

Representación cartográfica interactiva, 3D, dinámica y en línea, de patrones espaciotemporales de movilidad. Caso de estudio: el sistema de bicicleta compartida “Ecobici” de la Ciudad de México[☆]

3D online, dynamic, interactive cartography of spatiotemporal mobility patterns. Case study: the “Ecobici” bike sharing system of Mexico City

Francisco Javier Osorno-Covarrubias^{*a}, Manuel Suarez-Lastra^b, Luis Felipe Ruiz-Castañeda^c

^aLaboratorio de Análisis Geoespacial, Instituto de Geografía, UNAM

^bDepartamento de Geografía Económica, Instituto de Geografía, UNAM

^cLicenciatura en Ciencias de la Computación. Universidad Nacional Autónoma de México, Circuito Exterior, Ciudad Universitaria, Coyoacán, 04510, México.

Resumen

Dada su complejidad, la representación cartográfica de patrones espaciotemporales de movilidad presenta retos difíciles de resolver y comunicar a través de los métodos tradicionales, limitados por las características del sustrato en papel, inherentemente plano, estático y desconectado de los datos subyacentes. Si bien existe software comercial que ofrece herramientas para elaborar cartografía dinámica en 3 dimensiones, la difusión en línea de los productos generados permanece atada a costosas licencias en el servidor, y por ende están sujetas a una difusión limitada. En este artículo presentamos una metodología innovadora, basada en software libre y de código abierto, orientada a la generación de productos cartográficos que habilitan la exploración tridimensional, dinámica, interactiva y en línea de dichos patrones. Nuestro caso de estudio usa el acervo masivo de datos abiertos del sistema de bicicleta compartida ECOBICI, de la ciudad de México. Como prueba de concepto, se presentan cuatro escenas interactivas. Finalmente se discuten las ventajas y limitaciones de la metodología utilizada.

Palabras clave: mapas interactivos; cartografía dinámica en línea, escenas 3D, sistemas de bici compartida; movilidad ciclista; software libre, datos abiertos

Abstract

Given its complexity, the cartographic representation of spatio-temporal mobility patterns presents challenges which are difficult to solve and communicate using traditional methods that are also limited by its paper substrate, inherently flat, static and disconnected from the underlying data. Even though there is proprietary software that offers tools to create dynamic cartography in 3 dimensions, the generated products remain tied to the desktop GIS, disconnected from the World Wide Web and therefore, with limited reach. In this article we present an innovative methodology, based on free and open source software, that enables the online, three-dimensional, dynamic, interactive exploration of these patterns. Our study case uses the massive open data of the shared bike system Ecobici, from Mexico City. As proof of concept, four maps that allow exploring the dynamic behavior of the system, are presented. Finally, we argue about the possibilities and limitations of our methodology.

Keywords: interactive maps; dynamic maps; 3D maps; online cartography; mobility patterns; bicycle; bike sharing systems; Mexico City; Mexico

[☆] © F. J. Osorno-Covarrubias, M. Suarez-Lastra, L. F. Ruiz-Castañeda

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by-nc-sa/4.0/>), which permits non-commercial sharing of the work and adaptations, provided the original work is properly cited and the new creations are licensed under identical terms.

^{*}E-mail address: josorno@igg.unam.mx

1. Introducción

Existe un volumen creciente de registros de movilidad que provienen de una variedad de fuentes tales como los generados por sensores GPS en teléfonos móviles y patinetas eléctricas, los de arribo y retiro en puntos de anclaje de sistemas de bicicleta compartida o los generados por tarjetas inteligentes en

los torniquetes de estaciones de transporte público. Si bien estos flujos de datos están diseñados para satisfacer necesidades operativas específicas de los sistemas que los producen, al ser ofrecidos de manera pública y abierta constituyen una valiosa mina de indicios que pueden ser usados para mejorar la calidad de vida de los ciudadanos a través de decisiones de gestión urbana y políticas basadas en evidencia (Batty, 2013; Ferreira et al., 2013).

Se puede argumentar que el análisis visual de estos flujos masivos de datos espaciotemporales es pertinente en dos sentidos; en primer lugar, porque la representación visual de datos facilita la comunicación científica y la formulación de preguntas científicas pertinentes como un paso previo para escoger métodos computacionales apropiados y dirigir su trabajo (Ren et al., 2014). En segundo lugar, porque las técnicas de visualización científica son irremplazables –por lo menos a la fecha, en el estudio de procesos que se desarrollan en el espacio geográfico. La heterogeneidad de propiedades y relaciones que ocurren en dicho espacio, no pueden ser adecuadamente representadas para su uso en procesos enteramente automatizados dado que la extracción de evidencias depende aún de las capacidades de inferencia humana y del conocimiento tácito de las propiedades, relaciones y experiencias asociadas a espacios y lugares concretos (Andrienko & Andrienko, 2013; Keim et al., 2008).

El mapa —análogo o digital—, sigue siendo la herramienta de elección para la representación y comunicación de fenómenos espaciales. El rol del mapa se expandió con la tecnología SIG habilitando una integración estrecha de los procesos, que van desde la adquisición de los datos, su depuración, análisis, edición en productos cartográficos, hasta su comunicación a través de mapas impresos, o digitales y en línea.

En el caso específico de los fenómenos de movilidad, el desarrollo de la cartografía se ha mantenido al paso de la evolución tecnológica, de manera que es común encontrarla en aplicaciones de uso masivo, conectadas en tiempo real a bases de datos gigantescas. Tal es el caso, por ejemplo, de las aplicaciones de navegación vehicular y de transporte como servicio. Estos avances invitan a repensar la comunicación científica de fenómenos de movilidad, trascendiendo del paradigma del mapa bidimensional, estático, al de mapas interactivos, tridimensionales, dinámicos y conectados a fuentes de datos masivas. Existen desarrollos recientes en el ámbito comercial que habilitan este tipo de productos cartográficos (ArcGIS Desktop + ArcScene + ArcGIS Server). Sin embargo, estos productos son de código cerrado y la difusión permanece atada a costosas licencias.

Dado el hecho de que los fenómenos de movilidad nos atañen a todos por su impacto en la calidad de vida (Steg & Gifford, 2005), incluyendo los efectos sobre la salud y sus consecuencias económicas, conviene que estos estudios se puedan difundir en línea para ser consultados y discutidos con amplitud, idealmente por todos los actores involucrados: especialistas, tomadores de decisiones y público en general.

De las consideraciones anteriores, con el objeto de guiar nuestro trabajo, derivamos un conjunto de criterios metodoló-

gicos y de selección de herramientas a usar y/o desarrollar para la comunicación científica de patrones de movilidad. Así, nos propusimos usar y/o desarrollar métodos y herramientas que:

- sean adecuadas para el procesamiento de flujos masivos de datos espaciotemporales,
- habiliten la exploración de la dinámica espaciotemporal de redes de transporte, en sus nodos –por ejemplo, cicloestaciones, estaciones de metro, de Metrobús etc.; y arcos –por ejemplo, líneas de metro, carriles confinados de Metrobús, etcétera,
- faciliten la visualización/comunicación/comprensión, y permitan la exploración dinámica, interactiva, en escenas tridimensionales, de los patrones asociados al transporte,
- para fines de edición cartográfica, puedan integrarse de manera sencilla con un SIG de escritorio,
- por razones de costo, escalabilidad, e interés científico, estén basadas en software libre y de código abierto,
- por razones de economía del esfuerzo, partan de componentes existentes,
- maximicen la capacidad de difusión mediante su operación en línea.

En este estudio de carácter exploratorio ofrecemos una solución inédita, hasta donde alcanza nuestro conocimiento, que satisface las especificaciones referidas arriba. Esta solución formula un enfoque metodológico que integra los diferentes pasos desde la adquisición de datos de movilidad a partir de “streams” públicos masivos, hasta la publicación de escenas interactivas que permiten explorar en línea la dinámica de la red, utilizando herramientas de software libre.

El artículo está organizado en cuatro secciones, en la primera (sección 2) exploramos una selección de trabajos relacionados sobre temas de visualización de patrones de datos espaciotemporales y estudios de movilidad ciclista. En la sección 3 presentamos el caso de estudio y la metodología. En la sección 4 presentamos, a modo de prueba de concepto, tres escenas interactivas. En la sección 5 se discuten los resultados, se ofrecen algunas conclusiones y se delinea el trabajo a futuro.

2. Trabajos relacionados:

2.1. Visualización

El tema de la visualización de datos espacio temporales es un tópico de investigación activo y con amplia tradición. Una muestra de ello es el amplio estudio de Aigner et al. (2011), quienes realizan una revisión sistemática de 113 técnicas de visualización espacio temporal, categorizándolas de acuerdo con seis criterios referentes al carácter de los datos –marco de referencia, número de variables, el manejo del tiempo –disposición y primitivas temporales, y la naturaleza de la visualización

–dinamicidad y dimensionalidad. De este universo de técnicas¹ solo 15 incluyen las características de tener un marco de referencia espacial y tridimensional y solo una, descrita por Kapler & Wright (2005) es dinámica. Esta técnica que se implementa en el sistema Geotime, permite mostrar ítems de datos (por ejemplo, objetos, eventos, transacciones, flujos) en su contexto espacial y temporal. El sistema provee una versión dinámica e interactiva del concepto de cubo-espacio tiempo, en el que un mapa plano ilustra el contexto espacial y el tiempo se mapea en la dirección vertical. Otras variantes, adaptadas a la representación de series de tiempo involucran el uso de tornillos, espirales y poliedros, bajo el concepto de “helix icons” (Tominski et al., 2005) y el uso de “data vases” para la representación de series múltiples (Thakur & Hanson, 2010).

