

INSTITUTO TECNOLÓGICO AUTÓNOMO DE MÉXICO



Pronóstico de la Demanda y Balanceo Iterativo de un Sistema de Préstamo de Bicicletas con Estaciones Fijas

TESIS

QUE PARA OBTENER EL TÍTULO DE

MAESTRO EN CIENCIAS EN COMPUTACIÓN

PRESENTA

DAVID EDGARDO CASTILLO RODRÍGUEZ

ASESOR: DR. LUIS ANTONIO MONCAYO MARTÍNEZ

México, D.F.

2021

“Con fundamento en los artículos 21 y 27 de la Ley Federal del Derecho de Autor y como titular de los derechos moral y patrimonial de la obra titulada “**Pronóstico de la Demanda y Balanceo Iterativo de un Sistema de Préstamo de Bicicletas con Estaciones Fijas**”, otorgo de manera gratuita y permanente al Instituto Tecnológico Autónomo de México y a la Biblioteca Raúl Baillères Jr., la autorización para que fijen la obra en cualquier medio, incluido el electrónico, y la divulguen entre sus usuarios, profesores, estudiantes o terceras personas, sin que pueda percibir por tal divulgación una contraprestación”.

DAVID EDGARDO CASTILLO RODRÍGUEZ

2021

FIRMA

DEDICATORIA

Dedicada a todos mis familiares y amigos que me han apoyado en cada momento. También, dedicada a todos aquellos profesores y personas que conocí gracias a la MCC.

Índice general

1. Introducción	1
1.1. Sistemas de préstamos de bicicletas	3
1.2. El problema de investigación	4
1.3. Objetivo	5
1.4. Metodología de solución	5
1.5. Resumen de resultados	5
2. Revisión bibliográfica	7
2.1. Sistemas de préstamos de bicicletas y problemas identifi- cados	7
2.2. Métodos de solución	9
2.2.1. Estimación de demanda	9
2.2.2. Balanceo	11
3. Metodología de solución	13
3.1. Planteamiento del problema	13
3.1.1. Descripción de la información	16
3.2. Estimación de Oferta y Demanda	17
3.2.1. Índice de saturación	19
3.2.2. Cicloestaciones con escasos o saturación	21
3.2.3. Estimación de Oferta al tiempo t_i	26
3.2.4. Estimación de Demanda al tiempo t_i	29
3.2.5. Otra alternativa de estimación	30
3.3. Balanceo de la oferta y la demanda	32
3.3.1. Ejemplo del problema de transporte	33
3.4. El problema de ruteo	36

3.4.1.	Algoritmo de solución multi agente viajero . . .	43
3.4.2.	Ejemplo del multi-agente viajero	46
4.	Resultados y discusión	52
4.1.	Caso Ecobici	52
4.1.1.	Estaciones oferentes y demandantes en el tiempo t	52
4.1.2.	Porcentaje de llenado de las estaciones en el tiempo t	66
4.1.3.	Tiempo promedio del balanceo en el tiempo t . .	77
4.2.	Discusión	78
4.3.	Explicación de los resultados	78
4.4.	Posición de este trabajo en la literatura	79
4.5.	Deducciones	80
5.	Conclusiones	82
5.1.	Implicaciones del trabajo	82
5.2.	Importancia	83
5.3.	Lecciones aprendidas	84
5.4.	Replicabilidad de los resultados	85
5.4.1.	Implementación en un ambiente productivo . . .	87
	Bibliografía	90

Capítulo 1

Introducción

Hoy en día, muchas industrias o empresas de servicios se enfrentan al problema de la asignación eficiente de algún tipo de recurso, como la cantidad de dinero y el tiempo en el que un banco debe reabastecer algunos de sus cajeros automáticos, la ruta que debe seguir un camión repartidor para garantizar la entrega de un producto en un tiempo mínimo (o gastando la menor cantidad de combustible), el rebalanceo de un sistema de bicicletas compartidas para garantizar que la demanda sea cubierta, por ejemplo. Estos son sólo algunos problemas donde es conveniente analizar de qué manera debe hacerse la asignación de recursos (o de qué manera debe construirse una estrategia de asignación de recursos) para garantizar un buen servicio, o minimizar los costos asociados a la asignación de recursos.

La asignación eficiente de recursos puede plantearse mediante modelos de programación matemática, uno de los principales matemáticos que aportó y cuyo trabajo ha servido como base para el desarrollo de la teoría de tales modelos, fue George Bernard Dantzing (1914-2005), quien en [3] describió un modelado matemático para encontrar la ruta óptima para el envío de camiones de gasolina desde una central a varias estaciones de servicio. Dicho modelo ha servido como punto de partida para el desarrollo de nuevos modelos y algoritmos que puedan dar una solución óptima o una solución aproximada al problema de la asignación eficiente de recursos.

Los problemas de asignación de recursos pueden llegar a ser muy complicados, ya sea por el número de agentes que intervienen, el número de variables involucradas, o la aleatoriedad involucrada al tipo de bien que requiere asignarse, como es el caso de problemas que tanto su oferta y demanda son aleatorias. En estos últimos problemas, antes de plantear el problema de asignación, es necesario resolver el problema de estimación de la oferta y la demanda, de esta forma, una vez que se ha logrado estimar tanto la oferta como la demanda el siguiente paso es encontrar una ruta factible para hacer la reasignación eficiente de recursos.

Los problemas de reasignación se describen como: dado un conjunto de restricciones, encontrar la ruta óptima dentro de un grafo, o bien, dado un conjunto de restricciones encontrar un camino hamiltoniano, (nombrados así por el juego de los 20 caminos propuesto por el matemático Sir William R. Hamilton). Este último problema fue clasificado como *NP-completo* por el informático teórico Richard Karp [1], lo cual hace sumamente costoso encontrar soluciones exactas en tiempos eficientes para las necesidades del problema. De acuerdo a [2], un problema *NP-completo* se caracteriza por no tener aún un algoritmo que lo pueda resolver en tiempo polinomial, y estos pueden identificarse de acuerdo a las siguientes dos propiedades:

1. Que se pueda verificar en tiempo polinomial que una solución propuesta del problema en consideración, es en efecto, una solución del problema.
2. Otros problemas que pertenecen a la clase **NP** pueden ser reducidos en tiempo polinomial al problema en consideración.

En resumen, aún no se sabe si existen o no algoritmos eficientes (que encuentre soluciones exactas en tiempo polinomial) para los problemas clasificados como *NP-completos*, por ello, obtener soluciones exactas para dichos problemas es computacionalmente costoso, como se mencionó anteriormente. Sin embargo, gracias a los avances computacionales, hasta la fecha se han desarrollado una gran cantidad de algoritmos heurísticos los cuales si bien no encuentran la solución óptima exacta del pro-

blema, encuentran una solución aproximada de esta, de esta manera, se obtienen soluciones factibles en tiempos eficientes; algunos algoritmos de este tipo son el algoritmo de búsqueda tabú, algoritmos genéticos, búsqueda ambiciosa, por ejemplo.

Dicho lo anterior, en este trabajo se estudia un sistema de préstamo de bicicletas compartidas, la implementación de este tipo de sistemas han traído consigo una serie de problemas que pueden estudiarse para mejorar el servicio. Como se menciona en [6]-[18], estos problemas están relacionados con la estimación de la oferta y demanda (la cuál es estocástica), así como en el problema de rebalanceo. Para ello, en este capítulo se introducen estos sistemas así como la metodología de solución implementada en este trabajo. En el capítulo 2 se habla acerca de la revisión bibliográfica consultada para la elaboración de este trabajo. En el capítulo 3 se habla acerca de la metodología utilizada durante este trabajo. En el capítulo 4 se habla acerca de los resultados obtenidos. En el capítulo 5 se presentan las conclusiones obtenidas, también se habla brevemente acerca de una alternativa para implementar la metodología utilizada a lo largo de este trabajo en un ambiente productivo.

1.1. Sistemas de préstamos de bicicletas

Actualmente los sistemas de prestamos de bicicletas compartidas consisten en mediante el pago de una membresía, un usuario puede tomar una bicicleta de alguna estación de bicicletas preestablecida, y despues de cierto tiempo de uso, el usuario debe depositar la bicicleta que tomó en alguna estación de bicicletas del sistema. Estos sistemas han evolucionado a lo largo del tiempo, y a su vez, han sido catalogados en tres generaciones de acuerdo a las características tecnológicas que se han implementado para su operación. El primer sistema de préstamo de bicicletas del que se tiene registro, fue el sistema implementado en Alemania en 1965 [11], este sistema consistió en que un conjunto de bicicletas blancas fue puesto a disposición de manera pública a la ciudadanía, una persona podía tomar una bicicleta y dejarla estacionada en en cualquier parte de la ciudad para que otra persona pudiera también hacer uso de la bicicleta. Este sistema fracaso a los pocos días

por el robo o vandalismo de las bicicelas, ya que no había un control acerca de cómo cuidar las bicicletas. A este sistema se le ha catalogado como un sistema de primera generación. Posteriormente, a partir del año 1991, surgieron los servicios de segunda generación en Dinamarca, estos servicios se diferencian de los de primera generación al incorporar cicloestaciones en las cuales los usuarios podían utilizar una bicicleta al introducir una moneda. Aunque era un servicio mejor organizado, el robo de bicicletas continuo siendo un problema debido al anonimato de los usuarios. Gracias a los avances tecnológicos, al rededor del año 1996 se incorporaron tarjetas de cinta magnética en Inglaterra para el prestamo de bicicletas, lo que dio mayor control para la identificación de usuarios y a su vez dio origen a la tercera generación del servicio, la cual continua vigente hoy en día. Se puede consultar más acerca de la historia de los servicios de préstamos de bicicletas en [11].

1.2. El problema de investigación

La tercera generación de los sistemas de prestamos de bicicletas, además de brindar un mejor servicio, ha permitido a los investigadores el estudio cuantitativo de estos, ya que estos sistemas permiten recopilar información histórica de los mismos. Esta información puede ser analizada para encontrar características o patrones de interés en los datos recopilados. En particular, una área de investigación de estos sistemas está centrada en la estimación de la oferta y la demanda de bicicletas en algunas cicloestaciones, en particular, se han propuesto varia metodologías para realizar estas estimaciones como en [5]-[8], estas metodologías buscan explotar la estructura de correlación temporal del número de bicicletas. El problema de estimación de oferta y demanda, no es un problema sencillo, ya que además de ser un problema estocástico, y depender de las peculiaridades del sistema de bicicletas de estudio, también depende tanto de la calidad de información obtenida por el sistema, así como por las cualidades de la ciudad donde este opera, como el sistema estudiado en [5], por ejemplo.

Otra razón por la que es importante estimar tanto la oferta como la demanda, es porque a partir del conocimiento de estas cantidades es

posible rebalancear al sistema, analizar la demanda futura para conocer los niveles de servicio necesarios o para expandir el sistema, los cuales son algunas líneas actuales de investigación de los servicios de préstamo de bicicletas conforme se comenta en [6]-[18].

1.3. Objetivo

De acuerdo a lo mencionado en la sección 1.2, en este trabajo nos enfocamos al problema de la estimación de la oferta y demanda de un sistema de bicicletas compartidas, utilizando como caso de estudio el sistema de bicicletas compartidas de la ciudad de México, Ecobici. Para ello, mediante un análisis estadístico se identifican aquellas cicloestaciones que presenten algún problema de oferta o de demanda de bicicletas. Posteriormente, se realiza un ejercicio de rebalanceo considerando las estimaciones de oferta y demanda obtenidas.

1.4. Metodología de solución

Para la obtención de resultados se utilizó la información del año 2019 del servicio de bicicletas Ecobici, se puede consultar más detalle acerca de esta información en la sección 3.1. Posteriormente, se realizó un análisis estadístico para la identificación de cicloestaciones con problemas de oferta o demanda (saturación o escases de bicicletas), para ello se definió el índice de saturación explicado en la sección 3.2.1. Una vez identificadas aquellas estaciones con algún problema de oferta o demanda se procedió a estimar estas cantidades conforme lo descrito en las secciones 3.2.3 y 3.2.4. Una vez que se obtuvieron las estimaciones de oferta y demanda se procedió a realizar un ejercicio de rebalanceo conforme a lo descrito en la sección 3.4.

1.5. Resumen de resultados

En este trabajo se obtienen las estimaciones de oferta y demanda de bicicletas para el sistema de bicicletas compartidas ecobici, estas

esmitaciones tienen un error promedio de estimación del 10.5 % y del 12.9 % para la oferta y demanda, así como una desviación estándar del 6.7 % y 10 %, respectivamente. Posteriormente, se realizó un ejercicio de rebalanceo donde se logró atender la demanda entre un 70 % y 90 % dependiendo del día de la semana y la hora. En particular, estos resultados corresponden a los horarios de operación entre las 08:30-09:00 y las 18:30-19:00, ya que como se expone en [10], en estos horarios se presenta el mayor uso del sistema Ecobici.

Capítulo 2

Revisión bibliográfica

2.1. Sistemas de préstamos de bicicletas y problemas identificados

Los sistemas de préstamos de bicicletas han crecido alrededor del mundo, ya que como se menciona en [6], estos traen notables ventajas en términos de movilidad, ambientales y de salud. De acuerdo a [6], en 1960 eran unas cuantas las ciudades que tenían un sistema de préstamo de bicicletas, mientras que para el año 2016 son más de 800 ciudades las que han implementado un servicio de préstamo de bicicletas compartidas. En [6], también se exponen las principales líneas de investigación que se han derivado a partir de estos sistemas, las cuales son:

1. La necesidad del desarrollo eficiente de métodos de rebalanceo, lo cual es sumamente importante para brindar un servicio enfocado en una alternativa de movilidad sustentable.
2. Investigación de por qué las personas optan por no usar un servicio de préstamo de bicicletas, lo cual es sumamente importante para sistemas subutilizados, y también, para el crecimiento de los sistemas de bicicletas en general.
3. El comportamiento de los conductores ante los usuarios del sistema de bicicletas compartidas, lo cual es importante para implementar

medidas de seguridad, o para implementar caminos especiales para las bicicletas.

4. El desarrollo e investigación de métricas cuantitativas para medir el impacto de los sistemas de préstamos de bicicletas en términos de cambio climático, tránsito vehicular, y contaminación ambiental.

En la misma línea que en [6], en [18] se identifican cuatro áreas de investigación que han surgido gracias a los sistemas de préstamos de bicicletas, estas áreas son las siguientes:

1. Estrategias de diseño para implementar un sistema de préstamo de bicicletas, o para la expansión de un sistema ya establecido.
2. Análisis de la demanda futura que podría tener un sistema de préstamo de bicicletas para conocer los niveles de servicio necesarios.
3. Análisis del nivel de servicio enfocado a contemplar tanto el número de bicicletas en circulación así como el número de candados disponibles.
4. Operaciones de rebalanceo para preservar la sustentabilidad del sistema.

De las ideas expuestas en [6]-[18], es importante mencionar que un problema inherente al rebalanceo de bicicletas, es el problema de la estimación de la oferta y demanda de bicicletas, ya que si no se conocen estas cantidades, no es posible diseñar metodologías de rebalanceo. Es por ello que en este trabajo nos centramos en cómo obtener estas estimaciones para posteriormente realizar un ejercicio de rebalanceo.

En la siguiente sección se habla acerca de algunos métodos de solución que se han propuesto para obtener estimaciones tanto de oferta como de demanda de bicicletas, así como del problema de rebalanceo.

2.2. Métodos de solución

2.2.1. Estimación de demanda

En la literatura es posible encontrar varias propuestas para resolver el problema de la estimación de la oferta y demanda de bicicletas. En particular, cada propuesta es diseñada e implementada de acuerdo a las peculiaridades del sistema de préstamo de bicicletas estudiado. Por ejemplo, en [5], se analiza el sistema de préstamo de bicicletas tanto de la ciudad de Sevilla, como el de la ciudad de Barcelona, para ello, se propone utilizar modelos de series de tiempo para encontrar variables socio-demográficas que estén asociadas con la demanda de bicicletas en las distintas zonas de la ciudad. En esta estrategia se reduce la complejidad del problema al manejar la información de manera agregada por zona, y además, se utilizan modelos de series de tiempo para capturar la estructura de correlación inherente al flujo de los arribos o retiros de bicicletas. Algo que no es claro en el artículo, es por qué se decidió modelar el logaritmo del número estimado de bicicletas en cada zona, así como la corroboración de que los supuestos que este tipo de modelos tiene, como verificar que los errores del modelo siguen una ley de probabilidad normal, y que estos no cambian con el tiempo. Otra área de oportunidad que tiene este artículo es comprobar si la ingeniería de variables puede ayudar a mejorar los resultados, por ejemplo se usa tanto la variable número de **Estaciones por zona** y **Capacidad total por zona**, donde una variable tal vez más informativa sea el cociente entre estas variables. Por último, en los resultados obtenidos únicamente se muestran las pruebas t de cada coeficiente de cada variable, lo cual debe ser acompañado con un análisis más robusto, como la incorporación de la prueba F , para verificar que efectivamente los coeficientes de las variables que se reportan son significativos.

En la misma línea que en [5], en [7] se proponen tanto el uso de modelos de series de tiempo, como de redes bayesianas para obtener estimaciones del número de bicicletas disponibles en cada estación de la ciudad de Barcelona. Estos modelos reportan un error de predicción promedio entre el 8% y el 37%, con desviaciones estándar de entre el 12% hasta el 16%, donde el menor error y menor desviación estándar

corresponden al modelo de redes bayesianas. Algo que no es claro en este artículo es la forma en que se modela la red bayesiana, ya que utilizando las variables de tiempo, ventana temporal y número de bicicletas en la cicloestación, se modela el cambio del número de bicicletas en dicha estación para un intervalo de tiempo como la realización de una variable aleatoria gaussiana. En el artículo no se presenta ninguna prueba que indique que en efecto, este cambio siga una ley de probabilidad gaussiana. De acuerdo a los modelos de series de tiempo, estos reportan el mayor error y desviación estándar promedio, se utilizó un modelo de medias móviles, uno de estimación de tendencia, y un modelo que únicamente toma en cuenta el último valor observado de la serie, estos modelos en particular puede que no sean del todo adecuados, ya que no reflejan la alta volatilidad que es posible observar en un sistema de prestamos de bicicletas, además que debería considerarse un modelo propio para cada cicloestación y se le debe dar seguimiento conforme la información es generada, lo cual podría ser costoso de mantener.

Al igual que en [5] y en [7], en [8] se realiza el pronóstico del número de bicicletas disponibles en la ciudad de Barcelona mediante modelos de serie de tiempo ARMA, para ello, se ajustaron modelos con 10 parámetros en la componente autoregresiva así como 10 parámetros en la componente de media móvil, lo que origina modelos con 20 parámetros en total. El número de parámetros utilizados en estos modelos puede sugerir que los modelos estén sobre ajustados, y además, le dan nula interpretabilidad al modelo. También, en la práctica puede que sea algo complicado de implementar esta propuesta ya que se requiere ajustar un modelo con 20 parámetros para cada cicloestación.

