tomemos una variable aleatoria *Bernoulli* con parámetro  $p$ , entonces la entropía se define como:

$$H(p) = -p \log(p) - ((1-p) \log(1-p))$$

En el caso en que  $p$  tiende a un valor extremo 0 ó 1, entonces la solución tenderá a 0, indicando menor estocasticidad en el proceso. Por otro lado, la función de entropía alcanza su valor máximo en  $p = 0.5$ , donde se tiene el mayor nivel de incertidumbre posible, es decir, cuando las dos opciones tienen igual probabilidad de ocurrencia. Claramente, en cualquier problema de maximización de la entropía, la solución óptima es la distribución uniforme, es decir  $p_1 = p_2 = \dots = p_n = \frac{1}{n}$ , pero en el caso de tener información adicional del sistema, dicha distribución puede variar de acuerdo con los datos a priori obtenidos.

Matemáticamente, si se supone que se tiene  $g_1(x), \dots, g_m(x)$ ,  $m$  funciones de  $x$  con valor esperado conocido, notado como  $a_1, a_2, \dots, a_m$ , entonces la distribución determinada por el problema de máxima entropía con la información asociada con las funciones  $g_i(x)$  es:

$$\max H(x) = H(p_1, p_2, \dots, p_n) = - \sum_{i=1}^n p_i \log(p_i)$$

sueto a:

$$\sum_{i=1}^n p_i g_i(x_i) = a_j, j = 1, 2, 3, \dots, m$$

$$\sum_{i=1}^n p_i = 1$$

$$p_i \geq 0$$

Y a su vez, la solución del problema de optimización anterior determina una distribución consistente con las restricciones y que maximiza la incertidumbre. Este tipo de formulaciones suele presentar, en algunas situaciones, restricciones no lineales (López-Ospina, 2013). Cabe destacar que para este trabajo, por la naturaleza de los problemas de planificación de transporte, se consideran solo restricciones lineales.

### 1.6.9 Sistemas de Información Geográfica

El término de Sistema de Información Geográfica (SIG) hoy esta ampliamente difundido, especialmente entre los profesionales que trabajan en la planificación de transporte. El *National Center for Geographic Information and Analysis* tal como se cita en Puebla y Gould (1994), establece que un SIG, es un sistema de hardware, software y procedimientos diseñados para realizar la captura, almacenamiento, manipulación, análisis, modelización y presentación de datos referenciados espacialmente para la resolución de problemas complejos de planificación y gestión. Otro punto de vista, se trata de una extension del concepto de base de datos, tomado de Smitg tal como se cita en Puebla y Gould (1994), un SIG es un sistema de base de datos en el cual la mayor parte de los datos estan indexados espacialmente y que dispone de un conjunto de procedimientos para responder a cuestiones sobre las

entidades espaciales de la base de datos. De acuerdo con Campos (1991), los SIG son las nuevas herramientas de trabajo de los planificadores urbanos pues posibilitan la representación gráfica de modelos complejos mediante la organización y normalización de la información espacial.

## 1.7 Estructura del documento

Este trabajo está organizado de la siguiente manera:

En el capítulo 2 se describe la red Centro - La Hechicera y las operaciones del sistema de transporte masivo Trolebús Mérida; para luego introducir el modelo de estimación de la demanda de dicha red y las variables involucradas en el estudio.

El capítulo 3 presenta el modelo de generación donde se ajustan dos submodelos de regresión lineal para estimar la cantidad de viajes originados y atraídos a cada nodo de transporte; posteriormente se realizan las pruebas de verificación y validación a los mismos.

En el capítulo 4 se discute la teoría de la máxima entropía para obtener el modelo de distribución, donde se refleja el patrón de viajes de la red de transporte mediante su implementación en una matriz origen-destino.

En el capítulo 5 se analizan, a través de una aplicación web, algunos escenarios propuestos para la línea 2 del Trolebús Mérida.

Para finalizar, el capítulo 6 muestra las conclusiones del estudio y las recomendaciones del mismo.



## Capítulo 2

# Descripción y Análisis de la Red de Transporte

El análisis de la demanda de transporte es la base técnica para la planificación de las operaciones de la Línea 2 del Trolebús Mérida, pues este sistema deberá estar basado en las necesidades de sus usuarios. En este capítulo se describe la Línea 2 del Trolebús Mérida y se formula en su totalidad el modelo de demanda, describiendo y dividiendo en nodos la red de transporte en estudio para el posterior muestreo de las variables involucradas y almacenamiento en una base de datos espacial.

### 2.1 Sistema de Transporte Masivo Trolebús Mérida

Mérida fue la primera ciudad en Venezuela que adoptó el sistema de transporte trolebús desde su inauguración el 18 de Junio de 2007. Consiste en dos líneas de autobuses articulados de motorización dual: un motor eléctrico, alimentado por un sistema de cables, y un motor de combustible diesel; más una tercera línea de sistema de telecabinas o Trolcable y un sistema de buses Yutong. La administración total de este sistema corresponde a la compañía Tromerca (Trolebús Mérida C.A), empresa adscrita al Ministerio del Poder Popular para el Transporte Terrestre, encargada de financiar y ejecutar los programas de inversión en infraestructura y adquisición de equipos e instalaciones. El Sistema de Transporte Masivo Trolebús Mérida (STMTM)

comprende la Línea 1 Ejido - Centro que cubre la ruta que va desde el Terminal de Ejido hasta la estación Domingo Peña, en el sector Paseo de la Feria del centro de la ciudad (figura 2.1). La Línea 3 Centro - San Jacinto, un sistema de Trolcable desde el sector Paseo de la Feria, hasta el sector San Jacinto, en la cuenca del Chama; y finalmente la Línea 2 Centro - La Hechicera, aún en etapa de diseño y planeación, ruta que cubrirá el recorrido desde la estación Domingo Peña hasta el Núcleo Universitario Pedro Rincón Gutiérrez. Sólo hasta el 2016, a través de la la Línea 1, se han movilizado 24.4 millones de usuarios por la cual los índices de calidad de vida de los habitantes de Mérida se han incrementado significativamente y los indicadores de movilidad y accesibilidad han mejorado gracias a la reducción de tiempos de viaje y a la mayor circulación y calidad del transporte colectivo (Azuaje, 2006).

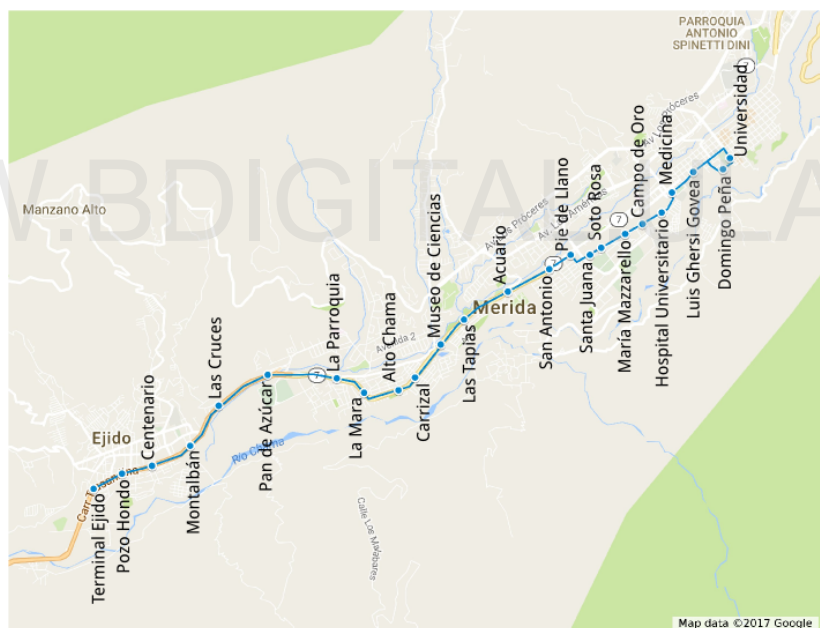


Figura 2.1: Red de rutas y estaciones del STMTM (Tromerca, 2017)

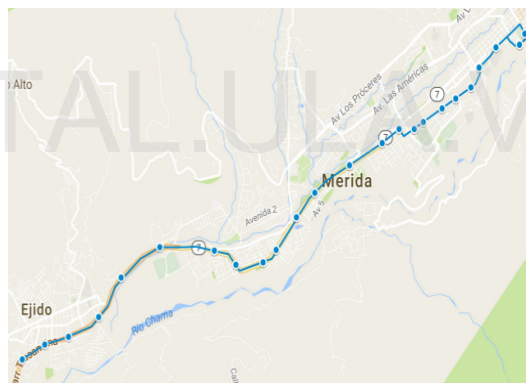
A continuación se estudia a grandes rasgos el funcionamiento de la Línea 1, siendo el único eje troncal actual operativo del STMTM.

### 2.1.1 Gestión de la Línea 1: Ejido - Centro

Con esta primera línea se da inicio al recorrido del STMTM desde las 6:00 am hasta las 8:30 pm, y se define como la banda longitudinal que permite la conexión urbana entre los municipios con mayor densidad de población del estado Mérida, siendo estos los municipios Campo Elías y Libertador. Cuenta con un recorrido total de 18,2 kilómetros por medio de un canal vial exclusivo (figura 2.2(a)) desde la estación Terminal de Ejido, en el municipio Campo Elías (en el sur del estado) hasta la estación Domingo Peña en el centro de la ciudad de Mérida (ubicada al norte de la misma) (figura 2.2(b)), lo que fortalece la movilidad entre las avenidas de mayor importancia que conforman la principal arteria vial propia de la ciudad, siendo estas: Centenario, Monseñor Chacón, Andrés Bello, Alexander Quintero, 16 de Septiembre, Don Tulio Febres Cordero, Calle 26 y Paseo Domingo Peña (Azuaje, 2006).



(a) Canal Exclusivo Trolebús



(b) Línea 1

Figura 2.2: Funcionamiento de la Línea 1 (Tromerca, 2017)

### 2.1.2 Nodos de transporte

Los nodos de transporte corresponden a los puntos de carga y descarga de usuarios del sistema. La Línea 1 presenta 28 estaciones o nodos, de las cuales dos son intermodales: estación Terminal de Ejido y estación Domingo Peña (figura 2.3(a)); una estación auxiliar: Andenes Mercado Periférico (Figura 2.3(b)); y las demás son estaciones típicas (figura 2.3(c)), siendo sólo una de ellas distinta (Estación Alto Chama), conservando

la disposición central de las 24 restantes (figura 2.3(d)).



Figura 2.3: Estaciones Línea 1 (Tromerca, 2017)

Entre sus aspectos funcionales se encuentran el acceso, espera, embarque y desembarque de pasajeros; cuentan con áreas de vestíbulos que permiten el desplazamiento eficiente de usuarios en su interior, de la misma manera presentan taquillas que permiten la compra-venta de los boletos, áreas de servicios básicos como salas sanitarias para los empleados del sistema y cuarto de tableros (figura 2.4).

Por lo general, cada estación presenta áreas constructivas de  $130m^2$ , con dimensiones en planta a lo largo de  $30m$  y anchos de  $4,20m$  totales, las alturas de piso a techo son de  $4.50m$  y cada estación se eleva  $60cm$  a partir de la sección vial; por medio de una plataforma rígida se garantiza el acceso directo sólo a los vehículos articulados, y una rampa de acceso a la estación que dirige cómodamente al usuario al recinto interno de espera (Azuaje, 2006).

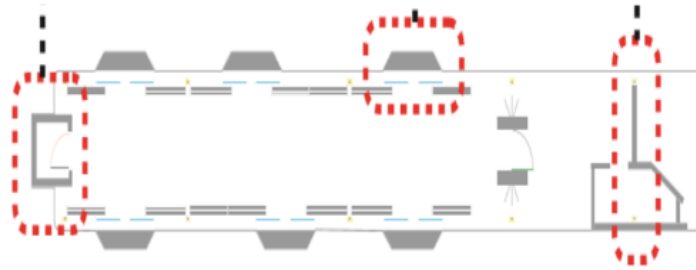


Figura 2.4: Infraestructura y acceso a nivel de estación (Tromerca, 2017)

### 2.1.3 Costo del transporte

La tarifa para el usuario de transporte público está relacionada con los costos operacionales y el nivel de subsidios (Wright et al., 2010). El STMTM cuenta con un modelo tarifario económicamente atractivo para la población merideña, además de un sistema automatizado de recaudación del pasaje donde, por medio de una tarjeta inteligente (figura 2.5) en sus dos versiones: usuario general y estudiantil, se permite agilizar el acceso a los servicios.

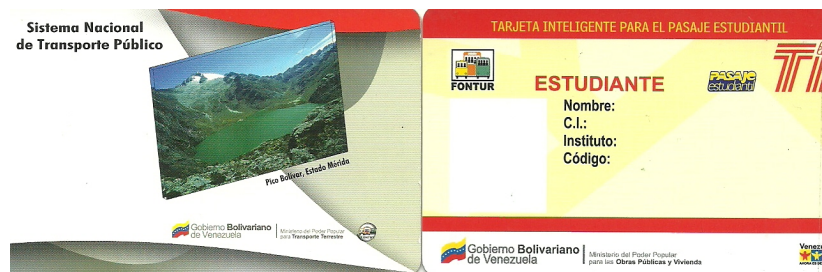


Figura 2.5: Tipos de tarjeta magnética inteligente (Tromerca, 2017)

Los usuarios de transporte son particularmente sensibles a los precios (Wright et al., 2010). La diferencia sustancial en niveles de tarifa de pasaje de casi 400 BsF (Mayo, 2017) entre los transporte públicos de la ciudad de Mérida y el STMTM ha generado un incremento en los niveles de demanda del Trolebús.

### 2.1.4 Unidades de transporte

El servicio de transporte de pasajeros entre los distintos nodos del sistema es prestado por medio de las tecnologías Trolebús y Autobús Yutong; actualmente para la Línea 1 están operativos un total de 45 y 141 unidades, respectivamente. Las características principales de los vehículos de transporte del sistema se resumen en la tabla 2.1.

Tabla 2.1: Tecnologías de transporte STMTM ([Tromerca, 2017](#))

Características	Trolebús	Yutong
Marca	Mercedes Benz	Mercedes Benz
Longitud	17.74m	13.75m
Ancho	2.50 m	2.30 m
Altura	3.00 m	2.50 m
Plataforma	0.72 m	0.70 m
Número de puertas	3 de 1.20m de ancho.	2 de 1.20m de ancho.
Configuración	2 cuerpos articulados	1 cuerpo
Capacidad aproximada	150 usuarios	85 usuarios
Asientos	42, 3 para discapacitados	33

### 2.1.5 Planificación de la Línea 2: Centro - La Hechicera

Con el fin de satisfacer la demanda de transporte en la mayor zona estudiantil del área metropolitana de la ciudad de Mérida, surge la necesidad de ampliar las rutas del STMTM con la puesta en marcha de la Línea 2 Centro - La Hechicera, la cual planifica extenderse por las vías arteriales más importantes de la ciudad, iniciando en el casco central, avenida Paseo Domingo Peña, tomando parte de la avenida Don Tulio Febres Cordero para luego atravesar la Calle 26 y Viaducto Campo Elías, donde hace cruce hacia la zona norte de la ciudad subiendo por la avenida Las Américas, continuando el recorrido que define toda la avenida Alberto Carnevalli llegando al núcleo estudiantil Pedro Rincón Gutiérrez de la Universidad de Los Andes, establecida como una banda longitudinal de disposición central, con una distancia total de 4.5 kilómetros, conformada por 8 estaciones (figura 2.6) que permitirán el acceso a las



distintas áreas presentes en el futuro escenario contiguo al tramo ([Azuaje, 2006](#)).

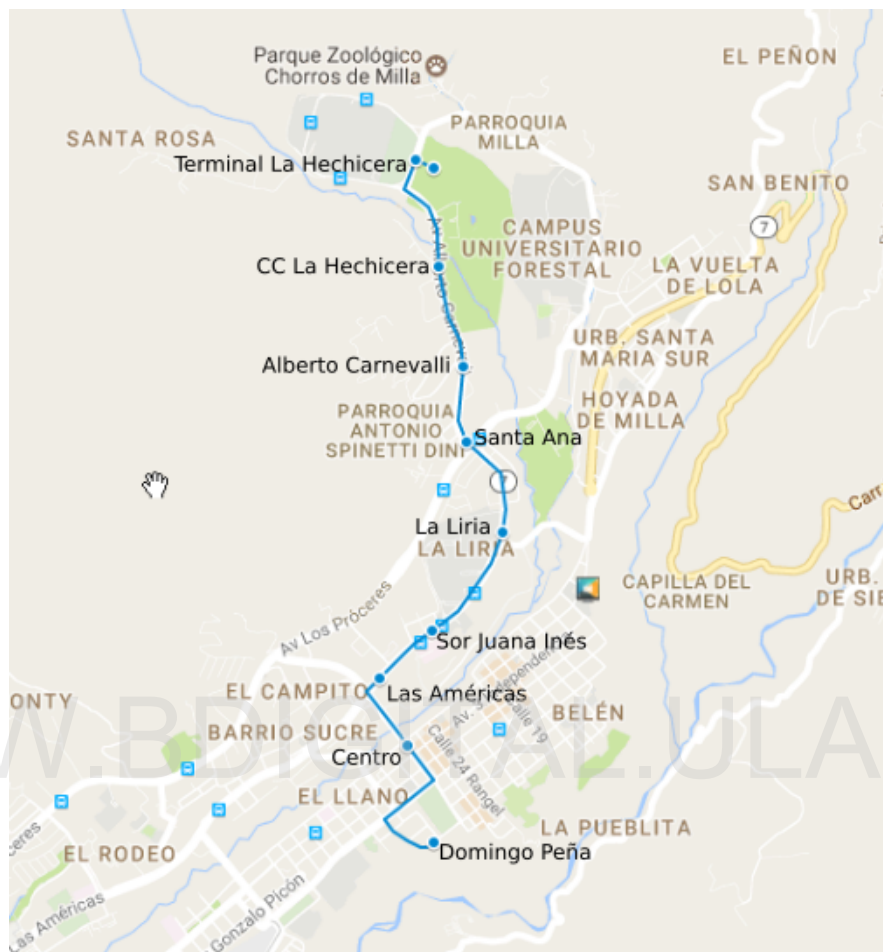


Figura 2.6: Red y nodos de transporte de la Línea 2 Centro - La Hechicera ([Tromerca, 2017](#))

La red de transporte estará compuesta por dos canales exclusivos para las tecnologías Trolebús y Yutong que recorrerán el eje central de las vías colectoras. Cada estación de carga y descarga de pasajeros se ubicará en el entorno próximo de cada parada de autobús de las principales áreas geográficas del recorrido (Centro, C.C. Yuan Lin, Hospital Sor Juana, Núcleo La Liria, Clínica Albarregas, Núcleo Pedro Rincon Gutierrez, entre otros). La localización geográfica y dimensiones que se estima poseerá cada estación se resumen en la tabla 2.2.

Como se observa en la tabla 2.2, las estaciones: La Hechicera, Alberto Carnevalli, La Liria y Sor Juana Inés corresponden a paradas tipo; la estación Las Américas, con

Tabla 2.2: Características de los nodos de la Línea 2 Centro - La Hechicera

Nodo	Nombre	Longitud	Latitud	Largo(m)	Ancho(m)
1	Centro	-71, 147090	8, 596296	56, 7	3, 82
2	Las Américas	-71, 148745	8, 599892	105, 89	5, 100
3	Sor Juana Ines	-71, 144606	8, 603385	64	3, 66
4	La Liria	-71, 142224	8, 607596	64	3, 66
5	Santa Ana	-71, 144010	8, 612131	En planificación	En planificación
6	Alberto Carnevali	-71, 144477	8, 617306	64	3, 66
7	C.C. La Hechicera	-71, 145582	8, 621906	64	3, 66
8	Terminal La Hechicera	-71, 152584	8.626684	En planificación	En planificación

una dimensión de  $540.039m^2$ , representa una estación triple pues tendrá capacidad para 1 Trolebús y 2 autobús Yutong, simultáneamente. Para las estaciones Santa Ana y Terminal La Hechicera aún no existe un bosquejo diseñado. Para la estación Terminal se esta programando desarrollar dos áreas separadas para el embarque y desembarque de usuarios de transporte, con áreas de  $318m^2$  y  $220m^2$ , respectivamente. La distancia vial en metros entre los centroides de cada una de las estaciones se visualiza en tabla 2.3.

Tabla 2.3: Distancia entre nodos de la Línea 2

Origen	Destino	Distancia (m)
Domingo Peña	Centro	975
Centro	Las Américas	525
Las Américas	Sor Juana Ines	390
Sor Juana Ines	La Liria	273
La Liria	Santa Ana	472
Santa Ana	Alberto Carnevalli	419
Alberto Carnevalli	C.C. La Hechicera	696
C.C. La Hechicera	Terminal La Hechicera	750
Total	Total	4500

Para la construcción de la nueva ruta del STMTM será necesario redimensionar los anchos de calle de todo el tramo, según normas establecidas para proyectos viales



de tal índole ([Azuaje, 2006](#)); cada canal exclusivo del STMTM contará con un ancho de calzada de 3 metros, respectivamente, y, adicionalmente, los bordillos se plantean con medidas mínimas de 20 centímetros, con los cuales se pretende dividir el canal exclusivo del trolebús en ambos sentidos con los canales de transporte convencional existentes. Será necesario además reducir de la misma manera el ancho de los canales que actualmente integran el recorrido vehicular de transporte convencional, estimando para cada banda lateral un ancho de 6 metros, garantizando el desplazamiento de dos vehículos paralelamente en un mismo sentido. Es importante también establecer la dimensión del área de circulación peatonal, para lo cual se estiman aceras de 2.40 metros de ancho, garantizando el desplazamiento de 4 personas paralelamente, para un ancho total de calle de 23.20 metros.



Figura 2.7: Características de la red vial de un sistema de transporte masivo Trolebús ([Wright et al., 2010](#))

Los beneficios directos que producirá el diseño del nuevo sistema de transporte para la zona en estudio de acuerdo con [Wright et al. \(2010\)](#) se presentan en la tabla 2.4.

En tal sentido, el desarrollo de la Línea 2 es considerado como una importante propuesta de transporte masivo, con la cual se estima garantizar la continuidad en materia de movilidad urbana y desarrollo de las distintas actividades practicadas en el centro y periferias al norte de la ciudad ([Azuaje, 2006](#)).

Tabla 2.4: Los beneficios de las iniciativas de proyectos de transporte público

Factor	Impactos / indicadores
Beneficio de ahorro de tiempo para los usuarios del transporte	Productividad laboral y calidad de vida
Beneficio de ahorro de tiempo para los vehículos de tráfico mixto	Productividad laboral y eficiencia en la entrega de bienes y servicios
Ahorro de combustible en el transporte	Reducción en gastos de combustible para los operadores de transporte público y reducción en la dependencia de combustible importado o uso reducido del suministro doméstico
Reducciones en el ruido y las vibraciones	Salud humana, ambiente construido y productividad laboral y educativa
Empleo en el sistema de transporte	Empleo en construcción y empleo operacional
Sector comercial	Valor de las propiedades
Beneficios sociales	Reducciones en el crimen e igualdad para los grupos de bajos ingresos
Imagen de la ciudad	Turismo

## 2.2 Modelo de demanda de transporte Centro - La Hechicera

En esta sección se describe el modelo de demanda propuesto; como el mismo está referido solamente al transporte público, se permitió también evaluar fácilmente la demanda del STMTM para diferentes escenarios propuestos (capítulo 5). Para construir el modelo de demanda de transporte es preciso referirnos a los trabajos realizados por Puebla et al. (2008), donde emprenden las herramientas de planificación del transporte desde una perspectiva espacial y de esta manera logran estudiar fácilmente el comportamiento de la demanda y su interacción con el sistema de actividades del territorio sobre la que se emplaza. En este marco de ideas, se propuso para este trabajo un modelo de estimación que se alimente de la información disponible de uso de suelos y las variables urbanísticas que caracterizan la red de transporte

Centro - La Hechicera, un modelo donde además se analice su relación con la demanda de transporte público en la zona, tomando como base el algoritmo de las cuatro etapas en sus dos primeras fases: Generación y Distribución (figura 2.8).