El cambio de enfoque del paradigma estrecho de la cartografía como el arte y técnica de hacer mapas en papel al de una cartografía digital y en línea conectada a infraestructuras de datos, lo aborda Basaraner (2016), quien revisa el rol y alcance de la cartografía a la luz de desarrollos recientes en los procesos de creación de información espacial, integración de bases de datos y la generación de información geográfica y mapas digitales en línea. Este enfoque resulta particularmente pertinente en nuestro caso de estudio, ya que se centra en la representación de datos masivos de movilidad.

Otro aspecto central de nuestro estudio es la visualización de flujos origen-destino en condiciones en las que la trayectoria real seguida en cada instancia se desconoce o es irrelevante. Su visualización resulta problemática dado el apiñamiento visual que crece con el cuadrado del número de puntos de conexión. Graser et al. (2019) proponen una técnica para la reducción del apiñamiento visual (“visual clutter”) en representaciones cartográficas bidimensionales, orientada a operar en software de escritorio. Nuestro enfoque aborda la misma problemática, pero desde la perspectiva de escenas tridimensionales interactivas.

Sandvick (2008) y más recientemente Zhu et al. (2018) abordan el problema de la representación dinámica de procesos espaciotemporales utilizando lenguajes de marcado (KML y CZML) para la visualización en línea a través de globos virtuales (Google Earth y Cesium, respectivamente). Este enfoque de visualización a través de lenguajes de marcado se ha visto frenado a raíz de que Google Earth descontinuó su plugin para los visualizadores web. Por su parte, la visualización científica bajo el paradigma de globo virtual basada en Cesium, también ha perdido ímpetu por las mismas razones. El paquete plotKML (Hengl et al., 2015) que depende de Google Earth, ofrece una alternativa interesante para generar archivos KML a partir de estructuras de datos disponibles en R.

Harder y Brown (2017) exploran el estado del arte en materia de visualización cartográfica interactiva en línea usando herramientas de software propietario para la edición de escenas tridimensionales interactivas (ArcGIS Desktop y ArcScene) y

difusión en línea (ArcGIS Server), ofreciendo ejemplos de aplicaciones en una variedad de campos.²

En el mapeo de resultados políticos es común usar cartogramas que distorsionan el tamaño de las circunscripciones para igualar la densidad y reflejar de manera más correcta la naturaleza de la distribución espacial de los votantes. Sin embargo, esta práctica dificulta la lectura e interpretación del mapa para el lector en general. Oruednik (2017) destaca la utilidad del mapeo interactivo en 3 dimensiones como un recurso que permite mantener el mapa topográfico base intacto, mientras se usa la tercera dimensión para expresar resultados electorales superpuestos sobre una superficie que refleja la densidad demográfica. Aunque el tema es muy diferente, este estudio tiene similitud con el nuestro por las herramientas de software que utiliza para la visualización interactiva 3D. También es similar en que usa la tercera dimensión para mapear atributos: en su caso la densidad es mapeada a una superficie ráster de elevación y los resultados electorales se superponen en ella. En nuestro caso usamos la altura, color, forma de rasgos tridimensionales para representar la intensidad de los atributos cartografiados. Otra diferencia fundamental es que nosotros añadimos mecanismos para la visualización dinámica de series de tiempo.

2.2. Big data y movilidad ciclista

El trabajo de O’Brien et al. (2014) es el primero en estudiar el fenómeno de los sistemas de bicicleta compartida (SBC) a nivel global. Para ello construyen un base de datos que abarca 38 sistemas de bicicleta compartida (SBC) en países de Europa, Oriente Medio, Asia, Australasia y América y proponen una clasificación de los sistemas SBC basada en la huella geográfica y las variaciones espaciotemporales en la tasa de ocupación.

Se han indagado sistemáticamente los retos teóricos y las aplicaciones del Big Data para tareas de gestión urbana (Thakuriah et al., 2017; Thakuriah et al., 2017a, 2017b). De manera más específica Romanillos et al. (2016) exploran el campo de investigación emergente de Big Data, aplicado al estudio de la movilidad ciclista. Estos autores revisan las técnicas, objetivos y hallazgos, de una amplia muestra de estudios sobre el tema, clasificándolos en tres categorías según el origen de los datos fuente: 1) Estudios basados en datos GPS 2) en puntos colectados en tiempo real a través de diversos sensores (por ejemplo, torniquetes o puntos de anclaje en sistemas de bicicleta compartida) y 3) datos históricos de viajes origen-destino. Encuentran en estos estudios una transición hacia datos masivos generados a través de aplicaciones en teléfonos inteligentes. Destacan el potencial de los acervos históricos de viajes para el análisis de tendencias y el reconocimiento de patrones, que generan información útil sobre los sistemas de movilidad mismos y de manera indirecta, para otros temas de salud, sociodemográficos y de transporte.

El tema de la visualización de flujos de movilidad ciclista en SBC lo abordan Romanillos et al. (2018) señalando la importancia de rastrear los patrones de flujo a un nivel más detallado

¹La página web del libro de Aigner, tiene este explorador interactivo de las técnicas descritas

²El libro interactivo con vínculos a los mapas Web puede verse aquí

que el nivel de estación. Para ello analizan un gran número de rutas GPS proporcionadas por el SBC de Madrid y proveen un análisis de cómo se distribuye el tráfico en la red de calles urbana y los patrones de variación en el tiempo.

3. Caso de estudio y metodología

El uso de la bicicleta en Ciudad de México ha tenido un auge en la última década, pasando de 100 mil viajes en 2007 (EOD2007), a un número que pudiera oscilar entre 250 y 290 mil, según la fuente. Ello significa que prácticamente se han triplicado los viajes en un lapso de 10 años. A partir de la elaboración de la Estrategia de Infraestructura Ciclista en 2007, el gobierno de la ciudad ha puesto en marcha tres componentes del sistema: infraestructura lineal (ciclovías, ciclocarriles y carriles compartidos), biciestacionamientos (masivos, semimasivos y de corta estancia), así como el sistema de bicicletas públicas ECOBICI. La ciudad, en 2019, cuenta con más de 242 km de ciclovías, y carriles compartidos. Se cuenta también con 4 biciestacionamientos masivos y semimasivos ubicados en estaciones de Metro de gran afluencia que brindan en conjunto, poco menos de mil espacios de estacionamiento. Adicionalmente la ciudad cuenta con más de 2 mil biciestacionamientos de corta estancia. En el caso de Ecobici, el sistema registra cerca de 30 mil viajes diarios según estadísticas oficiales.

De acuerdo con los estudios que se han realizado sobre el ciclismo urbano en CDMX, los viajes en bicicleta son primordialmente hechos con el propósito de ir a trabajar, seguido del regreso a casa y el ir de compras. A diferencia de los viajes que se realizan en otros modos, los viajes a la escuela parecen estar subrepresentados.

La distribución horaria sigue muy de cerca la de los viajes en general, aunque se ha probado que los viajes en bicicleta suelen ser más cortos (30 min) que aquellos que se realizan en otros modos. Los estudios han probado que la bicicleta es el modo de transporte más rápido en la ciudad en hora pico. Mientras que la de un automóvil es de 7 km/h, similar a la del Metro cuando se consideran los tramos de caminata y espera de sus usuarios, la de la bicicleta asciende a 14 km/h. No es sorprendente, entonces, que de acuerdo a la Encuesta de Movilidad Ciclista (Gobierno de la Ciudad de México, 2014) las dos primeras razones por las cuales los usuarios deciden usar la bicicleta sean la rapidez y el costo.

El sistema de bicicletas compartidas ECOBICI, brinda una serie de ventajas a los ciclistas, sobre todo el relativo al resguardo diario de la bicicleta; su propiedad y mantenimiento, aunque es limitativo en términos de que requiere de estaciones específicas en donde recoger y dejar la bicicleta. Dicho sistema ha reemplazado un buen número de viajes que antes se realizaban caminando (24 %), en colectivo (26 %), metro (15 %), y un reducido número en automóvil (6 %).

Debido al volumen de viajes que se realizan en ECOBICI, es claro que se ha convertido en un componente importante en la movilidad de CDMX, por lo que la comprensión del sistema y la dinámica de sus flujos es esencial para su planeación

futura. Derivado de lo anterior, este artículo se enfoca en la exploración de herramientas visuales cartográficas que permitan analizar el funcionamiento del sistema de una manera dinámica, para coadyuvar al desarrollo del uso de la bicicleta en Ciudad de México.