Siguiendo la misma línea que en [5]-[8], en [10] se propone un modelo de series de tiempo ARIMA para pronosticar el número total de retiros que enfrenta el sistema de bicicletas compartidas de la ciudad de México en un día de operación. Este modelo puede ser utilizado con fines de mantenimiento predictivo de todo el sistema, ya que le permite conocer a los administradores del servicio qué cantidad total de bicicletas se espera que se utilicen en algún momento del día. Una posible desventaja de este enfoque es que los coeficientes del modelos son complicados de

interpretar, ya que no es un modelo parsimonioso.

Otro punto a recalcar de [7], [8] y [10], es que al modelar el flujo de estaciones mediante un modelo de tipo ARMA o ARIMA, las predicciones a largo plazo de estos modelos tienden a regresar a la media, ocasionando muy pobres estimaciones de largo plazo, por lo que es preferible utilizar estos modelos para realizar predicciones únicamente en el corto plazo, por lo que el intervalo de predicción de dichos modelos estará influenciado por los intervalos de tiempo con los que fueron contruidos los modelos, por ejemplo en [7] se utilizan intervalos de tiempo de cinco minutos, mientras que en [8] se utilizan intervalos de tiempo de dos minutos, por lo que las estimaciones en una ventana temporal de más de treinta minutos podrían ser muy imprecisas, mientras que en [10] la ventana temporal utilizada fue de quince minutos, por lo que las predicciones tal vez sean imprecisas para intervalos de tiempo mayores a una hora.

2.2.2. Balanceo

En la literatura hay dos enfoques que se han implementado para atender el problema de balanceo, estos dos enfoques consisten en implementar un sistema de incentivos, o mediante el uso de camiones rebalanceadores. Por ejemplo, en [15] se incentiva a los usuarios a depositar las bicicletas en estaciones específicas mediante un sistema de pagos, donde cada cicloestación cuenta con la información del pago recibido si se deposita la bicicleta en esa cicloestación o en alguna otra cicloestación de los alrededores, para ello, las cicloestaciones deberán compartir sus estados así como los pagos calculados entre ellas. Aunque este enfoque es interesante, no se tiene en cuenta qué algoritmo de consenso es viable utilizar, así como algún protocolo de tolerancia de fallas. Tampoco se toma en cuenta que al implementar un sistema distribuido de este tipo, cuáles serán las latencias relacionadas al cálculo y distribución de los incentivos, así como un posible tiempo de espera extra que deba afrontar el usuario al esperar a que el sistema se actualice.

Por otro lado, en [18] se aborda el problema de rebalanceo a través

de encontrar rutas factibles, para ello los autores aproximan el algoritmo de la máxima estrella generada en un conjunto de vértices. Es interesante notar que los autores reportan tiempos mayores a una hora para el cálculo de algunas instancias del problema, lo que hace que esta aproximación no sea eficiente si se desea aplicar en un sistema de prestamo de bicicletas donde se requiera hacer un rebalanceo en algunos minutos. También, en esta referencia no se da ninguna información acerca de la demanda total atendida por el algoritmo implementado.

Capítulo 3

Metodología de solución

En este capítulo se explica la metodología implementada para resolver el problema de la estimación de la oferta y demanda de bicicletas, así como del problema de transporte y ruteo. Para ello, primero se describe el sistema de bicicletas compartidas de la ciudad de México (Ecobici), y se plantea el problema de escases y saturación en el servicio. Posteriormente, se expone la información con la que se cuenta para resolver el problema anterior. Luego, mediante un análisis exploratorio de datos, se explica cómo identificar cicloestaciones con problemas de saturación o escases, para después detallar cómo estimar la oferta o demanda de dichas cicloestaciones para un día entre semana. Por último, se describe el problema de transporte y ruteo.

A continuación, en la gráfica 3.1, se resume de manera esquemática el flujo de de trabajo correspondiente a la metodología de solución.

3.1. Planteamiento del problema

El sistema de bicicletas compartidas de la ciudad de México, llamado Ecobici, es un sistema de movilidad que comenzó operaciones en febrero del 2010. Este sistema consiste en que un usuario, mediante la adquisición de una membresía, puede utilizar alguna bicicleta de cualquiera de las 480 cicloestaciones (contabilizadas hasta agosto del 2020)



Gráfica 3.1: Flujo de trabajo.

que ofrece el sistema, y con ello desplazarse por la ciudad de México. Cuando el usuario decide finalizar el servicio, puede depositar la bicicleta en cualquier cicloestación que tenga algún lugar disponible; el usuario puede hacer uso de este servicio a cualquier hora entre las 05:00 horas y las 00:30 horas del siguiente día, y para ello a lo más puede utilizar una bicicleta durante 45 minutos. Es posible encontrar más información acerca del servicio en [4].

El servicio y el número de usuarios continua creciendo desde que el

sistema comenzó operaciones, esto tal vez se deba a que el sistema ha brindado una manera alternativa para desplazarse por la ciudad, y de esta manera, reducir el tiempo de traslado asociado a otras formas de movilidad, como el tránsito generado en algunas partes de la ciudad en horas pico, por ejemplo. El constante incremento del número de usuarios puede estar asociado a que en ciertas horas del día la demanda de bicicletas sea superior al número de bicicletas disponibles en algunas cicloestaciones, o que algunas cicloestaciones se encuentren saturadas y que los usuarios no puedan depositar las bicicletas en dichas cicloestaciones, lo cual puede ocasionar frustración en los usuarios, lo cual se traduce en mala calidad del servicio. Por ejemplo, en [12]-[13] se exhiben algunos problemas que presenta este servicio, como lo son que algunas cicloestaciones presentan un mayor número de retiros que arribos o viceversa, lo cual puede traducirse en un problema de oferta y demanda que enfrenta el servicio.

Una alternativa para mejorar la calidad del servicio es evitando que haya escases o saturación en las cicloestaciones a lo largo del día, para ello es indispensable conocer con anticipación el número de bicicletas que se demandan o que se ofrecen en cada cicloestación, lo cual es sumamente complicado ya que estas dos cantidades son aleatorias por la naturaleza misma del servicio. Aunque es complicado conocer exactamente estas cantidades, es posible estimarlas mediante métodos estadísticos, como con modelos de series de tiempo, promedios móviles, técnicas de regresión, mejor predictor lineal, por ejemplo. Aunque todos estos métodos estadísticos tienen tanto bondades como debilidades, también es cierto que para su implementación dependen de la información con la que se cuenta, lo cual podría ser un problema adicional, como para el caso de Ecobici de lo cual se hablará más adelante.

Una vez que es posible obtener estimaciones tanto de la oferta como de la demanda, es posible plantear el problema de rebalanceo para reasignar de manera eficiente bicicletas de cicloestaciones saturadas a cicloestaciones con escases de estas. La reasignación de bicicletas puede estudiarse ya sea a partir del problema de transporte, del problema de ruteo, o una combinación entre ambos. El primer problema consiste en

determinar de manera óptima de qué cicloestaciones se deben tomar bicicletas y a qué cicloestaciones deben ser llevadas, generando así un plan de distribución. El segundo problema consiste en determinar cuál debe ser la ruta óptima para recoger las bicicletas de las cicloestaciones saturadas y llevarlas a las cicloestaciones con escasos. Un factor constante en ambos problemas es que se requiere conocer con anticipación tanto la oferta como la demanda de bicicletas de cada cicloestaciones, ya que estas cantidades son el insumo principal para dichos problemas. Por lo que antes de plantear estos problemas debemos enfocarnos en cómo estimar tanto la oferta como la demanda de bicicletas.

Dicho lo anterior, en la sección 3.2 se habla acerca del problema de oferta y demanda, pero antes de ello, a continuación se menciona brevemente la información con la que se cuenta.

3.1.1. Descripción de la información

La información es pública y puede consultarse en el sitio web del servicio, [4]. Esta ha sido reportada de manera mensual desde Febrero del 2010 hasta hoy en día, por ejemplo, toda la información correspondiente al mes de abril es reportada hasta principios de mayo. La información reportada es a nivel servicio, y consta de las siguientes variables:

1. **Genero_Usuario:** esta es una variable nominal que toma el valor de M si el usuario es hombre y F si es mujer.
2. **Edad_Usuario:** esta es una variable numérica que reporta la edad del usuario.
3. **Bici:** esta es una variable nominal que define un identificador único para cada bicicleta.
4. **Ciclo_Estacion_Retiro:** esta es una variable nominal que define un identificador único para la cicloestación de la que la bicicleta fue tomada por el usuario al comenzar el servicio.

5. **Fecha_Retiro:** esta es una variable nominal que reporta la fecha de cuando la bicicleta fue tomada. Esta variable es reportada en el formato mes/día/año.
6. **Hora_Retiro:** esta es una variable numérica que reporta el tiempo en el que una bicicleta fue tomada. Esta variable es reportada en el formato horas::minutos::segundos.
7. **Ciclo_Estacion_Arribo:** esta es una variable nominal que define un identificador único para la cicloestación en la que la bicicleta fue devuelta por el usuario al terminar el servicio.
8. **Fecha_Arribo:** esta es una variable nominal que reporta la fecha de cuando la bicicleta fue devuelta por el usuario. Esta variable es reportada en el formato mes/día/año.
9. **Hora_Arribo:** esta es una variable numérica que reporta el tiempo en el que una bicicleta fue devuelta por el usuario. Esta variable es reportada en el formato horas::minutos::segundos.

3.2. Estimación de Oferta y Demanda

Como se mencionó en la sección 3.1.1, la información es reportada de manera mensual a nivel servicio, lo cual genera los siguientes dos problemas:

1. Que la información sea reportada a nivel mensual sugiere que no sea viable considerar un modelo de pronóstico que se alimente de la información del sistema al momento en que se está usando el servicio y que dé algún pronóstico para algún intervalo futuro de tiempo.
2. Que la información se reporte a nivel servicio ocasiona que no se conozcan el número de bicicletas disponibles en cada cicloestación durante el día, es decir, no se reporta la oferta de bicicletas, por lo que esta debe ser estimada con la información disponible.

Adicionalmente a los dos problemas anteriores, se identificó que algunas cicloestaciones presentan un problema de oferta o demanda distorsionada, es decir, los datos sugieren que en algunas estaciones es posible estacionar o retirar más bicicletas de la capacidad total de la cicloestación, lo cual no es posible. Este comportamiento inusual en los datos puede ayudarnos a identificar un conjunto de cicloestaciones que presenten un problema de saturación o escases, lo cual se detalla en la sección 3.2.1.

Los tres problemas identificados anteriormente sirven de guía para proponer una estrategia de estimación para la oferta y la demanda. El problema uno nos sugiere enfocarnos en estimar tanto la oferta como la demanda de manera frecuentista, es decir, identificar ciertos patrones en la información, y explotar dichos patrones con el fin de obtener estimaciones. El problema dos, nos sugiere que las estimaciones se realicen en intervalos de tiempo, donde es posible estimar tanto el número de bicicletas que llegan como las que son tomadas de cada cicloestación, teniendo así una estimación de la posible oferta y demanda. El problema tres nos puede proporcionar evidencia acerca de qué cicloestaciones presentan un problema de oferta o demanda en algún momento en el tiempo, para así encontrar un subconjunto de cicloestaciones que requieren ser rebalanceadas.

Dicho lo anterior, en este trabajo se utiliza un enfoque de estadística frecuentista para obtener estimaciones tanto de oferta como de demanda para una cicloestación en particular, para un día de la semana dado y para un periodo de tiempo en específico, lo cual se describe en las secciones 3.2.3 y 3.2.4. También, mediante el enfoque frecuentista, en este trabajo se construye un índice de saturación, el cual nos dará evidencia de que ciertas cicloestaciones tienen un problema de saturación o escases, para lograr esto, se estudiarán los patrones que prevalecen en este índice durante un año de observación.

A continuación se describe cómo se construye el índice de saturación así como su interpretación.

3.2.1. Índice de saturación

Como se mencionó en la sección 3.2, un problema asociado a la información con la que se cuenta es que en ningún momento se cuenta con el número de bicicletas que hay en cada ciclo estación, por lo que una posible alternativa de estimarlo es mediante el flujo de bicicletas a lo largo del día, es decir, contar con el número acumulado arribos menos retiros. Esta cantidad podría estar fuertemente sesgada, ya que no se conoce con cuántas bicicletas comienza cada ciclo estación. Por ejemplo, si el flujo acumulado a las 12:00 hrs es de 8, esto quiere decir que desde que inició el servicio hasta esa hora ocurrieron 8 arribos más de bicicletas que retiros, por lo que debería haber 8 bicicletas disponibles en esa ciclo estación en ese instante, lo cual es cierto únicamente si la ciclo estación comenzó con cero bicicletas, pero si la ciclo estación hubiera comenzado operaciones con 10 bicicletas, en realidad habrían 18 bicicletas a las 12:00 hrs.

Una posible manera de incorporar nuestro desconocimiento acerca del estado inicial de cada ciclo estación, es asumir que el número de bicicletas con las que comienza operaciones cada ciclo estación sigue una ley de probabilidad uniforme en el conjunto de posibles bicicletas que puedan estar estacionadas. Luego si dividimos este conjunto de posibles bicicletas entre el máximo número de bicicletas que se pueden estacionar, tendremos una variable aleatoria discreta, X , que toma valores entre cero y uno. Dicho lo anterior, para la ciclo estación c , se propone la construcción del siguiente índice:

$$I_{t_i}^c = \frac{\sum_{t_j \leq t_i} A_{t_j} - R_{t_j}}{k} + C_0, \quad (3.1)$$

donde A_{t_j} son los arribos de bicicletas registrados en el intervalo de tiempo de t_{j-1} a t_j , R_{t_j} son los retiros de bicicletas registrados en el intervalo de tiempo de t_{j-1} a t_j , k es el número de candados (cantidad máxima de bicicletas que se pueden estacionar) que puede tener la ciclo estación c , y C_0 es el nivel al inicio del día del índice, donde $E[X] = C_0$.

La idea intuitiva de este índice es monitorear el flujo de bicicletas para cada ciclo estación a lo largo de cada día de operación, y a

partir de ahí encontrar patrones que se mantienen durante un año de observación. El flujo de bicicletas para un intervalo de tiempo de 30 minutos, por ejemplo de 07:30 a 08:00 queda capturado mediante la diferencia de $A_{08:00} - R_{08:00}$. Ahora, para capturar el balance total desde que una cicloestación comienza operaciones hasta por ejemplo el tiempo $t_j = 08 : 00$ (y suponiendo una discretización del tiempo en intervalos de 30 minutos), se obtiene sumando $A_{05:00} - R_{05:00} + A_{05:30} - R_{05:30} + \dots + A_{08:00} - R_{08:00} = \sum_{t_j \leq 08:30} A_{t_j} - R_{t_j}$. Después, se normaliza el índice respecto al número de candados que hay en cada cicloestación, por ejemplo, para la cicloestación c se realiza el cociente de $\sum_{t_j \leq t_i} A_{t_j} - R_{t_j}$ entre k , donde k es el número de candados de dicha cicloestación. Por último se suma el nivel inicial $E[X] = C_0$ del índice, este nivel inicial representa la cantidad de bicicletas con las que comienza cada cicloestación en un día dado, por ejemplo si $C_0 = 0$ significaría que la cicloestación comienza con 0 bicicletas, mientras que si $C_0 = 0.5$ significaría que la cicloestación comienza el día con la mitad de su capacidad. Notemos que X es una variable aleatoria, y por conveniencia podemos utilizar el mejor estimador mínimo cuadrático mediante una constante, el cual es $E[X]$, y al tratarse de una variable aleatoria uniforme discreta entre cero y uno, esta esperanza será aproximada a 0.5

Notemos que sin intervención alguna, y fijando C_0 igual a cero, para cada cicloestación, el índice de saturación debería oscilar entre 0 y 1 a lo largo de un día normal de operación, por ejemplo, si $I_{t_i}^c \approx 1$ significaría que desde que el servicio comenzó operaciones hasta el momento del día t_i , la cantidad de arribos a la ciclo estación c ha sido mucho mayor que la cantidad de retiros, y por ello la cicloestación debería estar casi llena, mientras que si $I_{t_i}^c \approx 0$ significaría que el número de retiros ha sido mayor que el número de arribos y por ello la cicloestación debería estar casi vacía.

Aunque el índice también puede ser utilizado para obtener estimaciones tanto de oferta como de demanda, este debe ser tratado con cautela ya que como se indica en la sección 3.2, un problema encontrado en los datos es la distorsión de la oferta y la demanda, lo cual es descubierto mediante este índice, por ello en lugar de utilizarlo con fines

de estimación, es preferible usarlo como guía para detectar estaciones con problemas de saturación o escases. Lo anterior es debido a que los datos sugieren que este puede ser mayor a 1 o menor a 0 (aún cuando C_0 sea igual a cero o a uno), lo cual no tiene sentido ya que no se pueden estacionar (o retirar) más bicicletas que el número máximo de candados que tiene cada cicloestación (número máximo de bicicletas), este comportamiento tal vez pueda explicarse a que en determinados momentos del día se llevan (o se extraén) bicicletas de cicloestaciones que así lo requieran, y estos arribos o retiros no quedan registrados en las bases de datos disponibles, lo que ocasiona que aparentemente en algunas cicloestaciones se puedan extraer o depositar más biciletas que el número de candados disponibles. En este caso el índice sigue siendo un buen elemento para identificar cicloestaciones con problemas, pues de no existir intervención alguna por parte del sistema, el índice simplemente convergería a 1 o 0 según sea el caso.

A continuación se describe la construcción del índice de saturación promedio. También, se ejemplifica como usar dicho índice para encontrar patrones de escases o saturación de bicicletas para alguna cicloestación dada.