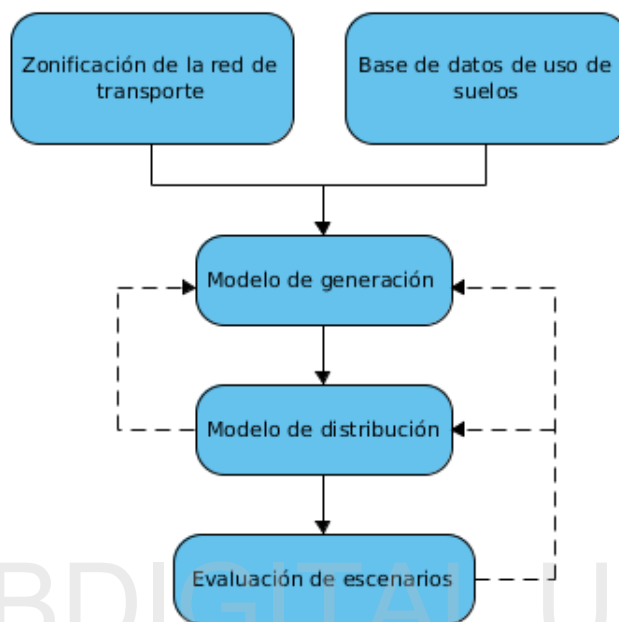


Figura 2.8: Propuesta de modelo de demanda de transporte para la red Centro - La Hechicera

En los últimos años, se ha dedicado mucho esfuerzo a construir complejos modelos de demanda de cuatro etapas (que se basan en numerosas encuestas y que asumen una perspectiva regional de la movilidad), pero existen muy pocos trabajos que logran explicar la movilidad de los orígenes y demandas de viajes de una red de transporte por medio de las características urbanas y uso de suelos (Cardozo et al., 2010). El enfoque anterior brinda la base para que los planificadores urbanos cuestionen la forma de hacer ciudad en Mérida, donde son notables, con respecto a Azuaje (2006), los graves problemas de congestionamiento de tránsito vehicular que ocurren a largo del día en el área metropolitana, que manifiesta una circulación inflexible sobre pocas calles longitudinales y donde coinciden sobre una misma zona diversas actividades de carácter comercial, financiero y de servicios, convirtiéndolas en áreas susceptibles a altas concentración de personas. Como se puede observar en la figura 2.8, las entradas del

modelo de demanda corresponden a la información de la base de datos que describe las características de uso de suelos de la red, siendo el uso de suelos la distribución espacial del sistema de actividades. El software QGIS versión 2.18.8 (un SIG libre y de código abierto) fue la herramienta esencial utilizada en este trabajo para generar, almacenar y procesar dichas entradas. La ciudad de Mérida se encuentra en la jurisdicción del municipio Libertador del estado Mérida y, a efectos de este estudio, queda excluida gran parte del área metropolitana para ser consideradas sólo las zonas entorno a la red de transporte Centro - La Hechicera (figura 2.9).

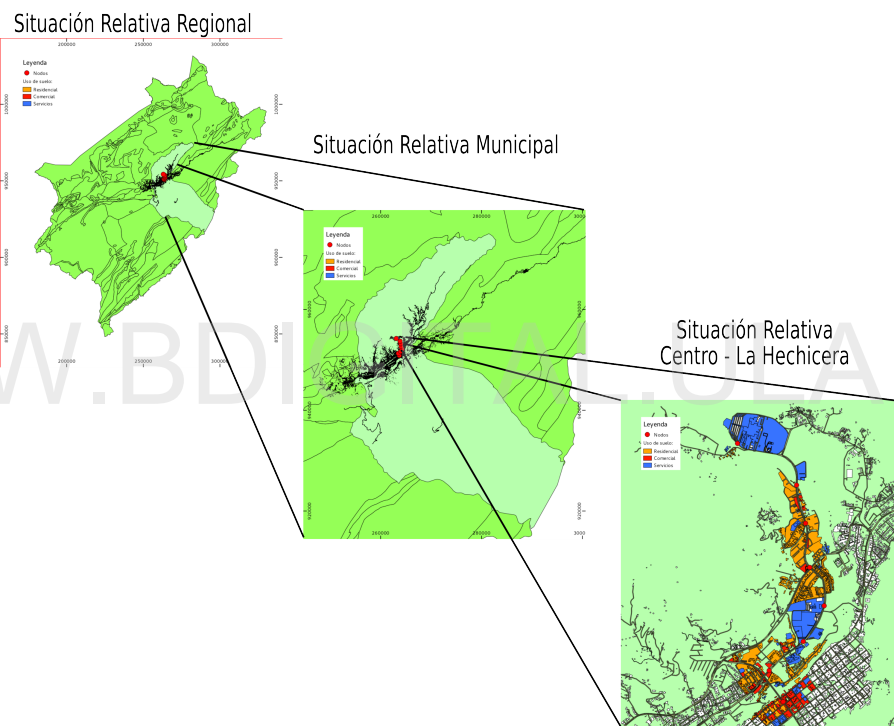


Figura 2.9: Situación relativa del área de estudio Centro - La Hechicera

El material cartográfico necesario para construir la base de datos para el modelo de demanda fueron los archivos vectoriales del plano de la ciudad de Mérida escala 1: 10.000, bases de datos cartográfico - digital de la red vial del municipio Libertador y en su defecto el plano análogo de Uso Urbano para la ciudad de Mérida 2010 (figura 2.10) encontrado en [Hernández y Molina \(2011\)](#), los cuales sirvieron como base para construir la cartografía digital inherente a los usos del suelo de la red Centro - La Hechicera. Dichos archivos en formato analógicos y digitales fueron suministrados por la Unidad

Mapoteca - Fototeca del Instituto de Geografía de la Universidad de Los Andes.

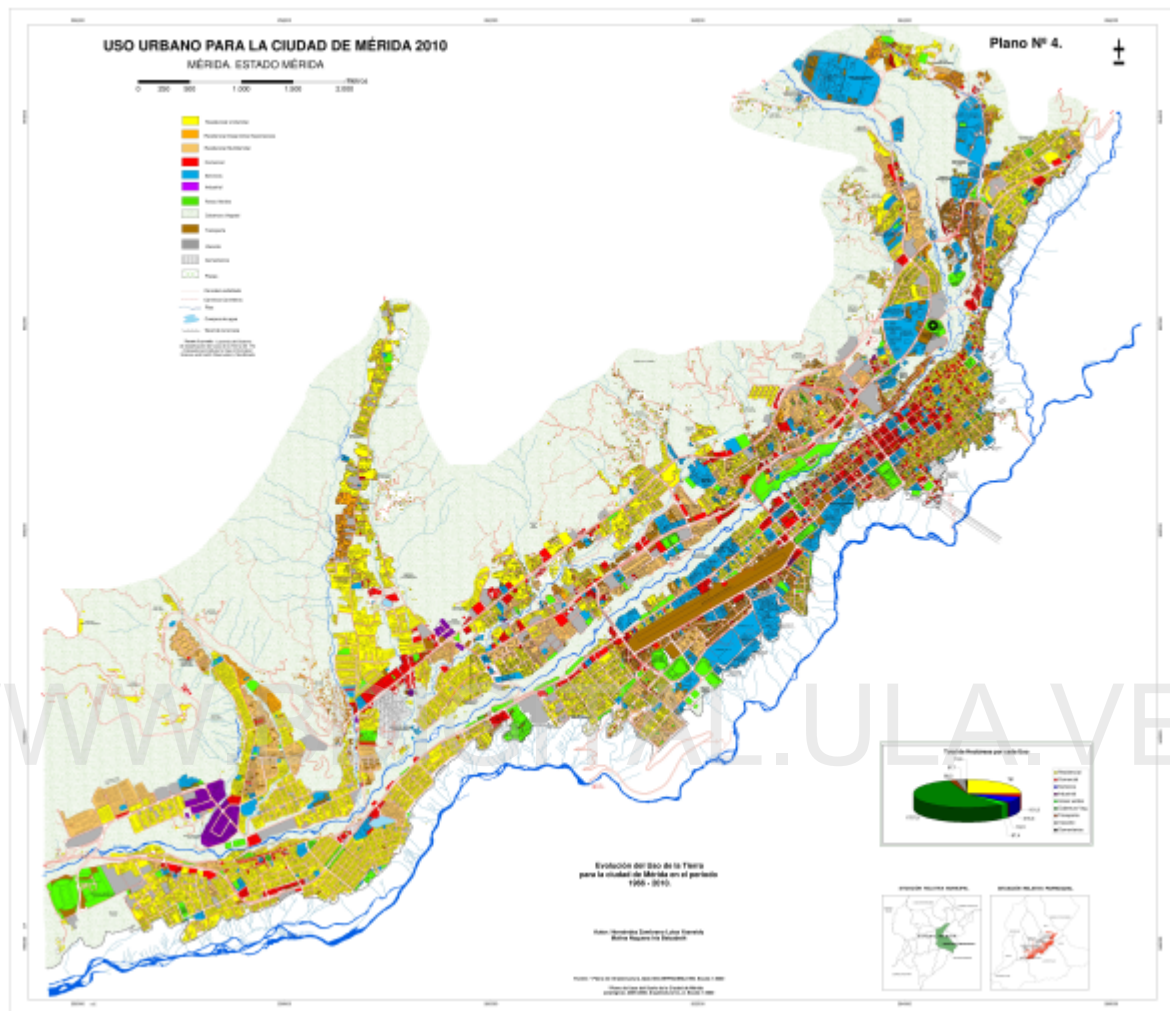


Figura 2.10: Plano de Uso Urbano de la ciudad de Mérida 2010 ([Hernández y Molina, 2011](#))

El uso de la tierra puede entenderse como la utilización efectiva que se le da a un determinado espacio o territorio. Respecto a este concepto, se puede hacer una primera clasificación en función del ámbito geográfico en donde se desarrolle: uso urbano y uso rural, siendo el uso urbano según Gelves tal como se cita en [Hernández y Molina \(2011\)](#), el que se caracteriza por el desarrollo de actividades típicas y de funciones urbanas, por ejemplo: comercial, de servicios y residencial. La leyenda del plano de la figura 2.10 es una referencia al sistema de clasificación del uso de la tierra (figura 2.11) que sugiere

el ITC (International Institute for Aerospace Survey and Earth Sciences).



Figura 2.11: Leyenda empleada en la cartografía temática de uso de suelos

La selección de los colores en las leyendas estuvo sujeta a convencionalismos cartográficos (Hernández y Molina, 2011). Igualmente es importante para analizar la red de transporte, estudiar a profundidad las características y espacios urbanos (tabla 2.5) que corresponden a cada uso o leyenda presentada en la figura 2.11.

WWW.BDIGITAL.ULA.VE

Tabla 2.5: Usos urbanos ([Hernández y Molina, 2011](#))

Uso	Características
Áreas vacantes	Comprende todos aquellos espacios no desarrollados o que han sido previamente desarrollados y que actualmente se encuentran sin uso en las márgenes de las áreas urbanas o dentro de ellas. No incluye los terrenos con uso agrícola y forestal evidente
Áreas verdes	Espacios ocupados por vegetación dentro de áreas urbanas y que por lo general son utilizados como lugares de esparcimiento. Se consideran parques, plazas y centro deportivos
Cementerios	Lugares que, bajo distintos criterios sanitarios, son aptos para la ubicación final de restos humanos
Comercial	Son edificaciones o inmuebles destinados única y exclusivamente a la actividad comercial, donde el producto no es objeto de transformaciones. Son incluidos los comercios mayoristas y minoristas. Los primeros generalmente son depósitos de víveres destinados a la venta al por mayor, en tanto que los segundos son edificaciones o locales destinados a la venta de artículos al por menor
Cobertura vegetal	Es definida como aquella capa de vegetación natural que se encuentra distribuida en el ámbito urbano Genéricamente comprende todas las biomásas con diferentes características fisionómicas: pastizales y bosques naturales
Industrial	Abarca aquellos espacios destinados a la actividad industrial en cualquiera de sus dimensiones locales: pequeña y mediana
Residencial	Espacios que albergan al conjunto de infraestructuras con una finalidad habitacional: residencial unifamiliar, residencial multifamiliar y desarrollos espontáneos
Servicios	Ajustándose al criterio de Angulo (1988) la categoría abarca, en conjunto, las áreas cuyos espacios físicos son destinados a satisfacer necesidades específicas relacionadas a: servicios gubernamentales, educación (en todos sus niveles), atención médico asistencial, diversión, servicios personales y del hogar, transporte y comunicaciones, entidades financieras, oficinas privadas y servicios al comercio, atención técnica especializada
Transporte	Comprende aeropuertos, terminales de transporte terrestre y estacionamientos



## 2.3 Zonificación de la red de transporte

Definir el área de estudio es el primer paso dentro de la planificación de transporte (Wright et al., 2010). En un sentido amplio, indica la división geográfica de la red Centro - La Hechicera en cierto número de nodos, para posteriormente construir el modelo de generación de viajes. Todos los datos de origen y destino van a ser muestreados y codificados según este sistema de zonificación. La red de transporte fue utilizada como una estructura sencilla, pero abstracta, donde los nodos estarán asociados a objetos geográficos de la ciudad (uso de suelos, paradas, terrenos, edificios) y, de manera similar, los arcos que conectan a los nodos representarán las carreteras, avenidas, entre otros. En la metodología de las cuatro etapas, estas zonas están basadas en sectores censales o subdivisiones políticas (parroquias, sectores, entre otros) que han sido usadas como base de información censal existente. Tal como se observa en la figura 2.12, la red de transporte recorre 4 de las principales parroquias del municipio Libertador, siendo estas: Antonio Spinetti Dini, Milla, Sagrario y El Llano.

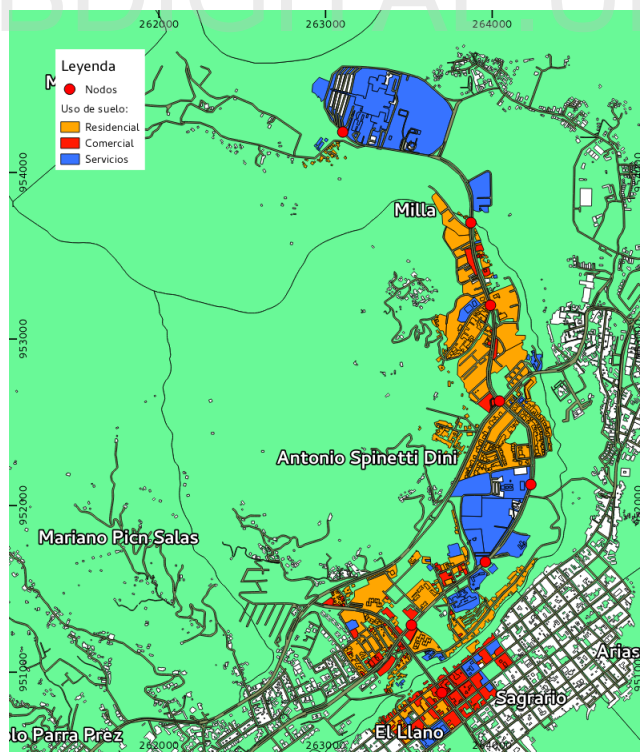


Figura 2.12: División político territorial de la red de transporte Centro - La Hechicera



### 2.3.0.1 Área de cobertura de transporte

La elección de la distancia para delimitar el área de influencia de un nodo de transporte es una cuestión relevante en cualquier estudio del uso del transporte público y debe ser justificada en términos de captación de demanda (Puebla et al., 2008). El paso previo antes de dividir en nodos la red de transporte es establecer el área de cobertura del transporte público, es decir, las zonas geográficas de usos urbanos adyacentes a la red de transporte que influyen sobre la demanda de viajes en dicha red. Tal como se observa en la figura 2.13, un umbral de distancia elevado considera áreas lejanas al nodo, que aportan pocos usuarios y pueden distorsionar los resultados finales. Mientras que un umbral muy bajo puede dejar fuera una gran parte de los usuarios del transporte público. Sin embargo, es común que en las investigaciones, para delimitar zonas de servicio de una red de transporte, se acepte una distancia estándar teórica que con respecto a Gutiérrez tal como citó en Puebla et al. (2008), corresponde desde 300 a 600 metros. Estos umbrales representan la máxima distancia que la mayoría de la población está dispuesta recorrer a pie para usar el transporte público.

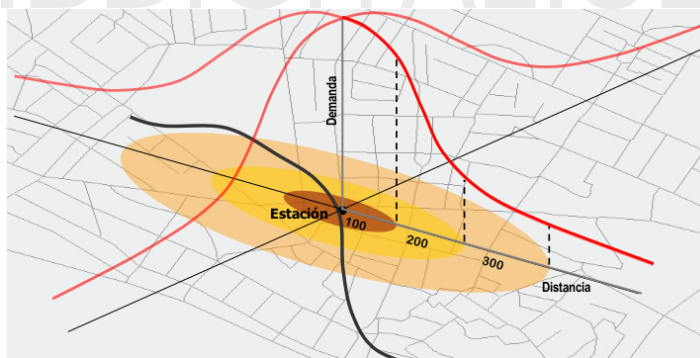


Figura 2.13: Área de influencia de transporte público (Gutiérrez et al., 2013)

Dado que el valor de los umbrales de influencia (300 a 600 metros) de la red de transporte no abarcan en su totalidad la demografía de cada una de las parroquias involucradas, se determinó que realizar una subdivisión política de nodos para la red de transporte en base a las parroquias que la conforman no sería lo adecuado; por lo tanto, para obtener una mejor captación de la demanda del sistema de transporte público con las características del espacio en particular, se dividió la red de transporte en un total

de 8 nodos que corresponden a los puntos de concentración de usuarios de transporte público mas importantes de la red, es decir, las áreas de embarque y desembarque (paradas de autobús de transporte público) que presentan mayor demanda debido al uso urbano del entorno de cada parada. Cabe aclarar que para la representación de los nodos en la red fue necesario realizar una simplificación de las paradas reales del transporte público, de manera que dos paradas existentes en ambos sentidos de la calle constituyen un solo nodo en el red. A partir de esta información se procedió a la construcción del grafo o red de transporte (figura 2.14) asignando un identificador a cada uno de los nodos que representan las diferentes paradas de transporte público y las 14 aristas que los conectan.

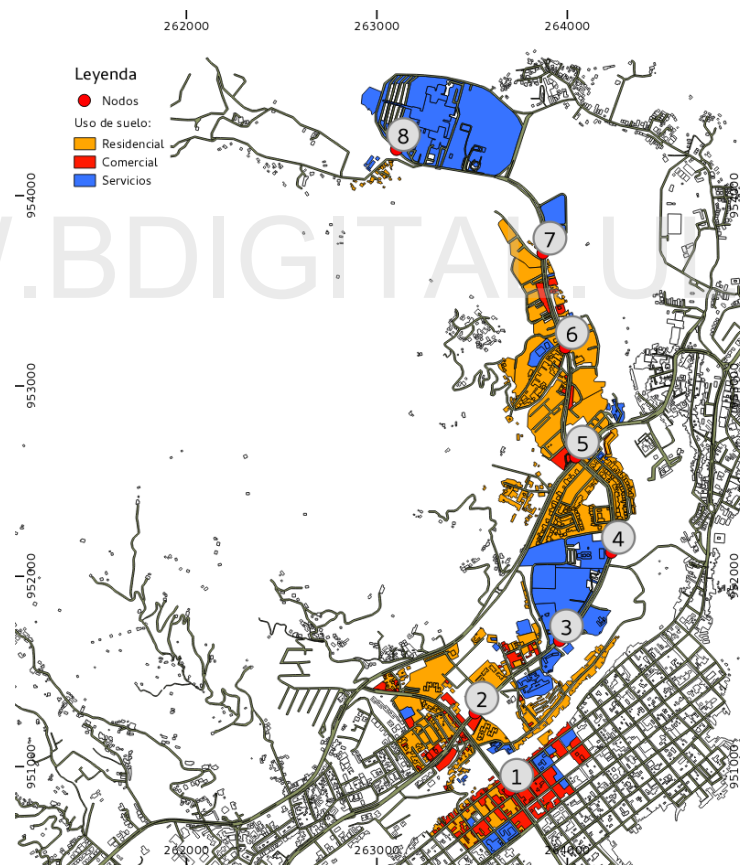


Figura 2.14: División en nodos de la red Centro - La Hechicera

La selección estratégica de los nodos para llevar a cabo el modelo y la medición de viajes va a determinar la medida en la que los resultados de la medición representan

la situación actual. Cabe mencionar además que cada uno de los nodos o paradas estudiadas se encuentran próximos a las futuras estaciones de la Línea 2 del STMTM, lo que permitirá así predecir fácilmente la demanda que presentará cada una de ellas. Para un contexto más amplio se presenta en la tabla 2.6 la descripción geográfica de cada zona o nodo de transporte asignado para la red Centro - La Hechicera presentados en la figura 2.14, junto con los espacios y características urbanas más importantes que definen su entorno.

Tabla 2.6: Identificación de los nodos y sus características geográficas y urbanas. Sistema de referenciación: WGS 84/UTM zona 18S

Nodo	Parada	Coord. <i>x</i>	Coord. <i>y</i>	Principales espacios de uso urbano
1	Centro	263609.0	950867.8	Centros comerciales, comercio minorista, Facultad de Odontología, Rectorado Administrativo, Gobernación del Estado, Liceo Libertador
2	Yuan Lin	263514.1	951281.8	CC Yuan Lin, CC Mamayeya, CC Viaducto; Residencias: Las Marias, Viaducto, Cardenal Quintero. Comercios minoristas. Barrio Santo Domingo
3	Faces	263958.7	951660.6	Hospital Sor Juana Inés, Barrio Pueblo Nuevo, UNEFA, UNA, Faces ULA, Sector Juan Bautista, UE Emilio Fuenmayor
4	Facijup	264231.3	952126.3	Facultad de Ciencias Políticas y Jurídicas, Plaza de Toros
5	Santa Ana Sur	264043.8	952627.6	Clínica Albarregas, Clínica Ocular, Urbanización Santa Ana Sur, Residencias Albarregas
6	Santa Ana Norte	263988.9	953203.0	Residencias Domingo Salazar, Urbanización Santa Ana Norte, UE Alberto Carnevalli
7	CC La Hechicera	263871.8	953698.5	CC La Hechicera, Residencias La Hechicera, Urbanización La Montaña, Fundacite
8	Núcleo Pedro Rincón Gutierrez	263102.2	954243.1	Facultades Ingeniería, Ciencias y Arquitectura

## 2.4 Modelo de datos

El modelo de datos de un SIG no es más que un modelo del mundo real, una representación digital en base a objetos discretos (Puebla y Gould, 1994). En esta sección se modelaron las variables de la zona Centro - La Hechicera, por medio de su georreferenciación en una base de datos espacial, que luego servirá como insumo de entrada al modelo generación de viajes, tal como se detalla en la figura 2.8.