3.1. El sistema de bicicleta compartida ECOBICI

El 16 de febrero de 2010 se puso en marcha el sistema de bicicletas públicas Ecobici que fue el primero automatizado en América Latina (Suárez-Lastra, 2018). El sistema inició con 1,114 bicicletas distribuidas en 90 cicloestaciones. A la fecha cuenta con 480 cicloestaciones distribuidas en un polígono que abarca 3800 ha. y cuenta con 6,100 bicicletas en operación. La distancia promedio en línea recta entre estaciones es de 150 m y las dos más alejadas se encuentran a 10.05 km una de otra.

De acuerdo con una compilación reciente de sistemas de bicicleta compartida en el mundo (Wikipedia, 2019), el sistema Ecobici de la ciudad de México ocupa el lugar 26 en términos del número de bicicletas. De estos, 19 están en ciudades chinas incluyendo los 11 primeros lugares.

En la figura 1 se muestra un mapa interactivo con la ubicación de las cicloestaciones, las fases de implementación de Ecobici, las líneas de metro y Metrobús. La herramienta de búsqueda permite localizar las estaciones por nombre o por número. La herramienta de medición permite medir áreas y distancias a lo largo de trayectos.

3.2. Datos

El sistema Ecobici ofrece públicamente tres elementos relevantes para conocer la estructura y estado de operación de la red (figura 2) (1) información estática sobre las cicloestaciones (tabla “cicloEstacion”) (2) información casi-en-tiempo-real sobre el estatus de operación de las cicloestaciones (tabla “statusCEstacion”) y (3) información anonimizada sobre cada uno de los viajes (tabla “viaje”), que se publica mensualmente.

El muestreo periódico del estatus de las cicloestaciones permite responder preguntas relevantes. Por un lado, para el usuario y por el otro lado para el operador del sistema. Desde la perspectiva del usuario interesa saber dónde está la estación en operación más cercana, si ésta tiene bicis disponibles y en su caso si tiene lugares de anclaje disponibles. Desde la perspectiva del operador del sistema interesa saber cuántas estaciones están activas y cuántas fuera de servicio, cuántas bicicletas y lugares de anclaje hay disponibles en cada una, cuáles están llenas y cuáles vacías.

La figura 2 presenta el modelo de datos para nuestro estudio. La entidad “viaje”, registra el momento y lugar de inicio y terminación de cada viaje. De aquí se obtiene la duración precisa. Se desconoce la trayectoria seguida y por ende la distancia y velocidad de cada viaje. Para fines comparativos y de animación se puede calcular la distancia asumiendo que el usuario sigue la trayectoria más corta. Para ello utilizamos los algoritmos de pgRouting, sobre la red de calles del OpenStreetMap, procesado para darle una topología rigurosa de red de arcos y nodos.

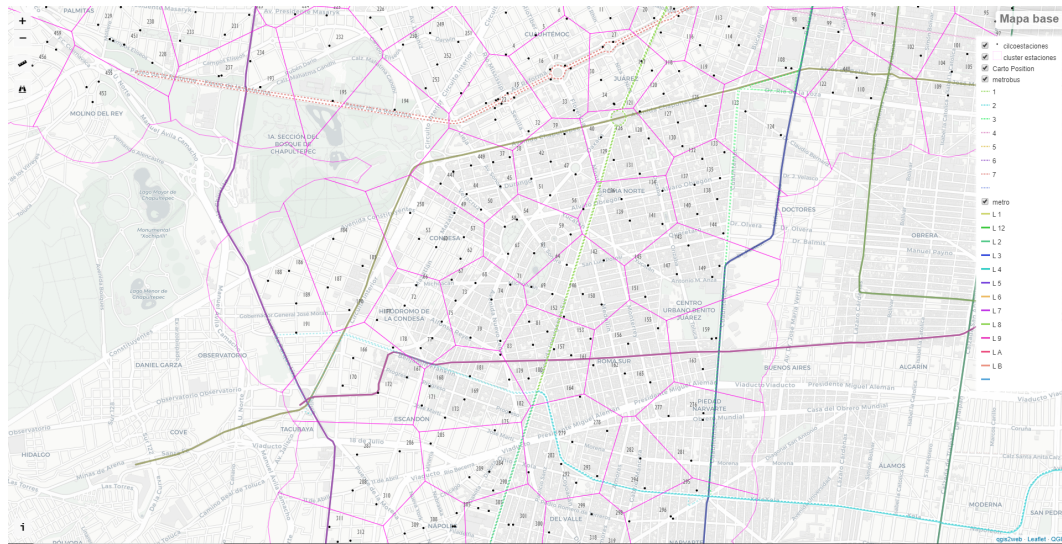


Figura 1: Mapa base: cicloestaciones de Ecobici, líneas de metro y metrobús. Fuente: elaboración propia a partir de datos abiertos de los tres sistemas. (Click en el hipervínculo para ver la versión interactiva del mapa) / Figure 1. Base map: Ecobici’s bike sharing stations, metro and metrobús lines. Source: prepared by the authors based on open data sources for the three systems. (Click to the link to see the interactive version of the map)

Nótese que la marca de tiempo en los datos fuente, permite agregarlos a diferentes niveles: por ejemplo, intervalos de 15 minutos, de una hora, un día, semana, mes, año, o todo el lapso de operación del sistema. Por otro lado, es posible georreferenciar la ubicación precisa de cada cicloestación y por ende los datos asociados a cada una de ellas, a nivel puntual o agregándolas de acuerdo con criterios de análisis espacial o criterios de inclusión dentro de particiones útiles, como las áreas geoestadísticas básicas y otras teselaciones. Los mapas que presentamos están basados en “vistas” SQL que combinan ambos criterios para conocer, por ejemplo, la dinámica horaria de las cicloestaciones a lo largo de un día, el balance mensual de arribos y retiros en las cicloestaciones, o el número de conexiones entre pares Origen-Destino, en un intervalo de tiempo determinado –hora pico matutino, hora pico vespertino etc.

3.3. Metodología de implementación

De acuerdo con Tominsky et al. (2012), los datos de trayectorias asociados a atributos pueden definirse formalmente de la siguiente manera: una trayectoria $d \in D$ es un conjunto ordenado de puntos de datos $d = (d_1, d_{l_d})$. Cada punto de datos $d_k : 1 < k < l_d$ tiene la forma $d_k \in (S^n \times T \times A_1, \dots, A_m)$, donde S^n define las coordenadas espaciales del punto, por ejemplo x , y para $n = 2$, o x , y , z para $n=3$, T define el tiempo y $A_i : 1 \leq i \leq m$ son los atributos cualitativos o cuantitativos del punto. Esta definición muestra la complejidad del problema: los datos incluyen componentes espaciales, temporales y también numéricos y/o categóricos. Si bien Tominsky et al. (2012) se centran en trayectorias en el espacio bidimensional y utilizan el tiempo como la tercera dimensión, nosotros generalizamos este esquema agregando también rasgos con geometría de punto y polígono y utilizando la tercera dimensión para representar los atributos de una o más variables y manejando la dimensión tem-

poral a través de animaciones interactivas. De esta manera, en comparación con la simbología de los rasgos en un mapa bidimensional, la representación de atributos asociados a los rasgos se enriquece: los puntos adquieren volumen, las líneas adquieren grosor y altura, los polígonos volumen. Las variaciones de atributos continuos pueden mapearse a variaciones simultáneas de color, anchura y altura. Las diferencias categóricas pueden mapearse a iconos volumétricos diferentes: esferas, conos, cilindros etc.

3.3.1. Arquitectura de la solución propuesta

En la figura 3 se muestran los elementos de una arquitectura que habilita la gestión de flujos de datos de movilidad, el análisis espacial, la edición de escenas 3D, su exportación a código web, y la modificación de éste para habilitar la visualización interactiva de procesos dinámicos.

Los elementos principales y sus interacciones son los siguientes:

- Stream de datos: Utilizando scripts de Python extraemos del API de Ecobici información sobre el estatus de las cicloestaciones. Esta operación se repite de manera periódica y durante lapsos de la duración que se requiera. Las secuencias de datos obtenidas permiten reproducir la dinámica del sistema a diferentes niveles de compactación. Estos datos en formato json se suben a la base de datos utilizando la librería de Python psycopg2, que implementa cursores de PostgreSQL. La información de los viajes del mes anterior se hace disponible al inicio de cada mes en archivos csv que se pueden importar directamente a la base de datos.
- Base de datos de PostgreSQL: La base de datos tiene las extensiones PostGIS, PL/PgSQL y PgRouting. A través