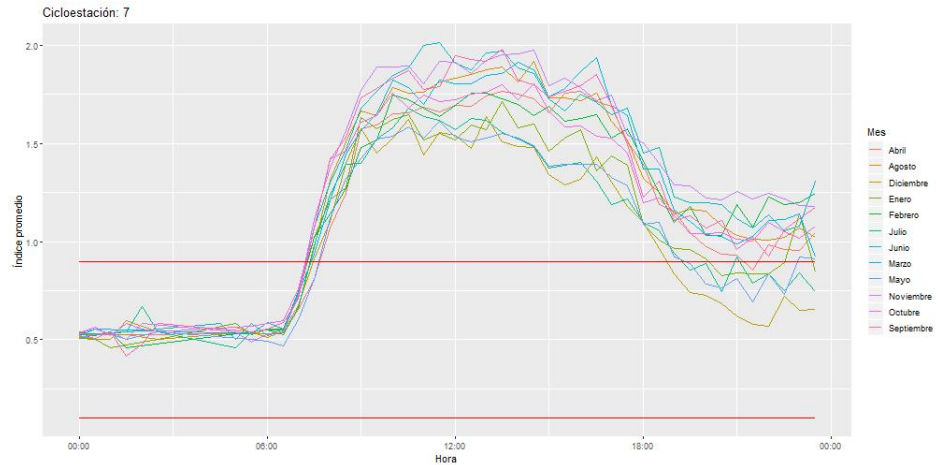
3.2.2. Cicloestaciones con escases o saturación

Como se mencionó anteriormente, se identificarán patrones que prevalecen en el índice de saturación a lo largo de un año de observación, es decir, durante todo 2019. Estos patrones los podemos observar si estudiamos el índice de saturación promedio en cada mes M , para cada cicloestación c , para cualquier tiempo t_i , es decir,

$$\bar{I}_{t_i}^{c,M} = \sum_{d \in M} \frac{I_{t_i}^{c,d}}{|M|} + C_0, \quad (3.2)$$

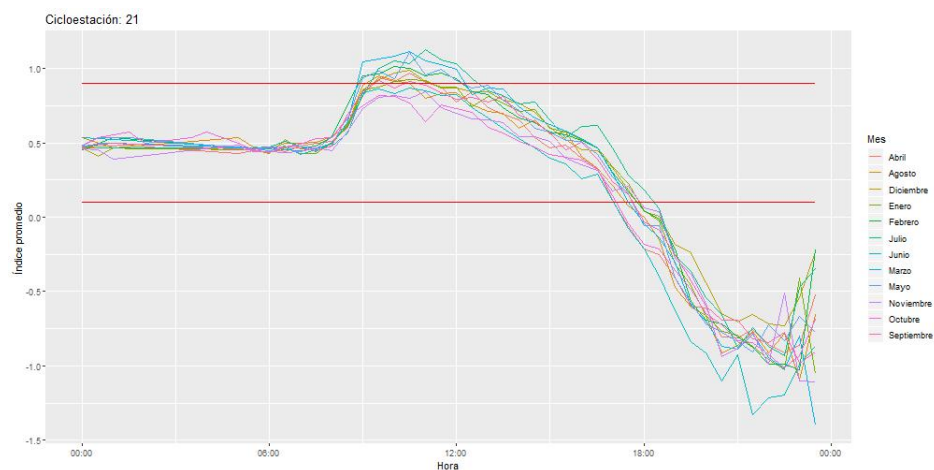
donde d representa algún día en particular del mes M , mientras que $|M|$ es el número de días correspondiente a dicho mes. A continuación se presentan algunos ejemplos de los patrones identificados mediante el índice promedio:

En la gráfica 3.2 se puede observar que para todos los meses del año 2019, para la cicloestación 7, el índice promedio alcanza valores mayores a 1 al rededor de las 08 : 00 hrs, lo que nos da evidencia para creer que esta cicloestación se satura al rededor de dicha hora.



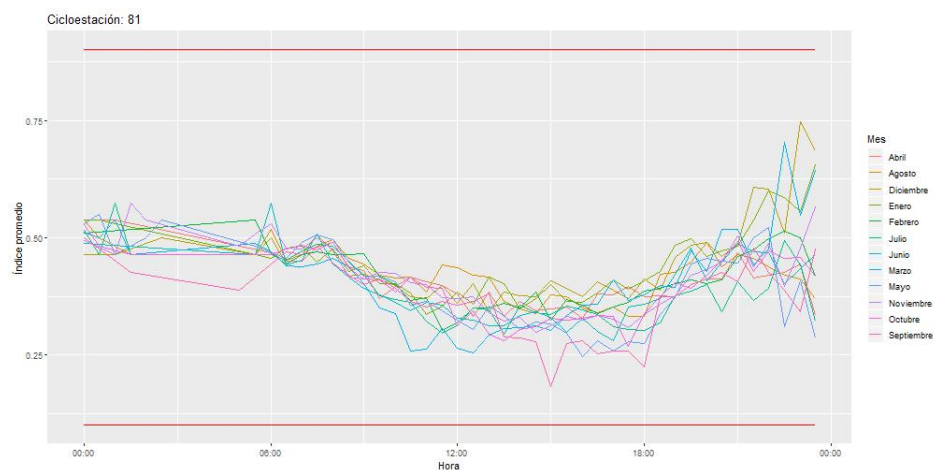
Gráfica 3.2: El patrón identificado de la cicloestación 7 nos da evidencia de una posible saturación de bicicletas al rededor de las 08 : 00 hrs.

En la gráfica 3.3 se puede observar que para todos los meses del año 2019, para la cicloestación 21, el índice promedio alcanza valores menores a 0 al rededor de las 18 : 00 hrs, lo que nos da evidencia para creer que esta cicloestación presenta escases al rededor de dicha hora.



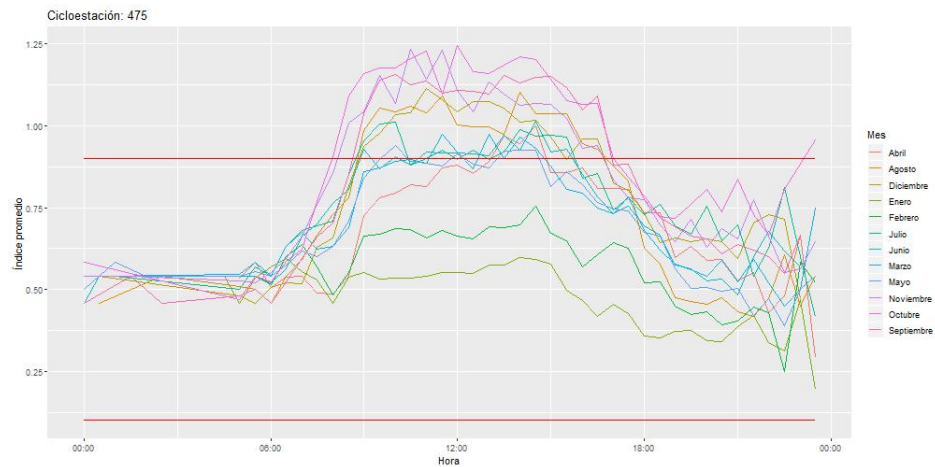
Gráfica 3.3: El patrón identificado de la cicloestación 21 nos da evidencia de una posible escases de bicicletas al rededor de las 18 : 00 hrs.

En la gráfica 3.4 se puede observar que para todos los meses del año 2019, para la cicloestación 81, el índice promedio permanece entre 0.15 y 0.75 a lo largo de todo el día, lo que nos da evidencia para creer que esta cicloestación no presenta ningún tipo de problema.



Gráfica 3.4: El patrón identificado de la cicloestación 81 nos da evidencia de una posible equilibrio de bicicletas durante el día.

En la gráfica 3.5 se observa que para la cicloestación 475 no hay un patrón lo suficientemente claro en todos los meses para determinar si esta cicloestación presenta algún tipo de problema o no, esto último también podría depender de qué tan riguroso sea nuestro criterio, ya que si permitimos que el índice promedio sea mayor a 1 en al menos cierta cantidad de meses como 10 meses, por ejemplo, entonces esta cicloestación podría presentar problemas de saturación al rededor de las 14 : 00 hrs, pero si nuestro criterio es que se debe observar el patrón en todos los meses, entonces esta cicloestación podría ser descartada bajo este criterio.



Gráfica 3.5: No hay un patrón claro en la cicloestación 475, por lo que no tenemos evidencia de que esta cicloestación presente un problema de saturación, escasas o que permanece en equilibrio.

Por último, en las tablas 3.1 y 3.2 se muestra aquellas cicloestaciones que de acuerdo al índice muestran evidencia de saturación o escases en algún momento del día.

5	6	7	9	11	13	14	15	16	17	18
19	22	23	24	25	26	27	28	30	36	37
38	39	47	51	55	63	88	116	117	118	134
136	141	150	197	198	199	200	201	202	205	206
207	209	214	217	219	224	225	226	227	232	233
235	236	239	242	248	249	250	253	261	290	295
301	302	303	304	305	315	316	326	335	420	436
452	454	455	458	459	460	463	468			

Tabla 3.1: Cicloestaciones identificadas con un problema de saturación de acuerdo al criterio del índice promedio de saturación.

21	41	43	48	52	53	75	77	84	85	86
119	138	140	143	151	153	158	161	162	164	165
166	167	168	169	171	172	173	174	176	181	183
186	188	189	191	192	193	194	204	208	210	211
228	237	238	254	255	256	259	264	265	266	267
269	271	272	273	274	282	285	286	288	320	321
329	330	331	351	384	448	450	465	466	480	

Tabla 3.2: Cicloestaciones identificadas con un problema de escases de acuerdo al criterio del índice promedio de saturación.

3.2.3. Estimación de Oferta al tiempo t_i

Como se mencionó en la sección 3.2, como no se conoce el número de bicicletas que hay en cada cicloestación en algún momento del tiempo estas deben ser estimadas, una manera de proponer una estimación es considerando el número de bicicletas que llegan en un intervalo de tiempo para un día en específico. Dicho lo anterior, si consideramos el número de bicicletas por cicloestación, para algún intervalo de tiempo

y para todos los días lunes del año, por ejemplo, podemos construir una distribución de frecuencias relativas y con ello estimar la probabilidad de que en alguna cicloestación, para algún día de la semana dado y para un intervalo de tiempo dado se observe algún número de arribos. Esta información la podemos utilizar para estimar la oferta, es decir, si se observó cierto número de arribos en un intervalo de tiempo, entonces la cicloestación debe tener por lo menos esa cantidad de bicicletas durante dicho intervalo.

Dicho lo anterior, definamos a $O|c_i, [t_j, t_{j+1}], d$ como la variable aleatoria que cuenta el número de arribos condicionada a la cicloestación c_i , el intervalo de tiempo de $[t_j, t_{j+1}]$ y al día de la semana d . La cual representa el número de arribos de bicicletas que observaríamos condicionado a que únicamente estamos considerando la cicloestación c_i en el intervalo de tiempo $[t_j, t_{j+1}]$ y en el día de la semana d , por ejemplo, la variable aleatoria $O|250, [8 : 00, 8 : 30], \text{Lu}$ representa el número de arribos que observaríamos dado que estamos considerando la cicloestación número 250 en el intervalo de tiempo de 08:00 hrs a 08:30 hrs, para el día de la semana Lunes.

Entonces, el mejor predictor mínimo cuadrático mediante una constante es $E[O|c_i, [t_j, t_{j+1}], d]$, la cual se calcula analíticamente como:

$$E[O|c_i, [t_j, t_{j+1}], d] = \sum_{k \geq 0} k * P(O = k|c_i, [t_j, t_{j+1}], d). \quad (3.3)$$

De la expresión anterior podemos estimar cada valor de $P(O = k|c_i, [t_j, t_{j+1}], d)$ si consideramos la frecuencia relativa de arribos observados durante un gran número de días, es decir:

$$\hat{P}(O = k|c_i, [t_j, t_{j+1}], d) = \frac{\sum_{d_i} k * I_{t \in [t_j, t_{j+1}], o_i = k}}{n}, \quad (3.4)$$

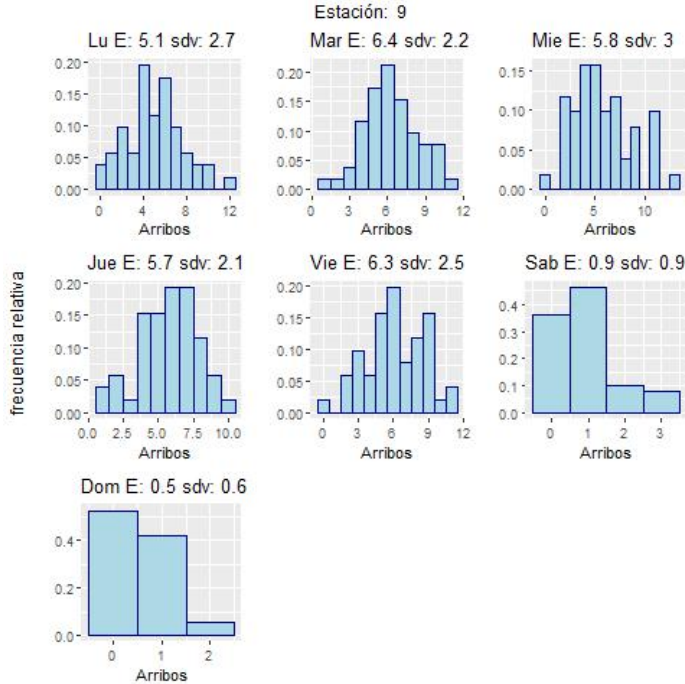
done d_i es el i -ésimo día d del año, n es el total de días d en el año, e $I_{t \in [t_j, t_{j+1}], o_i = k}$ es la función indicadora que toma el valor de uno cuando el número de arribos en el intervalo de tiempo $[t_j, t_{j+1}]$ es igual a k , y

cero en caso contrario. Con ayuda de la expresión 3.4, podemos estimar a $E[O|c_i, [t_j, t_{j+1}], d]$ como:

$$\hat{E}[O|c_i, [t_j, t_{j+1}], d] = \sum_{k \geq 0} k * \hat{P}(O = k|c_i, [t_j, t_{j+1}], d), \quad (3.5)$$

la cual denotaremos como $\hat{O}_{c_i, [t_j, t_{j+1}], d}$.

En la gráfica 3.6, podemos observar la función de probabilidad de oferta estimada para la cicloestación 9 en un intervalo de tiempo de 07:30 a 08:00 hrs, para todos los días de la semana, también se reporta la desviación estándar como *sdv*.



Gráfica 3.6: Función de probabilidad de la oferta estimada de la cicloestación 9, para el intervalo de tiempo de 07:30 a 08:00 hrs, para todos los días de la semana.

3.2.4. Estimación de Demanda al tiempo t_i

La estimación de la demanda es análoga a la de la oferta, solo que en lugar de considerar el número de arribos, se considera el número de retiros. De acuerdo a lo anterior, definamos a $D|c_i, [t_j, t_{j+1}], d$ como la variable aleatoria que cuenta el número de retiros condicionada a la cicloestación c_i , el intervalo de tiempo de $[t_j, t_{j+1}]$ y al día de la semana d . Entonces, el mejor predictor mínimo cuadrático mediante una constante es $E[D|c_i, [t_j, t_{j+1}], d]$, la cual se calcula analíticamente como:

$$E[D|c_i, [t_j, t_{j+1}], d] = \sum_{k \geq 0} k * P(D = k|c_i, [t_j, t_{j+1}], d). \quad (3.6)$$

De la expresión anterior podemos estimar cada valor de $P(D = k|c_i, [t_j, t_{j+1}], d)$ si consideramos la frecuencia relativa de retiros observados durante un gran número de días, es decir:

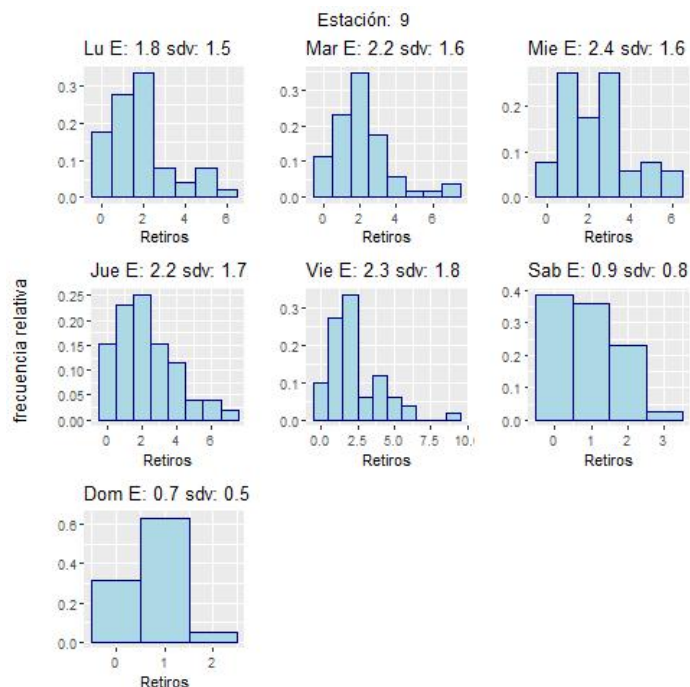
$$\hat{P}(D = k|c_i, [t_j, t_{j+1}], d) = \frac{\sum_{d_i} k * I_{t \in [t_j, t_{j+1}], r_i = k}}{n}, \quad (3.7)$$

done d_i es el i -ésimo día d del año, n es el total de días d en el año, e $I_{t \in [t_j, t_{j+1}], r_i = k}$ es la función indicadora que toma el valor de uno cuando el número de retiros en el intervalo de tiempo $[t_j, t_{j+1}]$ es igual a k , y cero en caso contrario. Con ayuda de la expresión 3.7, podemos estimar a $E[D|c_i, [t_j, t_{j+1}], d]$ como:

$$\hat{E}[D|c_i, [t_j, t_{j+1}], d] = \sum_{k \geq 0} k * \hat{P}(D = k|c_i, [t_j, t_{j+1}], d), \quad (3.8)$$

la cual denotaremos como $\hat{D}_{c_i, [t_j, t_{j+1}], d}$.

En la gráfica 3.7, podemos observar la función de probabilidad de demanda estimada para la cicloestación 9 en un intervalo de tiempo de 07:30 a 08:00 hrs, para todos los días de la semana, también se reporta la desviación estándar como sdv .



Gráfica 3.7: Función de probabilidad de la demanda estimada de la cicloestación 9 para el intervalo de tiempo de 07:30 a 08:00 hrs, para todos los días de la semana.

3.2.5. Otra alternativa de estimación

Como se mencionó anteriormente, el índice de saturación podría ser utilizado para obtener estimaciones tanto de oferta como de demanda, para ello deberá truncarse cuando este tome valores mayores a uno o menores a cero. Por ejemplo, para utilizarlo para obtener estimaciones de oferta, una vez que se han identificado aquellas cicloestaciones cuyo índice de saturación presenta el mismo patrón durante todo el año y además dicho patrón nos da evidencia de que hay un exceso de bicicletas o escases de ellas, podemos estimar la oferta de bicicletas si ahora

consideramos el índice de saturación promedio de todo el año, es decir, para la cicloestación c y el momento del día t_i considérese el siguiente índice promedio anual:

$$I\bar{P}A_{t_i}^c = \sum_{M \in N} \sum_{d \in M} \frac{I_{t_i}^{c,d}}{|M| \times |N|} + C_0. \quad (3.9)$$

Mediante este índice es posible obtener estimaciones de la oferta O_{t_i} de cada cicloestación c al tiempo t_i de la siguiente manera:

$$\hat{O}_{t_i} = \begin{cases} I\bar{P}A_{t_i}^c * k & \text{si } 0 \leq I\bar{P}A_{t_i}^c \leq 1 \\ k, & \text{si } I\bar{P}A_{t_i}^c > 1, \\ 0, & \text{si } 0 < I\bar{P}A_{t_i}^c \end{cases}$$

donde k es el número de candados de la cicloestación c .