Principalmente, para cada nodo se almacenaron los registros de viajes de transporte público (orígenes y demandas), se representa además en la base de datos la relación entre cada entidad nodo con los usos urbanos que comprenden su área de influencia (figura 2.14), especialmente para las entidades de uso de suelos tipo residencial, comercial y servicios, que son representados en el SIG por medio de las etiquetas de color amarillo, rojo y azul respectivamente. En la figura 2.15, se muestra el modelo entidad - relación de la red de transporte en estudio, así como cada uno de los atributos (variables urbanísticas) a ser tomados en cuenta para alimentar el modelo de la demanda.

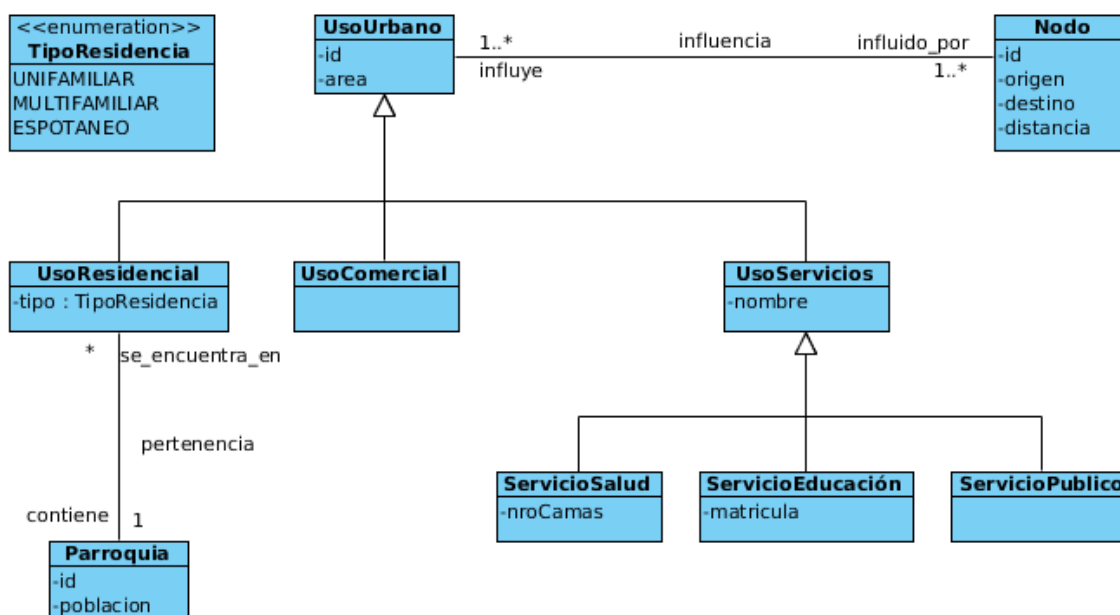


Figura 2.15: Modelo de datos de la red Centro- La Hechicera

Estos atributos o variables urbanas serán utilizadas para predecir y entender el comportamiento de la demanda de viajes en cada nodo propuesto. Los planificadores de transporte suelen emplear un conjunto de variables relevantes que describen a los usos de suelos, como el número de matrícula estudiantil, el número de habitantes, la posesión de automóvil, entre otros (Cardozo et al., 2010). Las principales variables urbanísticas consideradas en este trabajo para modelar la demanda son: el área medida en ( $m^2$ ) para los distintos usos de suelos: *Residencial*, *Comercial* y de *Servicios*, asimismo para cada uno de estos usos o entidades se determinaron las principales propiedades que los definen: *número de camas de hospital* en los usos de servicios de salud, *el número de matrícula estudiantil* para los usos de servicios educativos, e igualmente se calcularon y almacenaron las distancias en metros ( $m$ ) desde el centroide de cada nodo al centro de la ciudad, con el propósito de estudiar la influencia de dicha variable con la dinámica de viajes.

WWW.BDIGITAL.ULA.VE

# Capítulo 3

## Modelo de Generación

El modelo de generación es el primer paso en el proceso de planificación del transporte, tiene como propósito determinar y predecir el número total de viajes generados ( $O_i$ ) y demandados ( $D_j$ ) por cada nodo en un día típico de la semana, partiendo de las características urbanas de la red de transporte. Estos viajes pueden expresarse en viajes de personas o viajes de vehículos. Con respecto a los objetivos perseguidos en este trabajo se realizó la estimación del modelo de generación para viajes de personas, específicamente viajes en transporte público.

### 3.1 Sistemas de información de transporte

Según [Quintero et al. \(2011\)](#), los datos de generación de viajes han permitido determinar los requerimientos de acceso a usos específicos del suelo, estimar volúmenes de tráfico futuros con los cuales sustentar posibles mejoras a los sistemas de transporte y evaluar los impactos producto de cambios de usos del suelo. El *Institute of Transportation Engineers* (ITE) es el sistema de información de transporte mas utilizado en los Estados Unidos para estimar los viajes en una determinada zona; para ello, el ITE ha desarrollado el *Trip Generation*, publicación que reúne tasas de generación de viajes para distintos usos del suelo obtenidos de diversos estudios realizados principalmente en Norte America. De acuerdo con [Martin y McGuckin \(1998\)](#), en el formato del manual de viajes del ITE, las tasas se presentan como

viajes diarios promedio por vehículo en base a factores específicos del sitio de estudio como número de empleados, pies cuadrados de superficie bruta, unidades de viviendas, número de matrícula estudiantil, número de camas de centros de salud, entre otros. Estas tasas de viajes han sido usadas en diversos estudios de transporte en Latinoamérica y España, pero lamentablemente se evidencia en Rosas (2012), la gran diferencia existente con respecto a la morfología urbana de las ciudades venezolanas con las norteamericanas. Con base en lo anterior, es preciso referirse al trabajo de Rosas (2012), donde se comparan las tasas estimadas de generación de viajes para las zonas de hospitales en el Distrito Metropolitano de Caracas (DMC) con las producidas mediante el manual del ITE; los resultados obtenidos sobre el comportamiento de la demanda de transporte en Venezuela son marcadamente diferentes a los EEUU, es decir, las tasas de generación de viajes calculadas en el DMC arrojaron valores mucho mayores a las tasas de generación de viajes para hospitales del ITE, con una diferencia de hasta tres veces mayor cantidad de viajes en Venezuela. Esto podría sugerir la necesidad de levantar nuestros propios datos para determinar la demanda adaptada a la realidad local, ya que las diferencias de las características culturales, socioeconómicas, políticas y ambientales modifican los patrones de conducta de la población.

### 3.2 Formulación del modelo de origen y atracción de viajes

Dentro del modelo de generación de viajes es necesario estimar dos submodelos, denominados:

1. **Modelo de origen** ( $O_i$ ) : cantidad de viajes que se originan en el nodo  $i$ ,  $\forall i = 1, 2, 3, \dots, 8$
2. **Modelo de atracción** ( $D_j$ ) : cantidad de viajes que se destinan en cada nodo  $j$ ,  $\forall j = 1, 2, 3, \dots, 8$

Con el propósito de obtener una evaluación rápida y sencilla del modelo de generación, se ajustaron dos modelos de regresión lineal múltiple (MRLM) con ayuda

del SIG de la red de transporte, que contiene la base de datos de las variables explicativas: metros cuadrados de uso de suelos de tipo residencial, comercial y de servicios, matrículas estudiantiles, número de camas de hospital, entre otras, siendo éstas las características que definen la estructura urbana de la red Centro - La Hechicera. Profundizar en estos modelos permitirá indagar el papel que juegan las variables urbanas con la demanda de transporte, tal como se fundamenta en [Cardozo et al. \(2010\)](#), donde no sólo usan los modelos de regresión para estimar la demanda sino también valorar y evaluar el papel de los usos de suelos en la movilidad del transporte público de una ciudad. Como resumen, los MRLM tuvieron como objeto explicar el comportamiento de las variables dependientes ( $O_i$  y  $D_j$ ), mediante el conjunto de variables regresoras ( $x_i$ ) dadas las siguientes expresiones matemáticas:

$$O_i = f(x_1, x_2, \dots, x_5) \quad \forall i = 1, 2, 3, \dots, 8$$

$$D_j = f(x_1, x_2, \dots, x_5) \quad \forall j = 1, 2, 3, \dots, 8$$

Siendo las  $x_i$  :

$x_1$  :  $m^2$  de uso residencial del área de influencia  $\forall i = 1, 2, 3, \dots, 8$

$x_2$  :  $m^2$  de uso comercial del área de influencia  $\forall i = 1, 2, 3, \dots, 8$

$x_3$  :  $m^2$  de uso de servicios del área de influencia  $\forall i = 1, 2, 3, \dots, 8$

$x_4$  : total de matrículas estudiantiles comprendidas dentro del área de influencia  $\forall i = 1, 2, 3, \dots, 8$

$x_5$  : distancia en  $m$  al centro de la ciudad  $\forall i = 1, 2, 3, \dots, 8$

Las anteriores  $x_i$  fueron consideradas principalmente porque las mismas permiten ser proyectadas fácilmente, lo que contribuirá a estudiar la demanda y las acciones necesarias sobre la red de transporte en el futuro.

### 3.3 Recolección de la información

Fue necesario para este trabajo recolectar dos tipos de información, la primera en relación con los viajes originados y demandados a través de la red de transporte Centro



- La Hechicera, y la segunda, referentes a las variables urbanísticas comentadas en el apartado anterior. Para ajustar los modelos de regresión fue necesario muestrear y georreferenciar en diferentes capas de un formato SIG la siguiente información:

1. **Muestreo de movilidad de pasajeros:** se midió la información relativa a los viajes de transporte público originados y demandados (*variable dependiente*) en cada uno de los 8 nodos que conforman la red de transporte; dicho muestreo fue realizado durante los meses de Marzo y Abril de 2017 por el Servicio Comunitario de la Empresa Tromerca. Estos registros consisten en la medición desde las 7:00 AM a 6:00 PM en intervalos de 10 minutos de conteos manuales de subidas y bajadas de pasajeros en cada nodo de transporte, específicamente para los días Lunes y Miércoles. El resumen de los viajes totales diarios (7:00 AM a 6:00 PM) cuantificados para cada nodo se presenta en la tabla 3.1.

Tabla 3.1: Total de viajes diarios originados y atraídos por nodo de transporte

Nodo	Viajes de origen	Viajes de destino
1	607	1278
2	1848	1355
3	2329	3205
4	570	1356
5	1320	697
6	556	483
7	187	393
8	2254	3049

Para el total de 8 nodos de transporte, solo la red vial de un nodo, específicamente el nodo 8 (Núcleo Pedro Rincón Gutiérrez) es simple, es decir, los vehículos de transporte circulan por un solo canal compartido; para el resto de los nodos, el tráfico circula en ambos sentidos por canales distintos por lo que fue necesario muestrear un total de 16 paradas de uso de transporte público. Como se puede observar en los registros de viajes, se visualiza una movilidad total de casi 19500 viajes, siendo los nodos 3 : (Faces) y 8 : (Núcleo Pedro Rincón Gutiérrez), los que presentan mayor demanda con 28% y 27% del total de viajes de la red de

transporte. La visualización por medio del SIG de los resultados del muestreo de movilidad de pasajeros realizado en la red del transporte, facilitó el análisis espacial de la demanda (figura 3.1). En el apéndice C, se presenta un resumen estadístico detallado del muestreo, así como la distribución ajustada a los datos de viajes en cada nodo.

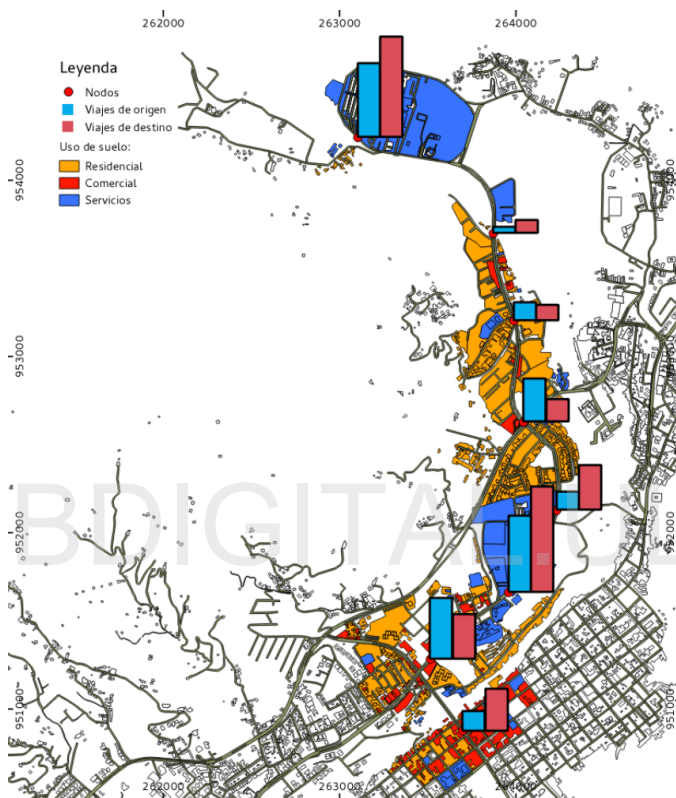


Figura 3.1: Visualización geográfica del conteo de viajes

2. **Registro de variables de planificación:** En el caso de las variables independientes se computaron los  $m^2$  correspondientes a cada uso de suelos por medio del SIG; se consideró además la información disponible de matrículas estudiantiles y número de camas correspondientes a cada institución educativa y de salud, respectivamente, que conforman el entorno de influencia de la red de transporte, algunos de los cuales se resumen en la tabla 3.2.

Aproximadamente 28000 estudiantes hacen vida en el entorno de la red Centro - La Hechicera, con lo que se confirma que esta red abastece la mayor población estudiantil

Tabla 3.2: Datos de matrícula estudiantil (ULA, 2016) y camas de hospitalización (INE, 2011)

Institución Educativa	Matrícula
Facultad Ciencias Económicas y Sociales	5160
Facultad de Humanidades y Educación	5145
Facultad Ciencias Jurídicas y Políticas	4869
Facultad de Ingeniería	4800
Facultad de Ciencias	1440
Facultad de Arquitectura y Diseño	1434
Faculta de Odontología	925
Universidad Nacional Experimental de las Fuerzas Armadas	1831
Unidad Educativa Alberto Carnevalli	435
Unidad Educativa Emilio Fuenmayor	315
Liceo Bolivariano Libertador	869
Institución de Salud	Número de camas
Hospital Sor Juana Ines	86
Clínica Albarregas	12

del área metropolitana (Azuafe, 2006), siendo el Nucleo Universitario La Liria y el Núcleo Universtiaro Pedro Rincón Gutierrez, las instituciones educativas que cuentan con mayores matrículas, por lo tanto se justifica que la mayor movilidad de pasajeros se visualice en los nodos 3 (Faces) y el nodo 8 (Núcleo Pedro Rincón Gutiérrez).

### 3.4 Metodología del análisis geográfico

Después de almanenar en la base de datos del SIG de la red Centro - La Hechicera la información de las variables con las que se tratará de explicar la demanda de transporte, primeramente se debe considerar cómo analizar espacialmente las zonas entorno a cada nodo: una primera tarea fue establecer el umbral de distancia donde influye el transporte público, la delimitación de esta área ya fue introducida en el capítulo 2, específicamente en la sección 2.3. A través de la herramienta *Buffer* (zonas de amortiguamiento) de QGIS y la superposicion de capas dentro del SIG se facilitó la tarea de calcular las respectivas variables urbanas (*variables independientes*) que

se encuentran dentro de un *buffer* de cobertura establecido para este trabajo de 300 metros desde el nodo hacia su entorno, en base a la metodología de trabajo llevado a cabo en [Cardozo et al. \(2010\)](#), donde se establece que bajo un comportamiento espacial lógico, en su recorrido hasta la parada del transporte el peatón seguirá el trazado de la red vial, por lo que considerar la métrica euclidiana para medir esta área de influencia constituye una aproximación poco realista. Por lo tanto, en este trabajo se optó por establecer el *buffer* midiendo la distancia sobre la red (figura 3.2) (denominada distancia de Manhattan) y no en base a la distancia euclideana, con el objetivo de obtener una mejor estimación. Cabe mencionar que fue establecido este umbral, puesto que valores superiores collevaban a un mayor solapamiento entre las zonas de influencia de cada nodo, esto debido a que la red de transporte sólo cuenta con 4.5 *km* de recorrido, además los MRLM se ajustaron significativamente para el umbral de 300 *m*.

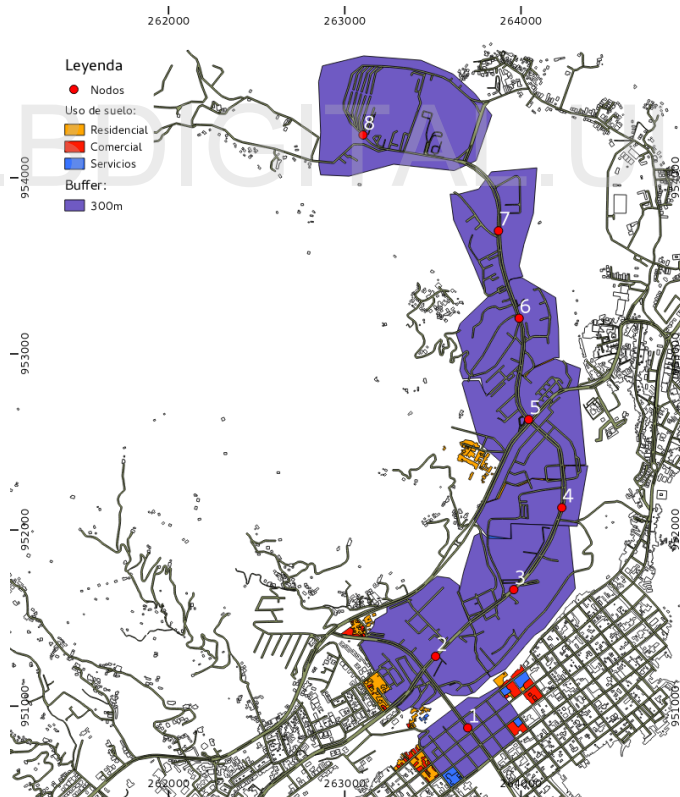


Figura 3.2: Buffer de influencia (300 *m*) para cada nodo de transporte

## 3.5 Ajustes de los modelos de regresión lineal múltiple

Como fue mencionado en la sección 3.2, el objetivo del modelo de generación es formular dos MRLM que permitan una predicción inmediata de la demanda de transporte, a partir de un conjunto de variables explicativas de las que se pueda tener información futura de manera sencilla, denominados estos dos submodelos como *modelos macro de demanda*. Puesto que se consideró el periodo usual de análisis en la planificación de transporte, siendo este de 1 día, en nuestro caso la predicción de la cantidad de viajes totales será desde las 7:00 AM a 6:00 PM. Se pretende reducir el número de variables independientes en base a su contribución a los modelos; para realizar esta tarea, se ha empleado el software QGIS para consultar la información de la base de datos espacial de la red y el lenguaje de programación estadístico R versión 3.3.3, para procesar y facilitar el análisis de datos.

### 3.5.1 Análisis de datos

El primer paso del análisis consistió en estandarizar y normalizar los datos obtenidos por medio del SIG, lo que permitió comparar adecuadamente los coeficientes; en este caso es evidente, puesto que las variables explicativas poseen unidades de medida diferentes, producto de que las mismas proceden de diversas fuentes y formas de captura. Para tener una idea de la relación que existe entre las variables independientes  $x_i$ , una vez tipificadas con la cantidad de viajes  $O_i$  y  $D_j$ , los datos obtenidos fueron graficados en el diagrama de dispersión de la figura 3.3, donde se refleja cierta relación no lineal entre los pares de variables  $(x_i, O_i)$  y  $(x_i, D_i)$ , lo que fundamenta la hipótesis de que los datos podrían ajustarse igualmente a un modelo no lineal. Además se puede concluir que cada par  $(x_i, O_i)$  y  $(x_i, D_i)$  se encuentran altamente correlacionados. Lo anterior se confirma observando los coeficientes de correlación bivariada *pearson* mostrados en la tabla 3.3 y tabla 3.4, respectivamente, para cada modelo (origen y atracción), donde la variable de uso de suelos  $x_2$  ( $m^2$  uso de servicios), presenta alta correlación positiva con respecto al fenómeno de transporte, con coeficientes por encima

de 60% para los viajes de origen y 80% para las atracciones de viajes. Igualmente, entre las variables urbanísticas destaca la relación que tiene la variable  $x_4$  (*matrículas estudiantiles*) con los viajes que se originan y atraen en la red, presentando coeficientes mayores a 67% y 92%, respectivamente. Cabe mencionar que todos los valores ( $r$ ) tabulados en las tablas 3.3 y 3.4, son considerablemente significativos (*coeficientes por encima de 0.5*), si tenemos en cuenta los numerosos factores que entran en juego en la dinámica de la demanda de transporte.