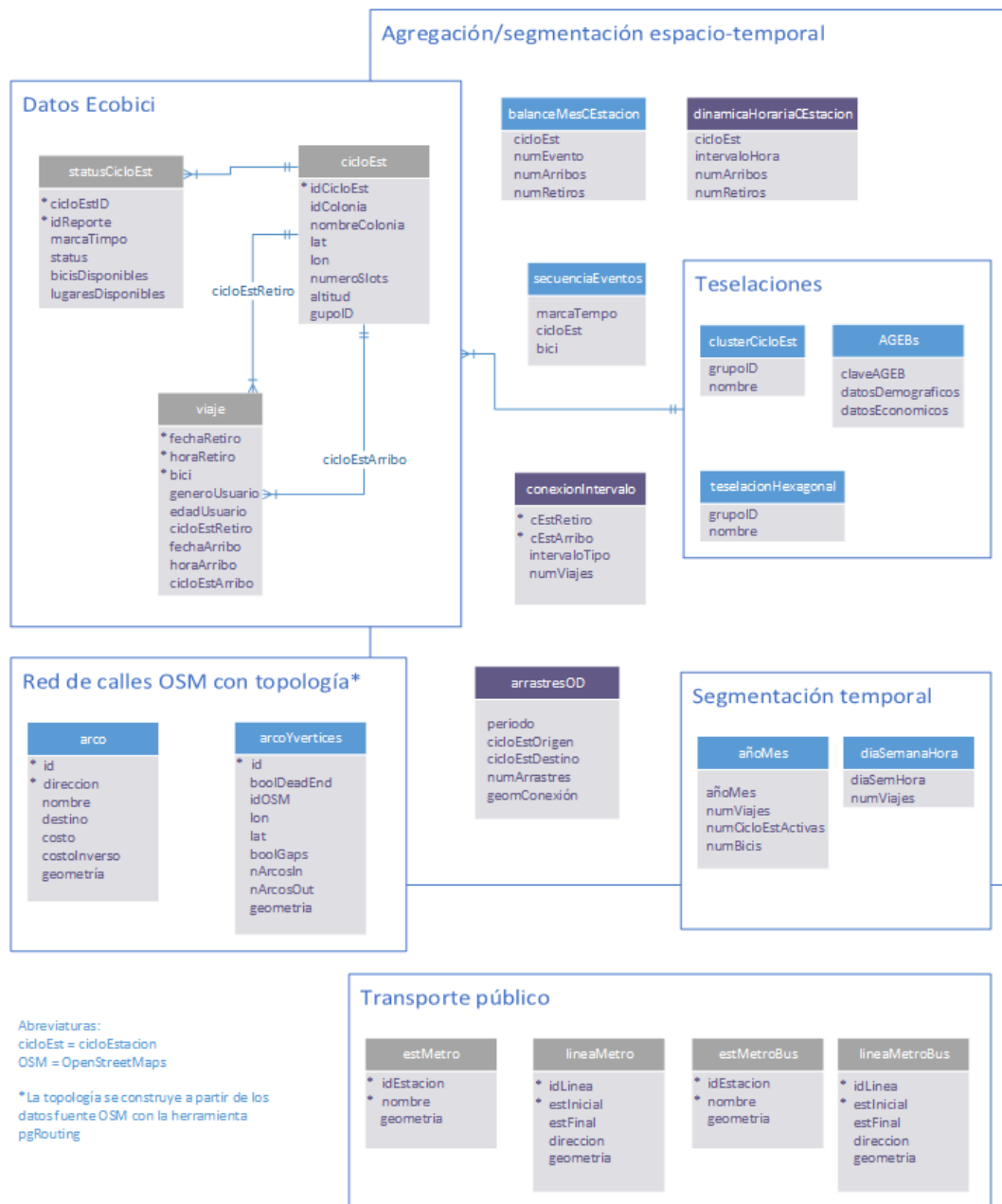


Figura 2: Modelo de datos para nuestro caso de estudio. Las entidades de color gris las obtenemos de fuentes de datos públicas y abiertas, las de color azul claro las calculamos a través de una variedad de procesos que incluyen el uso de consultas SQL y procedimientos almacenados (PGSQL). Por razones de claridad modificamos los nombres de las entidades y campos de manera que se auto expliquen. Las entidades de color púrpura participan directamente en algún mapa o gráfica. Nótese la presencia de marcas de tiempo (timestamp) en algunas entidades, lo cual permite analizar el comportamiento dinámico del sistema. Las entidades con campos de tipo geometría permiten realizar análisis espaciales. Se utiliza información complementaria de otras fuentes de datos abiertos como OpenStreetMaps y la geometría de las redes de transporte del metro y metrobús / Figure 2. Data model for our case study. The entities in gray are obtained from open data sources, the light blue ones are calculated through a variety of processes that may include SQL queries and stored procedures (PGSQL). For reasons of clarity we modified entities and field names to be self-explanatory. The entities in purple participates directly in some map or graph. Note the presence of timestamps in some entities, which allows analyzing the dynamic behavior of the system. Entities with geometry type fields allow spatial analysis. Complementary information is also used from other open data sources such as the street network derived from OpenStreetMap data, and the metro and metrobús transport networks

del cliente PGAdmin es posible conectarse de manera remota para realizar tareas de gestión de datos, realizar consultas espaciales y definir funciones, por ejemplo, para calcular la geometría 3D de la parábola que representa

un cierto número de viajes entre una estación de origen y una de destino.

- R Studio Server: La realización de estudios estadísticos finos y la elaboración de gráficas puede efectuarse a tra-

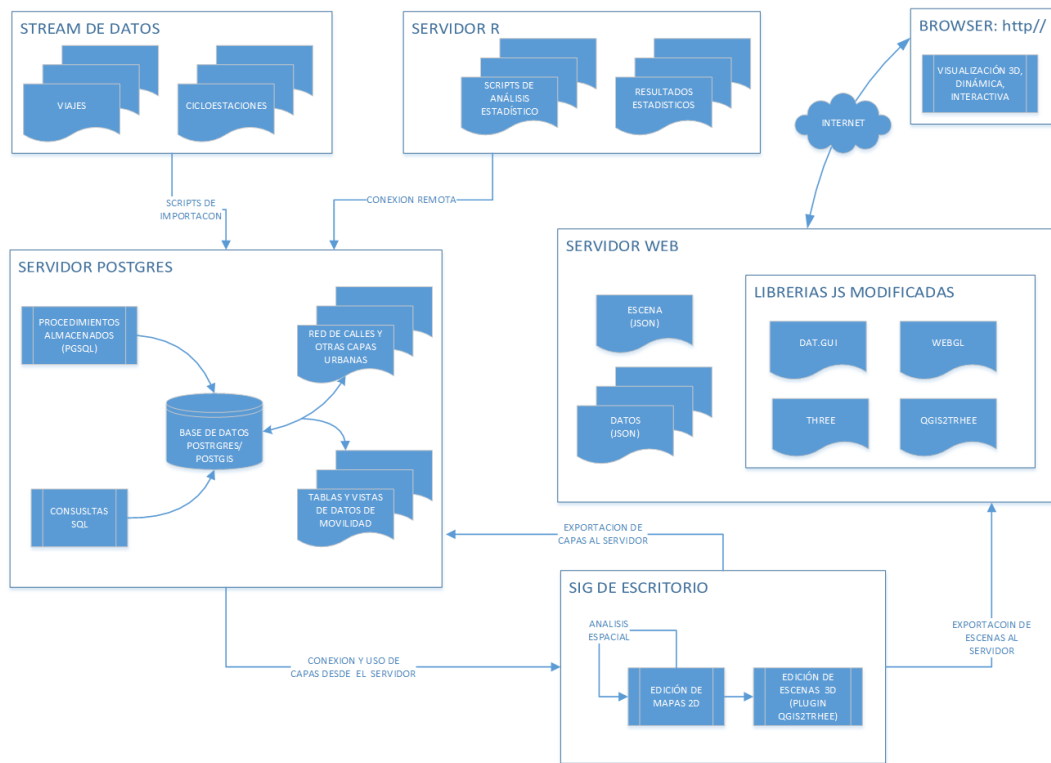


Figura 3: Principales componentes del sistema y sus interacciones. / Figure 3. System architecture: main components and its interactions

vés de R-Studio Server. El conector de R a PostgreSQL habilita la conexión directa a las tablas y vistas desde R. Esta herramienta también permite trabajar con los datos en forma remota, concurrente y distribuida.

- QGIS: El SIG de escritorio QGIS habilita el análisis y la edición de mapas utilizando los datos y vistas disponibles en la PostgreSQL. Esta conexión bidireccional permite, por ejemplo, calcular el voronoi generado a partir de los centroides geométricos de los clústers de cicloestaciones, obtenidos por un proceso de agregación por k-medias, y exportarlo a PostGIS. En sentido inverso; en Postgres se puede calcular la geometría 3D para un conjunto de conexiones origen-destino. Las vistas resultantes son accesibles desde QGIS.
- Edición y exportación de escenas: el plugin de QGIS "Qgis2Three.js", permite editar escenas 3D y exportarlas a un mapa interactivo formado por código HTML y scripts js que codifican la interacción. Alternativamente se pueden generar mapas interactivos 2D utilizando el plugin de QGIS "Qgis2Web".
- Edición del código y escenas para habilitar la visualización dinámica: A través de un entorno de programación, "Atom" por ejemplo, se realiza la edición de código para habilitar la visualización dinámica de los patrones espaciotemporales. En este paso realizamos modificaciones al código de las librerías `dat_gui.panel.js`, `Qgis2Three.js`

THREE.js para introducir los widgets y la lógica que habilita la exploración de secuencias temporales (iniciar/detener la secuencia, recorrerla paso a paso, etc.).

3.3.2. Estrategia de análisis

Los patrones de movilidad ciclista presentan variaciones temporales marcadas destacándose la presencia de pulsos a diferentes escalas relacionados con periodos de trabajo/ocio a lo largo del año, la semana y el día. También se ven afectados por procesos meteorológicos estacionales y/o episódicos.