Como se mencionó anteriormente, el caso en que $I\bar{P}A_{t_i}^c$ es mayor a 1 o menor a 0 en la práctica no tendrían sentido, pues implicaría que se estacionan más bicicletas que el número de candados disponibles o que se retiran más bicicletas de las que cuenta la cicloestación, lo cual es sólo posible si hay algún tipo de intervención. En caso de que esta intervención ocurra, valores mayores a 1 o menores a 0 seguirán informando que la cicloestación es muy concurrida y por ello debe considerarse para implementar estrategias de balanceo que ayuden a mejorar la calidad del sistema.

Para obtener estimaciones de la demanda D_{t_i} de alguna cicloestación c en el tiempo t_i puede estimarse de la siguiente manera:

$$\hat{D}_{t_i} = \sum_{M \in N} \sum_{d \in M} \frac{R_{t_i}^{c,d}}{|M| \times |N|}, \quad (3.10)$$

es decir, como el promedio anual de los retiros de la cicloestación c al tiempo t_i .

3.3. Balanceo de la oferta y la demanda

Como se mencionó en la sección 3.1, una alternativa para mejorar el servicio de prestamo de bicicletas es mediante la identificación de aquellas cicloestaciones que presenten un problema de saturación o escases. Una vez identificadas aquéllas cicloestaciones, y suponiendo que conocemos tanto su oferta como su demanda, el siguiente paso sería realizar un rebalanceo entre estas, es decir, llevar bicicletas de aquellas cicloestaciones cuya oferta de bicicletas sea superior a su propia demanda a cicloestaciones cuya demanda sea superior a su propia oferta. Lograr este rebalanceo es posible mediante la resolución del problema de transporte, el cual consiste en minimizar los costos asociados de la distribución de algún bien de una o mas fuentes de origen a uno o más destinos. El problema de transporte es usualmente modelado como un grafo dirigido, donde cada nodo ofrece o demanda cierta cantidad de algún bien en particular, y cada arista es asociada al costo de transportar el bien entre los nodos conectados por dicha arista.

Para nuestro caso, todas las cicloestaciones cuya oferta de bicicletas es superior a su propia demanda son los nodos oferentes de bicicletas, mientras que las cicloestaciones cuya demanda de bicicletas es superior a su propia oferta son los nodos demandantes de bicicletas, y el costo de transporte entre dos cicloestaciones (nodos) es la distancia que hay entre estas. De esta manera podemos adecuar nuestro problema de balanceo al problema de transporte.

El problema de transporte puede plantearse como un problema de programación lineal, donde nos importa minimizar los costos totales de transporte dado un conjunto de restricciones de oferta y demanda de cada nodo, es decir, lo podemos plantear analíticamente de la siguiente manera:

$$\text{Min} \sum_i \sum_j C_{i,j} x_{i,j} \quad (3.11)$$

s. a

$$\sum_{i:(i,j)} x_{i,j} + D_i = \sum_{i:(j,i)} x_{j,i} + O_i \text{ para todo nodo } i$$

$$x_{i,j} \geq 0,$$

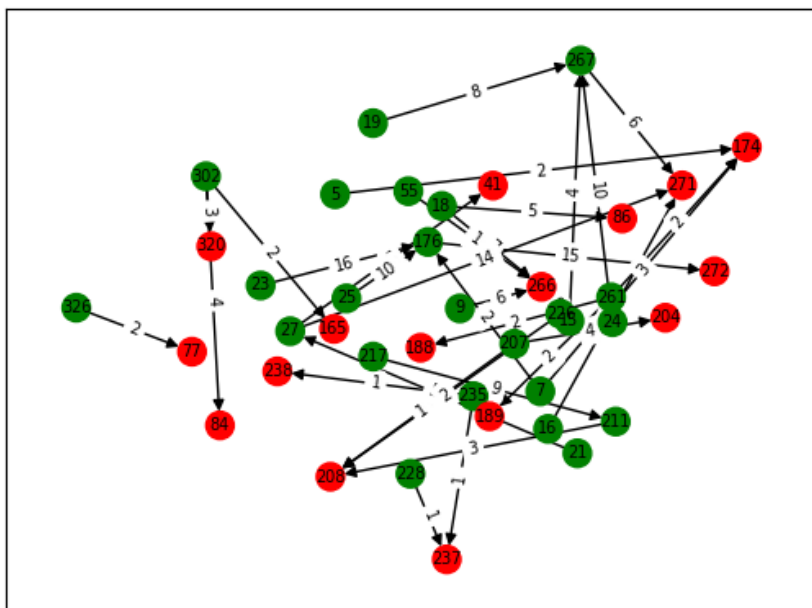
donde $x_{i,j}$ representa el número de bicicletas que son enviadas de la cicloestación i a la cicloestación j , la expresión $\sum_{i:(i,j)} x_{i,j}$ representa todas las bicicletas que salen de alguna ciclo estación i y llegan a la cicloestación j , mientras que la expresión $\sum_{i:(j,i)} x_{j,i}$ representa a todas las bicicletas que salen de la cicloestación j rumbo a la cicloestación i , de esta manera, la expresión $\sum_{i:(i,j)} x_{i,j} + D_i = \sum_{i:(j,i)} x_{j,i} + O_i$ significa que el total de bicicletas que salen de alguna cicloestación i mas sus respectivas demandas D_i , deben ser iguales a la cantidad de bicicletas que llegan a dicha cicloestación i más sus respectivas ofertas, es decir, la cantidad de bicicletas que salen de alguna cicloestación debe ser igual a la cantidad de bicicletas que recibe.

Un punto importante que hay que tener en consideración al implementar el problema de transporte, es que la oferta debe ser igual a la demanda, en caso de que este supuesto no se cumpla, es posible agregar variables de holgura, es decir, agregar nodos ficticios cuya oferta o demanda se igual al faltante necesario, y que tengan un costo de transporte igual a cero hacia cualquier otro nodo. para más detalles acerca del modelo de transporte, puede consultarse [14].

3.3.1. Ejemplo del problema de transporte

Por ejemplo, podemos identificar un conjunto de cicloestaciones con problemas de escases y saturación el día lunes a las 08:30 hrs mediante 3.2.2, y después, estimar sus respectivas ofertas y demandas mediante a lo explicado en las secciones 3.2.3 y 3.2.4, respectivamente. Entonces, tendríamos el problema de rebalancear el conjunto de cicloestaciones que aparece en la tabla 3.3. Las estimaciones anteriores son utilizadas para resolver el problema de transporte descrito en la sección 3.11, con lo cual obtendríamos el plan de transporte que aparece en la tabla 3.4.

En la gráfica 3.8 se muestra de manera visual el plan de transporte encontrado, los nodos color verde representan cicloestaciones oferentes de bicicletas o cicloestaciones de transbordo, mientras que los nodos color rojo representan cicloestaciones demandantes. Cada arco representa de qué cicloestación se deben tomar bicicletas y a cual llevar, mientras que cada número asociado a cada arco representa la cantidad de bicicletas que hay que llevar.



Gráfica 3.8: Vista gráfica de la solución del plan de transporte de la tabla 3.3

Cicloestación	Oferta	Demanda	Cicloestación	Oferta	Demanda
5	7	5	189	1	3
7	10	5	197	14	2
9	10	4	198	4	0
13	13	1	204	2	6
14	18	6	207	7	2
15	16	4	208	3	9
16	16	3	211	4	10
18	17	5	217	12	3
19	13	5	226	11	1
21	8	3	228	7	6
23	20	4	235	11	0
24	12	2	236	11	1
25	12	2	237	3	5
27	19	4	238	2	3
41	6	12	261	10	2
55	2	1	266	3	25
77	3	5	267	1	17
84	4	8	271	1	34
86	2	7	272	1	16
165	1	3	302	11	2
174	3	10	320	1	4
176	1	14	326	8	2
188	3	5			

Tabla 3.3: Oferta y Demanda de las cicloestaciones identificadas que requieren rebalanceo.

Tomar de	Llevar a	bicicletas	Tomar de	Llevar a	bicicletas
5	174	2	176	272	15
7	174	3	207	204	4
7	176	2	207	208	1
9	266	6	211	208	3
15	266	8	217	211	9
15	267	4	226	208	2
16	271	13	228	237	1
18	86	5	235	237	1
18	266	7	235	238	1
19	267	8	261	174	2
21	27	5	261	188	2
23	176	16	261	189	2
24	267	10	267	271	6
25	176	10	302	84	4
27	41	6	302	165	2
27	271	14	302	320	3
55	266	1	326	77	2

Tabla 3.4: Solución del plan de transporte de las cicloestaciones de la tabla 3.3

3.4. El problema de ruteo

En la sección anterior se habló acerca del problema de balanceo, mediante el cual se obtuvo el plan de distribución que minimiza los costos totales de transporte, o en nuestro caso la distancia total recorrida. El plan de distribución también nos indica de qué cicloestaciones oferentes tomar bicicletas y a qué cicloestaciones deben ser llevadas, para cumplir con la demanda de la mejor manera posible. Como se ilustra en la sección 3.4, el plan de distribución sugiere que la manera óptima de atender la demanda de las cicloestaciones 84, 165 y 320 es mediante el envío de 4, 2 y 3 de la cicloestación 302, respectivamente. Si tuviéramos disponibles 3 camiones que salen de la estación 302, el tiempo total que tardaría el balanceo para estas estaciones sería simplemente el tiempo

que se tarda un camión en alcanzar a la cicloestación más distante. Ahora, considerando el mismo escenario anterior, qué pasa si contamos con un único camión y queremos satisfacer el mismo plan de demanda, será conveniente mandar primero el camión a la cicloestación 84 luego a la 165 y por último a la 320, o será mejor comenzar con la 320, después ir a la 84 y por último a la 165. Esto último expone que para satisfacer un plan de demanda pueden existir distintas rutas y estrategias, por lo que naturalmente nos podemos preguntar si existe una estrategia de ruteo que sea óptima ante los distintos recursos de transporte con los que se cuente, como es el caso de camiones transportadores de bicicletas.

La pregunta que nos hicimos anteriormente puede responderse mediante alguna variante del problema del agente viajero, este problema consiste en encontrar un ciclo hamiltoniano en un grafo completo con n vértices, donde un ciclo hamiltoniano es un camino que une a todos los vértices, el cual conecta a cada nodo una única vez con cualquier otro nodo mediante una arista. El problema del agente viajero se puede estudiar como un problema de decisión, al incorporar una restricción de costo, es decir, se debe encontrar una ruta que una a todos los nodos y que conecte a un nodo con cualquier otro una única vez, sujeto a que esta ruta debe representar un costo total menor a k unidades, este costo total es el resultado de la suma de todos los costos resultantes de viajar de algún nodo i a un nodo j . Siguiendo la notación de [2], este problema se puede plantear de la siguiente manera:

$$\begin{aligned} TSP = \{ \langle G, c, k \rangle : & G = (V, E) \text{ es un grafo completo,} \\ & c \text{ es una función de } V \times V \rightarrow Z, \\ & k \in Z, \\ & G \text{ tiene una ruta del agente viajero de costo a lo más } k \} \end{aligned}$$

El problema del agente viajero es un problema **NP-completo**, la prueba se puede encontrar en [2]. Por ello se han propuesto heurísticas para encontrar soluciones factibles, como la heurística ambiciosa por ejemplo. Esta heurística consiste en seleccionar un vértice al azar, y

unirlo con el vértice que se encuentre más cercano, tomando como medida de distancia el costo asociado de la arista entre estos vértices. Una vez que se encontró el vértice más cercano, ahora se toma como punto de partida este nuevo vértice y se vuelve a unir con el vértice más cercano, de esta manera, se crea un camino entre el conjunto de vértices. Se puede consultar más acerca de este problema en [2].

Para nuestro problema de estudio es complicado que un único camión pueda cumplir con todo el plan de distribución en un intervalo de tiempo menor a media hora, lo cual podría ser necesario para mantener la calidad del servicio en momentos específicos del día, por ello, una generalización natural del problema del agente viajero es el problema del multi-agente viajero, el cual ahora considera que existen k agentes que deben recorrer un total de m ciudades, donde ningún agente puede visitar la ciudad que ya visitó otro agente, y además cada agente puede visitar cada ciudad a lo más una vez. También existe un costo positivo, $c_{i,j}$, asociado a ir de la i -ésima ciudad a la j -ésima. Dicho lo anterior, el objetivo es encontrar las k rutas que minimizan el costo total de visitar todas las ciudades.

Cuando el problema del multi-agente viajero se utiliza en un contexto de transporte, es decir, se asume que cada agente es un vehículo que carga o descarga cierta cantidad de un bien en cada ciudad visitada, a este problema se le conoce como el *Problema de enrutamiento de vehículos*, cuyo contexto se adecua a nuestro problema de encontrar la mejor ruta para transportar las bicicletas considerando que es posible contar con varios vehículos de transporte. Este problema también puede plantearse como una generalización del problema del agente viajero, ya que de acuerdo a [9], el problema de enrutamiento de vehículos consiste en encontrar una ruta mínima que contenga los subconjuntos A' , E' y V' , donde $A' \subset A$, $E' \subset E$ y $V' \subset V$, para los conjuntos de vértices, aristas y aristas dirigidas, V , E y A , respectivamente. Donde existe un peso para cada arista dirigida o no dirigida, y la ruta mínima se obtiene a través de la suma de pesos mínimos. De la definición anterior se obtiene el problema del agente viajero como un caso especial al considerar que E es un conjunto arbitrario, que $V = V'$ y que $A = A' = E' = \emptyset$.

En [19] se puede consultar más acerca del *Problema de enrutamiento de vehículos* y sus distintas variantes, en particular, en este trabajo se emplea como punto de partida la variante ahí presentada como *Problema de enrutamiento de vehículos con capacidad de carga*, la cual consiste en la distribución de cierta cantidad de un bien desde un único depósito denotado como punto cero, hacia un conjunto de n puntos, los cuales son llamados clientes. La cantidad del bien que debe llevarse al i -ésimo cliente es la demanda de dicho cliente, la cual está dada por la cantidad no negativa q_i . Se asume que hay una cantidad de K vehículos disponibles, donde cada uno tiene la misma capacidad Q para transportar los bienes desde el depósito hacia cualquier cliente, y los costos de operación de todos los vehículos son los mismos. Un vehículo que satisface la demanda de un conjunto de S clientes, comienza su operación desde el depósito, después visita a cada uno de los clientes que pertenecen al conjunto S para finalmente regresar al depósito, donde cada que el vehículo visita al j -ésimo cliente a partir del i -ésimo cliente se incurre en un costo de viaje denotado como $c_{i,j}$.

Siguiendo las ideas anteriores se puede definir una ruta como la secuencia $r = (i_0, i_1, \dots, i_s, i_{s+1})$ ($i_0 = i_{s+1} = 0$ para indicar que se parte del depósito y se regresa al depósito), donde $S = i_1, \dots, i_s$ es el conjunto de clientes visitados en esa ruta en particular. La ruta r tiene el costo $c(r) = \sum_{j=0}^s c_{i_j, i_{j+1}}$, y es factible si se mantiene la restricción de capacidad para el conjunto S , la cual puede formularse como $q(S) = \sum_{i \in S} q_i \leq Q$, y además ningún cliente es visitado más de una vez por algún vehículo. Entonces, una solución factible del *problema de enrutamiento de vehículos con capacidad de carga* con K vehículos consiste en encontrar K rutas factibles (una para cada vehículo). Por ejemplo, las rutas r_1, \dots, r_k que corresponden a los subconjuntos de clientes S_1, \dots, S_k , son una solución factible si todas las rutas son factibles, y los subconjuntos S_1, \dots, S_k forman una partición del conjunto total de clientes. Dicho lo anterior, encontrar soluciones factibles del *problema de enrutamiento de vehículos con capacidad de carga* consiste en resolver dos tareas interconectadas, es decir, se debe encontrar una partición factible de S_1, \dots, S_k subconjuntos de clientes, y para cada

vehículo k , debe encontrarse una ruta factible a través del subconjunto $0 \cup S_i$, donde 0 representa el depósito.

Encontrar soluciones para el *problema de enrutamiento de vehículos con capacidad de carga* es complicado, ya que se necesita resolver simultáneamente dos tareas que dependen una tanto de la otra. Sin embargo, notemos que si tuviéramos los subconjuntos S_1, \dots, S_k , podríamos encontrar rutas factibles dentro de cada subconjunto mediante la aplicación del problema del agente viajero.

Siguiendo las ideas anteriores y con el fin de encontrar un método que resuelva el problema anterior al tomar la mejor ruta de todas las posibles, para una instancia con N clientes (vértices) y K camiones, notemos que existen $\binom{N}{K}(N-K)^K$ formas de asignar N vértices en K subconjuntos S_1, S_2, \dots, S_k donde cada subconjunto contenga al menos un cliente (vértice), posteriormente, dentro de cada subconjunto debe resolverse el problema del agente viajero para encontrar la mejor ruta entre los elementos del subconjunto. Teniendo así que el problema de enrutamiento de vehículos con capacidad de carga para K vehículos es también un problema NP-Completo.

Encontrar soluciones exactas del problema del agente viajero es complicado, ya que como se comentó en secciones anteriores, no existe un algoritmo que en tiempo polinomial obtenga dichas soluciones. Dicho lo anterior, con el fin de encontrar soluciones subóptimas en tiempos eficientes, en este trabajo se utilizó una heurística ambiciosa, la cual consiste en que dado un vértice inicial, el siguiente vértice a seleccionar será aquel cuya distancia sea mínima respecto al vértice inicial, posteriormente este vértice recién seleccionado se convertirá en nuestro vértice de referencia, entonces el siguiente paso será actualizar el vértice de referencia lo cual se realiza simplemente al considerar nuevamente el vértice cuya distancia sea mínima al vértice de referencia y no haya sido seleccionado anteriormente, de esta manera, para un conjunto de N vértices simplemente hay que realizar n decisiones (selección de vértices), y de esta manera obtenemos un algoritmo que encuentra una solución subóptima en tiempo polinomial. La elección del vértice inicial

puede realizarse de manera aleatoria, o bien, se puede plantear realizar el ejercicio n veces, donde cada vez se selecciona un vértice inicial diferente, y al final nos quedamos con la solución que de la menor distancia mínima, o bien puede utilizarse un criterio experto, el cual consiste en seleccionar inicialmente algún vértice con una gran cantidad de bicicletas que ofrecer, ya que de esta manera el camión podría llenarse más rápidamente, para que así pase a la fase de descarga y con ello se reduzca el tiempo de operación. En este trabajo se realizó la selección del vértice inicial de acuerdo con el último criterio, es decir, el vértice donde cada camión comienza su ruta es un vértice con una gran cantidad de bicicletas.