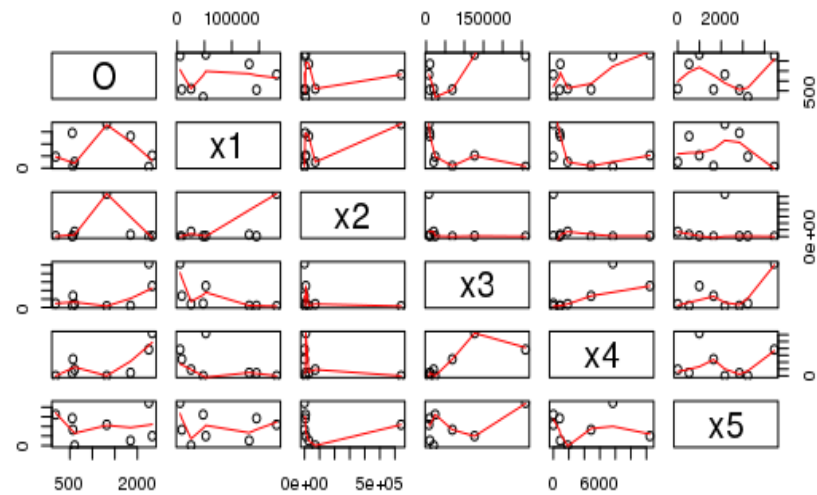
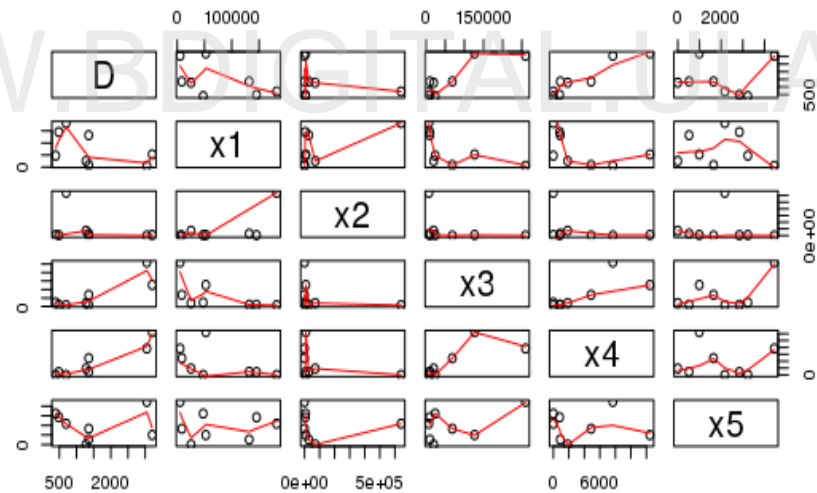
Tabla 3.3: Matriz de correlación *pearson* ( $x_i, O_i$ )

	$O_i$	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
$O_i$	1.00000000	0.01154156	0.03388558	0.6375743	0.67012002	0.01436294
$x_1$	0.01154156	1.00000000	0.63007394	-0.5907565	-0.52179754	-0.09262333
$x_2$	0.03388558	0.63007394	1.00000000	-0.3074456	-0.35195105	-0.02401281
$x_3$	0.63757429	-0.59075647	-0.30744557	1.00000000	0.74976926	0.51747491
$x_4$	0.67012002	-0.52179754	-0.35195105	0.7497693	1.00000000	0.01560783
$x_5$	0.01436294	-0.09262333	-0.02401281	0.5174749	0.01560783	1.00000000

Tabla 3.4: Matriz de correlación *pearson* ( $x_i, D_i$ )

	$D_i$	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
$D_i$	1.00000000	-0.50182043	-0.30539505	0.8486525	0.92822537	0.03780274
$x_1$	-0.50182043	1.00000000	0.63007394	-0.5907565	-0.52179754	-0.09262333
$x_2$	-0.30539505	0.63007394	1.00000000	-0.3074456	-0.35195105	-0.02401281
$x_3$	0.84865248	-0.59075647	-0.30744557	1.00000000	0.74976926	0.51747491
$x_4$	0.92822537	-0.52179754	-0.35195105	0.7497693	1.00000000	0.01560783
$x_5$	0.03780274	-0.09262333	-0.02401281	0.5174749	0.01560783	1.00000000

Las matrices de correlaciones bivariadas sirvieron además para reconocer problemas de multicolinealidad entre las variables antes de ser introducidas a los modelos de regresión, donde los valores más altos se observaron entre las variables  $x_3$  (*m<sup>2</sup> de uso de tipo servicios*) y  $x_4$  (*matrículas estudiantiles*), con 0.7497693. Para esta ocasión

(a) Diagrama de dispersión  $(x_i, O_i)$ (b) Diagrama de dispersión  $(x_i, D_j)$ Figura 3.3: Diagramas de dispersión  $(y \text{ vs } x_i)$ 

todas las variables independientes están debajo del nivel de riesgo de 0.8, esto según Clark y Hosking tal como se citó en [Gutiérrez et al. \(2013\)](#). Para obtener mejores resultados en el ajuste de los submodelos de regresión (orígenes y atracciones) de viajes



se utilizó el criterio de selección de variables *backward* (Pulido y De la Vara Salazar, 2008); por lo general, esta técnica de selección de variables tiene una alta demanda de cálculo. R incluye la función (*step*) para la selección de variables, que se basa en que un modelo con más predictores siempre será el mejor ajustado. Por lo tanto, se evaluaron los modelos en base al criterio de información de Akaike (*AIC*) (apéndice D.3), siendo este criterio una medida de ajuste que penaliza un modelo por tener más variables. A partir de lo anterior fueron obtenidos dos modelos sencillos pero con un alto poder explicativo que incorporan 3 y 2 variables independientes, respectivamente, para los modelos de origen y atracción de viajes con valores de *AIC* de -3.410981 y -2.888857. Para explicar las variables de respuesta con una muy elevada significación estadística, los modelos resultantes presentados en las ecuaciones (3.1) y (3.2) fueron depurados aplicando una transformación logarítmica a los datos procedentes de los viajes de origen y destino.

$$\log O_i = 6.8109 + 0.8253x_1 + 1.3105x_3 - 0.7068x_5 \quad (3.1)$$

$$\log D_j = 7.0466 + 0.8829x_3 - 0.5805x_5 \quad (3.2)$$

Se obtuvo un total de  $k_1 = 3$  variables regresoras  $x_1$ ,  $x_3$  y  $x_5$  para el modelo de orígenes con 4 grados de libertad; de manera análoga, se obtuvieron  $k_2 = 2$  variables regresoras  $x_3$  y  $x_5$  que resultaron ser sinificativas para el modelo de atracción con 5 grados de libertad, quedando demostrado así el papel que juegan los usos de suelos y la distancia del nodo al centro de la ciudad con el transporte público de la zona Centro - La Hechicera. Quedó excluida de los modelos la variable  $x_2$ , es decir que las estructuras urbanas de la red de transporte que son de uso exclusivo para actividades comerciales, no explican el fenómeno de viajes de la zona. Igualmente, si se observan los coeficientes para los dos modelos, se puede decir que cada variable regresora posee los signos esperados, es decir, los  $m^2$  de uso de suelos residencial y servicios tienen un impacto positivo sobre los orígenes o atracciones, específicamente para la ecuación (3.2), donde 88.29 es el incremento porcentual de los viajes atraídos en un nodo por cada variación en los  $m^2$  del tipo de uso de suelos servicios, mientras que para la ecuación (3.1), se refleja un efecto negativo para los viajes originados entre la distancia



en  $m$  de los nodos al centro de la ciudad, por lo que parece observarse mayor demanda de pasajeros en los nodos centrales y una menor necesidad o tendencia de uso del transporte público en los nodos periféricos de la ciudad.

### 3.5.2 Pruebas de hipótesis

En [Montgomery \(2005\)](#) se establece que, después de obtener un modelo de regresión, es necesario aplicar ciertas pruebas de hipótesis acerca de los parámetros ajustados ( $\hat{\beta}_i$ ), con el fin de asegurarse de que los mismos proporcionen una aproximación adecuada del verdadero sistema, en este caso, la demanda de transporte. A partir de lo anterior, se presentan en esta sección de manera resumida las pruebas de hipótesis más importantes. Es fundamental definir previamente que estos procedimientos requieren la suposición de que los errores  $\varepsilon_i$  siguen una distribución normal e independiente con media  $\mu$  cero y varianza constante  $\sigma^2$ , es decir, se requiere que  $\varepsilon_i \sim NID(0, \sigma^2)$  con  $Cov(\varepsilon_i, \varepsilon_j) = 0 \forall i \neq j$ . Una consecuencia de estos supuestos es que las observaciones del fenómeno de transporte  $O_i$  y  $D_j$  siguen una distribución normal e independiente con media  $\beta_0 + \sum_{j=1}^k \beta_j x_{ij}$  y varianza  $\sigma^2$ .

#### 3.5.2.1 Prueba de significación de la regresión

Para [Pulido y De la Vara Salazar \(2008\)](#), la hipótesis global más importante sobre un MRLM consiste en ver si la regresión es significativa, es decir, se evalúa si existe realmente una relación lineal entre las variables respuesta ( $\log O_i$  y  $\log D_j$ ) y el subconjunto de regresores tipificados  $(x_1, x_3, x_5)$  y  $(x_3, x_5)$ , respectivamente. Esto se comprobó por medio de la siguiente hipótesis:

$$H_0 : \beta_j = \dots \beta_k = 0, \forall j = 1, \dots, k$$

$$H_A : \beta_j \neq 0, \text{ para al menos un } j = 1, \dots, k$$

Siendo  $k = k_1 = 3$  para el caso de los orígenes de viajes y  $k = k_2 = 2$  para las atracciones de viajes, la variable global  $k$  será usada para designar la cantidad de variables explicativas incluidas en cada modelo. Según [Montgomery \(2005\)](#), el rechazo de  $H_0$  implica que al menos una de las variables regresoras  $x_i$  contribuye de

manera significativa; esta hipótesis fue evaluada a continuación para los dos modelos de regresión (*orígenes y atracciones*): el procedimiento incluyó un análisis de varianza donde se descompone la suma total de cuadrados ( $S_T$ ) en una suma de cuadrados debido a la regresión ( $SC_R$ ) y una suma de cuadrados debido a los errores ( $SC_E$ ), es decir:  $S_T = SC_R + SC_E$ . A partir de esto, el estadístico de prueba para la significancia del modelo de regresión lineal consistió en calcular  $F_0 = \frac{SC_R/k}{SC_E/(n-k-1)} = \frac{CM_R}{CM_E}$  y en rechazar  $H_0$  si  $F_0$  excede a  $F(\alpha, k, n-k-1)$ , de esta manera, con un nivel de confianza de 95% se procede a rechazar la hipótesis nula para ambos casos (modelo de origen y atracción), puesto que en el valor de los estadístico de prueba  $F_0$  obtenidos para cada modelo (81.99 y 85.41, respectivamente) son mayores a los tabulados  $F(\alpha, k, n-k-1) : (6.59, 6.94)$ . Igualmente, el *valor-p* =  $P(F > F_0)$  asociado a cada modelo (0.0004775, 0.0001364) denota la elevada significación estadística que poseen, por lo que se concluye que al menos una variable contribuye de manera significativa en el modelo de orígenes y atracción de viajes.

### 3.5.2.2 Evaluación del coeficiente de determinación

En el diseño de experimentos, el coeficiente de determinación  $R^2$  es una medida del porcentaje de variabilidad de los datos que son explicados por el modelo. Sin embargo, un valor alto de  $R^2$  no implica necesariamente que un modelo de regresión sea adecuado, puesto que el  $R^2$  siempre se incrementará cuando se agreguen variables al modelo sin importar si esta sea significativa estadísticamente o no (Montgomery, 2005). Por lo tanto es preferible usar el valor de  $R_{aj}^2$ ; ambos estadísticos están dados por las siguientes ecuaciones:

$$R^2 = \frac{SC_R}{SC_T} = 1 - \frac{SC_E}{SC_T}$$

$$R_{aj}^2 = \frac{SC_T/(n-1) - CM_E}{SC_T/(n-1)} = \frac{CM_{total} - CM_E}{CM_{total}} = 1 - \frac{CM_E}{CM_{total}}$$

Se debe cumplir que  $0 < R_{aj}^2 \leq R^2 < 1$ ; en general, para hablar de un modelo que tiene un ajuste satisfactorio es necesario que ambos coeficientes tengan valores superiores a 0.7 (Pulido y De la Vara Salazar, 2008). La bondad del ajuste de la ecuaciones de regresión obtenidas (ecuaciones (3.1) y (3.2)) presentan coeficientes de

determinación  $R^2$  y  $R_{aj}^2$  de 0.972 y 0.984 para el modelo de origen de viajes y un  $R^2$  y  $R_{aj}^2$  de 0.962 y 0.9716 para el modelo de atracción de viajes. Por lo tanto, respectivamente un total del 97% y 96% de la variación observada para los viajes originados y atraídos en cada nodo son explicadas por los modelos. De acuerdo con Pulido y De la Vara Salazar (2008), cuando en un modelo hay términos que no contribuyen de manera significativa a éste, el  $R_{aj}^2$  tiende a ser mucho menor que el  $R^2$ ; verificando esto último se puede concluir que la calidad del ajuste es satisfactorio para ambos modelos.

### 3.5.2.3 Pruebas sobre coeficientes individuales del modelo

En Pulido y De la Vara Salazar (2008), un aspecto clave en un análisis de regresión múltiple, es valorar qué tanto contribuye cada término a la explicación de la variable respuesta y, de esta forma, proseguir a eliminar los que tienen una contribución poco importante ( $\beta_j = 0$ ). Las hipótesis para probar la significancia de cualquier coeficiente individual ( $\beta_j$ ) se definen como:

$$H_0 : \beta_j = 0$$

$$H_A : \beta_j \neq 0 \dots j = 0, 1, \dots, k$$

El estadístico de prueba considerado para esta hipótesis es  $t_0 = \frac{\hat{\beta}_j}{\sqrt{CM_E C_{j+1,j+1}}}$ , donde  $C_{j+1,j+1}$  representa el elemento de la diagonal de la matriz  $\mathbf{X}'\mathbf{X}^{-1}$  correspondiente al parámetro  $\hat{\beta}_j$  (véase capítulo 1), en este caso se procede a rechazar  $H_0$  si  $|t_0| > t_{(\alpha/2, n-k-1)}$  o equivalentemente si  $\text{valor-}p = P(t > |t_0|) < \alpha$ . En las tablas 3.5 y 3.6 se muestra el resumen obtenido mediante el software R de la prueba anteriormente descrita para cada modelo de regresión (origen y atracción).

Como se muestra en la tablas 3.5 y 3.6, los valores tabulados del estadístico  $t_0$  para cada coeficiente y su *valor-p* correspondiente, cada uno tiene un fuerte poder explicativo, de tal manera que se rechaza la hipótesis nula para todos los coeficientes de los modelos, figurando entre las más significativas (a un nivel de significación de 0.05) la variable  $x_3 : m^2$  de uso de suelo de tipo servicios para el modelo de orígenes y atracción de viajes.

Tabla 3.5: Estadístico  $t_0$  vs *valor-p* de cada coeficiente (modelo de orígenes)

Coefficientes	Error estándar	valor $t_0$	$P(t >  t_0 )$
$\beta_0$	0.05233	130.16	0.00000000209
$\beta_1$	0.07289	11.32	0.000347
$\beta_3$	0.08481	15.45	0.000102
$\beta_5$	0.06873	-10.28	0.000504

Tabla 3.6: Estadístico  $t_0$  vs *valor-p* de cada coeficiente (modelo de atracciones)

Coefficientes	Error estándar	valor $t_0$	$P(t >  t_0 )$
$\beta_0$	0.05479	128.61	0.000000000539
$\beta_3$	0.06845	12.90	0.0000499
$\beta_5$	0.06845	-8.48	0.000375

#### 3.5.2.4 Diagnóstico de los supuestos del modelo

Por último, para evaluar la validez de los ajustes y de las conclusiones obtenidas en cada una de las hipótesis anteriores, se comprobaron los supuestos de los modelos de regresión presentados en la sección 3.5.2, por medio del análisis gráfico y pruebas formales de hipótesis que son útiles en la regresión. En la tabla 3.7 se muestra la cantidad viajes de origen y destino estimados por ambos modelos, junto con los residuos asociados a cada uno de los nodos.

Realizando una comparación del total de viajes capturados en el muestreo de movilidad y los valores ajustados para cada nodo de la tabla 3.7 una vez aplicado el antilogaritmo, las observaciones estimadas presentan errores relativos medios de  $\pm 8.79\%$  para la cantidad de viajes originados y  $\pm 7.68\%$  en la cantidad de viajes atraídos. El ajuste es satisfactorio en los viajes que se originan en los nodos 1 y 5, igualmente para los viajes que se atraen en los nodos 4 y 8 y la mayor presencia de error se encuentra en los nodos 2 y 3 (para el modelo de orígenes) con errores de  $17.3\%$  y  $14.4\%$  respectivamente, y en el nodo 5 (para el modelo de atracción) donde el mismo alcanza el  $14.2\%$  de error. También se observa en la tabla 3.7 cierta aleatoriedad en la sobreestimación y subestimación de cada uno de los viajes, no detectando así la

Tabla 3.7: Valores observados vs predichos y residuos  $\varepsilon_i$  para cada modelo  $O_i$  y  $D_j$ 

Nodo $i$	$O_i$	$\hat{O}_i$	$O_i - \hat{O}_i$	$D_i$	$\hat{D}_i$	$D_i - \hat{D}_i$
1	667	647	20	1278	1572	-294
2	1848	1584	264	1355	1164	191
3	2329	2733	-404	3205	3105	100
4	570	498	72	1356	1346	10
5	1320	1282	38	697	598	99
6	556	618	-62	483	471	12
7	187	194	-7	393	465	-72
8	2254	2148	106	3049	3042	7

presencia de autocorrelación; el supuesto anterior se patentó al comprobar que el *valor-p* del test de *Durbin-Watson* (apéndice D.2) para ambos modelos  $O_i, D_j : (0.273, 0.5062)$  es superior a la significancia 0.05, claro está, que la estimación de las variables  $O_i$  y  $D_j$  será más precisa en la medida que las variables  $x_i$  se encuentren dentro del intervalo de los valores con los que se realizó el ajuste de los modelos. Es fundamental verificar igualmente el supuesto de homogeneidad de las varianzas de los residuos, puesto que si no se cumple este supuesto las pruebas estadísticas aplicadas en la sección 3.5.2 no tendrán validez (Pulido y De la Vara Salazar, 2008), evaluando los resultados del *valor-p* : (0.369, 0.2432) del test de *Breusch-Pagan* (apéndice D.4) para ambos modelos ( $O_i, D_j$ ), respectivamente; es preciso concluir que las varianzas son constantes a lo largo de las muestras. Para finalizar con el diagnóstico de los supuestos del ajuste y terminar de contrastar las conclusiones anteriores, se analizan a continuación para cada modelo en particular las figuras 3.4 y 3.5, las gráficas de los residuos en papel de probabilidad normal, las gráficas de los residuos contra los predichos y por último se visualizan los residuos contra algunas de las variables regresoras.

Con respecto a las figuras 3.4(a) y 3.5(a), para los dos modelos se verifica que el supuesto de normalidad de los residuos se cumple satisfactoriamente, debidamente lo anterior fue comprobado al evaluar los residuos de cada modelo por medio de métodos analíticos, entre ellos destacan las pruebas de bondad de ajuste *Kolmogorov-Smirnov*

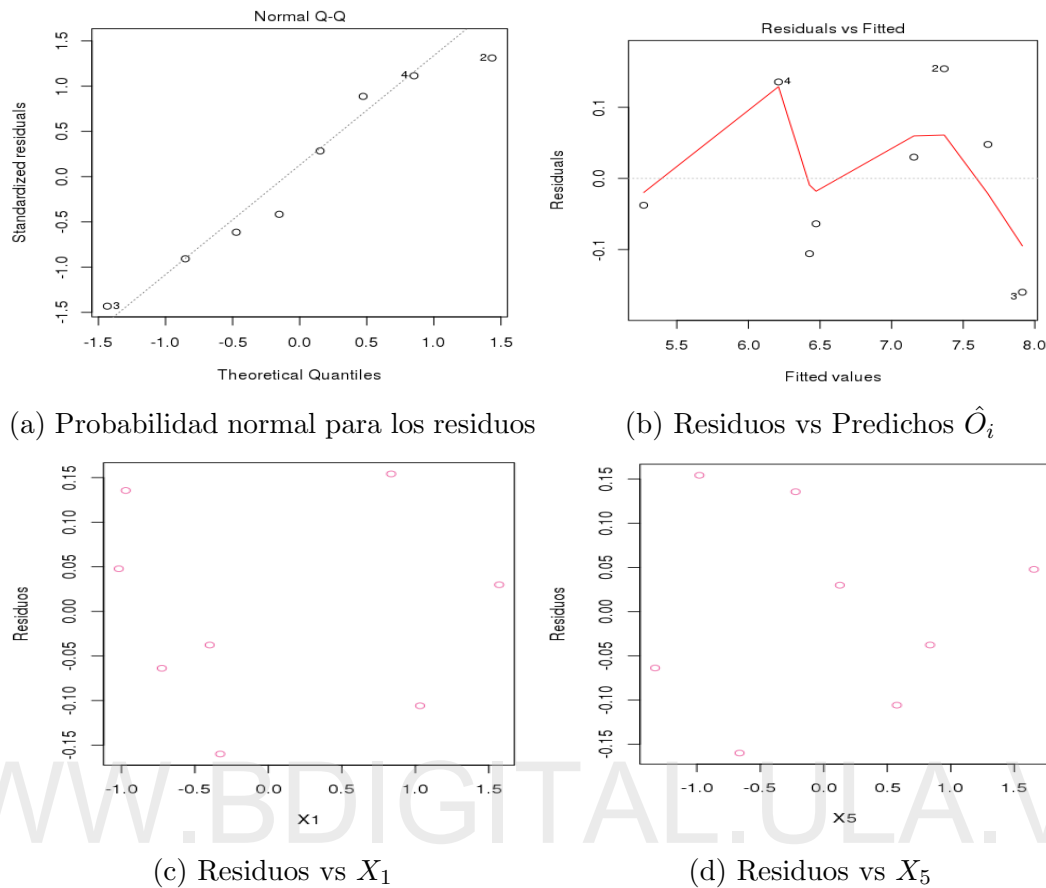


Figura 3.4: Verificación de supuestos del modelo de orígenes de viajes ( $O_i$ )

y la prueba de *Shapiro-Wilk* (apéndice D.1), siendo esta última la más recomendada en estos casos (Pulido y De la Vara Salazar, 2008). De acuerdo con los *valores-p* obtenidos para las pruebas anteriormente mencionadas: *Kolmogorov-Smirnov* : 0.06421, 0.06439 y *Shapiro-Wilk* : 0.796, 0.2222 se acepta la hipótesis de normalidad, siendo todas mayores que la significancia. Para un mejor entendimiento e interpretación de los residuos se procedió a analizar los resultados de las figuras 3.4(b) y 3.5(b) para cada modelo, de la misma manera indican un mayor error de ajuste para los nodos 2 y 3 de la variable de respuesta  $O_i$  y en el nodo 5 para la variable  $D_j$ ; igualmente, la distribución de los puntos no parece seguir ninguna tendencia clara, por ello, se puede considerar que el supuesto de varianza constante se cumple aceptablemente para ambos casos. Para el resto de las figuras, se muestran los residuales contra los  $m^2$  de uso de suelos tipo residencial y servicios, además la distancia en  $m$  al centro de la ciudad, los cuales

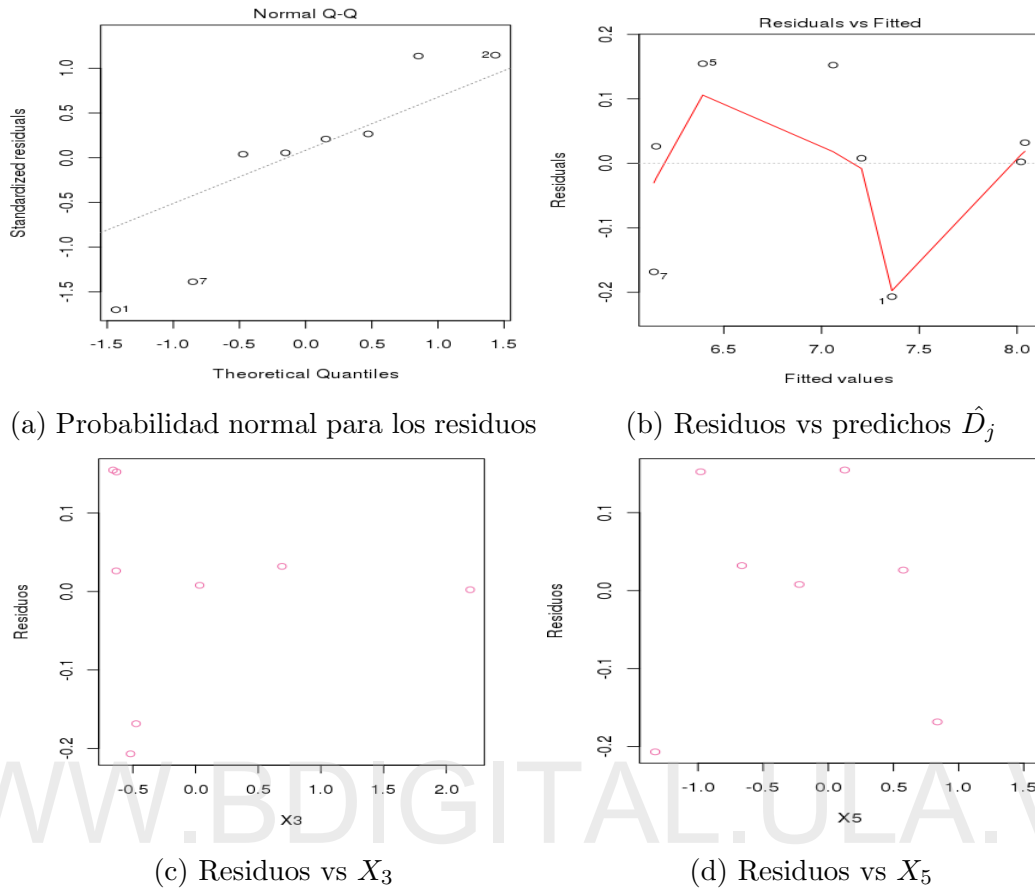


Figura 3.5: Verificación de supuestos del modelo de atracciones de viajes ( $D_j$ )

no reflejan ningún patrón fuerte con lo que se pudiera incumplir los supuestos. De acuerdo con el análisis de datos realizado para verificar la adecuación de los ajustes, es propio decir que los modelos matemáticos de las ecuaciones (3.1) y (3.2) que explican la cantidad de viajes originados y atraídos en cada nodo son ajustes aceptables, en general, no violan ninguno de los supuestos, además se podran emplear como potentes herramientas para predecir la demanda de transporte en la red Centro - La Hechicera. Los resultados obtenidos en la estimación de la demanda actual de viajes en cada nodo puede visualizarse por medio del SIG de la red de transporte (figura 3.6). Como se puede observar en la tabla 3.4 el modelo de transporte está desequilibrado, puesto que  $\sum_{i=1}^n O_i \neq \sum_{j=1}^n D_j$ . En el siguiente capítulo se mostrará el procedimiento para balancear los orígenes y demandas de la red de transporte.