Para tener una perspectiva completa de estos patrones de movilidad los analizamos partiendo de mayor a menor agregación, en el tiempo, en el espacio y en ambas dimensiones.

4. Resultados

4.1. Tendencias a largo plazo y patrones estacionales

4.1.1. Expansión de Ecobici

En la figura 4a se muestra el proceso de expansión del sistema Ecobici en términos de estaciones, bicicletas y superficie cubierta. Los datos se derivan de los registros de viajes acumulados desde febrero de 2010 a febrero de 2019. Los escalones en el número de estaciones en la gráfica 4b corresponden con los periodos de despliegue de estaciones en los polígonos de operación. Nótese que al entrar en operación el polígono IV no hay un incremento en el número de bicicletas, proporcional al incremento en la superficie. Presumiblemente esto se debe a que el

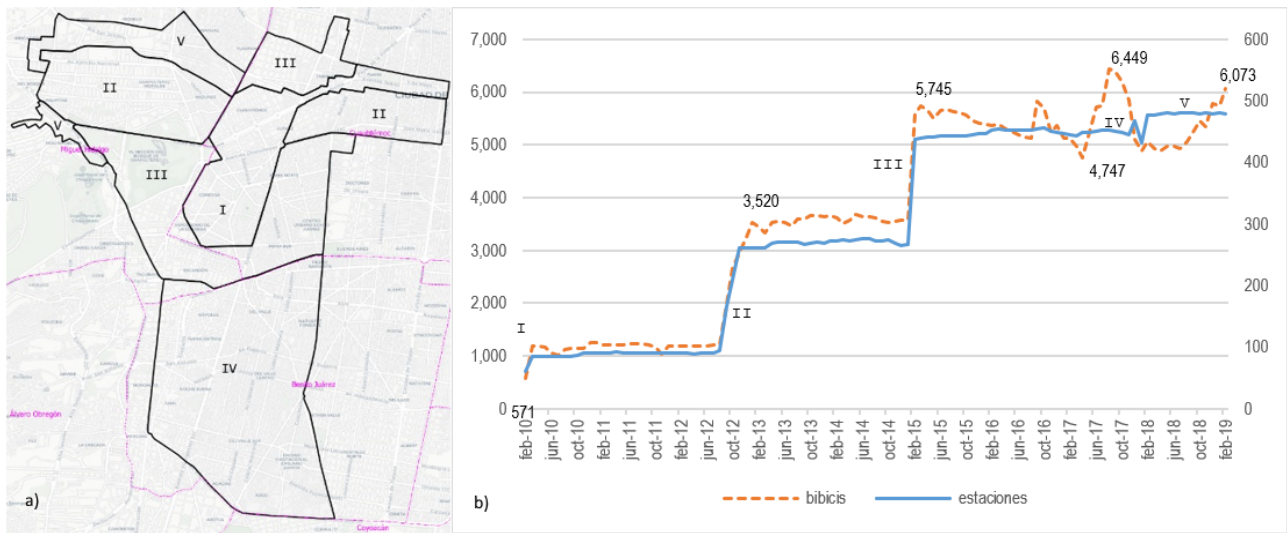


Figura 4: Evolución de Ecobici desde febrero de 2010 a febrero de 2019 a) polígonos de operación incorporados durante cada fase b) número de estaciones y bicicletas en operación (cortes mensuales). Fuentes: para el diagrama 1) Ecobici, para el diagrama 2) elaboración propia a partir de los datos acumulados de viajes publicados por Ecobici. / Figure 4. Evolution of Ecobici from February 2010 to February 2019 a) new areas incorporated during each phase b) number of bike stations and bicycles in operation (monthly cuts). Sources: for diagram a) Ecobici, for diagram b) author’s elaboration based on accumulated data of origin-destination trips, published by Ecobici.

número de estaciones que entran en operación es apenas mayor que el número de estaciones fuera de servicio. Nótese también que desde el inicio de la fase III se observa un patrón de picos y valles quizás asociado con fases de despliegue de nuevas bicis seguidas del retiro gradual de las que presentan condiciones deficientes.

El número de viajes acumulado a febrero de 2019 es de 58’237,050 de los cuales solo el 25.99 % ha sido realizado por mujeres. La duración promedio es de 00:13:43 (elaboración propia a partir del registro de viajes).

4.1.2. Evolución del número de viajes

En la figura 5 se presenta la tendencia registrada en el número de viajes desde febrero de 2010. A modo de referencia se incluye la variación mensual en el número de bicicletas. Nótese la disminución significativa y recurrente en el número de viajes, durante los meses de diciembre. Esto podría atribuirse a periodos de frío, pero también coincide con periodos de baja actividad laboral por vacaciones. Descontando las variaciones estacionales a partir de octubre de 2017 se aprecia una tendencia decreciente, que comienza a revertirse en julio de 2018, posiblemente por la introducción de nuevos sistemas de bicicletas sin anclaje –Mobike y VBike, Dzeba, y los sistemas de scooters compartidos –Grin, Lime y Bird entre otros.

La figura 6 presenta el número promedio de viajes diarios en cortes semanales durante el periodo de enero de 2018 a febrero de 2019. El rango de variación va de 32,611 viajes por día en la cuarta semana de febrero de 2018 a un mínimo de 11,638 en la cuarta semana de diciembre. En esta gráfica se aprecian disminuciones recurrentes en el número de viajes que posiblemente coinciden con puentes largos y periodos vacacionales. Ello sugiere que existen una dinámica asociada a ciclos laborales. Suá-

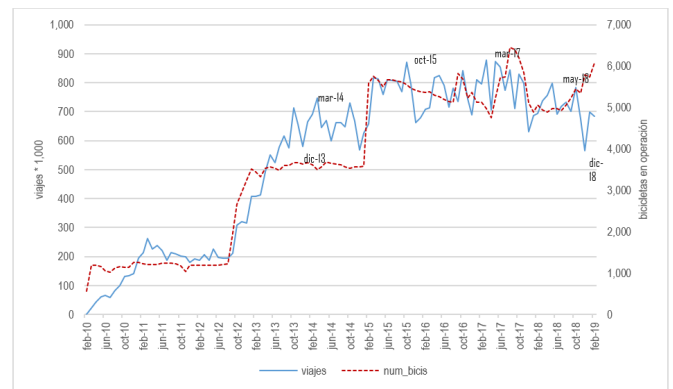


Figura 5: Viajes por mes entre febrero de 2010 y febrero de 2019. El pico máximo, 878,369 viajes por mes, se observa en marzo de 2017. En diciembre de 2018 se presenta un valor de 565,708 que es el valor más bajo registrado desde junio de 2013 (524,102). Fuente: elaboración propia a partir del registro de viajes publicado por Ecobici. / Figure 5. Trips per month between February 2010 and February 2019. The maximum peak, 878,369 trips per month is observed in March 2017. In December 2018, 565,708 where recorded, which is the lowest value since June of 2013 (524,102). Source: own elaboration based on trips records published by Ecobici.

rez et al. (2018) encontraron que existía una disminución en la proporción de viajes en Ecobici tanto en días fríos como en los lluviosos, lo que puede sugerir una combinación de factores que influyen en el uso diario del sistema.

4.2. Patrones de movilidad periódicos semanales y diarios

En la figura 7 se muestra el patrón de actividad por hora durante la semana del sábado 13 de mayo a las 5:00 am al sábado 20 a la misma hora. Durante los días laborables el número promedio de viajes por día es de 27,985; durante los fines de

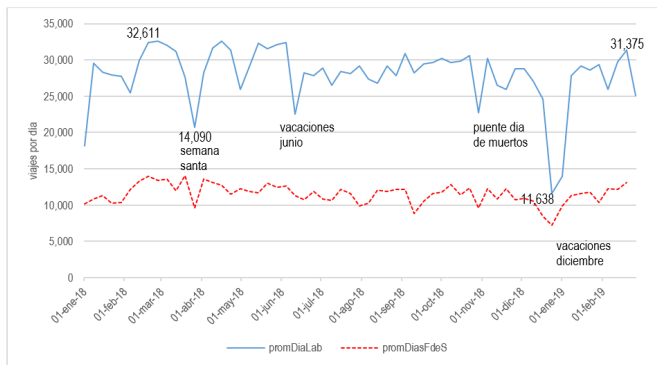


Figura 6: Promedio diario de viajes entre enero de 2018 y febrero de 2019 inclusive. Los datos se presentan en cortes semanales. La línea azul corresponde a días entre semana (lunes a viernes) la línea roja a sábados y domingos. / Figure 6. Daily average trips between January 2018 and February 2019 inclusive. Data is presented in weekly cuts. The blue line corresponds to weekdays (Monday to Friday) the red line to Saturdays and Sundays

semana el promedio es de 11,542, lo cual represente el 41 % de los registrados en días laborables. En la tabla 1 se muestra la duración, distancia y velocidad. Se tiene completa certeza de la duración de cada viaje. La distancia que se presenta es la distancia en línea recta entre el origen y el destino (OD), por ende, el valor reportado subestima la velocidad real, sin embargo, el sesgo registrado es consistente para todos los viajes independientemente del propósito y resulta útil para fines comparativos. Nótese que la velocidad media para desplazamientos durante el “ocio-dominical” es significativamente menor a la de los desplazamientos durante los días laborables.