De manera intuitiva, el algoritmo implementado en este trabajo consiste en que una vez que se especifica el número de camiones con el cual realizar el balanceo, es decir, K camiones, a cada camión se le asigna un vértice inicial, donde se selecciona este vértice inicial de tal manera que el camión comience su ruta con una gran cantidad de bicicletas. Posteriormente, cada camión selecciona el conjunto de estaciones (vértices) de las cuales tomará bicicletas al resolver el problema del agente viajero con heurística ambiciosa descrito anteriormente. Adicionalmente a este procedimiento, el camión cuenta con dos criterios de paro, el primero es que se llene. El segundo criterio de paro es que ya no haya cicloestaciones cercanas (vértices) que visitar, es decir, cicloestaciones que en caso de ser visitadas hagan que el vehículo tenga que recorrer más de diez kilómetros (este parámetro se puede ajustar dependiendo de las peculiaridades del problema), este criterio se utiliza debido a que se requiere que el vehículo pueda realizar el balanceo en un intervalo de tiempo de treinta minutos, y viajar diez kilómetros supone que el vehículo tarda los treinta minutos disponibles, suponiendo que la velocidad máxima a la que puede viajar el vehículo es de veinte kilómetros por hora. Posteriormente, una vez que al camión ya no le es posible cargar bicicletas entra en fase de descarga, donde nuevamente resuelve el problema del agente viajero con heurística ambiciosa para agregar a su ruta aquellas cicloestaciones a las cuales debe atender su demanda, para ello, toma como vértice inicial la última cicloestación de la cual recogió bicicletas. Dicho lo anterior, en el peor de los casos, para una instancia con

K camiones se debe resolver $2K$ veces el problema del agente viajero con heurística ambiciosa, con lo cual se obtiene un algoritmo capaz de encontrar soluciones subóptimas en tiempo polinomial. En la siguiente sección se describe el algoritmo implementado en este trabajo.

3.4.1. Algoritmo de solución multi agente viajero

El problema del multiagente viajero es una generalización natural del problema del agente viajero, este problema se puede plantear como encontrar las rutas más cortas en que un conjunto de k agentes deben visitar n ciudades, donde cada agente debe visitar una ciudad una única vez, y no puede visitar las ciudades que ya han sido visitadas por otros agentes. Al igual que el problema del agente viajero, este es un problema NP completo, por lo que se han propuesto heurísticas para su solución, como la heurística ambiciosa, la cual ha sido utilizada en este trabajo tomando en cuenta algunas peculiaridades del sistema de préstamo de bicicletas a analizar. A continuación se describe el algoritmo implementado en forma de pseudocódigo:

Código principal:

1. *Start Supply, Demand*
2. $V_1, \dots, V_k = \text{vehiculo}(1, C_j), \dots, \text{vehiculo}(k, C_k).$
3. $\text{vehiculos}, \text{vehiculos_Aux} = \text{set}\{1, \dots, k\}, \text{set}\{1, \dots, k\}$
4. *while*($\text{vehiculos} \neq \emptyset$)
5. $i = \text{random}(\text{vehiculos})$
6. $V_i.\text{PickUP}()$
7. *if* $V_i.\text{status} = \text{Delivery}$
8. $\text{vehiculos} = \text{vehiculos} - \{i\}$
9. *end while*
10. *while*(vehiculos is not $\{1, \dots, k\}$)
11. $i = \text{random}(\text{vehiculos_Aux})$
12. $V_i.\text{Delivery}()$
13. *if* $V_i.\text{status} = \text{Done}$

14. $vehiculos = vehiculos \cup \{i\}$
15. $vehiculos_Aux = vehiculos_Aux - \{i\}$
16. *end while*

Método *PickUP*

1. $Available = Set_Diff(Supply, V_i.route)$
2. $min, C_i = min_dist(V_i.route, Available)$
3. *if* $V_i.km + min < DT$ & $V_i.bikes + C_i.bikes < CC$
4. $V_i.km = V_i.km + min$ & $V_i.bikes = V_i.bikes + C_i.bikes$
5. $Supply = Set_Diff(Supply, C_i)$
6. $V_i.route.append(C_i)$
7. *else*
8. $V_i.status = Delivery$

Método *Delivery*

1. $min, C_i = min_dist(V_i.route, Demand)$
2. *if* $V_i.km + min < DT$ & $V_i.bikes - C_i.bikes \leq 0$
3. $V_i.km = V_i.km + min$ & $V_i.bikes = V_i.bikes - C_i.bikes$
4. $Demand = Set_Diff(Demand, C_i)$
5. $V_i.route.append(C_i)$
6. *else*
7. $V_i.status = Done$

En la línea 1 del código principal se inicializan los conjuntos referentes a las estaciones oferentes y a las estaciones demandantes como Supply y Demand, respectivamente. En la línea 2 se inicializan m vehículos, cada vehículo tiene un índice así como una cicloestación C_k de partida, al inicializar el vehículo, se carga con las bicicletas que haya en tal cicloestación. En la línea 3 se crea un conjunto con los índices de los vehículos. En la línea 4 se entra a un ciclo while cuya condición de salida es que el conjunto vehiculos sea vacío, dentro de este ciclo, primero se selecciona un índice de manera aleatoria del conjunto de vehículos (línea 5), posteriormente se ejecuta el método PickUp del vehículo cuyo índice fue generado de manera aleatoria previamente. Cada vez que es invocado el método PickUp, de algún vehículo, este encuentra la cicloestación más cercana de la cicloestación actual, si es posible visitar la cicloestación más cercana respetando las restricciones de distancia máxima recorrida por el vehículo, y capacidad total del vehículo, entonces el vehículo se dirige a esta cicloestación actualizando su estado actual así como su distancia recorrida, y como el número de bicicletas que ha cargado, en caso de que al vehículo ya no le sea posible visitar ninguna otra ciclo estación o cargar bicicletas, entonces entra en estado de Delivery. En caso de que el vehículo entre al estado Delivery, este es removido del conjunto de vehículos (línea 8), pues ya no le es posible cargar más bicicletas o recoger bicicletas según la restricción de distancia recorrida. Una vez que todos los vehículos ya no pueden avanzar más kilometros o recoger más bicicletas, es momento de que los vehículos atiendan la demanda, para ello inicializamos el ciclo while de la línea 10. Nuevamente, en la línea 11 se genera un índice de manera aleatoria correspondiente a algún vehículo. Posteriormente, en la línea 12 se inicia el método Delivery, del vehículo cuyo índice fue generado de manera aleatoria previamente. El método Delivery funciona de manera análoga al método PickUp, en este método primero se busca la cicloestación demandante más próxima, si es posible que el vehículo viaje a esta cicloestación manteniendo la restricción de distancia entonces el vehículo se desplaza a esta cicloestación. Una vez que al vehículo ya no le es posible avanzar más kilometros o se quedó sin bicicletas para atender la demanda de alguna cicloestación, el vehículo pasa a estado Done, y el índice del vehículo es agregado al conjunto vehículo, y eliminado del

conjunto `vehiculo_Aux` (actualizar el primer conjunto es necesario para finalizar el ciclo `while`, mientras que actualizar el segundo conjunto es necesario para ya no tomar en cuenta este vehículo).

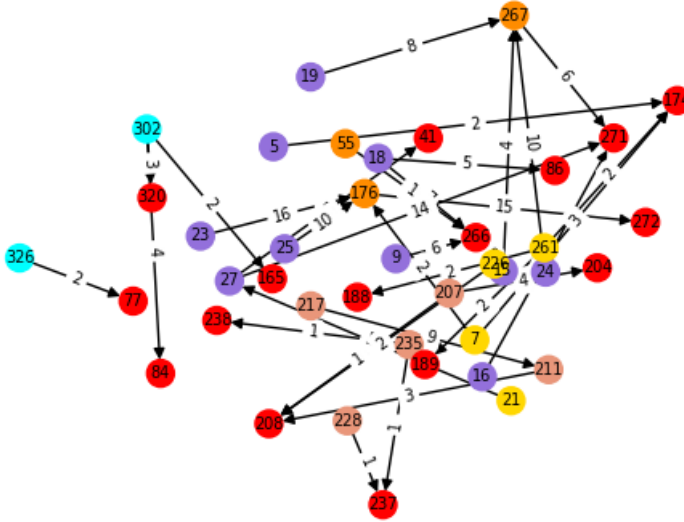
El método `PickUP`, en la línea 1 se obtienen las cicloestaciones disponibles para visitar. En la línea 2 se selecciona la cicloestación con distancia mínima del conjunto de cicloestaciones disponibles obtenidas en el paso anterior, dicha cicloestación está dada por la salida C_i , mientras que la distancia para llegar a esta cicloestación está dada por la salida `min`. Posteriormente, en la línea 3 comprueba si es posible viajar a la cicloestación con distancia mínima siempre y cuando se cumplan las restricciones de distancia y capacidad. En caso de que se cumplan las restricciones de distancia y capacidad, en las líneas 4-6 se actualizan el número de kilómetros viajados por el vehículo, la capacidad del vehículo y se elimina de las cicloestaciones oferentes la cicloestación recién visitada por el vehículo. Si ya no es posible visitar ninguna cicloestación, en la línea 8 el vehículo pasa a estado `Delivery`.

El método `Delivery` es muy similar al método `PickUP`, en la línea 1 se calcula la distancia mínima de las cicloestaciones demandantes (argumento `Demanda`), de estas cicloestaciones se selecciona aquella cuya distancia es mínima, la cual corresponde a la salida C_i , también se obtiene la distancia para llegar a esta cicloestación, la cual corresponde a la salida `min`. Posteriormente, en la línea 2 comprueba si es posible viajar a la cicloestación con distancia mínima siempre y cuando se cumplan las restricciones de distancia y capacidad. En caso de que se cumplan las restricciones de distancia y capacidad, en las líneas 3-5 se actualizan el número de kilómetros viajados por el vehículo, la capacidad del vehículo y se elimina de las cicloestaciones demandantes la cicloestación recién visitada por el vehículo. Si ya no es posible visitar ninguna cicloestación, en la línea 7 el vehículo pasa a estado `Done`.

3.4.2. Ejemplo del multi-agente viajero

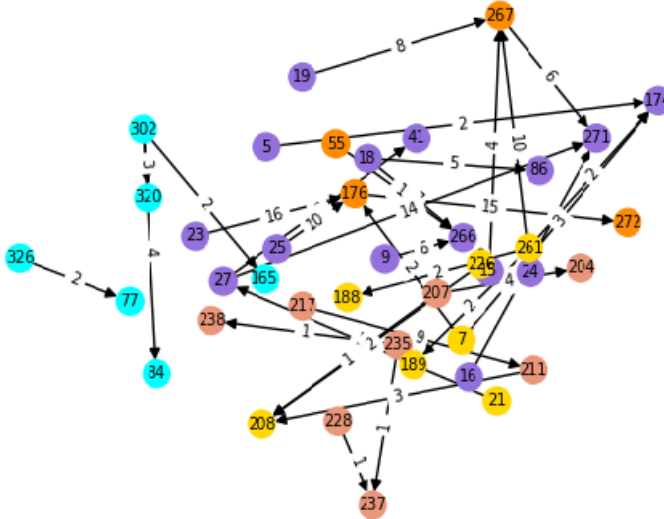
Si implementamos el algoritmo del multi-agente viajero con el fin de cumplir con el plan de distribución de la tabla 3.4, tenemos que consi-

derar con cuántos camiones podemos contar, la capacidad máxima de carga de cada camión, el tiempo máximo en que un camión tardará en cumplir con una ruta asignada, así como adecuar una estrategia de carga, es decir, si un camión alternadamente carga y descarga bicicletas, o si primero carga y después descarga. La estrategia a construir depende completamente del número de recursos disponibles, así como de las restricciones del servicio, por ejemplo, para el caso del sistema de bicicletas se esperaría atender la demanda de cierta hora en un intervalo no muy grande de tiempo, lo cual llevaría a considerar tanto el uso de varios camiones como la distancia máxima que un camión puede recorrer, pues esta última es proporcional al tiempo que tardará en completar la ruta. Para cumplir con el plan de distribución de la tabla 3.4, notemos que es conveniente adoptar una estrategia donde los camiones primero cargen bicicletas y después las transporten a donde es requerido, por ejemplo, si de la cicloestación 302 se requiere llevar bicicletas a las cicloestaciones 84 y 165, entonces por la desigualdad del triángulo, es menor $d(302, 84) + d(84, 165)$ que $d(302, 84) + d(84, 302) + d(302, 165)$. Una vez dicho lo anterior, podemos proponer atender el plan de distribución de la tabla 3.4 si consideramos cinco camiones que comienzan sus rutas en distintas cicloestaciones, con lo cual obtenemos el siguiente enrutamiento durante la fase de carga:



Gráfica 3.9: Rutas obtenidas mediante la fase de carga. El color rojo representa las cicloestaciones que demandan bicicletas, mientras que cada uno de los otros colores representa la ruta de cada uno de los cinco camiones, respectivamente.

Notemos que mediante la estrategia de recoger primero bicicletas, condicionamos a que en la fase de descarga los camiones se tienen que dirigir hacia las cicloestaciones que indica el plan de distribución de la tabla 3.4, donde para cada camión podemos aplicar localmente el algoritmo del agente viajero para que así cada camión encuentre la mejor ruta de las cicloestaciones demandantes de bicicletas, lo cual se ilustra en la figura 3.10.



Gráfica 3.10: Rutas obtenidas mediante la fase de descarga. Cada color representa la ruta total de cada uno de los cinco camiones, respectivamente.

La estrategia anterior puede ser algo restrictiva al considerar que los vehículos deben decidir sus movimientos de acuerdo al plan de distribución, y con ello no considerar otras mejores rutas, por ello a continuación se expone otra alternativa, la cual consiste en que cada camión se puede mover libremente hacia cualquier cicloestación que no haya sido visitada todavía por ningún otro camión, para ello, cada camión contará con una capacidad de carga máxima de 40 bicicletas, el camión comienza dirigiéndose hacia las cicloestaciones oferentes hasta que alcance su capacidad máxima de carga o haya recorrido diez kilómetros. Una vez que alguna de las dos condiciones anteriores sean satisfechas, el camión visitará a los cicloestaciones demandantes de bicicletas con el fin de atender su demanda, donde el camión terminará su recorrido si se queda sin bicicletas o si supera los 10 kilómetros de distancia total recorrida.

En las tablas 3.5 a 3.9 se muestran las rutas obtenidas de 5 vehícu-

los al aplicar la estrategia anterior. El renglón C_i hace referencia a la cicloestación, B al número de bicicletas a tomar o dejar según sea el caso, DA es la demanda atendida, mientras que DT es la distancia total recorrida. Cada columna representa el número de bicicletas que se deben tomar o dejar en la cicloestación correspondiente. Por último, es importante mencionar que las cicloestaciones demandantes requieren un total de 151 bicicletas, y mediante esta estrategia se consiguen transportar 134 bicicletas, por lo que la demanda atendida representa un 88.7 %, por lo que a diferencia de cumplir con el 100 % de la demanda como indica el plan de distribución, se logra satisfacer un porcentaje menor, pero con la ventaja de que es posible satisfacer condiciones adicionales como el número de kilómetros máximo que debe viajar cada vehículo, y además se reduce el número de camiones necesarios, ya que de requerir un camión para cada cicloestación demandante, se requieren únicamente cinco camiones en total.

	Tomar	Tomar	Tomar	Tomar	Dejar	Dejar	Dejar	Dejar
C_i	13	261	7	14	86	41	267	266
B	12	8	5	12	5	6	16	10
DA:	37							
DT:	6.27 km							

Tabla 3.5: Resumen de la ruta del vehículo 1.

	Tomar	Tomar	Tomar	Tomar	Tomar	Dejar	Dejar
C_i	16	15	24	5	9	174	271
B	13	12	10	2	6	7	30
DA:	40						
DT:	5.1 km						

Tabla 3.6: Resumen de la ruta del vehículo 2.

	Tomar	Tomar	Dejar	Dejar	Dejar	Dejar
C_i	236	235	237	208	189	77
B	10	11	2	6	2	2
DA:	12					
DT:	7.62 km					

Tabla 3.7: Resumen de la ruta del vehículo 3.

	Tomar	Tomar	Dejar	Dejar	Dejar	Dejar
C_i	226	217	211	204	188	165
B	10	9	6	4	2	2
DA:	14					
DT:	9.55 km					

Tabla 3.8: Resumen de la ruta del vehículo 4.

	Tomar	Tomar	Tomar	Dejar	Dejar	Dejar
C_i	21	23	25	176	272	84
B	5	16	10	13	15	4
DA:	31					
DT:	8.43 km					

Tabla 3.9: Resumen de la ruta del vehículo 5.

Capítulo 4

Resultados y discusión

A continuación se presentan los resultados del balanceo del sistema ecobici para cualquier día de la semana para los horarios de 08:30 a 09:00, y de 18:30 a 19:00, estos horarios fueron seleccionados, ya que como se menciona en [10], en estos horarios es cuando se registra mayor movilidad en el sistema, los cuales coinciden con el horario habitual laboral de la ciudad de México. Las estimaciones tanto de la oferta como de la demanda fueron realizadas de acuerdo a lo expuesto en las secciones 3.2.3 y 3.2.4, respectivamente. Estas estimaciones se realizaron con el 80 % de la información disponible, posteriormente con el 20 % de la información restante se calculó tanto el error promedio de estimación así como la desviación estándar, los cuales fueron aproximadamente de 10.5 % y 6.7 % para la oferta, y de 12.9 % y 10 % para la demanda. Para realizar el ejercicio de rebalanceo, se implementó el problema del multi-agente viajero descrito en la sección 3.4.2. Se seleccionó esta estrategia estrategia, debido a que esta nos permite encontrar mejores rutas factibles durante un intervalo de tiempo no mayor a media hora.

4.1. Caso Ecobici

4.1.1. Estaciones oferentes y demandantes en el tiempo t

A continuación, en las tablas 4.1 y 4.2 se muestran las cicloestaciones que fueron identificadas con algún problema de oferta o demanda

de acuerdo a lo expuesto en la sección 3.2.1. Las estimaciones de oferta para el horario de 08:30 a 09:00 pueden consultarse en la tabla 4.3, las estimaciones de demanda para el horario 08:30 a 09:00 pueden consultarse en la tabla 4.4. Las estimaciones de oferta para el horario de 18:30 a 19:00 pueden consultarse en la tabla 4.5, las estimaciones de demanda para el horario de 18:30 a 19:00 pueden consultarse en la tabla 4.6

5	7	9	13	14	15	16	18	19	21	23
24	25	27	41	55	77	84	86	165	174	176
188	189	197	198	204	207	208	211	217	226	228
235	236	237	238	261	266	267	271	272	302	320
326										

Tabla 4.1: Estaciones identificadas en el horario de 08:30 a 09:00 hrs.