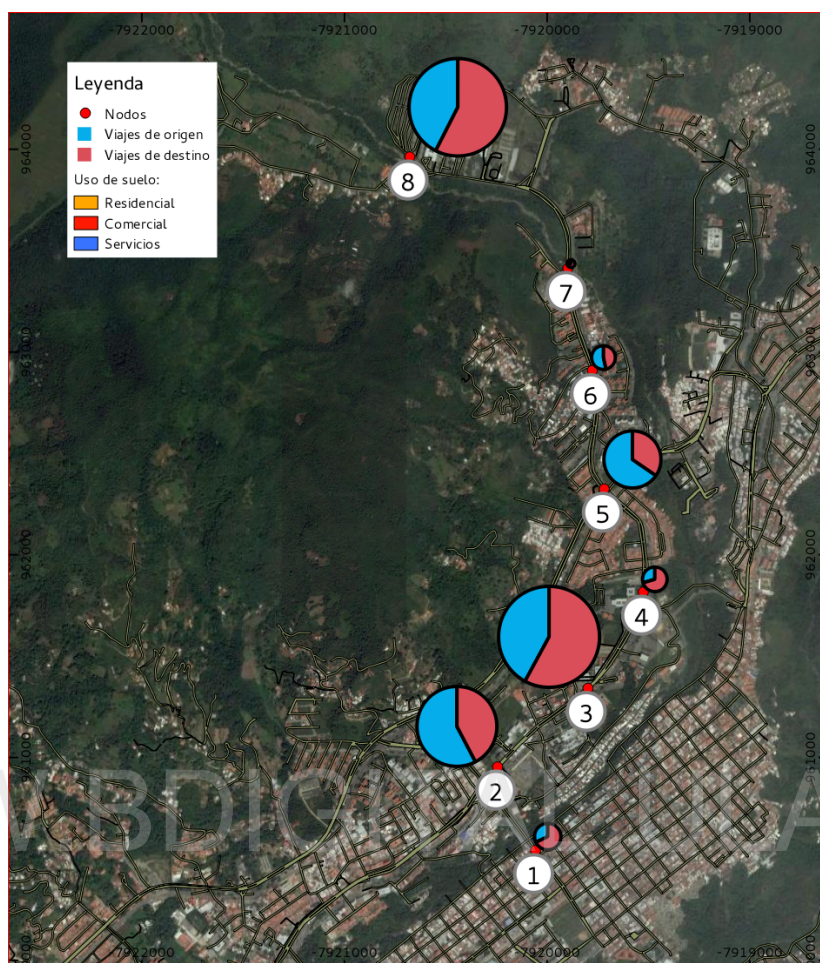


Figura 3.6: Visualización geográfica de los valores ajustados  $\hat{O}_i$  y  $\hat{D}_j$



## Capítulo 4

# Modelo de Distribución

En este capítulo se propone la formulación de un modelo de distribución de viajes basado en la teoría de máxima entropía, con la cual se construye la matriz Origen-Destino que representa la cantidad de viajes  $V_{ij}$  entre un nodo origen  $i$  y un nodo destino  $j$  de la red de transporte Centro - La Hechicera. Adicionalmente, se obtiene la formulación de un modelo de optimización sujeto a la información de orígenes y atracciones de viajes en cada nodo calculados en el capítulo 3.

### 4.1 Balance entre orígenes y atracciones

Debido a que la distribución de los viajes se realiza usando la estructura general del modelo de transporte presentada en el capítulo 1, para obtener una solución factible se requiere que el total de viajes originados sea igual al total de viajes atraídos; dicho de otra manera, el modelo debe estar balanceado, garantizando :  $\sum_{i=1}^n O_i = \sum_{j=1}^n D_j$ . En este caso, el modelo viola la condición anterior, puesto que  $9671 < 11816$ . En [Martin y McGuckin \(1998\)](#) se expone que, en realidad, el total de estimaciones de los orígenes de viaje y atracciones no será exactamente igual, en este caso porque las estimaciones realizadas de ambos fenómenos son el resultado de modelos diferentes; ante este escenario siempre se podrá restaurar el balance por medio de la aplicación de un factor de equilibrio ([de Dios Ortúzar y Willumsen, 2008](#)). En la práctica habitual, para escalar el total de orígenes y demandas son utilizadas las ecuaciones (4.1) y (4.2).

$$f_{eqO} = \frac{V}{\sum_{j=1}^n D_j} = \frac{\sum_{i=1}^n O_i}{\sum_{j=1}^n D_j} \quad (4.1)$$

$$f_{eqD} = \frac{V}{\sum_{i=1}^n O_i} = \frac{\sum_{j=1}^n D_j}{\sum_{i=1}^n O_i} \quad (4.2)$$

El procedimiento consiste en equilibrar la demanda de viajes por medio de la aplicación de sólo una de las dos ecuaciones presentadas. Hay casos en que se usa la ecuación (4.1), donde las atracciones de viajes ( $\sum_{j=1}^n D_j$ ) se escalan para igualar los totales de orígenes ( $\sum_{i=1}^n O_i$ ) en la red de transporte; de manera alterna ocurre con el factor de equilibrio de la ecuación (4.2), es decir, se escalan los orígenes de viajes, de modo que se cumpla :  $\sum_{i=1}^n O_i = \sum_{j=1}^n D_j$ . La decisión de qué factor de equilibrio se requiere implementar queda a voluntad del planificador de transporte y de la naturaleza del problema. En este caso se optó por aplicar un ajuste a los totales de viajes originados en la red ( $f_{eqD}$ ), esto como consecuencia de que el modelo de atracción de viajes ajustado en el capítulo 3 posee el mayor poder explicativo con el mínimo de error; una vez balanceados los viajes (origen y destino), ahora pueden ser utilizados como entradas en un modelo de distribución, el mismo puede ser modelado desde el enfoque de un modelo gravitacional o por medio de un modelo de maximización de la entropía (de Dios Ortúzar y Willumsen, 2008), este último se presenta y desarrolla a detalle en la siguiente sección.

## 4.2 Formulación del modelo de optimización

Como ya fue introducido en el Capítulo 1, el concepto de entropía es muy conocido y ampliamente utilizado en la teoría de la información, puesto que su objetivo es generar distribuciones de probabilidad a partir de información dada. Con la finalidad de descubrir los patrones de viajes para cada arco( $i, j$ ) de la red de transporte Centro - La Hechicera, se implementa el modelo de transporte en una matriz origen - destino ( $O-D$ ): en Wright et al. (2010) consiste esencialmente en una matriz bidimensional de celdas (tabla 4.1) donde las filas y columnas representan cada uno de los nodos de transporte del área de estudio, las celdas de cada fila  $i$  contienen los viajes originados en ese nodo que tiene como destinos los nodos en las columnas correspondientes y la

diagonal principal corresponde a los viajes intrazonales. Por lo tanto  $V_{ij}$  es el número de viajes entre el origen  $i$  y el destino  $j$ ; la matriz total es  $V$  que corresponde a la condición de factibilidad  $V = \sum_{i=1}^n O_i = \sum_{j=1}^n D_j$ .

Tabla 4.1: Estructura de matriz de viajes O-D (Wright et al., 2010)

Orígenes	Destinos								
	1	2	3	4	5	6	7	8	$\sum_{i=1}^n O_i$
1	$V_{11}$	$V_{12}$	$V_{13}$	$V_{14}$	$V_{15}$	$V_{16}$	$V_{17}$	$V_{18}$	$O_1$
2	$V_{21}$	$V_{22}$	$V_{23}$	$V_{24}$	$V_{25}$	$V_{26}$	$V_{27}$	$V_{28}$	$O_2$
3	$V_{31}$	$V_{32}$	$V_{33}$	$V_{34}$	$V_{35}$	$V_{36}$	$V_{37}$	$V_{38}$	$O_3$
4	$V_{41}$	$V_{42}$	$V_{43}$	$V_{44}$	$V_{45}$	$V_{46}$	$V_{47}$	$V_{48}$	$O_4$
5	$V_{51}$	$V_{52}$	$V_{53}$	$V_{54}$	$V_{44}$	$V_{46}$	$V_{47}$	$V_{48}$	$O_4$
6	$V_{51}$	$V_{52}$	$V_{53}$	$V_{54}$	$V_{55}$	$V_{56}$	$V_{57}$	$V_{58}$	$O_5$
7	$V_{61}$	$V_{62}$	$V_{63}$	$V_{64}$	$V_{65}$	$V_{66}$	$V_{67}$	$V_{68}$	$O_6$
8	$V_{81}$	$V_{82}$	$V_{83}$	$V_{84}$	$V_{85}$	$V_{86}$	$V_{87}$	$V_{88}$	$O_8$
$\sum_{j=1}^n D_j$	$D_1$	$D_2$	$D_3$	$D_4$	$D_5$	$D_6$	$D_7$	$D_8$	$V$

Dado que los dos modelos (generación y distribución) son independientes uno del otro, es necesario establecer una serie de consistencia entre las etapas: para este caso, en base a la tabla 4.1, se debe cumplir  $\sum_i^n V_{ij} = D_j$  y  $\sum_j^n V_{ij} = O_i$ .

El objetivo de un modelo de distribución es estimar el número de viajes en cada celda de la matriz  $V_{ij}$ , según la información disponible del sistema. Entendiendo la entropía como probabilidad y relacionada con la incertidumbre respecto a la información de que dispone un analista, se pueden distinguir 3 estados de información (Moura Berodia et al., 2006).

**Estados micro:** viajes en particular de un individuo desde un origen  $i$  a un destino  $j$ .

**Estados meso:** la cantidad de viajes realizados desde el nodo  $i$  al nodo  $j$   $V_{ij}$ , que a diferencia del anterior representa una base mas agregada de la información pero no tan detallada.

**Estados macro:** número de viajes originados  $O_i$  y atraídos  $D_j$  por cada nodo de la red. Este último estado representa el nivel superior de información agregada que se puede tener del sistema. De acuerdo con [de Dios Ortúzar y Willumsen \(2008\)](#), las estimaciones relativas a modelos de demanda son, por lo general, a descripciones de estados macro a causa de la incertidumbre asociada a la previsión en niveles más desagregados

De esta manera, un estado macro esta compuesto por un conjunto de estados meso, los que cada uno a su vez se componen de un conjunto posible de estados micro. Tal como se ilustra en la figura 4.1

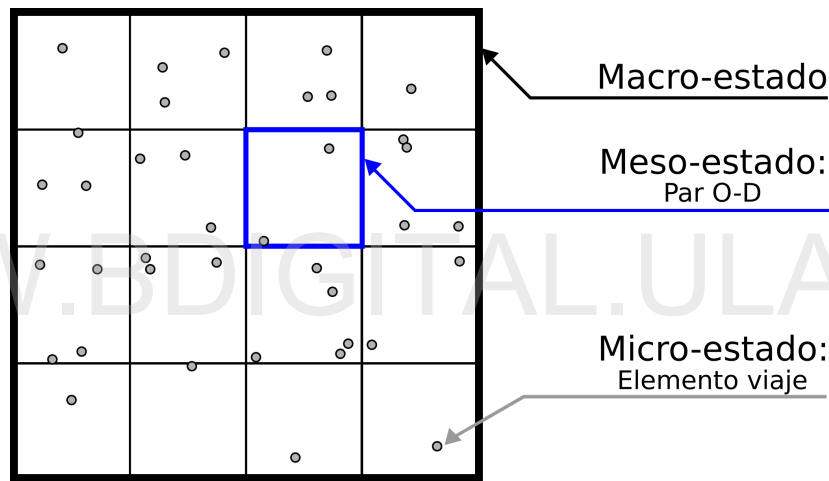


Figura 4.1: Representación de los 3 de estados de información del sistema

Cabe destacar que la mayor parte de la información actual disponible sobre el modelo de demanda de transporte para este trabajo, es justamente al nivel del estado macro. Sabiendo esto, en Romero tal como se citó en [López-Ospina \(2013\)](#), la distribución más probable de viajes entre dos pares de nodos  $(i,j)$  esta dada por:  $V_{ij} = \mathbf{V}p(i,j)$ , donde la matriz  $\mathbf{V}$  es la suma total de viajes originados o atraídos y  $p(i,j)$  es la probabilidad de viajar entre el par origen y destino. Para abordar el problema anterior se propone el metodo de maximización de la entropía para la distribución de viajes, que se basa en aceptar que todos los micro-estados consistentes en una información de estados más agregados (meso ó macro-estados), son equiprobables ([de Dios Ortúzar y Willumsen, 2008](#))), en otras palabras, cada elemento

(viaje individual) de la figura 4.1 tiene la misma probabilidad de pertenecer a cada una de las categorías  $(V_{ij})$ . En base a la hipótesis anterior, Wilson demostró, tal como se citó en de Dios Ortúzar y Willumsen (2008), que el número de micro-estados, denotado como  $W(V_{ij})$ , asociados a un meso-estado  $V_{ij}$  está dado por la ecuación (4.3)

$$W(V_{ij}) = \frac{V!}{\prod_{ij} V_{ij}!} \quad (4.3)$$

Retomando el concepto de entropía de una distribución introducido en el capítulo 1, aquella matriz O-D que cumpla con la distribución uniforme de los viajes  $V_{ij}$ , satisfaciendo la información conocida del sistema (restricciones del estado macro  $O_i$  y  $D_j$ ) será la matriz de viajes más probable, debido a que ésta es la que genera mayor nivel de incertidumbre. En otras palabras, de acuerdo con de Dios Ortúzar y Willumsen (2008), el meso-estado más probable sería aquel que tuviera un mayor numero de micro-estados asociados y dado que todos los micro-estados son equiprobables, la matriz de viajes factible es la que maximiza  $W$ . Ahora se procede a obtener el modelo de optimización: para facilitar los cálculos, se ha aplicado una transformada monótona creciente (logaritmo en este caso) a la función  $W$ , puesto que maximizar  $W$  es lo mismo que maximizar  $\log W$ , ya que ambas funciones presentan los mismos máximos, teniendo en cuenta la aproximación de *Stirling* :  $\log x! = x \log x - x$ , se obtiene:

$$\begin{aligned} \log(W(V_{ij})) &= \log\left(\frac{V!}{\prod_{ij} V_{ij}!}\right) \\ \log(W(V_{ij})) &= \log V! - \sum_{ij}^n \log(V_{ij}!) \\ \log(W(V_{ij})) &= \log V! - \sum_{ij}^n (V_{ij} \log V_{ij} - V_{ij}) \end{aligned}$$

A continuación, son eliminados de la ecuación los términos que no son optimizables: dado que el término  $\log V!$  representa el valor de una constante, puede ser omitido del problema, con lo que se resulta la ecuación (4.4) que generalmente se conoce como función de entropía.

$$\log(W(V_{ij})) = - \sum_{ij}^n (V_{ij} \log V_{ij} - V_{ij}) \quad (4.4)$$

Finalmente, para obtener la matriz de viajes más probable en la red de transporte, se maximiza  $\log(W(V_{ij}))$  teniendo en cuenta las restricciones del macro-estado disponibles para el modelo de demanda, en este trabajo se adaptó el siguiente problema de programación matemática. Este tipo de modelos con restricciones de origen y destino balanceados se denominan modelos de entropía doblemente acotado; cabe notar que al resolver estos modelos de optimización se deriva el modelo gravitacional (de Dios Ortúzar y Willumsen, 2008), el problema presentado a continuación posee tantas restricciones de origen como de destinos para los 8 nodos de transporte de la red Centro - La Hechicera.

$$\max - \sum_{ij}^8 (V_{ij} \log V_{ij} - V_{ij})$$

sa :

$$\sum_{j=1}^8 V_{1j} = 742, \sum_{i=1}^8 V_{i1} = 1278$$

$$\sum_{j=1}^8 V_{2j} = 2258, \sum_{i=1}^8 V_{i2} = 1355$$

$$\sum_{j=1}^8 V_{3j} = 2846, \sum_{i=1}^8 V_{i3} = 3205$$

$$\sum_{j=1}^8 V_{4j} = 697, \sum_{i=1}^8 V_{i4} = 1356$$

$$\sum_{j=1}^8 V_{5j} = 1613, \sum_{i=1}^8 V_{i5} = 697$$

$$\sum_{j=1}^8 V_{6j} = 679, \sum_{i=1}^8 V_{i6} = 483$$

$$\sum_{j=1}^8 V_{7j} = 228, \sum_{i=1}^8 V_{i7} = 393$$

$$\sum_{j=1}^8 V_{8j} = 2753, \sum_{i=1}^8 V_{i8} = 3049$$

$$\sum_{i=1}^8 V_{ii} = 0$$

Como última observación, tenemos la restricción  $\sum_{i=1}^8 V_{ii} = 0$ , puesto que no existen los viajes intrazonales, esto según la naturaleza geográfica de los viajes para la red de transporte Centro-La Hechicera.

### 4.3 Algoritmo de estimación de la matriz O-D

Aunque el problema de optimización derivado en la sección anterior se puede resolver con las técnicas clásicas de la programación matemática, la estructura general de la matriz de viajes permite desarrollar un algoritmo de cómputo que permite estimar con simplificados cálculos el modelo de distribución de viajes más probable de una red de transporte (de Dios Ortúzar y Willumsen, 2008). A continuación, se muestra la estimación de la matriz de viajes del modelo de distribución por medio de su aplicación directa en cada paso del algoritmo.

1. El primer paso es definir la matriz O-D del modelo de demanda previamente equilibrado (en este caso fueron equilibrados los viajes de origen), a partir de la información disponible de la red de transporte (orígenes, demandas, entre otros).

Orígenes	Destinos								$O_i$
	1	2	3	4	5	6	7	8	
1	0								742
2		0							2258
3			0						2846
4				0					697
5					0				1613
6						0			679
7							0		228
8								0	2753
$D_j$	1278	1355	3205	1356	697	483	393	3049	11816

2. EL segundo paso es establecer la matriz semilla del problema, usualmente se fija como semilla aquella matriz, a cuyas incógnitas  $V_{ij}$  le corresponde el valor de 1.

Orígenes	Destinos								$O_i$
	1	2	3	4	5	6	7	8	
1	0	1	1	1	1	1	1	1	742
2	1	0	1	1	1	1	1	1	2258
3	1	1	0	1	1	1	1	1	2846
4	1	1	1	0	1	1	1	1	697
5	1	1	1	1	0	1	1	1	1613
6	1	1	1	1	1	0	1	1	679
7	1	1	1	1	1	1	0	1	228
8	1	1	1	1	1	1	1	0	2753
$D_j$	1278	1355	3205	1356	697	483	393	3049	11816

3. Si la matriz viola las restricciones macro de origen, en este caso ( $7 < O_i$ ), se debe obtener el factor de expansión por filas  $f_{expF} = \frac{O_i}{\sum_{j=1}^8 V_{ij}}$ .

Orígenes	Destinos								$O_i$	$f_{expF}$
	1	2	3	4	5	6	7	8		
1	0	1	1	1	1	1	1	1	742	106
2	1	0	1	1	1	1	1	1	2258	322.57
3	1	1	0	1	1	1	1	1	2846	406.47
4	1	1	1	0	1	1	1	1	697	99.57
5	1	1	1	1	0	1	1	1	1613	230.42
6	1	1	1	1	1	0	1	1	679	97
7	1	1	1	1	1	1	0	1	228	32.57
8	1	1	1	1	1	1	1	0	2753	393.28
$D_j$	1278	1355	3205	1356	697	483	393	3049	11816	

4. A nivel de cada fila se multiplica el valor  $V_{ij}$  por su correspondiente factor de expansión. Como la matriz resultante viola las restricciones macro de destino, se deben obtener los factores de expansión por columna  $F_{expC} = \frac{D_j}{\sum_{i=1}^8 V_{ij}}$ .



Orígenes	Destinos								
	1	2	3	4	5	6	7	8	$O_i$
1	0	106	106	106	106	106	106	106	742
2	322.57	0	322.57	322.57	322.57	322.57	322.57	322.57	2258
3	406.47	406.47	0	406.47	406.47	406.47	406.47	406.47	2846
4	99.57	99.57	99.57	0	99.57	99.57	99.57	99.57	697
5	230.42	230.42	230.42	230.42	0	230.42	230.42	230.42	1613
6	97	97	97	97	97	0	97	97	679
7	32.57	32.57	32.57	32.57	32.57	32.57	0	32.57	228
8	393.28	393.28	393.28	393.28	393.28	393.28	393.28	0	2753
$D_j$	1278	1355	3205	1356	697	483	393	3049	11816
$F_{expC}$	0.807	0.992	2.501	0.853	0.478	0.303	0.237	2.355	

5. A nivel de cada columna multiplicamos el valor  $V_{ij}$  por su correspondiente factor de expansión.

Orígenes	Destinos								
	1	2	3	4	5	6	7	8	$O_i$
1	0	105.152	265.106	90.418	50.668	32.118	25.122	249.630	742
2	260.313	0	806.747	275.152	154.188	97.738	76.449	759.652	2258
3	328.021	403.218	0	346.718	194.292	123.160	96.333	957.236	2846
4	80.352	98.773	249.024	0	45.594	30.169	23.598	234.487	697
5	185.948	228.576	576.380	196.548	0	69.817	54.609	542.639	1613
6	78.279	90.272	242.597	82.741	46.366	0	22.989	228.435	679
7	26.283	32.309	81.457	27.782	15.568	9.868	0	76.702	228
8	317.376	390.133	983.593	335.467	187.987	119.139	93.207	0	2753
$D_j$	1278	1355	3205	1356	697	483	393	3049	11816

6. La base del método es que siempre se alcanzará la convergencia. En este caso la matriz no converge para la primera iteración, puesto que aún se violan las restricciones para los orígenes y destinos, por lo tanto se debe regresar al paso 3.