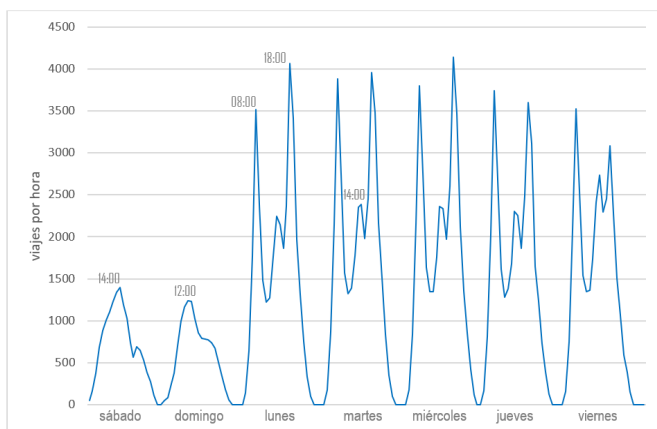


Figura 7: Número de viajes iniciados por hora durante la semana del 13 al 19 de mayo de 2017. / Figure 7. Number of trips by hour during the week of May 13 to 19, 2017.

4.3. Dinámica horaria de arribos y retiros en las cicloestaciones

El explorador interactivo (figura 8), permite examinar la dinámica de inicio y finalización de viajes para cada cicloestación (retiros y arribos de bicicletas, respectivamente). El patrón ge-

neral de actividad exhibe los picos detectados en la figura 6, y revela patrones espaciales interesantes.

Durante el pico matutino (entre las 8 y 9 am) se presenta una concentración de inicios de viaje en las cicloestaciones cercanas a la estación del tren ligero de Buenavista (figura 8b). El tope se presenta en las estaciones 271 (Carlos J. Meneses-Av. Jesús García) y la contigua 266 con 106 y 61 viajes iniciados respectivamente. Entre las dos suman 167 viajes iniciados en la misma cuadra. En un distante tercer lugar está la estación 107 (Tolsá-Balderas) con 36 viajes, lo cual representa apenas un tercio de los registrados en la 271. Durante el mismo periodo se registra un atractor de arribos en las cicloestaciones a lo largo del corredor de edificios de oficinas en la avenida Reforma entre las glorietas de río Misisipi y la glorietta de Niza (figura 8c). La estación más ocupada es la 14 (Reforma-Río de la plata) que registra 52 arribos, seguida de la 27 (Reforma-Havre) con 45 arribos.

Durante el pico vespertino (6 a 7 pm) el patrón se invierte. Los inicios de viaje se distribuyen en forma relativamente equitativa en el mismo corredor de Avenida Reforma. Se registra un tope de 39 (retiros / inicios de viaje) en la estación 27 (Reforma-Havre) y 37 en la siguiente estación que es la 16 (Reforma-Río Tíber) (figura 8d). En ese mismo rango horario, el atractor principal de arribos es la zona de Buenavista. El máximo se registra en la estación 266 (Av. Jesús García-Carlos J. Meneses) con 85 arribos. La estación 267 (en la misma cuadra y acera), registra 78. Entre las dos suman 163. La siguiente más atareada es la 43 (Juárez y Revillagigedo) con 35 arribos (figura 8e).

En síntesis: las estaciones cercanas a la terminal del tren ligero de Buenavista dominan los inicios y finales de viaje durante el pico matutino y vespertino respectivamente. El corredor sobre avenida Reforma, extendiéndose a la zona de Polanco, constituye un atractor y emisor difuso durante los picos matutino y vespertino respectivamente.

4.4. Conexiones y flujos dominantes en las horas pico

Un problema bien conocido de la representación cartográfica de conexiones origen-destino (OD), consiste en que el número de conexiones direccionales posibles para N estaciones, es N^2 . Para el sistema Ecobici estas son $(480)^2 = 230,400$. En la figura 9 se presentan las 93,703 conexiones direccionales que ocurrieron durante el mes de julio de 2018. Como puede apreciarse a medida que crece el número de estaciones se hace necesario establecer un compromiso entre legibilidad y representatividad del mapa, es decir: si se incluyen todas las instancias de la entidad de interés, el mapa es exhaustivo, pero se vuelve ilegible. Por otro lado, si se presentan pocas instancias el mapa se vuelve legible, pero puede ser muy poco representativo. Esta aseveración también es aplicable a mapas interactivos en línea, dado que entran en juego factores adicionales de la usabilidad/tiempo-de-respuesta para escenas interactivas, considerando que la velocidad de conexión y/o las capacidades del hardware desde el navegador “cliente” pueden ser factores limitativos adicionales.

Tabla 1: Duración, distancia y velocidad media de los desplazamientos durante los intervalos de mayor actividad.
 Table 1. Duration, distance and average speed of trips during intervals of peak activity.

Desplazamiento tipo	Intervalo	Duración Media	Distancia Media O-D ² (m)	Velocidad media O-D ³ (km/h)	Número acumulado de viajes ⁴
Hogar-trabajo ¹	6:30-9:30	12:13:22 AM	1930	9.4	265130
Trabajo-hogar ¹	16:30-19:30	12:16:15 AM	2074.7	8.4	314806
Ocio dominical	11:00-14:00	12:21:32 AM	1825.5	6.8	25837

¹Lunes a viernes. ²Distancia O-D es la distancia origen destino medida en línea recta. ³Velocidad medida considerando la velocidad observada y la distancia en línea recta. ⁴Se descartan los "viajes" que tienen una duración menor a 2 minutos.

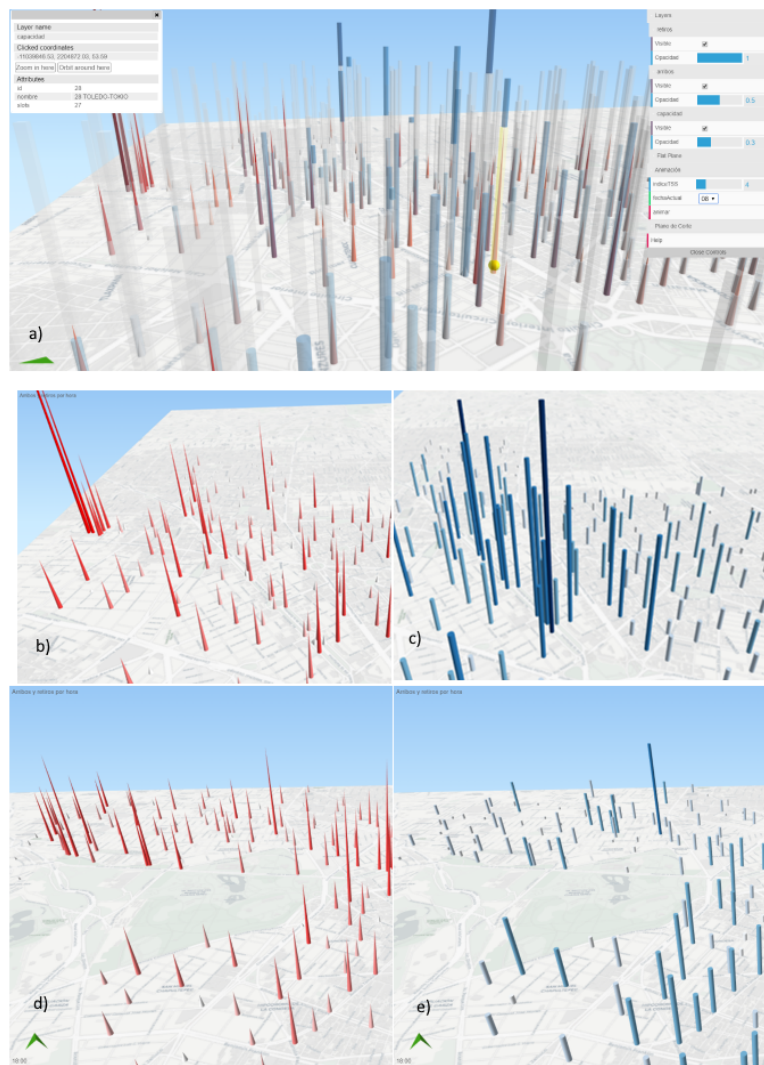


Figura 8: a) Exploración interactiva de la dinámica horaria de las cicloestaciones durante el lunes 31 de julio de 2017. Se muestra el número de arribos (cilindros azules), retiros (conos rojos) y la capacidad (columnas cuadrangulares) de cada ciclo estación. En b y c se muestran los retiros y arribos en el intervalo de 8 a 9. En d y e arribos y retiros durante el intervalo de 18 a 19. Nótese que si el $|(arribos - retiros)| = capacidad$, entonces la estación se vacía (o llena). El acarreo de bicis en camiones desde las estaciones llanas a las vacías y viceversa, permite mantener un nivel de arribos o retiros superior a la capacidad de la estación. / Figure 8.