5	6	7	9	11	13	14	15	16	17	18
19	22	23	24	25	26	27	28	30	36	37
38	39	41	47	48	51	53	55	63	75	77
84	85	86	88	116	117	118	134	136	138	140
141	143	150	153	158	161	162	165	167	168	169
171	172	173	174	176	181	183	186	188	189	191
197	198	199	200	201	202	204	205	206	207	208
209	210	211	214	217	219	224	225	226	227	232
233	235	236	239	242	248	249	250	253	254	255
256	259	261	264	265	266	267	269	271	272	273
274	282	285	286	288	290	301	302	303	304	305
315	316	320	326	329	330	331	335	351	384	420
436	448	450	452	454	455	458	459	460	463	465
466	468									

Tabla 4.2: Estaciones identificadas en el horario de 18:30 a 19:00 hrs.

Tabla 4.3: En esta tabla se muestra la oferta estimada de las cicloestaciones identificadas con saturación de bicicletas en el intervalo de tiempo de 08:30-09:00. La columna C_i hacer referencia a la i -ésima cicloestación identificada, mientras que las columnas lunes, martes, miércoles, jueves, viernes, sábado y domingo hacen referencia a la oferta estimada en cada uno de estos días, respectivamente.

C_i	lunes	martes	miércoles	jueves	viernes	sábado	domingo
5	7	8	8	8	8	1	1
7	10	11	12	11	13	2	0
9	10	11	12	10	11	1	1
13	13	15	15	13	12	1	1
14	18	20	20	20	18	2	2
15	16	17	17	17	16	2	1
16	16	21	18	20	16	1	2
18	17	19	18	19	18	2	2
19	13	15	15	16	16	1	1
21	8	8	9	8	6	0	0
23	20	20	21	19	19	2	1

C_i	lunes	martes	miércoles	jueves	viernes	sábado	domingo
24	12	15	13	13	12	2	2
25	12	13	13	14	12	2	2
27	19	17	15	17	14	2	2
41	6	6	5	6	5	1	1
55	2	3	2	2	2	1	1
77	3	3	3	4	3	1	1
84	4	5	4	5	5	1	0
86	2	3	3	3	4	1	1
165	1	2	2	2	1	1	0
174	3	4	4	3	3	2	2
176	1	0	1	0	0	0	0
188	3	3	2	3	2	0	0
189	1	1	1	1	1	0	0
197	14	14	16	16	15	1	1
198	4	4	4	4	4	1	0
204	2	2	2	2	2	1	1
207	7	8	8	7	7	1	1
208	3	3	3	3	4	1	1
211	4	5	4	4	5	1	0
217	12	12	13	12	12	2	1
226	11	11	11	9	8	1	1
228	7	8	9	8	7	1	1
235	11	11	11	11	10	1	1
236	11	11	11	10	10	1	0
237	3	4	4	4	4	0	1
238	2	2	2	2	2	1	1
261	10	12	10	11	12	1	0
266	3	3	3	3	3	1	1
267	1	1	1	0	0	1	1
271	1	1	1	1	1	1	1
272	1	1	0	0	0	0	0
302	11	13	12	12	11	2	2
320	1	1	1	1	0	0	1
326	8	10	9	10	9	1	1

Tabla 4.4: En esta tabla se muestra la demanda estimada de las cicloestaciones identificadas con escasos de bicicletas en el intervalo de tiempo de 08:30-09:00. La columna C_i hace referencia a la i -ésima cicloestación identificada, mientras que las columnas lunes, martes, miércoles, jueves, viernes, sábado y domingo hacen referencia a la demanda estimada en cada uno de estos días, respectivamente.

C_i	lunes	martes	miércoles	jueves	viernes	sábado	domingo
5	5	6	5	5	5	1	1
7	5	6	5	6	5	1	1
9	4	4	4	4	3	1	1
13	1	1	1	1	1	0	1
14	6	6	6	6	5	1	2
15	4	5	4	4	4	1	1
16	3	3	3	3	3	1	1
18	5	6	6	5	5	1	2
19	5	6	6	5	4	1	1
21	3	4	4	4	2	1	2
23	4	4	4	4	3	1	2
24	2	3	3	3	3	0	1
25	2	2	2	1	2	1	1
27	4	4	4	4	3	1	2
41	12	12	13	13	11	2	3
55	1	1	1	1	1	0	0
77	5	6	6	6	6	1	1
84	8	9	8	9	8	2	1
86	7	6	8	6	8	2	1
165	3	3	3	3	3	2	1
174	10	10	10	8	7	2	2
176	14	11	19	14	17	1	1
188	5	5	4	5	5	1	1
189	3	5	4	5	5	2	1
197	2	3	3	3	2	1	0
198	0	0	1	1	0	0	1
204	6	6	6	8	8	1	1
207	2	2	2	2	2	0	0

C_i	lunes	martes	miércoles	jueves	viernes	sábado	domingo
208	9	9	9	8	10	3	2
211	10	10	10	13	11	2	1
217	3	3	3	3	2	0	1
226	1	2	2	2	2	0	0
228	6	6	6	5	5	1	1
235	0	1	1	0	0	0	0
236	1	1	1	1	1	1	1
237	5	5	6	4	4	1	1
238	3	4	4	4	3	1	0
261	2	2	1	1	2	0	1
266	25	28	33	27	21	2	1
267	17	15	18	18	18	1	0
271	34	30	36	36	36	5	3
272	16	18	21	12	16	5	1
302	2	2	1	2	1	0	0
320	4	4	4	4	4	2	1
326	2	2	2	2	2	0	1

Tabla 4.5: En esta tabla se muestra la oferta estimada de las cicloestaciones identificadas con saturación de bicicletas en el intervalo de tiempo de 18:30-19:00. La columna C_i hacer referencia a la i -ésima cicloestación identificada, mientras que las columnas lunes, martes, miércoles, jueves, viernes, sábado y domingo hacen referencia a la oferta estimada en cada uno de estos días, respectivamente.

C_i	lunes	martes	miércoles	jueves	viernes	sábado	domingo
5	3	3	3	3	3	1	1
6	4	4	5	5	3	2	2
7	4	4	4	4	4	1	2
9	2	3	3	3	3	1	2
11	3	3	3	2	3	1	1
13	2	1	2	2	1	1	1
14	4	4	5	4	3	2	1
15	6	6	6	5	5	3	2
16	4	4	3	4	2	1	1

C_i	lunes	martes	miércoles	jueves	viernes	sábado	domingo
17	2	3	3	3	2	1	2
18	6	6	7	6	5	1	1
19	2	3	3	3	2	2	1
22	2	2	2	2	2	1	1
23	2	3	3	2	2	2	1
24	5	6	6	6	4	1	1
25	4	5	5	6	4	2	1
26	5	5	5	6	4	2	1
27	15	15	15	15	12	4	4
28	9	8	9	9	5	2	1
30	4	4	4	5	5	3	2
36	9	10	10	10	8	2	2
37	5	6	5	5	4	2	1
38	8	9	9	8	6	2	1
39	3	3	4	3	3	1	1
41	9	11	10	9	12	2	2
47	8	7	8	8	6	2	2
48	4	4	4	4	3	1	1
51	5	5	5	5	4	2	2
53	7	7	7	5	4	2	2
55	3	4	4	4	3	1	1
63	4	4	4	5	4	2	2
75	3	4	4	4	3	1	1
77	6	6	5	6	4	1	1
84	7	7	8	7	4	1	1
85	5	6	7	7	4	2	1
86	6	6	5	6	5	1	2
88	1	1	2	1	2	1	1
116	5	5	5	6	4	2	2
117	4	4	4	4	3	2	1
118	4	4	4	3	3	1	1
134	5	6	7	6	5	2	2
136	7	7	7	7	6	2	2
138	4	5	4	4	3	1	1
140	4	4	4	3	3	2	1

C_i	lunes	martes	miércoles	jueves	viernes	sábado	domingo
141	6	7	6	7	7	2	2
143	3	3	4	3	2	1	1
150	7	9	8	7	4	1	1
153	4	4	4	4	3	1	1
158	11	11	12	12	8	1	1
161	2	2	2	2	2	1	1
162	6	5	6	6	5	2	1
165	4	5	4	5	4	2	1
166	3	4	3	4	3	1	1
167	4	4	5	4	3	1	1
168	4	3	3	3	3	1	1
169	5	5	4	4	3	2	2
171	4	3	4	3	2	1	1
172	4	4	4	4	3	1	1
173	3	3	3	3	2	1	1
174	13	13	13	14	10	2	2
176	14	11	14	15	12	1	1
181	4	4	5	4	3	1	1
183	6	7	6	7	5	2	2
186	5	5	5	4	3	2	1
188	2	3	2	2	2	1	1
189	2	2	2	2	2	1	1
191	2	3	3	3	2	1	1
197	2	3	2	3	3	1	1
198	1	1	1	1	1	1	0
199	1	1	1	0	0	0	0
200	2	2	2	2	2	1	1
201	2	2	2	2	1	1	0
202	1	1	1	1	1	1	1
204	3	3	3	3	1	0	0
205	3	3	2	3	2	1	1
206	4	4	3	4	3	2	1
207	2	2	2	2	2	1	1
208	8	8	8	8	5	1	2
209	3	3	3	2	2	0	1

C_i	lunes	martes	miércoles	jueves	viernes	sábado	domingo
210	3	2	3	3	2	1	0
211	8	8	8	10	5	1	1
214	1	1	1	1	1	1	1
217	8	8	8	8	5	2	1
219	2	2	2	2	1	1	0
224	1	1	1	1	1	0	1
225	1	2	2	1	1	1	0
226	2	2	3	2	1	1	0
227	2	1	1	1	1	0	0
232	1	2	1	2	1	1	1
233	1	2	2	2	2	1	1
235	1	2	2	1	1	1	0
236	1	1	0	1	1	1	0
239	2	2	2	2	1	1	1
242	5	4	5	5	3	1	1
248	2	2	2	2	1	1	1
249	1	1	1	1	1	0	1
250	2	2	2	2	1	1	0
253	4	4	4	4	2	0	1
254	9	8	9	8	5	1	1
255	6	6	6	6	4	1	2
256	5	5	6	6	4	1	1
259	3	4	4	4	3	1	1
261	1	1	1	1	1	1	1
264	6	5	5	4	4	1	1
265	10	9	9	8	6	1	1
266	33	37	35	36	32	3	2
267	36	36	36	36	29	3	2
269	7	6	5	3	2	1	1
271	25	24	25	21	16	5	4
272	18	18	19	21	11	2	1
273	20	19	16	15	9	1	1
274	11	11	11	10	6	1	2
282	3	3	3	3	2	1	1
285	3	3	3	3	2	1	1

C_i	lunes	martes	miércoles	jueves	viernes	sábado	domingo
286	3	3	3	3	2	1	1
288	2	3	3	3	2	1	1
290	2	2	2	2	1	1	1
301	2	2	2	3	2	1	1
302	2	3	2	3	2	1	1
303	2	2	2	2	1	1	1
304	2	1	1	1	1	1	1
305	5	5	5	4	2	1	1
315	1	1	1	1	1	1	0
316	2	2	1	2	2	1	1
320	3	2	3	3	2	1	1
326	2	2	2	2	2	1	1
329	4	4	4	4	3	1	1
330	1	1	1	1	1	1	1
331	2	2	2	2	2	1	0
335	1	1	1	1	1	1	1
351	5	4	5	4	2	1	1
384	6	5	5	5	3	1	0
420	1	1	1	1	1	1	1
436	1	1	1	1	1	1	1
448	6	5	5	5	4	1	1
450	5	5	4	5	4	1	1
452	3	3	3	2	1	1	1
454	1	1	0	0	1	1	1
455	1	1	1	1	1	1	0
458	1	1	1	1	1	0	0
459	1	1	1	1	1	1	0
460	2	2	2	2	1	1	0
463	2	2	2	2	1	1	1
465	4	3	3	3	2	1	1
466	3	3	3	2	2	0	0
468	1	1	1	1	1	1	1

Tabla 4.6: En esta tabla se muestra la demanda estimada de las cicloestaciones identificadas con escasos de bicicletas en el intervalo de tiempo de 18:30-19:00. La columna C_i hace referencia a la i -ésima cicloestación identificada, mientras que las columnas lunes, martes, miércoles, jueves, viernes, sábado y domingo hacen referencia a la demanda estimada en cada uno de estos días, respectivamente.

C_i	lunes	martes	miércoles	jueves	viernes	sábado	domingo
5	4	6	5	4	3	1	2
6	5	5	5	5	4	2	1
7	7	8	8	7	5	1	1
9	7	8	7	7	4	1	1
11	5	5	6	6	4	1	1
13	9	11	12	13	5	1	1
14	5	6	7	5	5	2	1
15	7	7	8	6	7	2	2
16	7	8	7	8	5	2	1
17	7	8	9	8	5	1	1
18	9	11	11	10	8	2	2
19	5	7	6	5	5	2	1
22	13	11	13	11	5	1	2
23	7	8	7	7	5	2	2
24	7	8	9	9	6	2	1
25	6	6	8	9	7	2	2
26	5	4	5	5	4	2	2
27	17	16	17	17	12	4	3
28	10	10	11	10	6	2	1
30	7	8	7	8	6	2	2
36	13	15	14	13	8	2	2
37	8	8	8	7	4	2	1
38	10	11	10	10	8	1	1
39	5	4	5	4	3	2	1
41	7	6	6	6	5	3	3
47	9	9	9	9	6	2	2
48	2	3	3	3	2	2	1
51	7	6	7	6	5	2	2

C_i	lunes	martes	miércoles	jueves	viernes	sábado	domingo
53	5	5	5	4	4	2	2
55	2	2	3	2	2	1	1
63	6	6	5	5	4	1	2
75	3	4	3	4	3	2	2
77	5	6	5	6	3	2	1
84	5	6	5	6	4	1	1
85	4	4	4	4	3	1	1
86	4	5	5	5	3	2	1
88	2	2	2	2	2	1	1
116	4	5	4	5	3	2	2
117	4	5	5	5	4	1	1
118	4	5	4	3	3	1	1
134	5	6	6	5	4	2	2
136	6	8	6	6	5	2	2
138	2	2	2	2	2	1	1
140	3	3	3	2	2	1	1
141	5	6	5	5	5	2	2
143	2	2	2	2	2	1	1
150	3	3	3	3	3	2	1
153	2	3	3	3	2	1	1
158	3	3	3	3	3	2	2
161	2	2	2	2	2	1	1
162	3	3	4	3	3	1	1
165	3	3	2	2	2	2	1
166	2	3	3	3	2	1	1
167	3	4	4	3	2	2	1
168	2	2	2	2	2	1	1
169	2	2	2	2	2	1	1
171	2	2	2	2	2	1	1
172	2	3	3	2	2	1	1
173	2	2	2	2	2	1	1
174	5	6	6	6	5	2	2
176	0	0	0	0	0	0	0
181	2	3	3	2	2	1	1
183	3	3	3	3	2	1	1

C_i	lunes	martes	miércoles	jueves	viernes	sábado	domingo
186	4	4	4	4	3	1	1
188	2	2	3	2	2	1	1
189	2	2	2	2	2	1	0
191	2	3	3	2	2	1	1
197	3	4	4	4	3	2	2
198	2	2	2	2	3	1	1
199	6	6	6	6	3	1	1
200	5	4	4	4	2	1	1
201	2	2	2	2	2	0	1
202	5	5	5	5	3	1	1
204	2	2	2	2	2	1	1
205	5	6	4	5	2	1	1
206	5	3	4	5	3	2	1
207	7	8	8	8	4	1	1
208	8	8	8	8	6	1	1
209	5	4	4	4	2	1	1
210	3	3	4	3	2	1	1
211	6	7	8	7	5	1	1
214	3	4	4	4	2	1	1
217	13	13	13	13	8	2	1
219	3	3	3	2	2	1	1
224	3	4	4	4	2	1	1
225	3	4	4	4	2	1	1
226	3	4	4	3	2	1	1
227	2	2	2	2	2	1	1
232	5	5	5	4	3	1	1
233	4	5	4	4	3	1	1
235	5	6	7	6	4	1	1
236	6	5	8	7	4	0	1
239	4	4	3	3	2	1	1
242	5	6	6	6	4	2	2
248	4	5	5	4	3	1	0
249	3	3	4	3	2	1	1
250	5	6	6	6	3	1	1
253	6	5	5	4	4	1	1

C_i	lunes	martes	miércoles	jueves	viernes	sábado	domingo
254	2	2	2	2	2	1	2
255	1	2	2	2	2	1	0
256	3	3	3	3	3	1	1
259	2	1	2	1	1	1	1
261	9	10	8	9	4	1	0
264	0	1	0	1	0	1	0
265	0	1	1	1	1	1	1
266	1	1	1	1	1	1	1
267	0	0	0	0	0	0	0
269	0	0	0	0	0	0	0
271	3	3	4	3	3	3	2
272	2	2	2	2	1	0	1
273	0	0	0	0	0	0	0
274	1	2	1	1	1	1	1
282	2	2	2	2	2	1	1
285	1	2	2	1	1	1	1
286	4	4	4	4	3	1	1
288	2	2	2	1	2	1	1
290	4	4	4	4	2	0	0
301	5	5	4	5	4	1	1
302	6	7	6	6	4	1	1
303	5	4	4	5	3	1	1
304	4	4	5	5	3	1	1
305	7	8	7	7	4	1	1
315	3	3	4	3	2	1	1
316	4	4	4	3	2	1	1
320	2	2	2	2	1	1	1
326	4	4	3	4	2	1	1
329	3	3	3	3	2	1	1
330	1	1	1	1	1	1	1
331	2	2	3	2	1	0	1
335	4	5	5	4	2	1	1
351	3	3	3	2	2	1	1
384	6	5	5	5	3	1	1
420	3	4	4	3	2	1	1

C_i	lunes	martes	miércoles	jueves	viernes	sábado	domingo
436	3	3	4	4	2	1	1
448	1	1	1	1	1	1	1
450	1	1	1	1	1	1	1
452	3	3	4	3	2	1	0
454	3	4	4	5	1	1	1
455	2	3	2	2	1	1	1
458	1	2	2	2	1	1	1
459	3	3	3	2	1	2	1
460	4	4	4	3	1	1	1
463	3	3	3	3	2	0	1
465	2	2	2	2	1	0	0
466	1	2	1	2	2	1	1
468	4	4	4	4	4	0	0

4.1.2. Porcentaje de llenado de las estaciones en el tiempo t

En las tablas 4.7-4.16 se muestra los resultados obtenidos al balancear las cicloestaciones identificadas en la tabla 4.1.1. Estas tablas muestran la ruta, las bicicletas cargadas o descargadas, así como la demanda atendida y el número total de kilómetros recorridos por cada vehículo, donde V_i se refiere al i -ésimo vehículo utilizado, B se refiere al número de bicicletas cargadas o descargadas, DA se refiere a la demanda total atendida y KM al número total de kilómetros recorridos. En las tablas 4.17 y 4.18 se muestra la demanda total atendida como porcentaje de la demanda total requerida para cada día de la semana tanto a las 08:30 como a las 18:30, respectivamente. La columna OT se refiere a la oferta total, la columna DT se refiere a la demanda total, las columnas OD y DA se refieren a la oferta ofrecida y a la demanda requerida, es decir, estas dos cantidades resultan de a cada cicloestación restarle su demanda propia, con ello, OD se refiere a la demanda disponible después de atender la demanda propia de cada cicloestación, mientras que DR se refiere a la demanda requerida después de considerar la oferta propia de cada cicloestación. La columna DA se refiere a la demanda atendida del total de demanda requerida, mientras que la columna $\%DA$ se refiere a

la demanda atendida en relación a la demanda requerida.