Después de 4 iteraciones para este caso, se muestra en la tabla 4.2 la matriz de viajes más probable, es decir, aquella que posee la mayor cantidad de estados micros

y que al mismo tiempo satisface las restricciones de origen y destino. En este tipo de problemas, las variables de viajes estimados  $V_{ij}$  sólo tendrá sentido real si su valor es entero; para estos casos especiales se suele redondear el valor del viaje y, aunque existe esta posibilidad, no hay garantía de que la solución siga siendo óptima, pues se sobreestima la cantidad de viajes que se esperarían entre cada arco  $(i,j)$ . Para conservar la factibilidad a partir de la cota inferior de todas variables se evaluó el espacio de soluciones, modificando iterativamente el valor de los  $V_{ij}$  de manera que se cumplan las restricciones de origen y destino, de esta manera de obtuvo una solución que está cerca del óptimo, con estimaciones de viajes entre nodos mucho mejores que las obtenidas mediante el redondeo. Cabe mencionar que la distribución de viajes está sujeta a cambios en sus parámetros, es decir se pueden calibrar matrices de viajes para distintos valores de orígenes y demandas correspondientes a intervalos de tiempo distintos, con el fin de poder modelar el patron de viajes de la red de transporte para distintos períodos del día, lo cual dependerá de la necesidad y precisión del estudio; cabe notar que los valores de orígenes y demandas aquí presentados corresponden a los viajes totales capturados desde la 7:00 am a 6:00 pm. El modelo de distribución desarrollado en este capítulo ofrece la base para poder realizar una planificación del sistema de transporte de la Línea 2 Centro- La Hechicera.

Tabla 4.2: Matriz de viajes con máxima entropía de la red de transporte Centro - La Hechicera

Orígenes	Destinos								$O_i$
	1	2	3	4	5	6	7	8	
1	0	91.41136	247.84037	79.36193	43.22360	28.183136	22.16591	229.66656	742
2	234.58529	0	767.85819	248.56922	137.39902	87.369750	68.35689	713.4480	2258
3	375.80018	455.84956	0	397.76538	219.55735	142.551234	111.54167	1143.29910	2846
4	71.25596	85.28374	234.51977	0	41.16469	26.139251	20.12711	218.37719	697
5	157.40664	189.43910	515.62189	166.39575	0	58.251659	46.24552	479.34347	1613
6	65.23931	78.26634	212.48711	68.23633	37.15146	0	19.12221	198.35772	679
7	21.20121	26.24213	72.58962	22.20623	12.11995	7.079397	0	66.50789	228
8	352.51141	428.50778	1154.08304	373.46517	206.38393	133.425573	105.44070	0	2753
$D_j$	1278	1355	3205	1356	697	483	393	3049	11816

# Capítulo 5

## Estudio de Escenarios

El siguiente capítulo propone distintos escenarios de transporte de la Línea 2 Centro - La Hechicera, donde se estudian los patrones de viajes de la red con el fin de brindar a los planificadores herramientas sólidas para tomar las decisiones en materia de transporte, garantizando un sistema eficiente y sostenible. Primeramente, los escenarios estudiados siempre deben estar adaptados a las circunstancias locales de la red y el modelo de distribución de viajes presentando en el capítulo anterior sintetiza las bases necesarias para lograr tal adaptación; sin embargo, para propósitos mas específicos, se decidió evaluar la información disponible desagregada de viajes, logrando estimar diferentes matrices O-D para diferentes períodos del día, con el propósito de evaluar los flujos de pasajeros, detectar el período de tiempo donde ocurre la mayor demanda, entre otros. Para lograr esto, se desarrolló una aplicación web, utilizando el framework para R (*Shiny*) (apéndice A.1) y la biblioteca (*igraph*) (apéndice A.2) para modelar los resultados en un grafo de transporte del Centro - La Hechicera. La principal función de dicha herramienta interactiva es la estimación y monitoreo de diferentes patrones de viajes que puedan presentarse durante cada hora del día en la Línea 2 Centro - La Hechicera a partir del concepto de maximización de la entropía (capítulo 4). Cabe señalar que el código fuente está disponible de forma pública en un repositorio de *Github*, en el apéndice B se presenta parte de la estructura del código fuente y sus principales componentes.

A fin de lograr un enfoque de análisis que permita un primer acercamiento

descriptivo del fenómeno de viajes en la Línea 2, se evaluó el panorama de cuántos pasajeros se proyectan en cada nueva estación, con ayuda de los viajes de origen y destino capturados en las paradas de transporte público (tabla 2.6), que se encuentran próximas a las futuras estaciones del Trolebús (tabla 2.2). De las cuáles se pretende además obtener los flujos de viajes entre dichas infraestructuras, especialmente para las horas pico, puesto que en estas horas suele ocurrir una saturación del transporte público. Estos escenarios parten del principio de que el sistema de transporte Trolebús no abarcará toda la demanda de transporte público capturada en la red Centro - La Hechicera; igualmente, se espera un incremento del cambio modal por parte de las personas que consideren más atractivo y económico viajar mediante el sistema de transporte Trolebús. En vista a lo anterior, a través de la aplicación web se permite el diseño de los patrones de viajes dados los diferentes parámetros (% de transporte público cubierto e incremento del cambio modal), lo que permite el fácil ajuste de las matrices O-D ante cualquier escenario (figura 5.1). Resulta importante destacar que la Línea 2 aun se encuentra en su etapa exploratoria, por esta razón los escenarios seleccionados solo brindan nociones descriptivas futuras de la red.

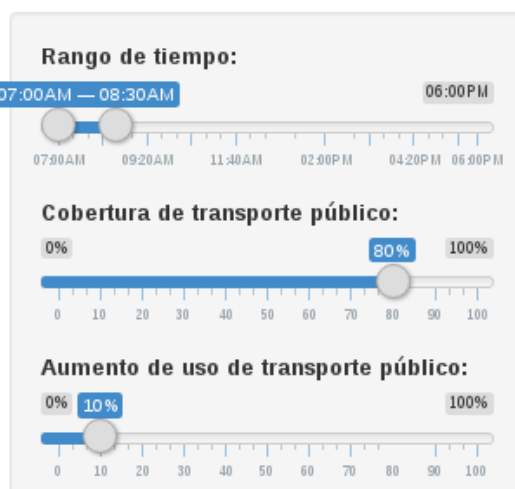


Figura 5.1: Parámetros de la matriz de viajes en la aplicación web

## 5.1 Escenario 1

Distribución de viajes en las horas pico mañana desde las 7:00 AM a 8:30 AM, cubriendo el 80% del transporte público y ajustando un incremento de 10% en el cambio modal

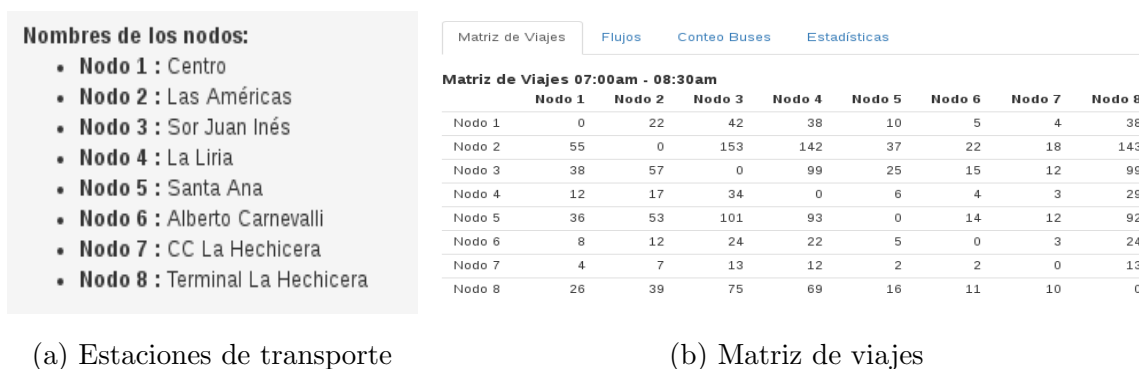


Figura 5.2: Matriz de viajes estimada. Hora: 7:00 AM - 8:30 AM

A partir de la matriz O-D estimada mediante la aplicación web (figura 5.2), se obtuvo la tabla 5.1, donde se muestra los principales resultados del flujo de viajes estimados bajo este escenario para cada arco de la red de transporte, en un recorrido completo de la Linea 2, es decir en sentido Centro - La Hechicera - Centro, de los cuales se observa que, para los arcos Las Américas - Sor Juana y Sor Juana - La Liria se espera transportar el mayor flujo de pasajeros en el sentido Centro - la Hechicera con 645 y 700 personas, respectivamente. Además la mayor afluencia de pasajeros en este sentido se observa en el nodo Las Américas con 509 viajes, siendo el nodo Terminal La Hechicera la estación que atrae más viajes con 431 pasajeros en total, por otra parte, la estación CC La Hechicera es la única estación que presenta una demanda casi insignificante con valores entre 9 y 13 pasajeros que usarán el Trolebús en sentido Centro - La Hechicera y La Hechicera - Centro respectivamente. Este comportamiento de la demanda durante la hora pico mañana se puede explicar en base a las características de uso de suelos del entorno de cada estación donde, tal como se discutió en el capítulo 3, las zonas residenciales afectan los orígenes de viajes, mientras que las zonas de servicios, principalmente estudiantiles, afectan las atracciones, por lo tanto es evidente este patrón de viajes para la hora establecida. En la figura 5.3(a) se

observa una mayor movilización de la población desde las zonas del Centro a la zona norte de la ciudad (La Hechicera), especialmente desde la estación Las Américas en la cual 429 viajes que se originan en esa zona, se dirigen hacia las diversas estaciones que abarcan la mayor zona estudiantil del área metropolitana, siendo estas las estaciones Sor Juana Ines, La Liria y Terminal La Hechicera. Es importante destacar en la tabla 5.1, que para los arcos correspondientes al sentido Hechicera - Centro, se observan menor afluencia de pasajeros, nuevamente se observan que la mayor cantidad de viajes terminan en las estaciones La Liria y Sor Juana Ines. Las conclusiones anteriores se respaldan al observar los intervalos de confianza del 95% para los viajes originados y atraídos que se esperan en cada estación de la red cada 10 minutos dentro del período 7:00 AM - 8:30 AM (tabla 5.2).

Tabla 5.1: Flujo de viajes entre cada arco de la red Centro - La Hechicera. Hora: 7:00 AM - 8:30 AM

Origen $i$	Destino $j$	Pasajeros en el arco $(i,j)$	Pasajeros que suben en $i$	Pasajeros que bajan en $j$
Centro	Las Américas	158	158	22
Las Américas	Sor Juana Inés	645	509	192
Sor Juana Inés	La Liria	700	247	275
La Liria	Santa Ana	467	42	77
Santa Ana	Alberto Carnevalli	508	118	62
Alberto Carnevalli	C.C. La Hechicera	470	24	52
C.C. La Hechicera	Terminal La Hechicera	431	13	431
Terminal La Hechicera	C.C. La Hechicera	242	242	9
C.C. La Hechicera	Alberto Carnevalli	273	40	12
Alberto Carnevalli	Santa Ana	334	73	23
Santa Ana	La Liria	590	279	195
La Liria	Sor Juana Inés	457	62	245
Sor Juana Inés	La Américas	305	93	183
La Américas	Centro	177	55	177

Tabla 5.2: Intervalos de confianza para la media de los viajes. Hora: 7:00 AM - 8:30 AM

Identificador	Nodo	Viajes Origen	Intervalo de confianza ( $O_i$ )	Viajes Destino	Intervalo de confianza ( $D_j$ )
1	Centro	88	3.96 ; 15.59	180	11.73 ; 28.27
2	Las Américas	309	3.96 ; 15.59	209	19.81 ; 26.63
3	Sor Juana Inés	189	16.36 ; 25.64	440	38.05 ; 59.73
4	La Liria	59	2.81 ; 10.3	472	45.44 ; 59.45
5	Santa Ana	219	14.47 ; 34.19	104	8.92 ; 14.19
6	Alberto Carnevali	58	4.68 ; 8.21	79	6.45 ; 11.11
7	C.C. La Hechicera	32	1.41 ; 5.7	65	3.83 ; 10.61
8	Terminal La Hechicera	138	9.84 ; 20.82	442	36.77 ; 61.45

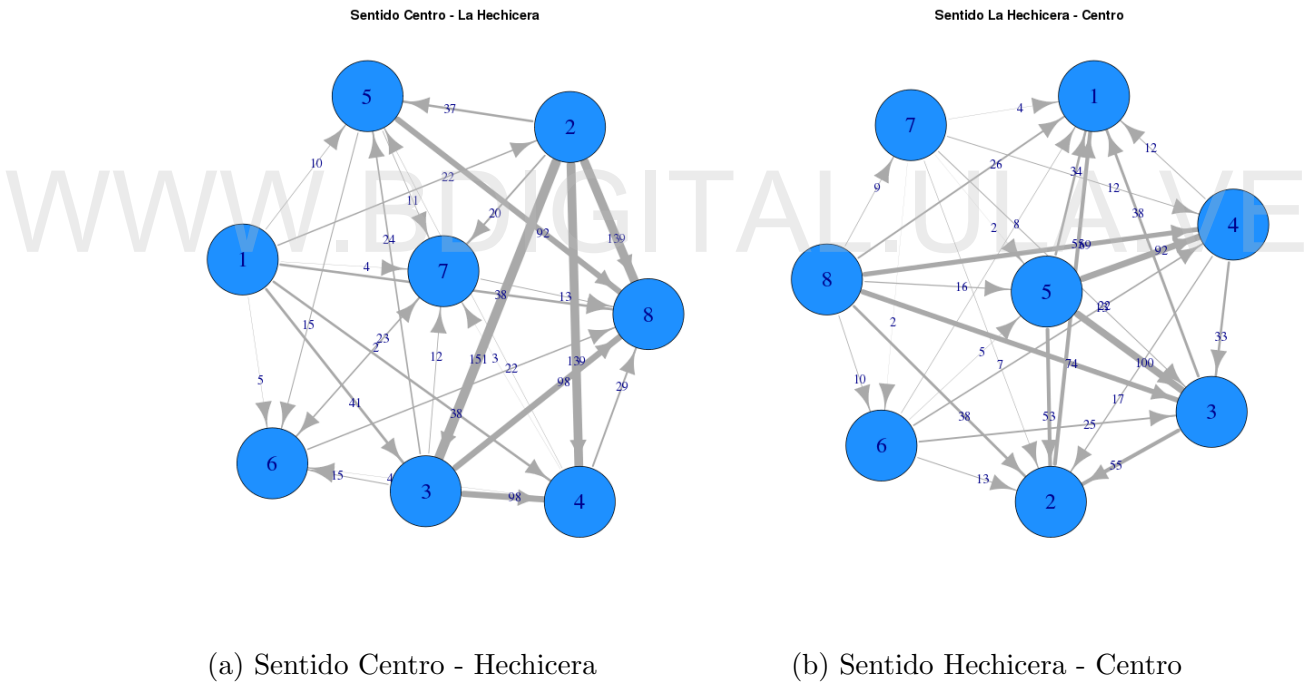


Figura 5.3: Flujo entre cada arco  $(i,j)$ . Hora: 7:00 AM - 8:30 AM

## 5.2 Escenario 2

Distribución de viajes en las horas pico tarde desde las 12:00 M - 2:00 PM, abarcando el 90% del transporte público y ajustando un incremento de 15% en el cambio modal

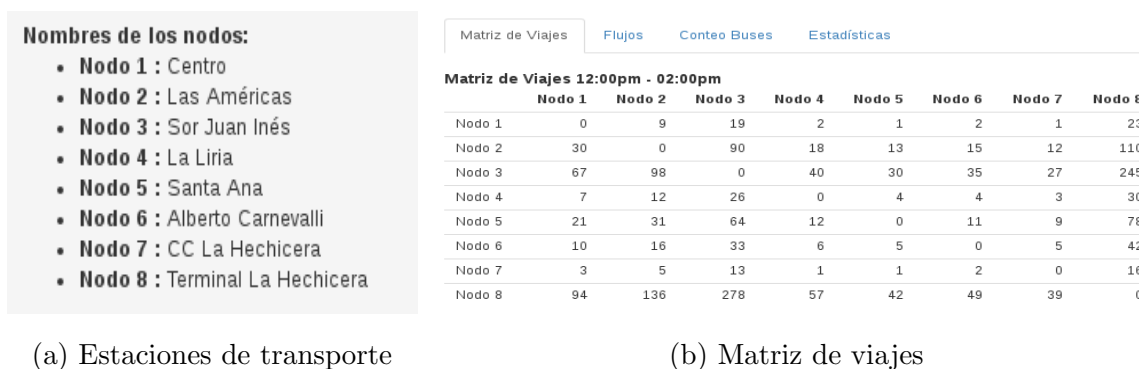


Figura 5.4: Matriz de viajes estimada. Hora: 12:00 M - 2:00 PM

Para este escenario en particular, ocurre el caso contrario al escenario anterior: el patrón de viajes para las horas pico 12: 00 M a 2: 00 PM consiste en la movilización de la población desde la zona Norte La Hechicera al centro de la ciudad, esto se observa en la tabla 5.3 obtenida mediante la matriz de viajes estimada (figura 5.4), donde los flujos entre cada arco oscilan entre 544 y 780 pasajeros. Según (Wright et al., 2010), en Bogotá, el sistema de transporte Trolebús logró descongestionar los carriles de tráfico mixto y el impacto del cambio modal en la primera fase estuvo por encima del 10% de usuarios de vehículos privados al sistema Trolebús; por este motivo se consideró un escenario de cambio modal del 15%. Para este caso, la estación Terminal La Hechicera es el principal punto atractor de viajes con 544 pasajeros, asimismo se estima una gran movilización desde la estación Terminal la Hechicera con un total de 695 pasajeros que suben, siendo los principales destinos en el sentido Hechicera - Centro, las estaciones Sor Juana Ines y Las Américas (figura 5.5(b)). Los intervalos de confianza entre períodos de 10 min (tabla 5.4) para los pasajeros que suben en la estación Terminal La Hechicera se encuentran entre 48 y 79 personas, igualmente para las estación Sor Juana Inés, se esperan entre 42 y 58 pasajeros cada 10 minutos aproximadamente. Es claro que las estaciones anteriormente mencionadas son altamente demandadas, puesto que están



ubicadas cerca de universidades y colegios. En base a esto, los planificadores de Tromerca estarán en mayor capacidad de evaluar si los diseños propuestos para las estaciones Sor Juana y Terminal La Hechicera van a tener la suficiente capacidad de manejar la demanda proyectada.

Tabla 5.3: Flujo de viajes entre cada arco de la red Centro - La Hechicera. Hora: 12:00M - 2:00 PM

Origen $i$	Destino $j$	Pasajeros en el arco $(i,j)$	Pasajeros que suben en $i$	Pasajeros que bajan en $j$
Centro	Las Américas	57	57	9
Las Américas	Sor Juana Inés	306	258	109
Sor Juana Inés	La Liria	574	378	60
La Liria	Santa Ana	555	41	48
Santa Ana	Alberto Carnevalli	605	98	67
Alberto Carnevalli	C.C. La Hechicera	585	47	57
C.C. La Hechicera	Terminal La Hechicera	544	16	544
Terminal La Hechicera	C.C. La Hechicera	695	695	39
C.C. La Hechicera	Alberto Carnevalli	681	25	51
Alberto Carnevalli	Santa Ana	700	70	48
Santa Ana	La Liria	780	128	70
La Liria	Sor Juana Inés	749	45	414
Sor Juana Inés	La Américas	500	165	298
La Américas	Centro	232	30	232

Tabla 5.4: Intervalos de confianza para la media de los viajes. Hora: 12:00 M - 2:00 PM

Identificador	Nodo	Viajes Origen	Intervalo de confianza ( $O_i$ )	Viajes Destino	Intervalo de confianza ( $D_j$ )
1	Centro	69	(3.4 ; 8.1)	237	(11.82 ; 27.68)
2	Las Américas	316	(19.09 ; 33.58)	312	(19.8 ; 32.2 )
3	Sor Juana Inés	590	(41.1 ; 57.23)	530	(39.19 ; 49.14)
4	La Liria	100	(6.19 ; 10.47)	143	(7.92 ; 15.91)
5	Santa Ana	248	(14.89 ; 26.44)	104	(5.61 ; 11.72 )
6	Alberto Carnevalli	135	(5.15 ; 17.35 )	127	(4.99 ; 16.18)
7	C.C. La Hechicera	53	(2.24 ; 6.59)	104	(5.21 ; 12.12)
8	Terminal La Hechicera	759	(48.32 ; 78.18 )	553	(30.31 ; 61.86)

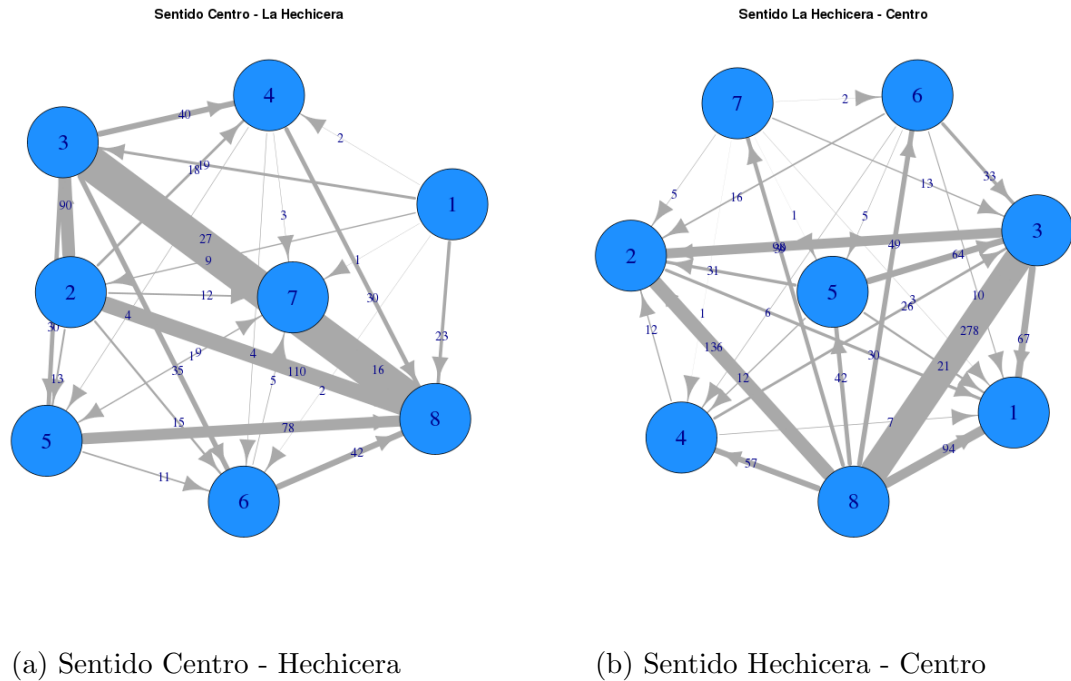


Figura 5.5: Flujo entre cada arco  $(i,j)$ . Hora: 12:00 M - 2:00 PM

### 5.3 Escenario 3

Distribución de viajes en las horas pico tarde desde las 4:00 M - 6:00 PM, abarcando el 90% del transporte público, ajustando un incremento de 20% en el cambio modal

En Wright et al. (2010), para incluir un posible cambio modal desde vehículos privados, usualmente es suficiente añadir un 20 % a la demanda, puesto que será una tasa razonable a la esperada después que el sistema de transporte Trolebús entre en funcionamiento. Para este escenario se muestra una menor cantidad en el flujo de pasajeros en cada uno de los arcos (tabla 5.5) que para las horas consideradas en el escenario 1 y 2. Este tipo de resultados resaltan la importancia de las matrices de viajes para la estimación de la congestión en diferentes puntos de la red, lo que ayudará a priorizar la cantidad de unidades de transporte necesarias para el sistema en las distintas horas del día; como es de esperarse, se observan mayores flujos de pasajeros en dirección Hechicera - Centro (figura 5.6), esto debido a que en las horas

pico de 4:00 PM a 6:00 PM finalizan las jornadas laborales. En la tabla 5.5, la estación Terminal La Hechicera posee la mayor cantidad de viajes originados, con un total de 312 pasajeros y el nodo Centro será el principal destino de los viajes con un total de 288, por otra parte, la estación CC La Hechicera presenta la menor demanda con tan sólo 12 pasajeros que tendran como destinos las estaciones en el sentido La Hechicera - Centro.