Interactive viewer of the bike station hourly dynamics during July 31, 2017. b) arrivals (blue cylinders), number of trip starts (red cones) and capacity for each station are shown. In b) and c) arrivals and trip starts are shown for the 8 to 9 am interval. In d) and e) are shown for the 6 to 7 pm interval. Notice that if, for an interval $|(arribos - retiros)| = capacidad$ then the station is full or empties. Bike trawling from full stations to empty ones allow to sustain a flow higher than the capacity of the station.

4.4.1. Agrupación de cicloestaciones por el método de *k*-medias

Una forma razonable de minimizar el número de conexiones es agrupar las estaciones muy cercanas. En las figuras 9b y 9c se aprecian dos ejemplos de un patrón recurrente de distribución de cicloestaciones en las zonas de más alta demanda; es común encontrar varias estaciones contiguas, bien sea a unos cuantos pasos, cruzando la acera o la vuelta de la esquina. Por ejemplo, para un usuario que inicia o termina un viaje cerca de la estación de Buenavista debe ser relativamente indiferente iniciar o terminar su viaje en cualquiera de las 9 estaciones que pueden encontrarse en un tramo de 150 metros (figura 9b). Lo mismo puede decirse para las 8 estaciones cerca de la glorieta de la palma en Paseo de la Reforma (figura 9c).

Con la finalidad de agrupar estaciones sin perder representatividad zonal:

- 1) Generamos 96 clústers contiguos por el método de *k*-medias. Escogimos 96 centroides siguiendo una heurística de Pareto ya que 96 es el 20 % de 480,
- 2) Se calcularon los centroides geométricos de cada clúster,
- 3) Se calculó el voronoi de estos centroides para obtener 96 zonas de proximidad, cortándolo con un buffer que se extiende 250 metros alrededor de las estaciones en la periferia. En la figura 10 se muestran los clústeres resultantes,
- 4) Para la representación 3D se calculó una trayectoria parabólica para cada conexión de manera que la altura es proporcional al número de viajes.

Algunos parámetros de la segmentación resultante son los siguientes: Distancia media observada entre los centroides de clúster: 542.8 m. Distancia media esperada: 447.2 m. Índice de vecino más cercano: 1.2. Número de puntos: 96. Z-Score: 4.0

A través de este mecanismo redujimos el número de conexiones a 9,216, que representa el 4 % de las conexiones originales (962/4802).

Para mapear las conexiones durante las fases de mayor actividad, seleccionamos solamente los viajes alrededor de la hora pico matutina (8:00), vespertina (18:30) y la hora pico durante el ocio dominical (3:00) con un rango de más menos \pm una hora y media (figura 11). La altura de los arcos varía en forma proporcional con el número de viajes. El grosor depende de la clase a la que pertenece de acuerdo con cortes de Jenks.

4.4.2. Análisis de Pareto

Con el objetivo de medir la representatividad del mapa, realizamos un análisis de Pareto buscando maximizar la proporción del total del fenómeno explicada con un número mínimo de rasgos (figura 12). Este procedimiento complementa el método expuesto en el inciso anterior, para ello:

- 1) Separamos los arribos y los retiros para comparar el grado de concentración/dispersión durante las horas pico,

- 2) Acomodamos las estaciones en orden decreciente según el número de viajes,
- 3) Generamos los diagramas de Pareto,
- 4) Calculamos las curvas de regresión para tener un indicador de la bondad de ajuste y medir la proporción del total de los viajes explicada por un número dado de rasgos.

En 162 cicloestaciones, que representan el 36.1 %, inician 63.9 % de los viajes (figura 12a).

En 146 cicloestaciones (32.4 %), termina el 67.6 % de los viajes (figura 12b). En la tabla 2 nótese la enorme disparidad entre los tres primeros orígenes y los siguientes. Esta misma concentración se observa, aunque de manera menos aguda, en los destinos. Estos resultados confirman lo que puede apreciarse de manera directa en el mapa interactivo.

4.5. Balance y promedio diario de arrastres en cada cicloestación

Un factor determinante para la operación eficiente del SBC durante las horas de demanda pico, es la estrategia de arrastrar bicicletas de las estaciones que se llenan con rapidez a las estaciones que se vacían, con objeto de satisfacer la demanda de bicis en estas últimas y lugares de anclaje en las primeras (ver tabla 3). Dado el costo que representa mover las bicis de un lado a otro, se plantea un problema de optimización interesante que consiste en maximizar el flujo en la red minimizando el número de arrastres y la distancia recorrida. En la figura 13c, se muestra un camión con capacidad para transportar 50 bicicletas. Ecobici cuenta también con camiones para 16 bicicletas a los que se adjuntan remolques con capacidad para arrastrar 16 bicis más.

Dada la tabla de viajes (figura 2), es posible obtener de manera indirecta, información sobre los arrastres que sufre cada bicicleta de la siguiente manera: para cada viaje *i* de la bici *b* (${}_bV_i$) las tuplas ordenadas por *bici*, *fechaRetiro*, *horaRetiro*, permiten establecer la siguiente secuencia de cambios de posición de esa bici, de una estación a la siguiente:

$${}_bV_i[\text{fechaRetiro}_i, \text{horaRetiro}_i, \text{cicloEstacionRetiro}_i, \\ \text{fechaArribo}_i, \text{horaArribo}_i, \text{cicloEstacionArribo}_i]$$

$${}_bV_{i+1}[\text{fechaRetiro}_{i+1}, \text{horaRetiro}_{i+1}, \text{cicloEstacionRetiro}_{i+1}, \\ \text{fechaArribo}_{i+1}, \text{horaArribo}_{i+1}, \text{cicloEstacionArribo}_{i+1}]$$

...

$${}_bV_{i+n}[\text{fechaRetiro}_{i+n}, \text{horaRetiro}_{i+n}, \text{cicloEstacionRetiro}_{i+n}, \\ \text{fechaArribo}_{i+n}, \text{horaArribo}_{i+n}, \text{cicloEstacionArribo}_{i+n}]$$

En esta secuencia, si la estación de inicio del siguiente viaje de la bici *b* es diferente de la estación destino del viaje anterior,

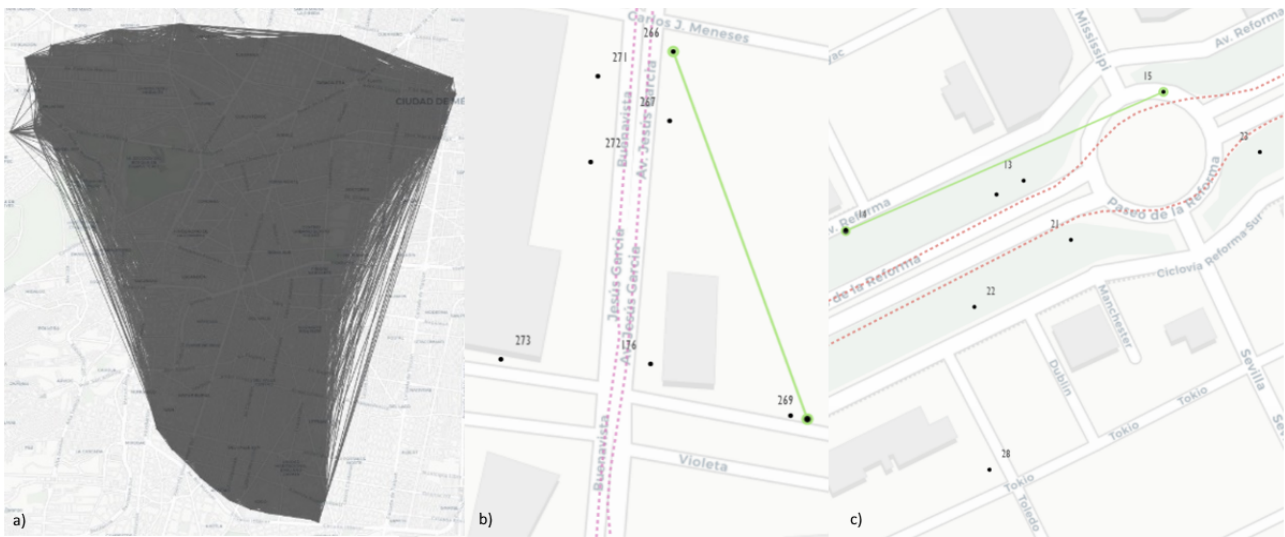


Figura 9: Necesidad y pertinencia de reducir el número de conexiones: a) 93,703 conexiones OD (líneas de color gris) ocurridas durante junio de 2018, la mayoría de las cuales son muy raras, es decir se presentan una sola vez o tienen una frecuencia muy baja b) y c) dos ejemplos de concentración de estaciones a una distancia menor a 200 metros. Los puntos numerados son cicloestaciones; en ambas imágenes la línea verde indica una distancia lineal de 200 metros. / Figure 9. Need and relevance of reducing the number of connections: a) 93,703 OD connections (gray lines), occurred during June 2018, most of which are very rare, that is, they occur only once or have a very low frequency b) and c) two examples of concentration of bike-sharing stations (numbered points) at a distance of less than 200 meters (green lines).

eso implica que fue arrastrada de la estación *cicloEstacionArribo_i* a la estación *cicloEstacionRetiro_{i+1}*, en algún momento durante el intervalo comprendido entre:

$$\begin{aligned} & fechaArribo_i + horaArribo_i \quad y \\ & fechaArribo_{i+1} + horaArribo_{i+1} \end{aligned}$$

Para tener una idea de la distribución espacial de este fenómeno, generamos el registro exhaustivo de arrastres durante el mes de enero de 2019. De aquí se deriva el promedio diario de arrastres de entrada y salida para cada cicloestación (ver el mapa interactivo asociado a la figura 13). En este periodo se observan 110,029 arrastres, que representan un 15.7 % de los 697,279 viajes registrados.