V_1	C_i	13	261	7	86	41	267		
	B	12	8	5	-5	-6	-16		
	DA	25							
	KM	5.09							
V_2	C_i	16	14	9	5	18	174	176	271
	B	13	12	6	2	12	-7	-13	-33
	DA	45							
	KM	5.89							
V_3	C_i	236	235	237	208	189	77		
	B	10	11	-2	-6	-2	-2		
	DA	12							
	KM	7.62							
V_4	C_i	226	217	211	204	188	84		
	B	10	9	-6	-4	-2	-2		
	DA	16							
	KM	8.73							
V_5	C_i	23	24	25	266	272			
	B	16	10	10	-22	-15			
	DA	36							
	KM	3.19							

Tabla 4.7: Lunes 08:30

V_1	C_i	13	261	7	86	41	267		
	B	14	10	5	-3	-6	-14		
	DA	23							
	KM	5.09							
V_2	C_i	16	14	9	5	18	174	176	271
	B	18	14	7	2	13	-6	-11	-29
	DA	46							
	KM	5.89							
V_3	C_i	236	235	238	208	189	77		
	B	10	10	-2	-6	-4	-3		
	DA	15							
	KM	7.62							
V_4	C_i	226	217	211	204	188	84		
	B	9	9	-5	-4	-2	-4		
	DA	15							
	KM	8.73							
V_5	C_i	23	24	25	266	272			
	B	16	12	11	-25	-17			
	DA	39							
	KM	3.19							

Tabla 4.8: Martes 08:30

V_1	C_i	13	261	7	86	41	267		
	B	14	9	7	-5	-8	-17		
	DA	30							
	KM	5.09							
V_2	C_i	16	14	9	5	18	174	176	271
	B	15	14	8	3	12	-6	-18	-28
	DA	52							
	KM	5.89							
V_3	C_i	236	235	237	238	204	188		
	B	10	10	-2	-2	-4	-2		
	DA	10							
	KM	6.98							
V_4	C_i	226	217	211	208	189	77		
	B	9	10	-6	-6	-3	-3		
	DA	18							
	KM	6.49							
V_5	C_i	23	24	25	266	272			
	B	17	10	11	-30	-8			
	DA	38							
	KM	3.19							

Tabla 4.9: Miércoles 08:30

V_1	C_i	13	261	7	86	41	267		
	B	12	10	5	-3	-7	-17		
	DA	27							
	KM	5.09							
V_2	C_i	16	14	9	5	18	174	176	271
	B	17	14	6	3	14	-5	-14	-35
	DA	54							
	KM	5.89							
V_3	C_i	236	235	238	208	189	77		
	B	9	11	-2	-5	-4	-2		
	DA	13							
	KM	7.62							
V_4	C_i	226	217	211	204	188			
	B	7	9	-9	-6	-1			
	DA	16							
	KM	8.73							
V_5	C_i	23	24	25	266	272			
	B	15	10	13	-24	-12			
	DA	36							
	KM	3.19							

Tabla 4.10: Jueves 08:30

V_1	$\frac{C_i}{\text{B}}$	13	261	7	86	41	267		
	$\frac{\text{DA}}{\text{KM}}$	11	10	8	-4	-6	-18		
		28							
		5.09							
V_2	$\frac{C_i}{\text{B}}$	16	14	9	5	18	174	176	271
	$\frac{\text{DA}}{\text{KM}}$	13	13	8	3	13	-4	-17	-29
		50							
		5.89							
V_3	$\frac{C_i}{\text{B}}$	236	235	208	204	188	84		
	$\frac{\text{DA}}{\text{KM}}$	9	10	-6	-6	-3	-3		
		18							
		9.54							
V_4	$\frac{C_i}{\text{B}}$	226	217	211	189	77	165		
	$\frac{\text{DA}}{\text{KM}}$	6	10	-6	-4	-3	-2		
		15							
		8.34							
V_5	$\frac{C_i}{\text{B}}$	23	24	25	266	272			
	$\frac{\text{DA}}{\text{KM}}$	16	9	10	-18	-16			
		34							
		3.19							

Tabla 4.11: Viernes 08:30

V_1	C_i	158	162	153	150	84	63	51	37	232	199
	B	8	3	2	4	2	-2	-2	-3	-4	-5
	DA	16									
	KM	9.74									
V_2	C_i	176	269	41	30	25	22	468	214		
	B	14	7	2	-3	-2	-11	-3	-2		
	DA	21									
	KM	10.36									
V_3	C_i	266	19	39	36	207	209				
	B	32	-3	-2	-4	-5	-2				
	DA	16									
	KM	9.5									
V_4	C_i	267	11	23	253	225	235				
	B	36	-2	-5	-2	-2	-4				
	DA	15									
	KM	9.39									
V_5	C_i	271	27	24	248	205	460				
	B	22	-2	-2	-2	-2	-2				
	DA	10									
	KM	9.74									
V_6	C_i	272	264	18	13	249	224	454			
	B	16	6	-3	-7	-2	-2	-2			
	DA	16									
	KM	9.18									
V_7	C_i	273	17	261	250	200					
	B	20	-5	-8	-3	-3					
	DA	19									
	KM	6.89									
V_8	C_i	274	450	48	16	7	239	233	236		
	B	10	4	2	-3	-3	-2	-3	-5		
	DA	16									
	KM	9.5									
V_9	C_i	265	256	183	255	9	38	217	202		
	B	10	2	3	5	-5	-2	-5	-4		
	DA	16									
	KM	8.91									

Tabla 4.12: Lunes 18:30

V_1	C_i	41	450	265	19	11	18	14	249	214	
	B	5	4	8	-4	-2	-5	-2	-2	-3	
	DA	17									
	KM	9.73									
V_2	C_i	158	162	165	150	169	290	305	304	303	
	C_i	315	316	326	420						
	B	8	2	2	6	3	-2	-3	-3	-2	
	B	-2	-2	-2	-3						
	DA	19									
	KM	10.97									
V_3	C_i	266	17	13	250	217	224	235			
	B	35	-5	-10	-4	-5	-3	-4			
	DA	31									
	KM	8.49									
V_4	C_i	267	16	261	239	232	233				
	B	36	-4	-9	-2	-3	-3				
	DA	21									
	KM	7.1									
V_5	C_i	271	176	5	7	22	468	200			
	B	21	11	-3	-4	-9	-3	-2			
	DA	21									
	KM	9.77									
V_6	C_i	272	269	23	30	36	207	202			
	B	16	6	-5	-4	-5	-6	-4			
	DA	22									
	KM	10.26									
V_7	C_i	273	9	38	63	301	302	335			
	B	19	-5	-2	-2	-3	-4	-4			
	DA	19									
	KM	9.36									
V_8	C_i	274	264	24	28	248	225	236			
	B	9	4	-2	-2	-3	-2	-4			
	DA	13									
	KM	9.51									
V_9	C_i	259	256	183	255	85	47	37	226	205	199
	B	3	2	4	4	2	-2	-2	-2	-3	-5
	DA	15									
	KM	10.76									

Tabla 4.13: Martes 18:30

V_1	$\frac{C_i}{B}$	41	450	265	19	11	15	28	301	
	$\frac{B}{DA}$	4	3	8	-3	-3	-2	-2	-2	
	$\frac{DA}{KM}$	12								
	$\frac{KM}{KM}$	9.74								
V_2	$\frac{C_i}{C_i}$	158	162	165	150	84	290	305	304	303
	$\frac{C_i}{B}$	315	316	326						
	$\frac{B}{B}$	9	2	2	5	3	-2	-2	-4	-2
	$\frac{B}{DA}$	-3	-3	-4						
	$\frac{DA}{KM}$	20								
	$\frac{KM}{KM}$	7.68								
V_3	$\frac{C_i}{B}$	266	27	30	51	248	468	199		
	$\frac{B}{DA}$	34	-2	-3	-2	-3	-3	-5		
	$\frac{DA}{KM}$	18								
	$\frac{KM}{KM}$	10.73								
V_4	$\frac{C_i}{B}$	267	18	14	249	217	224	214	460	
	$\frac{B}{DA}$	36	-4	-2	-3	-5	-3	-3	-2	
	$\frac{DA}{KM}$	22								
	$\frac{KM}{KM}$	9.32								
V_5	$\frac{C_i}{B}$	271	17	23	37	207	200			
	$\frac{B}{DA}$	21	-6	-4	-3	-6	-2			
	$\frac{DA}{KM}$	21								
	$\frac{KM}{KM}$	8.24								
V_6	$\frac{C_i}{B}$	272	176	16	13	250	225	233	235	454
	$\frac{B}{DA}$	17	14	-4	-10	-4	-2	-2	-5	-4
	$\frac{DA}{KM}$	31								
	$\frac{KM}{KM}$	10.01								
V_7	$\frac{C_i}{B}$	274	269	5	7	22				
	$\frac{B}{DA}$	10	5	-2	-4	-11				
	$\frac{DA}{KM}$	15								
	$\frac{KM}{KM}$	5.84								
V_8	$\frac{C_i}{B}$	273	25	24	36	205	202			
	$\frac{B}{DA}$	16	-3	-3	-4	-2	-4			
	$\frac{DA}{KM}$	16								
	$\frac{KM}{KM}$	9.23								
V_9	$\frac{C_i}{B}$	259	256	183	255	85	254	9	261	232
	$\frac{B}{DA}$	2	3	3	4	3	7	-4	-7	-4
	$\frac{DA}{KM}$	22								
	$\frac{KM}{KM}$	9.59								

Tabla 4.14: Miércoles 18:30

V_1	C_i	158	162	165	290	305	304	303	315
	B	9	3	3	-2	-3	-4	-3	-2
	DA	14							
	KM	5.32							
V_2	C_i	176	269	450	23	37	207	199	209
	B	15	32	4	-5	-2	-6	-6	-2
	DA	21							
	KM	10.07							
V_3	C_i	266	25	38	217	202	214	235	
	B	35	-3	-2	-5	-4	-3	-5	
	DA	22							
	KM	10.36							
V_4	C_i	267	16	7	232	233	454		
	B	36	-4	-3	-2	-2	-5		
	DA	16							
	KM	8.88							
V_5	C_i	271	17	261	250	200			
	B	18	-5	-8	-4	-1			
	DA	18							
	KM	6.99							
V_6	C_i	272	264	18	13	249	224		
	B	19	3	-4	-11	-2	-3		
	DA	20							
	KM	6.49							
V_7	C_i	273	27	24	36	205			
	B	15	-2	-3	-3	-2			
	DA	10							
	KM	7.81							
V_8	C_i	274	265	259	11	9	248	225	236
	B	9	7	3	-4	-4	-2	-3	-6
	DA	19							
	KM	8.21							
V_9	C_i	266	19	30	22	468			
	B	35	-2	-3	-9	-3			
	DA	17							
	KM	7.24							

Tabla 4.15: Jueves 18:30

V_1	C_i	41	450	19	13	217		
	B	7	3	-3	-4	-3		
	DA	10						
	KM	5.9						
V_2	C_i	158	162	165	141	25	38	207
	B	5	2	2	2	-3	-2	-2
	DA	7						
	KM	8.85						
V_3	C_i	174	254	259	18	261	236	
	B	5	3	2	-3	-3	-3	
	DA	9						
	KM	7.04						
V_4	C_i	176	17	22	303	304	305	
	B	12	-3	-3	-2	-2	-2	
	DA	12						
	KM	8.86						
V_5	C_i	266	16	253	202			
	B	31	-3	-2	-2			
	DA	7						
	KM	7.75						
V_6	C_i	267	15	250	199			
	B	29	-2	-2	-3			
	DA	7						
	KM	7.62						
V_7	C_i	271	23	248	235			
	B	13	-3	-2	-3			
	DA	8						
	KM	8.68						
V_8	C_i	272	24	468	198			
	B	10	-2	-3	-2			
	DA	7						
	KM	8.7						
V_9	C_i	273	14	232	301			
	B	9	-2	-2	-2			
	DA	6						
	KM	10.39						

Tabla 4.16: Viernes 18:30

	OT	DT	OD	DR	DA	%DA
Lunes	344	271	224	151	136	90 %
Martes	377	281	242	146	138	94 %
Miércoles	368	306	241	179	148	82 %
Jueves	364	283	239	158	146	92 %
Viernes	343	270	230	157	145	92 %

Tabla 4.17: Porcentaje de llenado 08:30

	OT	DT	OD	DR	DA	%DA
Lunes	687	603	275	191	145	75 %
Martes	698	652	260	214	178	83 %
Miércoles	700	653	264	217	177	81 %
Jueves	687	620	264	197	157	79 %
Viernes	518	438	184	104	73	70 %

Tabla 4.18: Porcentaje de llenado 18:30

4.1.3. Tiempo promedio del balanceo en el tiempo t

A continuación en los cuadros 4.19 y 4.20 se muestra tanto la distancia promedio como el tiempo promedio que requiere para completarse cada una de las rutas mostradas en los cuadros 4.7-4.16.

	KM	T
Lunes	6.1	18.3
Martes	6.18	18.3
Miércoles	5.5	16.5
Jueves	6.1	18.3
Viernes	6.4	19.2

Tabla 4.19: Distancia y tiempo promedio 08:30

	KM	T
Lunes	9.2	27.7
Martes	9.5	28.6
Miércoles	8.9	26.7
Jueves	8.0	24.0
Viernes	8.1	24.5

Tabla 4.20: Distancia y tiempo promedio 18:30

4.2. Discusión

4.3. Explicación de los resultados

Los resultados obtenidos son referentes a lo que se esperaría observar en promedio en un día de la semana de operación en cierto intervalo de tiempo en particular. Como se menciona en la sección 3, estos resultados fueron obtenidos tras el análisis de un año de información histórica, es decir, del año 2019. Se ha considerado esta información, y no la más actualizada, como 2020 o inicios del 2121 porque los patrones de movilidad detectados en estos años son casi nulos, lo cual se explica a la pandemia a la que el mundo se enfrenta.

En las tablas 4.1 y 4.2 se muestra tanto la demanda total esperada así como la oferta total esperada de todas las cicloestaciones detectadas con algún problema de oferta o demanda de acuerdo al índice de saturación definido en la sección 3.2. Por ejemplo, se muestra que para el día lunes a las 08:30 hrs, se espera que las cicloestaciones mostradas en la tabla 4.1, tengan una oferta total de 344 bicicletas (OT) y una demanda total de 271 bicicletas (DT).

Posteriormente, para no crear un problema de desabasto, primero se atiende la demanda de cada cicloestación con su propia oferta, es decir, por ejemplo, como se muestra en las tablas 4.3 y 4.4, si la cicloestación 5 tiene una oferta estimada de 7 bicicletas, y una demanda estimada de 5 para el día lunes a las 08 : 30 : 00 hrs, la diferencia de estas dos canti-

dades nos indicará la cantidad de bicicletas que esta cicloestación puede aportar para el rebalanceo, o la cantidad de bicicletas que requieren ser enviadas a esta cicloestación, en este caso en particular, la cicloestación 5 puede aportar dos bicicletas. El número promedio de bicicletas que pueden ser utilizadas para el rebalanceo es reportado en la columna OD (oferta ofrecida) de las tablas 4.17 y 4.18, mientras que el número total de bicicletas requeridas es reportado en la columna DR (demanda requerida). Por ejemplo, para el día lunes a las 08:30 hrs en promedio hay 224 bicicletas que pueden ser reubicadas, mientras que en promedio, 151 bicicletas son requeridas entre las distintas cicloestaciones.

Por último, en las tablas 4.17 y 4.18, también se muestra el total de la demanda promedio atendida (DA), esta demanda resulta del número de bicicletas que es posible tomar de las cicloestaciones oferentes (OD) y llevarlas a las cicloestaciones que las requieran (DR). Por ejemplo, para el día lunes a las 08:30 hrs, en promedio, es posible reasignar 136 bicicletas de un total de 224 bicicletas, para atender una demanda total de 151 bicicletas, es decir, se atendió el 90 % de la demanda requerida (%DA). Es importante mencionar, que el porcentaje de demanda atendida depende tanto del número de camiones disponibles, así como las restricciones temporales o de distancia. En las tablas 4.19 y 4.20 se muestra tanto el tiempo promedio como los kilómetros promedio que se requieren para atender la demanda reportada en las tablas 4.17 y 4.18, respectivamente. Por ejemplo, para el día lunes a las 08:30 cada vehículo requiere recorrer 9.2 KM en promedio, lo que tal vez se pueda realizar en 18.3 minutos, si suponemos que la velocidad a la que un coche puede circular en la ciudad de México a esa hora es de 20 km/h.

4.4. Posición de este trabajo en la literatura

Los sistemas de préstamos de bicicletas continúan en crecimiento gracias a las ventajas que estos ofrecen. Dichos sistemas traen consigo patrones de movilidad que pueden ser estudiados con el fin de mejorarlos. En el caso de México, de acuerdo a las fuentes revisadas para la realización de este trabajo, son escasos los trabajos donde el punto central sea la estimación de la oferta y demanda del sistema, lo cual es

sumamente importante y necesario para los sistemas de préstamos de bicicletas, ya que el conocimiento de estas cantidades son la base para otro tipo de análisis, como la proyección de la demanda para determinar niveles de servicio, o para tomar decisiones acerca de cómo expandir el sistema, o acerca de qué áreas de una ciudad presentan mayor concurrencia, por ejemplo.