Tabla 5.5: Flujo de viajes entre cada arco de la red Centro - La Hechicera. Hora: 4:00PM - 6:00 PM

Origen $i$	Destino $j$	Pasajeros en el arco $(i,j)$	Pasajeros que suben en $i$	Pasajeros que bajan en $j$
Centro	Las Américas	135	135	15
Las Américas	Sor Juana Inés	310	190	192
Sor Juana Inés	La Liria	321	203	47
La Liria	Santa Ana	296	22	59
Santa Ana	Alberto Carnevalli	248	10	56
Alberto Carnevalli	C.C. La Hechicera	208	17	41
C.C. La Hechicera	Terminal La Hechicera	170	3	170
Terminal La Hechicera	C.C. La Hechicera	312	312	12
C.C. La Hechicera	Alberto Carnevalli	326	26	16
Alberto Carnevalli	Santa Ana	401	9	27
Santa Ana	La Liria	443	69	22
La Liria	Sor Juana Inés	497	76	320
Sor Juana Inés	La Américas	360	183	118
La Américas	Centro	288	46	288

El análisis anterior permitió obtener una estimación de la demanda probable en cada estación de la Línea 2 del Trolebús Mérida. La aplicación web desarrollada, reúne los dos modelos ya presentados (generación y distribución), lo cuales podrán ser usados como herramientas de predicción de la demanda de transporte en un amplio sentido (orígenes, destinos y patrón de viajes) para cualquier período del día; además brinda las bases para generar un modelo de simulación de la red de transporte con lo cual se podrán realizar planes de servicio para la Línea 2 del Trolebús Mérida. La herramienta permitirá a los planificadores evaluar proyecciones para otros escenarios, pues el patrón de viajes suele ser distinto cuando se consideran las horas pico (donde la mayoría de

Tabla 5.6: Intervalos de confianza para la media de los viajes. Hora: 4:00PM - 6:00 PM

Identificador	Nodo	Viajes Origen	Intervalo de confianza ( $O_i$ )	Viajes Destino	Intervalo de confianza ( $D_j$ )
1	Centro	176	(9.23 ; 20.11)	288	(15.91 ; 32.09 )
2	Las Américas	308	(19.93 ; 31.4 )	133	(8.26 ; 13.9)
3	Sor Juana Inés	504	( 35.41 ; 48.59)	512	(34.74 ; 50.59)
4	La Liria	127	(8.11 ; 13.06)	69	(2.71 ; 8.79 )
5	Santa Ana	111	(6.04 ; 12.46)	91	(4.5 ; 10.67 )
6	Alberto Carnevali	142	(5.72 ; 17.94 )	73	( 3.71 ; 8.45 )
7	C.C. La Hechicera	39	( 1.49 ; 5.01 )	53	(2.26 ; 6.57)
8	Terminal La Hechicera	411	(22.63 ; 45.87 )	172	(8.62 ; 20.04)

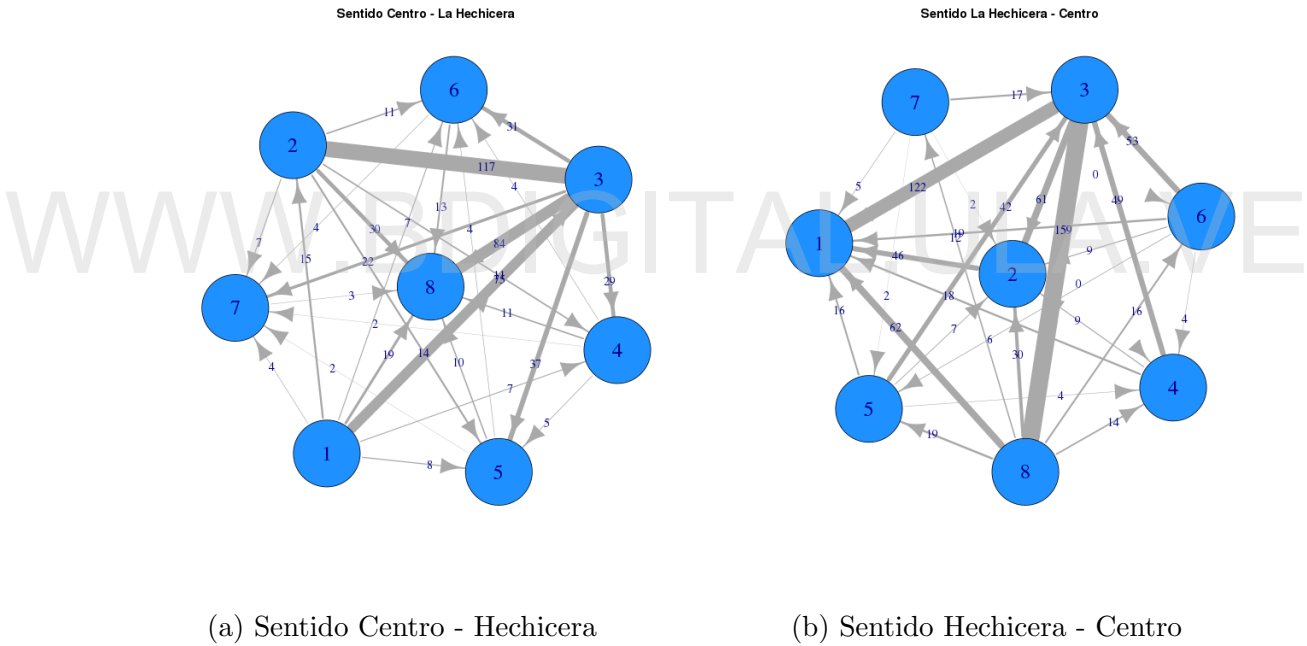


Figura 5.6: Flujo entre cada arco  $(i,j)$ . Hora: 4:00 PM - 6:00 PM

viajes son basados en el trabajo y colegios) y los períodos valle (donde la mayoría de viajes son de compras); además de evaluarse la necesidad de construcción de algunas de las estaciones que no presentan una demanda significativa, por ejemplo, la estación CC La Hechicera, esto según los escenarios de horas pico evaluados; por lo que se sugiere su reubicación a un punto geográfico con mayor concentración de pasajeros.

# Capítulo 6

## Conclusiones

Se ha propuesto un modelo de estimación de la demanda del sistema de transporte, combinando las diversas herramientas estadísticas y de optimización. A diferencia del método tradicional de las cuatro etapas, que se alimenta de múltiples encuestas domiciliarias, en este trabajo se logra estimar la demanda de transporte de la Línea 2 Centro - La Hechicera, demostrando ser una alternativa más rápida y menos compleja. Se han dividido los procesos de modelización de la demanda en dos etapas: el modelo de generación y modelo de distribución de viajes. Para estimar la etapa de generación se han empleado dos modelos de regresión lineal múltiple, donde la variable dependiente es el número de viajes que se originan y atraen en cada nodo de la zona Centro - La Hechicera, previamente estructurada como un grafo de transporte; las variables independientes, a su vez, representan las características urbanas (uso de suelos) de los nodos en su áreas de influencia. El SIG proporcionó las herramientas de cálculo necesarias para delimitar el área de estudio y almacenar la base de datos de las variables independientes. Los modelos obtenidos contribuyen a explicar la relación espacial entre la tipología urbana y la demanda de transporte público. Los resultados obtenidos poseen un ajuste considerablemente elevado con un  $R^2$  mayor a 0.97, lo que confirma que las variables de uso de suelos (*m<sup>2</sup> de uso residencial y de servicios*) contribuyen en la explicación de los incrementos observados en la demanda para cada nodo de la red de transporte, mientras que la variable distancia del nodo al centro de la ciudad tiende a reducir la cantidad de viajes observados y explica que existe mayor congestionamiento

en la zonas centrales y menor cantidad de viajes en las zonas perifericas de la ciudad. Estos dos ajustes ofrecen una nueva perspectiva al planificador urbano, donde se puede evaluar la demanda de una zona geográfica en particular en función de los cambios de su entorno (nuevas infraestructuras residenciales o áreas de servicios) sin necesidad de realizar encuestas domiciliarias. El hecho de combinar herramientas potentes y flexibles como R y QGIS en la planificación de transporte, permitió el análisis, visualización y la estimación rápida de la demanda que se presentará en cada estación futura del Trolebús Mérida, además de conocer la relación del volumen de pasajeros estimados con las variables urbanas del entorno de cada estación. La segunda etapa, el modelo de distribución, fue la base para entender el comportamiento de los flujos de viajes en la red. Por medio del uso del concepto de entropía, se estimó la matriz de viajes más probable de viajes dada la información conocida del sistema. Este modelo fundamenta el análisis de escenarios de la Linea 2 Centro - La Hechicera, pues al conocer el patrón de viajes de la población se podrán desarrollar planes a largo plazo en base a las necesidades de las diversas zonas de transporte que conforman la red.

## 6.1 Recomendaciones

El modelo presentado en este trabajo se basa en la estimación del fenómeno de la demanda para el transporte público. Sin embargo, para estudiar y planificar nuevas políticas de infraestructura y operación de un sistema de transporte masivo, los países desarrollados utilizan un modelo completo del sistema de transporte, por lo que valdría la pena (una vez disponibles los recursos y datos) desarrollar un modelo de cuatro etapas para este tipo de proyectos en la ciudad. No obstante, los resultados obtenidos mediante el modelo presentado en este trabajo son muy útiles para obtener un primer ajuste de la demanda.

La elección de las variables urbanísticas utilizadas en este trabajo se justifica, pues la intención es conocer la relación de los uso de suelos de la zona con su dinámica de viajes. Se deja abierta la posibilidad de ajustar otras variables para predecir la cantidad de viajes de origen y destino en cada nodo respectivamente. En la literatura se ha evaluado la influencia de la información de variables como la posesión de automóvil,

el número de habitantes, el número de empleados, entre otros,(Cardozo et al., 2010) sobre la demanda de transporte.

En relación con el modelo de entropía, se utilizó el supuesto para el factor de equilibrio ( $f_{eqD} = \frac{\sum_{j=1}^n D_j}{\sum_{i=1}^n O_i}$ ), lo que pudo haber introducido error en los resultados de las matrices origen - destino, por lo que sería interesante, por medio de encuestas origen - destino en la zona Centro - La Hechicera, comparar la calidad de las predicciones del modelo de distribución. Por último, pueden calibrarse diferentes modelos de distribución para el área de estudio, usando el modelo gravitacional u otras formulaciones de la entropía.

Como última recomendación para la Empresa Tromerca, debido a que existen diversas zonas residenciales ubicadas a lo largo de las vías secundarias de la red de transporte, es esencial prestarle servicio a estas áreas a través de rutas alimentadoras con el fin de construir un sistema que beneficie a toda la comunidad de la zona Centro - La Hechicera.

WWW.BDIGITAL.ULA.VE

# Bibliografía

Azuaje, W. A. (2006). *La estación de Trolebús Tromerca, en los escenarios de nodo de transporte público urbano, como proyecto arquitectónico interpretador del futuro de su espacio contiguo*. Tesis de Doctorado, Facultad de Arquitectura y Diseño. Universidad de Los Andes Venezuela.

Burnham, K. P. y Anderson, D. R. (2004). Multimodel inference understanding aic and bic in model selection. *Sociological methods & research*, 33(2):261–304. Disponible en [http://www.sortie-nd.org/lme/Statistical%20Papers/Burnham\\_and\\_Anderson.2004\\_Multimodel\\_Inference.pdf](http://www.sortie-nd.org/lme/Statistical%20Papers/Burnham_and_Anderson.2004_Multimodel_Inference.pdf). Fecha de consulta: 14 de Enero de 2017.

Campos, A. P. (1991). El sistema de información geográfica: un instrumento para la planificación y gestión urbana. *Geographica*, (28):175–192. Disponible <https://dialnet.unirioja.es/descarga/articulo/59785.pdf>. Fecha de consulta: 22 de Abril de 2017.

Cardozo, O. D., Gutiérrez, J., y García, J. C. (2010). Influencia de la morfología urbana en la demanda de transporte público: análisis mediante sig y modelos de regresión múltiple. *GeoFocus. Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, (10):82–102. Recuperado de [http://geofocus.rediris.es/2010/Articulo4\\_2010.pdf](http://geofocus.rediris.es/2010/Articulo4_2010.pdf). Fecha de consulta: 22 de Enero de 2017.

Cunat, J. F. S. (2006). Distribuciones de máxima entropía en espacios de probabilidad transformados. Recuperado de <http://exordio.qfb.umich.mx/archivos%20pdf%20de%20trabajo%20umsnh/Leer%20escribir%20PDF%202014/Tesis%20espa%C3%B1olas%202012/entropia.pdf>. Fecha de consulta: 29 de Febrero de 2017.



de Arce, R., Mahía, R., y Definición, I. (2001). Conceptos básicos sobre la heterocedasticidad en el modelo básico de regresión lineal tratamiento con e-views. Disponible en [https://www.uam.es/personal\\_pdi/economicas/anadelsur/pdf/heterocedasticidad.pdf](https://www.uam.es/personal_pdi/economicas/anadelsur/pdf/heterocedasticidad.pdf). Fecha de consulta: 19 de Marzo de 2017.

de Dios Ortúzar, J. y Willumsen, L. (2008). *Modelos de transporte [versión de Google Books]*, volumen 1. Ed. Universidad de Cantabria, España. Disponible en [https://books.google.co.ve/books?hl=es&lr=&id=T5ghp12eCGQC&oi=fnd&pg=PA15&dq=modelos+de+transporte&ots=GPBpFuGzHG&sig=KFPPh1WNgYR94Amb1p6SlKjh1Uw&redir\\_esc=y#v=onepage&q=modelos%20de%20transporte&f=false](https://books.google.co.ve/books?hl=es&lr=&id=T5ghp12eCGQC&oi=fnd&pg=PA15&dq=modelos+de+transporte&ots=GPBpFuGzHG&sig=KFPPh1WNgYR94Amb1p6SlKjh1Uw&redir_esc=y#v=onepage&q=modelos%20de%20transporte&f=false). Fecha de consulta: 17 de Enero de 2017.

Galán, J. (2005). Determinantes de la demanda por transporte público y privado en el área metropolitana de Monterrey. *Ciencia UANL*, 8(4):495–501. Disponible en [http://eprints.uanl.mx/1645/1/art\\_transporte.pdf](http://eprints.uanl.mx/1645/1/art_transporte.pdf). Fecha de consulta: 16 de Enero de 2017.

Gutiérrez, J., Cardozo, O., y García, J. (2013). Estimación directa de la demanda de transporte a nivel de estación mediante el uso de la regresión geográficamente ponderada. *Congreso Chileno de Ingeniería de Transporte*, (16). Recuperado de <http://revistas.uchile.cl/index.php/CIT/article/viewFile/28422/30111>. Fecha de consulta: 18 de Enero de 2017.

Hernández, L. y Molina, I. (2011). *Evolución del uso de la tierra en la ciudad de mérida en el período 1966-2010*. Tesis de Doctorado, Facultad de Ciencias Forestales y Ambientales. Escuela de Geografía. Universidad de Los Andes - Venezuela.

INE (2011). Instituto nacional de estadística - informe censo mérida. <http://www.ine.gov.ve/>. Fecha de consulta: 02 de Abril de 2017.

Lieberman, G. J. y Hillier, F. S. (2010). *Introducción a la Investigación de operaciones*. McGraw-Hill Companies.

López-Ospina, H. (2013). Modelo de maximization de la entropía y costos generalizados intervalares para la distribución de viajes urbanos. *Ingeniería y Universidad*,

- 17(2):391–407. Disponible <http://www.scielo.org.co/pdf/inun/v17n2/v17n2a08.pdf>. Fecha de consulta: 21 de Abril de 2017.
- Martin, W. A. y McGuckin, N. A. (1998). *Travel estimation techniques for urban planning*, volumen 365. National Academy Press Washington, DC. Disponible en [http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp\\_rpt\\_365.pdf](http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_365.pdf). Fecha de consulta: 12 de Abril de 2017.
- Mauttone Vidales, A. D. (2002). Modelos de demanda de transporte. *Reportes Técnicos 02-17*. Disponible en <https://www.colibri.udelar.edu.uy/bitstream/123456789/3484/1/TR0217.pdf>. Fecha de consulta: 24 de Abril de 2017.
- Monge, J. A. (2011). Planificación del transporte. 2(19). Recuperado de [http://www.lanamme.ucr.ac.cr/sitio-nuevo/images/boletines/boletin\\_prita\\_19\\_planificacion\\_transporte.pdf](http://www.lanamme.ucr.ac.cr/sitio-nuevo/images/boletines/boletin_prita_19_planificacion_transporte.pdf). Fecha de consulta: 13 de Marzo de 2017.
- Montgomery, D. C. (2005). Diseño y análisis de experimentos.
- Moreno, E. (2006). Análisis comparativo de la modelación del autotransporte: carga vs pasajeros. *Publicación Técnica (300)*. Instituto Mexicano del Transporte. Disponible en <http://imt.mx/archivos/Publicaciones/PublicacionTecnica/pt300.pdf>. Fecha de consulta: 13 de Enero de 2017.
- Moura Berodia, J. L., Ibeas Portilla, Á., Dell'Olio, L., Vega Zamanillo, Á., et al. (2006). Dos modelos de distribución de viajes: Modelo de entropía (hyman) versus triproporcional a través de una aplicación en la ciudad de concepción (chile). Consultado en <https://repositorio.unican.es/xmlui/handle/10902/1998>. Fecha de consulta: 27 de Febrero de 2017.
- Ornés, S. (2009). El urbanismo, la planificación urbana y el ordenamiento territorial desde la perspectiva del derecho urbanístico venezolano\* urban development, urban planning and. *Politeia*, 32(42):197–225. Recuperado de [https://www.researchgate.net/profile/Sandra\\_Ornes/publication/237025793\\_El\\_urbanismo\\_la\\_planificacin\\_urbana\\_y\\_el\\_ordenamiento\\_territorial\\_desde\\_la\\_perspectiva\\_del\\_derecho\\_urbanstico\\_](https://www.researchgate.net/profile/Sandra_Ornes/publication/237025793_El_urbanismo_la_planificacin_urbana_y_el_ordenamiento_territorial_desde_la_perspectiva_del_derecho_urbanstico_)

- venezolano/links/54078ee40cf23d9765ac2254.pdf. Fecha de consulta: 30 de Marzo de 2017.
- Puebla, J., Cardozo, O., y García, J. (2008). Modelos de demanda potencial de viajeros en redes de transporte público: aplicaciones en el Metro de Madrid. *IV Seminario de Ordenación Territorial*, 1(4). Recuperado de: [http://m.bdigital.uncu.edu.ar/objetos\\_digitales/3242/gutierrezproyeccion4.pdf](http://m.bdigital.uncu.edu.ar/objetos_digitales/3242/gutierrezproyeccion4.pdf). Fecha de consulta: 13 de Enero de 2017.
- Puebla, J. G. y Gould, M. (1994). *SIG: Sistemas de información geográfica*. Síntesis.
- Pulido, H. G. y De la Vara Salazar, R. (2008). *Análisis y diseño de experimentos*. McGraw-Hill.
- Quintero, A., Angulo, C., y Guerrero, J. (2011). Determinación de tasas de generación de viajes para conjuntos residenciales ubicados en la ciudad de Mérida, Venezuela. *Ciencia e Ingeniería*, 32(1):45–54. Recuperado de: <http://erevistas.saber.ula.ve/index.php/cienciaeingenieria/article/view/3252/3156>. Fecha de consulta: 13 de Enero de 2017.
- Rosas, A. M. (2012). *Estimación de tasas de generación de viajes para hospitales en el Distrito Metropolitano de Caracas*. Tesis de Doctorado, Universidad Simón Bolívar - Postgrado en Transporte Urbano.
- Salas, J. d. D. O. y García, C. R. (2003). Análisis y predicción de la demanda de transporte de pasajeros. Recuperado de [https://www.researchgate.net/profile/Raquel\\_Espino/publication/38182699\\_Analisis\\_y\\_prediccion\\_de\\_la\\_demanda\\_de\\_transporte\\_de\\_pasajeros\\_una\\_aplicacion\\_al\\_estudio\\_de\\_dos\\_corredores\\_de\\_transporte\\_en\\_Gran\\_Canaria/links/542d2f8d0cf29bbc126d1f6b.pdf](https://www.researchgate.net/profile/Raquel_Espino/publication/38182699_Analisis_y_prediccion_de_la_demanda_de_transporte_de_pasajeros_una_aplicacion_al_estudio_de_dos_corredores_de_transporte_en_Gran_Canaria/links/542d2f8d0cf29bbc126d1f6b.pdf). Fecha de consulta: 28 de Marzo de 2017.
- Taha, H. A. (2004). *Investigación de operaciones*. Pearson Educación.
- Tromerca (2017). Tromerca, oficina de tecnología de información y comunicación. <http://www.tromerca.gob.ve/>. Fecha de consulta: 11 de Abril de 2017.

- ULA (2016). Informe de gestión 2015 planes proyección 2016. [http://www.ula.ve/images/pdf/2016/informe\\_de\\_gestion\\_ula\\_2015.pdf](http://www.ula.ve/images/pdf/2016/informe_de_gestion_ula_2015.pdf). Fecha de consulta: 02 de Abril de 2017.
- Vozzi, L. y Acquaviva, L. (2011). Modelización del sistema de transporte de Rosario. *Asociación Argentina de Mecánica Computacional*, 30:2891–2903. Recuperado de <http://www.cimec.org.ar/ojs/index.php/mc/article/viewFile/3959/3876>. Fecha de consulta: 16 de Enero de 2017.
- Wright, L., Hook, W., y Pardo, C. F. (2010). Guía de planificación de sistemas brt. *Institute for Transportation & Development Policy, New York*. Disponible en <https://www.itdp.org/wp-content/uploads/2014/07/21.-BRT-Guide-Spanish-Parte-1.pdf>. Fecha de consulta: 11 de Enero de 2017.

WWW.BDIGITAL.ULA.VE

# Apéndice A

## Bibliotecas de R

A continuación, se documentan las principales bibliotecas de R utilizadas en este trabajo para la generación de los resultados y validaciones estadísticas de los modelos, además de la herramienta Shiny de Rstudio, con la que se facilitó el diseño de la interfaz web.

### A.1 Paquete: shiny v.1.0.3

Título: Framework para aplicaciones web para R.

Descripción: Biblioteca que hace que sea increíblemente fácil el proceso de construir aplicaciones web interactivas con R, bajo el concepto de enlace automático reactivo entre entradas y salidas.

Licencia: GPL-3.

Dependencias: R ( $\geq 3.0.0$ ).

URL: <http://shiny.rstudio.com>.

Mantenedor: Winston Chang [winston@rstudio.com](mailto:winston@rstudio.com).

Repositorio: CRAN.

Fecha publicación: 2017-04-26 17:13:08 UTC.

## A.2 Paquete: igraph v.1.0.1

Título: Análisis y visualización de redes de flujo.

Descripción: Rutinas para gráficos y análisis de redes. Proporciona funciones para generar gráficos aleatorios y regulares, métodos de centralidad, entre otros.

Licencia: GPL 2.

Dependencias: gmp, libxml2.

URL: <http://igraph.org>.

Mantenedor: Gabor Csardi [csardi.gabor@gmail.com](mailto:csardi.gabor@gmail.com).

Repositorio: CRAN.

Fecha publicación: 2015-06-26 11:13:24.

## A.3 Paquete: lmtest v.0.9-35

Título: Prueba de modelos de regresión lineal.

Descripción: Una colección de pruebas, conjuntos de datos y ejemplos para la verificación y diagnóstico en modelos de regresión lineal. Además, se proporcionan algunas herramientas genéricas para la inferencia en modelos paramétricos.

Licencia: GPL-2, GPL-3.

Dependencias: R ( $\geq 2.10.0$ ), stats, zoo.

URL: <http://igraph.org>.

Mantenedor: Achim Zeileis [achim.zeileis@r-project.org](mailto:achim.zeileis@r-project.org).

Repositorio: CRAN.

Fecha publicación: 2017-02-11 10:43:09.

# Apéndice B

## Codigo Fuente: Matriz O-D

Con respecto a la aplicación web desarrollada en este trabajo, para generar matrices de viajes origen - destino a partir de la teoría de máxima entropía, se brinda a continuación una visión general de la estructura de la misma y las funciones de su código. Como fue mencionado en el capítulo 5, el código fuente está disponible de forma pública en un repositorio de *Github*, para tener acceso: (<https://github.com/silva1221/distribucion-viajes-linea2-trolebus-merida.git>).

### B.1 Archivos de código fuente

**ui.R**: archivo fuente donde se encuentra la definición de la interfaz de usuario de la aplicación.

**server.R**: lado servidor de la aplicación donde se define la lógica de la aplicación.

**viajes.csv** : contiene la base de datos de los viajes entre nodos, los cuales serán usados para obtener la matriz origen - destino.

**inicializar\_nodos.R**: inicializaciones de algunos componentes necesarios para graficar los grafos de la red de transporte.

**calculo\_rangos.R**: contiene la definición de una función que calcula los índices de los datos requeridos en la base de datos, tomando como entrada el intervalo de horas del día seleccionado por el usuario.

**matriz\_viajes.R** : contiene el algoritmo de estimación de matriz origen - destino, siguiendo la metodología de la maximización de la entropía. Los parametros de entrada se definen como el total de viajes originados y atraídos en cada nodo en el intervalo de tiempo solicitado por el usuario.

**dibujar\_grados.R**: archivo que genera todos los graficos de la red de transporte con el apoyo de la biblioteca *igraph*.

La figura B.1, muestra la vista principal de la aplicación web de *Shiny*, la misma presenta el resultado para la matriz de distribución de viajes en el intervalo de tiempo: 7:00 AM a 6:00 PM, se visualizan además los *input sliders* que permiten cambiar los parámetros del transporte público cubierto por el Trolebús y el incremento ajustado en el cambio modal.



Figura B.1: Pantalla principal de la aplicación Matriz O-D

El siguiente código del archivo `matriz_viajes.R`, corresponde a la función de maximización de la entropía usado para estimar las matrices de viajes. El mismo se basa en el algoritmo presentado en la sección 4.3.

```
matrizViajes <- function(Origen, Destino){

  #Declaracion de Matriz Semilla de 1nos y restricciones de
  diagonal nula
```



```
NroNodos<-length(Origen)
MatrizViajes <- matrix(1,NroNodos,NroNodos)
for (i in 1:NroNodos){
  MatrizViajes[i,i] = 0

#Inicializacion de vectores de expansion de iteraciones
ExpansionFilas <- c(rep(1, NroNodos))
ExpansionColumna <- c(rep(1, NroNodos))
ExpansionFilasAnt <- c(rep(0, NroNodos))
ExpansionColumnaAnt <- c(rep(0, NroNodos))
SumasFilas <- c()
SumasColumnas <- c()

#Repetir hasta lograr convergencia
while (sum(MatrizViajes) != sum(Origen)) {
  for(i in 1:NroNodos)
    SumasFilas[i] <- sum(MatrizViajes[i,])

  for(i in 1:NroNodos)
    if((SumasFilas[i])==0)
      ExpansionFilas[i] <- 0
    else
      ExpansionFilas[i]<-Origen[i]/SumasFilas[i]

  for(i in 1:NroNodos)
    MatrizViajes[i,]<-(MatrizViajes[i,]*ExpansionFilas[i])

  for(i in 1:NroNodos)
    SumasColumnas[i] <- sum(MatrizViajes[,i])

  for(i in 1:NroNodos)
    if((SumasColumnas[i])==0)
      ExpansionColumna[i] <- 0
```

```
else
  ExpansionColumna[i] <- Destino[i] / SumasColumnas[i]

for(i in 1:NroNodos)
  MatrizViajes[,i] <- (MatrizViajes[,i] * ExpansionColumna[i]
  )

if (all(ExpansionFilas == ExpansionFilasAnt) & all(
  ExpansionColumna == ExpansionColumnaAnt))
  break

MatrizViajes <- trunc(MatrizViajes)
ExpansionFilasAnt <- ExpansionFilas
ExpansionColumnaAnt <- ExpansionColumna
}

MatrizViajes <- trunc(MatrizViajes)

for(i in 1:NroNodos){
  for(j in 1:NroNodos){
    if (i == j)
      next
    if(sum(MatrizViajes[i,]) == Origen[i])
      break

    if((sum(MatrizViajes[i,]) < Origen[i]) & (sum(
      MatrizViajes[,j]) < Destino[j]))
      MatrizViajes[i,j] = MatrizViajes[i,j] + 1
    }
  }
  return(MatrizViajes)
}
```

## Apéndice C

# Estadísticas y Datos para Simulación

A fin de poder estudiar el sistema del transporte de la Línea 2 bajo otros criterios, se presentan a continuación los datos de los viajes capturados durante el muestreo de movilidad para cada nodo de transporte público, además se muestra el ajuste obtenido para la distribución de los datos de origen y destino para la ruta Centro - La Hechiera, los mismo sirven como base para generar un modelo de simulación a fin de contrastar las estimaciones desarrolladas en este trabajo y formular estrategias de operación de la Línea 2 del Trolebús Mérida. Cabe notar que las mismas fueron separadas en periodos mañana (tablas C.1 y C.2) y tarde (tablas C.3 y C.4), pues en ambos intervalos la demanda difiere en su comportamiento en las diferentes paradas, tal como se refleja evidentemente en las paradas de Sor Juana Inés y La Liria. En la tabla C.5 se muestran las principales estadísticas de cada nodo en los viajes origen y destino observados en intervalos de 10 min desde las 7:00 AM a 6:00 PM.

Tabla C.1: Ajustes de distribuciones sentido Centro - La Hechicera. Periodo mañana: 7:00 AM - 12:00 M

Parada	Viajes origen	Viajes destino
Centro	$Gamma(\alpha = 1.201; \beta = 5.5217)$	$Gamma(\alpha = 1.683; \beta = 5.5978)$
Yuan-Lin	$Gamma(\alpha = 4.7624; \beta = 3.1828)$	$Gamma(\alpha = 3.3705; \beta = 2.3501)$
Sor Juana Inés	$Gamma(\alpha = 6.3484; \beta = 2.2862)$	$Gamma(\alpha = 6.2617; \beta = 3.8118)$
Facijup	$Gamma(\alpha = 1.0421; \beta = 3.2323)$	$Gamma(\alpha = 1.3011; \beta = 4.5712)$
Santa Ana Sur	$Gamma(\alpha = 1.0018; \beta = 3.3362)$	$Gamma(\alpha = 0.96865; \beta = 4.6456)$
Santa Ana Norte	$Gamma(\alpha = 0.59754; \beta = 7.1345)$	$Gamma(\alpha = 0.63793; \beta = 4.7027)$
La Hechicera	$Gamma(\alpha = 0.61238; \beta = 2.6643)$	$Gamma(\alpha = 0.88507; \beta = 4.2815)$
Núcleo La Hechicera	$Gamma(\alpha = \alpha : 1.6137; \beta = 15.247)$	$Gamma(\alpha = 2.4145; \beta = 5.6567)$

Tabla C.2: Ajustes de distribuciones sentido La Hechicera - Centro. Periodo mañana: 7:00 AM - 12:00 M

Parada	Viajes origen	Viajes destino
Centro	$Gamma(\alpha = 0.2376; \beta = 6.5346)$	$Gamma(\alpha = 1.5944; \beta = 6.0077)$
Yuan-Lin	$Gamma(\alpha = 3.3705; \beta = 2.3501)$	$Gamma(\alpha = 2.1465; \beta = 4.291)$
Sor Juana Inés	$Gamma(\alpha = 5.4312; \beta = 4.7968)$	$Gamma(\alpha = 16.376; \beta = 0.96741)$
Facijup	$Gamma(\alpha = 2.7; \beta = 1.6667)$	$Gamma(\alpha = 1.4868; \beta = 2.0531)$
Santa Ana Sur	$Gamma(\alpha = 2.8413; \beta = 4.2141)$	$Gamma(\alpha = 2.5989; \beta = 1.8359)$
Santa Ana Norte	$Gamma(\alpha = 0.78442; \beta = 6.609)$	$Gamma(\alpha = 1.0162; \beta = 3.6255)$
La Hechicera	$Gamma(\alpha = 0.52158; \beta = 3.1282)$	$Gamma(\alpha = 0.73452; \beta = 3.2961)$
Núcleo La Hechicera	$Gamma(\alpha = 1.8838; \beta = 12.384)$	$Gamma(\alpha = 1.2476; \beta = 13.879)$

Tabla C.3: Ajustes de distribuciones sentido Centro - La Hechicera. Periodo tarde: 12:00 M - 06:00 PM

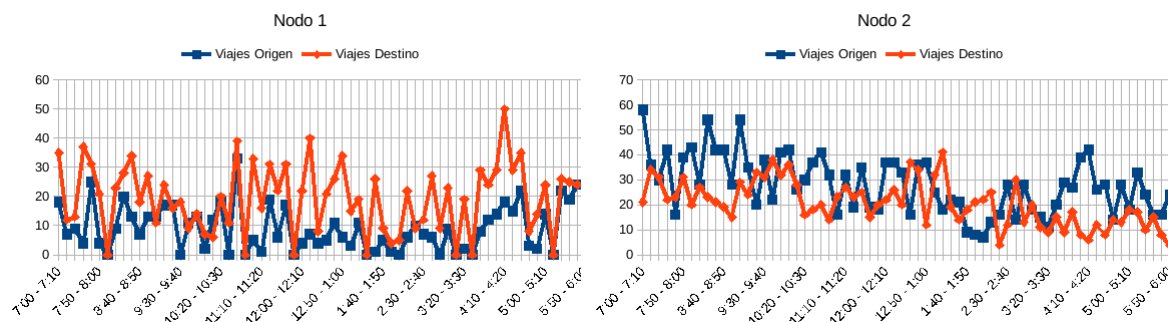
Parada	Viaje origen	Viajes destino
Centro	$Gamma(\alpha = 0.89901; \beta = 6.8223)$	$Gamma(\alpha = 1.2598; \beta = 7.5408)$
Yuan-Lin	$Gamma(\alpha = 7.76; \beta = 3.6006)$	$Gamma(\alpha = 9.5777; \beta = 1.4304)$
Sor Juana Inés	$Gamma(\alpha = 1.7371; \beta = 4.1063)$	$Gamma(\alpha = 7.811; \beta = 5.5307)$
Facijup	$Gamma(\alpha = 0.9728; \beta = 2.981)$	$Gamma(\alpha = 3.1648; \beta = 8.0889)$
Santa Ana Sur	$Gamma(\alpha = 1.6994; \beta = 3.0795)$	$Gamma(\alpha = 1.6618; \beta = 4.1922)$
Santa Ana Norte	$Gamma(\alpha = 1.078; \beta = 3.4324)$	$Gamma(\alpha = 16237; \beta = 3.5106)$
La Hechicera	$Gamma(\alpha = 0.33749; \beta = 3.7532)$	$Gamma(\alpha = 0.4516; \beta = 4.8715)$
Núcleo La Hechicera	$Gamma(\alpha = 3.3887; \beta = 8.6266)$	$Gamma(\alpha = 3.3887; \beta = 8.6266)$

Tabla C.4: Ajustes de distribuciones sentido La Hechicera - Centro. Periodo tarde: 12:00 M - 06:00 PM

Parada	Viajes origen	Viajes destinos
Centro	$Gamma(\alpha = 0.74031; \beta = 5.8083)$	$Gamma(\alpha = 1.906; \beta = 5.2816)$
Yuan-Lin	$Gamma(\alpha = 1,976; \beta = 32051)$	$Gamma(\alpha = 5.5925; \beta = 1.955)$
Sor Juana Inés	$Gamma(\alpha = 4.906; \beta = 4.3927)$	$Gamma(\alpha = 5.3013; \beta = 2.949)$
Facijup	$Gamma(\alpha = 4.7298; \beta = 1.6287)$	$Gamma(\alpha = 1.3828; \beta = 6.6051)$
Santa Ana Sur	$Gamma(\alpha = 4 - 0519; \beta = 5.084)$	$Gamma(\alpha = 1.3799; \beta = 4.01)$
Santa Ana Norte	$Gamma(\alpha = 3.5465; \beta = 1.1801)$	$Gamma(\alpha = 1.4579; \beta = 1.9663)$
La Hechicera	$Gamma(\alpha = 0.69048; \beta = 1.4483)$	$Gamma(\alpha = 0.89834; \beta = 3,7103)$
Núcleo La Hechicera	$Gamma(\alpha = 0.95687; \beta = 10.486)$	$Gamma(\alpha = 1.8149; \beta = 19.468)$

Tabla C.5: Principales estadísticas de los viajes observados

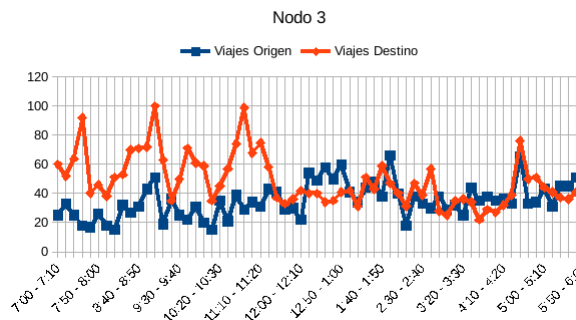
Nodo	Viaje	Mediana	Media	Varianza	Desv. estándar	Coef. variación	Curtosis	Mínimo	Máximo
1	Orígenes	7	9.196	61.975	7.872	0.717	-0.097	0	33
	Destinos	21	19.363	136.050	11.664	0.059	-0.506	0	50
2	Orígenes	28	28	135.076	11.622	0.347	-0.351	7	58
	Destinos	20	20.530	78.591	8.865	0.242	-0.580	4	41
3	Orígenes	34	35.287	143.254	11.968	0.497	-0.007	15	66
	Destinos	42.5	48.560	298.496	17.277	1.101	1.011	22	100
4	Orígenes	9	8.636	22.696	4.764	0.162	0.116	0	23
	Destinos	15.5	20.533	364.006	19.078	1.313	0.812	0	76
5	Orígenes	19	20	133.907	11.571	1.132	1.573	2	56
	Destinos	10.5	10.560	39.942	6.320	0.614	-0.038	0	29
6	Orígenes	6	8.424	51.755	7.194	1.343	1.914	0	33
	Destinos	6	7.318	29.327	5.415	1.573	4.097	0	29
7	Orígenes	3	2.833	7.248	2.692	0.797	-0.202	0	10
	Destinos	5	5.954	21.859	4.672	0.531	-0.667	0	17
8	Orígenes	28	34.151	503.330	22.435	1.063	1.250	4	115
	Destinos	38.5	46.196	968.283	31.117	1.127	0.906	5	139



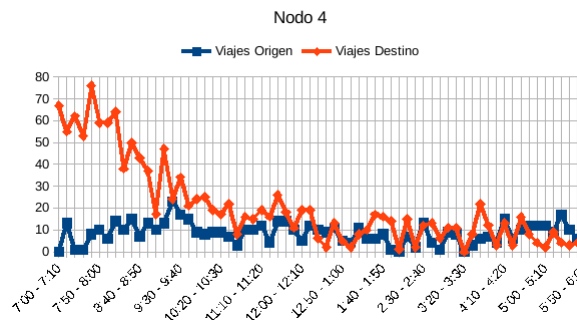
(a) Captura de viajes. Zona Centro

(b) Captura de viajes. Zona Yuan Lin

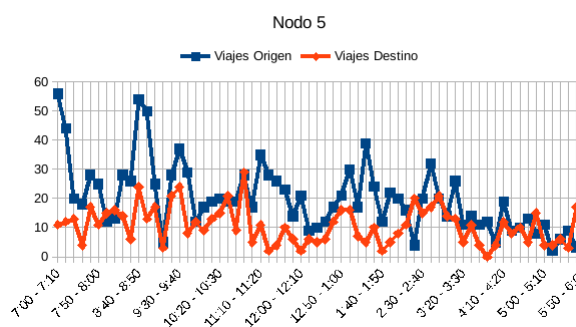
Figura C.1: Captura de viajes 7:00 AM - 6:00 PM



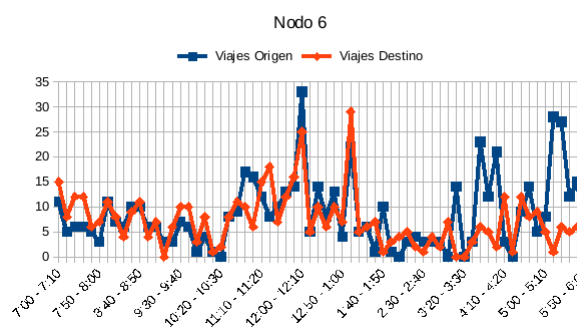
(a) Captura de viajes. Zona Sor Juana Inés



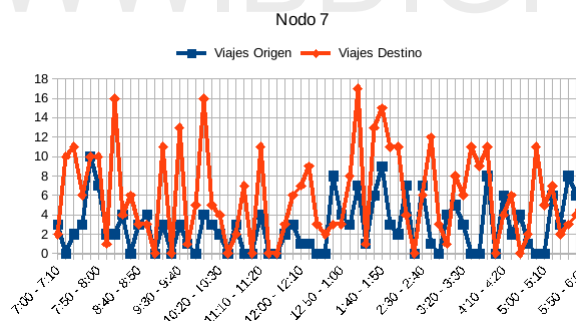
(b) Captura de viajes. Zona FACIJUP



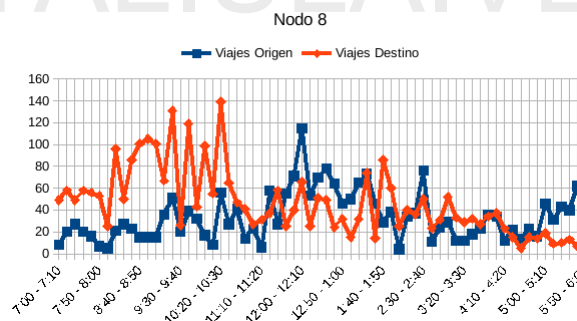
(c) Captura de viajes. Zona Santa Ana Sur



(d) Captura de viajes. Zona Santa Ana Norte



(e) Captura de viajes. Zona C.C. La Hechicera



(f) Captura de viajes. Zona Núcleo La Hechicera

Figura C.2: Captura de viajes 7:00 AM - 6:00 PM

# Apéndice D

## Pruebas Estadísticas en la Validación de Modelos

### D.1 Shapiro - Wilks

En Pulido y De la Vara Salazar (2008), para la prueba de Shapiro - Wilks, se considera una muestra aleatoria de datos  $x_1, x_2, x_3, \dots, x_n$ , que proceden de cierta distribución desconocida denotada como  $F(x)$ , se desea verificar si dichos datos fueron generados por un proceso normal, mediante la siguiente hipótesis:

$$H_0 : F(x) \text{ es normal}$$

$$H_A : F(x) \text{ no es normal}$$

El estadístico a comprobar en este caso, está dado por:

$$w = \frac{1}{(n-1)s^2} \left[ \sum_{i=1}^k a_i (x_{(n-i+1)} - x_{(i)}) \right]^2$$

donde  $S^2$ , es la varianza muestral,  $x_{(1)}, x_{(2)}, x_{(3)}, \dots, x_{(i)}$ , son los datos ordenados de menor a mayor,  $a_1, a_2, \dots, a_k$  son los coeficientes de tablas donde  $k$  es aproximadamente  $n/2$ . A partir de lo anterior, se acepta la hipótesis nula si el valor del estadístico es mayor a su nivel crítico al nivel  $\alpha$ .



## D.2 Durbin - Watson

Según [Pulido y De la Vara Salazar \(2008\)](#), esta prueba permite diagnosticar la presencia de autocorrelación entre los residuos consecutivos (ordenados en el tiempo), lo cual contradice el supuesto de independencia, sea  $\rho = (\varepsilon_t, \varepsilon_{t-1})$ ;  $t = 1, 2, 3, \dots, n$  el parámetro que representa la correlación entre residuos consecutivos. La prueba de Durbin - Watson es:

$$H_0 : \rho = 0$$

$$H_A : \rho > 0$$

En base a  $\varepsilon_i, i = 1, 2, 3, \dots, n$ , definidos como los residuos ordenados en el tiempo, se puede definir el estadístico Durbin - Watson como:

$$d = \frac{\sum_{i=2}^n (\varepsilon_i - \varepsilon_{i-1})^2}{\sum_{i=1}^n (\varepsilon_i)^2}$$

La decisión sobre la hipótesis esta dada por:

$$d < d_L, \text{ se rechaza } H_0$$

$$d > d_U, \text{ se acepta } H_0$$

$$d_L \leq d \leq d_U, \text{ no hay decisión}$$

Siendo  $d_L$  y  $d_U$ , estadísticos tabulados a partir de los  $n$  residuos, el nivel de significancia prefijado  $\alpha$ , y el número de variables explicativas del modelo. La prueba de Durbin Watson tiene el inconveniente de que no detecta autocorrelación entre residuos no consecutivos en el tiempo que también violan los supuestos del modelo de regresión lineal.

## D.3 Criterio de información de Akaike

Para [Burnham y Anderson \(2004\)](#), el AIC, proporciona un medio para la selección del modelo, basado en la teoría de la información. En el caso general, la AIC está dado por:

$$AIC = 2k - \ln(L)$$

donde  $k$ , es el número de parámetros en el modelo, y  $L$  es el valor máximo de la función de verosimilitud para el modelo estimado. La forma de la función AIC, permite penalizar el aumento en el número de parámetros de un modelo, desalentando el sobreajuste. El mejor modelo, por lo tanto, será aquel que posea menor valor AIC.

## D.4 Breusch - Pagan

En [de Arce et al. \(2001\)](#), la finalidad de la prueba de Breusch - Pagan, es analizar si la varianza estimada de los residuos de una regresión dependen de los valores de las variables independientes. Suponiendo que se ha estimado el siguiente modelo lineal:  $(y = \beta_0 + \beta_1 x + \varepsilon)$ , para obtener el conjunto de valores para  $\hat{\varepsilon}$ , se contruye la siguiente ecuación donde se componen los errores por el valor de la constante  $\beta_0$  y el conjunto de variables regresoras  $x$ :

$$(\hat{\varepsilon}_i)^2 = \gamma_0 + \gamma_1 x + v$$

La base de la prueba es suponer que las variables independientes no son significativas, por lo tanto la varianza de los errores no depende de las variables independientes, la estimación de la varianza se obtiene a partir de la media de los valores al cuadrado. Si la suposición no fuera correcta, entonces se puede rechazar la hipótesis nula de homocedasticidad.