Dicho lo anterior, este trabajo contribuye en ser uno de los pocos trabajos que se centran en estimar la oferta y la demanda del sistema de bicicletas compartidas de la ciudad de México, a través de un análisis estadístico de la información. También, en este trabajo se identifica un problema de oferta y demanda desbalanceada gracias al índice de saturación construido, ya que se identifica que en la información reportada por el sistema, existen cicloestaciones y momentos del día donde se registran más bicicletas disponibles que las que pueden estar ancladas, lo cual no es posible, y fue tomado en cuenta para obtener estimaciones más precisas. También, en este trabajo se realiza un ejercicio de rebalanceo, tomando como caso base el problema de transporte, es decir, este escenario nos da el mínimo número de camiones así como el mínimo número de kilómetros que se requieren para realizar un rebalanceo perfecto, lo cual es ineficiente en términos de costos operativos. Posteriormente, mediante el problema del agente viajero múltiple es posible obtener rutas factibles sujetas a restricciones de distancia, las cuales pueden ser traducidas en restricciones temporales asumiendo una velocidad máxima a la que puede desplazarse un vehículo rebalanceador.

4.5. Deducciones

De este trabajo podemos deducir que el sistema de bicicletas compartidas presenta un área de oportunidad en su operación, esta área de oportunidad es originada por el problema de oferta o demanda de bicicletas, la cual puede ser estudiada a partir de la información histórica que es capturada por el sistema. También, de acuerdo a los patrones estudiados, podemos deducir que solamente un conjunto de cicloestaciones presentan este problema, ya que aproximadamente doscientas cicloestaciones parecen estar balanceadas a lo largo del día, por lo que podrían

no requerir intervención alguna o abrir nuevas oportunidades, en el sentido de estudiar si estas cicloestaciones están siendo subutilizadas, y tal vez puedan aprovecharse de mejor manera si son reposicionadas.

Capítulo 5

Conclusiones

5.1. Implicaciones del trabajo

En este trabajo se realizaron estimaciones de oferta y demanda con la información histórica disponible en el sitio web del sistema de bicicletas compartidas Ecobici, esta metodología de solución es resultado de la complejidad inherente a la información disponible, ya que la información es reportada al final de cada mes, esto pone una restricción al tipo de modelo a utilizar, por ejemplo, sería complicado plantear modelos de series de tiempo como en [5]-[8], ya que no es posible acceder a la información en tiempo real para dar pronósticos futuros. Dicho lo anterior, para propósitos de estimar la oferta y demanda, mientras que las características acerca de cómo es reportada la información del sistema no cambien, se sugiere emplear un modelo frecuentista.

Otro punto a considerar es que se reporta el número de servicios de cada bicicleta, más no el número de bicicletas disponibles en cada cicloestación, por lo que con el fin de conocer tanto la oferta, el número de bicicletas en cada cicloestación debe estimarse. Una posibilidad es mediante el estudio del flujo de bicicletas, es decir, mediante el estudio de los arribos menos los retiros de cada cicloestación, el punto en contra de este enfoque es que debe suponerse que cada cicloestación comienza operaciones con una cantidad fija de bicicletas, lo cual tal vez sea posible si durante el tiempo en que el sistema no está en servicio se

hace un rebalanceo, para que al momento de que el sistema comience operaciones, se inicialice con estas cantidades. Otro problema detectado con este enfoque es que la información disponible muestra un problema de demanda desbalanceada, es decir, es posible obtener un flujo de bicicletas donde el número de retiros o arribos sobrepase el número de bicicletas disponibles en una cicloestación, lo cuál no es posible. Para solucionar estos problemas, en este trabajo se estudia el flujo de bicicletas de manera parcial, es decir, en un intervalo de tiempo fijo de treinta minutos, de esta forma, si el número de arribos sobrepasa a los retiros en x unidades, podemos deducir que el número de bicicletas disponibles en este momento es de por lo menos la x bicicletas. Dicho lo anterior, para el estudio del número de bicicletas en cada cicloestación se sugiere estudiar el flujo en determinados subintervalos de tiempo.

Con respecto al rebalanceo, pueden existir muchas estrategias y factores a considerar, desde el número de vehículos, la capacidad máxima de carga que estos pueden tener, así como el tiempo máximo disponible para realizar el balanceo, lo cual podría ser complicado en una ciudad tan concurrida como la ciudad de México. En este trabajo se logró encontrar rutas factibles en meos de treinta minutos asumiendo una velocidad máxima de veinte kilómetros por hora, para ello se utilizaron cinco vehículos en la mañana y nueve en la tarde. Aunque el tránsito de la ciudad de México es realmente un problema para realizar un rebalanceo eficiente, tal vez existan medidas que puedan considerarse para eficientizar el movimiento de los camiones rebalanceadores, por ejemplo, estudiar aquellas vialidades que tengan un servicio exclusivo de metrobus o trolebus, para coordinar la operación de los camiones rebalanceadores con estos servicios, sin comprometer su operación. De esta manera, los camiones rebalanceadores podrían tener una movilidad más eficiente.

5.2. Importancia

Actualmente existe muy poca literatura acerca del estudio del sistema de préstamos de bicicletas de la ciudad de México, algunas fuentes son [13], [12] y [10], por ejemplo. De esta literatura, no hay un trabajo

relacionado a la estimación de la oferta y demanda de cada cicloestación de una manera frecuentista, es decir, que permita obtener estimaciones mediante la identificación de patrones relevantes en los datos del sistema. De lo revisado, tampoco existe un ejercicio de rebalanceo que permita jugar tanto con restricciones de distancia así como con el número de camiones a utilizar.

Por otro lado, este trabajo permite estudiar un sistema de préstamo de bicicletas a nivel servicio, es decir, dónde únicamente se cuenta con la información del viaje que ha realizado una bicicleta, más no con la información del número de bicicletas en cada cicloestación.

Por último, algo que también es importante considerar es que durante este trabajo se implementó y justificó una metodología simple, pero poderosa para identificar patrones relevantes, la cual no recae únicamente en un ejercicio de simulación como algunos modelos más complejos que se han propuesto, como por ejemplo el sistema de incentivos simulados propuesto en [15].

5.3. Lecciones aprendidas

Los sistemas de prestamos de bicicletas son sistemas sumamente complejos que involucran el análisis de varios componentes estocásticos que podrían influir en su comportamiento, como factores geográficos, meteorológicos, comportamiento de los usuarios, por ejemplo. En general es complicado tratar con toda esta aleatoriedad debido a la poca información que es posible estudiar de estos sistemas, por lo que depende de la creatividad del investigador implementar alguna metodología útil para poder reconocer y explotar ciertos patrones.

Con respecto al sistema de préstamos de bicicletas de la ciudad de México, dada la información con la que se cuenta, no es posible construir modelos de predicción en tiempo real, es decir, modelos que se alimenten de la información recién generada y puedan dar una predicción futura. Esto tal vez pueda cambiar con el avance de la tecnología, para ello se necesita que el sistema reporte y brinde la información del estado del

sistema con un tiempo considerablemente mínimo de retraso.

Otro punto a considerar, es que al igual que otros sistemas del mundo, tanto los horarios de mayor operación así como las zonas más concurridas, parecen concordar tanto con los horarios laborales de la ciudad de México, así como con las zonas donde se concentran las oficinas, como se menciona en [10].

Por último, con respecto al rebalanceo, tal vez un algoritmo no sea suficiente para obtener resultados óptimos, ya que también se debe considerar factores como el tránsito, el clima, posible cierre de vialidades, rutas que usan otros servicios de transporte como el metrobus o trolebus, que tal vez puedan emplearse. Con este último punto, me parece que sería útil y complementario realizar un análisis acerca del tiempo promedio de desplazamiento que es posible obtener al utilizar estas rutas.

5.4. Replicabilidad de los resultados

Para poder replicar o implementar tanto el análisis estadístico como los códigos desarrollados a lo largo de este trabajo en un ambiente productivo se requerirá tener en consideración los siguientes puntos:

1. La obtención de las predicciones de oferta y demanda fueron realizadas mediante el análisis de la información histórica de todo el año 2019. Este análisis se realizó con ayuda del software estadístico R [17]. Para la replicabilidad de este análisis deben ejecutarse los siguientes códigos:
 - a) `meses_acumulados.r`: este código tiene como entrada la información disponible en el sitio web ecobici, en particular, se toman las bases de datos 2019-01.csv a 2019-12.csv, las cuales contienen la información de todos los arribos y retiros de bicicletas de todas las cicloestaciones para cada día de cada mes; con la información anterior, se calcula el índice de saturación reportado en la sección 3.2.1 para cada mes y

para cada cicloestación. La información del índice se guarda en los archivos `AcumProm_30mins2019-01.csv` a `AcumProm_30mins2019-12.csv`.

- b)* `Estaciones_sat.r`: este código toma como entrada las bases generadas en el paso anterior. En este código se generan las gráficas del índice de acumulación promedio para todas las cicloestaciones (algunos ejemplos de estas gráficas son las gráficas 3.2, 3.3, 3.4 y 3.5).
 - c)* `Bases_desagregadas.r`: en este código se realiza la lectura y limpieza de la información proporcionada por el sitio web de ecobici. Adicionalmente, en este código se generan las bases `PorDia_30min2019-1.csv`:`PorDia_30min2019-12.csv`, las cuales contienen el número de arribos y retiros de cada cicloestación agrupado en intervalos de treinta minutos para todas las cicloestaciones.
 - d)* `Bases_día.r`: este código toma como entrada las bases generadas en el código previo, posteriormente se calculan las funciones de probabilidad estimadas de los arribos y retiros de cada cicloestación para algún intervalo de tiempo de treinta minutos dado y los histogramas de estas (como ejemplo pueden verse las gráficas 3.6 y 3.7). También, aquí se obtienen las predicciones de oferta y demanda para un conjunto de cicloestaciones dado para el intervalo de tiempo especificado.
2. Una vez identificadas las cicloestaciones que requieren un rebalanceo de acuerdo a la metodología explicada en este trabajo, y calculadas sus respectivas ofertas y demandas esperadas, se procede a generar el ejercicio de rebalanceo, para el cual se utilizó el lenguaje de programación Phyton [20]. Para la replicabilidad de este análisis deben ejecutarse los siguientes códigos:
- a)* `main.py`: este código toma como entrada el catálogo de cicloestaciones así como la latitud y longitud de cada cicloestación para calcular la distancia que hay entre cada par de cicloestaciones. La distancia de cada par de cicloestaciones es guardada en una matriz de distancias. También, en este

código se ejecuta el algoritmo del plan de transporte descrito en la sección 3.3.1.

- b) grafos.py: con la salida del plan de transporte del código anterior, en este código se generan los gráficos y las rutas descritas en la sección 3.4.2.
- c) greedy2.py: este código toma como entrada los archivos de predicción generados por el código Bases_día.r, para posteriormente instanciar un número de vehículos especificados por el usuario, para posteriormente ejecutar el procedimiento descrito en la sección 3.4.1.

Los códigos anteriores, así como los archivos utilizados en este trabajo, y las salidas obtenidas por los distintos códigos pueden consultarse en <https://github.com/David-Castillo-R/Demand-forecasting-and-iterative-balancing-of-a-bicycle-loan-system-with-xed-stations>.

En la siguiente sección se habla brevemente acerca de qué consideraciones y pasos futuros se pueden realizar para implementar la metodología descrita en este trabajo en un ambiente productivo.

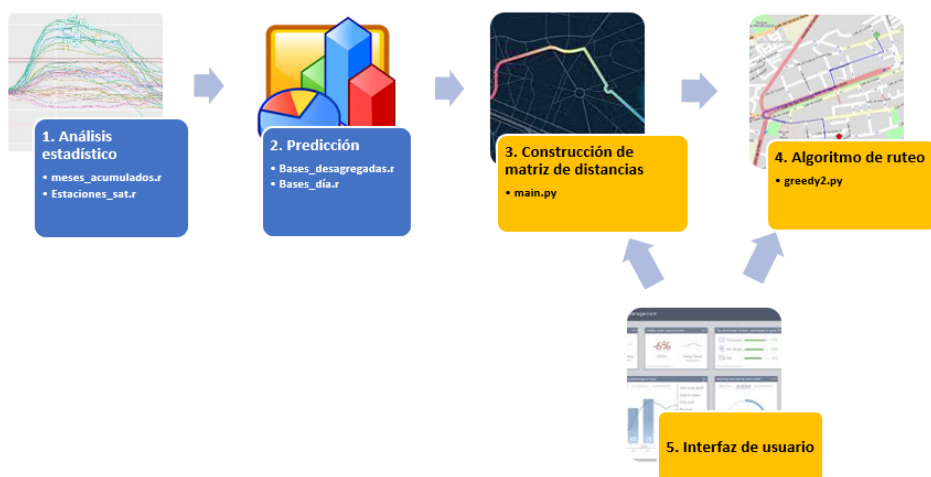
5.4.1. Implementación en un ambiente productivo

Para implementar la metodología anterior en un ambiente productivo, puede recurrirse a los servicios en la nube como Amazon Web Services o Microsoft Azure, por ejemplo. Para ello, puede considerarse los pasos de la arquitectura que se ilustran en la figura 5.1, dichos pasos se explican a continuación:

1. Como primer paso debe realizarse la lectura de la información, al igual que el cálculo del índice de saturación, lo cual es posible obtener con los códigos meses_acumulados.r y Estaciones_sat.r.
2. Posteriormente, y una vez ejecutado el paso anterior se procede a obtener la estimación tanto de la oferta como de la demanda de las cicloestaciones, lo cual es posible obtener mediante los códigos Bases_desagregadas.r y Bases_día.r.

3. El tercer paso se puede ejecutar al concluir los pasos anteriores, o bien, se puede ejecutar de manera independiente, ya que en este paso se calcula la matriz de distancias entre cualquier par de cicloestaciones.
4. Para la ejecución de este paso, es indispensable haber ejecutado primero tanto el paso número dos así como el paso anterior, en este paso se calculan las rutas que deben seguirse durante el balanceo, lo cual es posible mediante el código `greedy2.py`.
5. Este paso consiste en la elaboración de una aplicación de usuario, la cual únicamente requeriría ejecutar el paso número tres y número cuatro, ya que al usuario únicamente le podría interesar conocer el balanceo de un subconjunto de bicicletas dado, por lo que únicamente el usuario podría dar las coordenadas de las cicloestaciones que le interesa, o inclusive, dar un conjunto de coordenadas arbitrario para atender otro tipo de problemas de ruteo. Una vez escogido el subconjunto de cicloestaciones o dado un conjunto de coordenadas arbitrario (y su respectiva demanda y oferta), el usuario procedería a especificar el número de vehículos para realizar el balanceo, así como las restricciones de capacidad y distancia requeridas.

Adicionalmente a los pasos explicados anteriormente, se requiere contar con un orquestador que ejecute los pasos anteriores en el orden dicho, como orquestador se podría utilizar la paquetería Luigi de python, pero para ello será necesario que los códigos del paso uno y dos sean recodificados en python. Otro punto a considerar es que los pasos uno y dos, solo requieren ser ejecutados una única vez, ya que estos códigos consideran el análisis de todo una año de información, por lo que este análisis puede ser recalibrado cada año para el caso de ecobici.



Gráfica 5.1: Arquitectura a implementar en un ambiente productivo.

Bibliografía

- [1] Review by: Ronald V. Book. Reducibility among combinatorial problems. by richard m. karp; raymond e. miller; james w. thatcher. *Journal of Symbolic Logic*, 40:618–619, 1975.
- [2] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms, Third Edition*. The MIT Press, 3rd edition, 2009.
- [3] J. H. Dantzig, G. B.; Ramser. The truck dispatching problem. *Management Science*, 6:80–91, 1959.
- [4] Gobierno de la ciudad de México. Ecobici, 2010.
- [5] Ahmadreza Faghieh-Imani, Robert Hampshire, Lavanya Marla, and Naveen Eluru. An empirical analysis of bike sharing usage and rebalancing: Evidence from barcelona and seville. *Transportation Research Part A: Policy and Practice*, 97:177–191, 2017.
- [6] Elliot Fishman. Bikeshare: A review of recent literature. *Transport Reviews*, 36(1):92–113, 2016.
- [7] Jon Froehlich, Joachim Neumann, and Nuria Oliver. Sensing and predicting the pulse of the city through shared bicycling. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence, IJCAI’09*, page 1420–1426, San Francisco, CA, USA, 2009. Morgan Kaufmann Publishers Inc.
- [8] Andreas Kaltenbrunner, Rodrigo Meza, Jens Grivolla, Joan Codina, and Rafael Banchs. Urban cycles and mobility patterns: Ex-

- ploring and predicting trends in a bicycle-based public transport system. *Pervasive and Mobile Computing*, 6(4):455–466, 2010. Human Behavior in Ubiquitous Environments: Modeling of Human Mobility Patterns.
- [9] J. K. Lenstra and A. H. G. Rinnooy Kan. Complexity of vehicle routing and scheduling problems. *Networks*, 11(2):221–227, 1981.
- [10] David C. Luis M. Data visualization to forecast withdrawals in a bikesharing system: Mexico city case. *Challenge and Innovative Solutions for Multimodalitu in Global Transport Networks*, pages 65–73, 2019.
- [11] LLC MetroBike. The bike-sharing blog, 2004.
- [12] Luis A. Moncayo-Martínez. Analysing data set of the bike-sharing system demand with r scripts: Mexico city case. In Yaxin Bi, Rahul Bhatia, and Supriya Kapoor, editors, *Intelligent Systems and Applications*, pages 90–105, Cham, 2020. Springer International Publishing.
- [13] L. A. Moncayo-Martínez and A. Ramirez-Nafarrate. Visualization of the mobility patterns in the bike-sharing transport systems in mexico city. In *2016 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, pages 1851–1855, 2016.
- [14] Luis A Moncayo-Martínez. *Algoritmos especiales: el problema de transporte*, pages 149–188. 01 2014.
- [15] J. Pfrommer, J. Warrington, G. Schildbach, and M. Morari. Dynamic vehicle redistribution and online price incentives in shared mobility systems. *IEEE Transactions on Intelligent Transportation Systems*, 15(4):1567–1578, 2014.
- [16] J. Pfrommer, J. Warrington, G. Schildbach, and M. Morari. Dynamic vehicle redistribution and online price incentives in shared mobility systems. *IEEE Transactions on Intelligent Transportation Systems*, 15(4):1567–1578, 2014.

- [17] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020.
- [18] J. Schuijbroek, R.C. Hampshire, and W.-J. van Hoes. Inventory rebalancing and vehicle routing in bike sharing systems. *European Journal of Operational Research*, 257(3):992–1004, 2017.
- [19] Paolo Toth and Daniele Vigo, editors. *The Vehicle Routing Problem*. Society for Industrial and Applied Mathematics, USA, 2001.
- [20] Guido Van Rossum and Fred L. Drake. *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA, 2009.