En la tabla 4 se puede apreciar que el promedio diario de arrastres con salida de la estación 266 es de 161.2 bicis. Esto multiplica por 5 la capacidad de dicha estación que es de 32 bicicletas, lo cual implica realizar por lo menos 3 viajes a lo largo del día en camiones de 50 bicis y uno en camión de 16 bicis. Nótese que todas las estaciones de la lista se concentran

en la zona aledaña a la estación Buenavista y Reforma. Esto se puede apreciar también en el mapa interactivo a través de la capa de arrastres origen-destino (“arrastres OD”).

En el mapa (figura 13) se puede apreciar que la distribución espacial es notablemente asimétrica: muy pocas estaciones concentran casi todos los arrastres. Esto se explica porque no es rentable trasladar bicis de una estación a otra a menos que estas estén llenas o vacías.

5. Discusión y trabajo a futuro

Como se ilustra en el inciso 4.4 la cartografía de flujos direccionales resulta problemática. El método para organizar paquetes de conexiones (*edge bundling*) propuesto por Graser et al. (2019), funciona bien para mapas bidimensionales. Nuestro enfoque de representación basado en escenas interactivas añade dos dimensiones que permiten 1) mostrar con claridad la jerarquía de las conexiones codificándola por ejemplo, en la altura de la parábola y grosor de los iconos; 2) el uso de iconos direccionales (por ejemplo, conos), indica con toda claridad la

Tabla 2: Diez orígenes y destinos dominantes durante el pico matutino.
Table 2. Ten dominant origins and destinations during the morning peak.

origen	Id estación	272	266	271	1	41	19	36	43	21	267
	viajes	773	756	718	382	351	333	331	318	312	311
destino	Id estación	16	18	27	23	17	13	21	19	261	28
	viajes	653	613	526	521	519	451	442	401	374	362

Tabla 3: Diez estaciones con mayor desbalance considerando el promedio mensual de arribos y retiros
 Table 3. Ten most imbalanced stations considering its monthly average of arrivals vs. withdraws

CICLOESTACIÓN	REFORMA-DUBLÍN	NEWTON-HORACIO	AV. CENTRAL-J. MENESES	RUBÉN DARÍO-REFORMA ANDRÉS	BELLO-ELLIOT	ÁLVARO OBREGÓN-TONALÁ	EULER-AV. HORACIO	JUÁREZ-REVILLA-GIGEDO	AV. JESÚS GARCÍA-CARLOS J. MENESES	AV. JESÚS GARCÍA-CARLOS J. MENESES
Id estación	21	211	271	193	238	136	217	43	267	266
balance (arribos-retiros) (arrivals-withdraws)	-1034	-875	-634	-630	-629	588	690	768	2523	2674
	estaciones “sumidero” (sink stations)					estaciones “surtidor” (pump stations)				

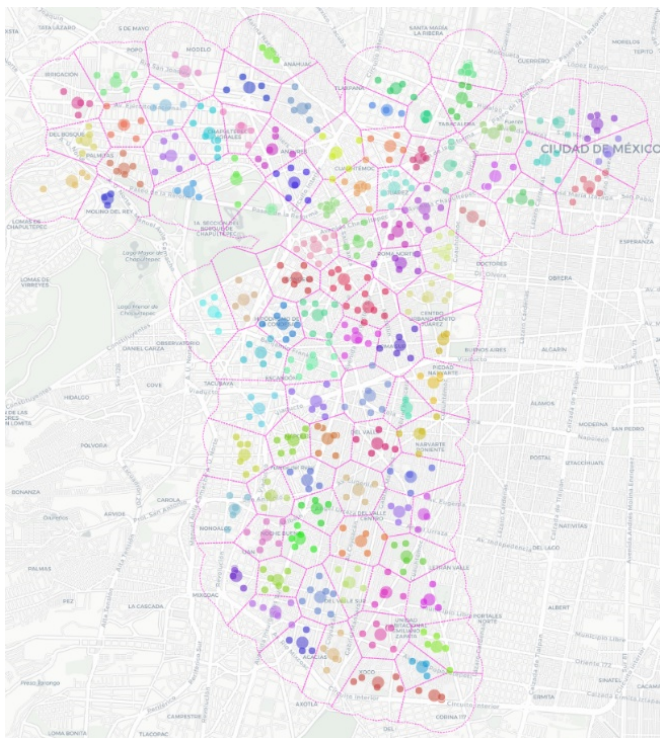


Figura 10: 96 clúster de estaciones generado por el método de k-medias. Los polígonos de Voronoi se derivan de los centroides correspondientes./ Figure 10. 96 clusters generated by k-means. Voronoi polygons are derived from the corresponding centroids.

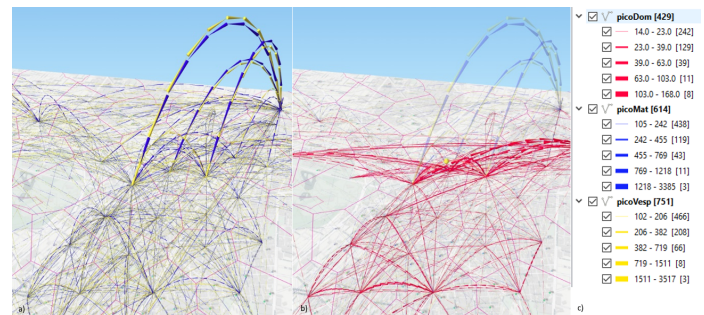


Figura 11: Conexiones direccionales dominantes medido en términos del número promedio de viajes por hora entre zonas: a) durante el pico matutino (azul) y vespertino (amarillo) en días laborables y b) durante el pico dominical (rojo); se incluye el patrón matutino como referencia (azul atenuado). Entre corchetes se indica el número de conexiones en cada rango. Nótese la simetría asíncrona de los flujos en a. En b nótese la diferencia muy notable en el número (altura) y longitud de las conexiones durante el periodo de ocio dominical. / Figure 11. Dominant directional connections measured in average trips per hour between zones: a) during the morning (blue) and evening (yellow) peaks on weekdays and b) during the Sunday peak (red); the morning pattern is included as a reference (attenuated blue). The number of connections in each range is indicated in brackets. Note the asynchronous symmetry of the flows in (a). In (b) note the sharp difference in the number (height) and length of connections, during the Sunday leisure period.

dirección del flujo; 3) el carácter interactivo del mapa permite al usuario controlar el grado de saturación de información, manipulando la visibilidad y opacidad de las capas y permitiéndole navegar en la escena; 4) la cuarta dimensión codificada en las marcas de tiempo (*timestamp*), habilita el examen de las variaciones espaciotemporales del fenómeno de flujo estudiado. Pensamos que estas características agregadas, facilitan exami-

nar el conjunto sin perderse en los detalles y, cambiando de escala, examinar los detalles sin perder el conjunto. Por supuesto nuestro enfoque es vulnerable también al crecimiento cuadrático en el número de nodos, por lo que el enfoque de Graser et al. 2019 es complementario. Como trabajo a futuro se sugiere explorar las posibilidades de extender este último método a 4 dimensiones.

Otro recurso de simplificación que utilizamos es el de agregar estaciones contiguas por el método de k-medias. Este recurso reduce cuadráticamente el número de conexiones a expensas de perder detalle al interior de las zonas conectadas: sabemos que el viaje inició, por ejemplo, en la vecindad de la estación

Tabla 4: Diez estaciones con más arrastres de entrada y salida. Se muestran totales y promedios diarios para el mes de enero de 2019
 Table 4. Ten stations with the most entry and exit trawls. Totals and daily average for January 2019

ESTACIÓN	ARRASTRES SALIDA	PROMEDIO DIARIO	ESTACIÓN	ARRASTRES ENTRADA	PROMEDIO DIARIO
266 AV. JESÚS GARCIA-CARLOS J. MENESES	4997	161.2	271 AV. CENTRAL-J. MENESES	4346	140.2
271 AV. CENTRAL-J. MENESES	3709	119.6	266 AV. JESÚS GARCIA-CARLOS J. MENESES	2411	77.8
267 AV. JESÚS GARCÍA-CARLOS J. MENESES	3620	116.8	272 AV. CENTRAL-J. MENESES	2165	69.8
272 AV. CENTRAL-J. MENESES	1853	59.8	23 REFORMA-PRAGA	1746	56.3
16 REFORMA-RÍO TIBER	1571	50.7	27 REFORMA-HAVRE	1468	47.4
18 REFORMA-RIO RHIN	1460	47.1	211 NEWTON-HORACIO	1448	46.7
23 REFORMA-PRAGA	1376	44.4	21 REFORMA-DUBLIN	1334	43
14 REFORMA-RIO DE PLATA	1296	41.8	16 REFORMA-RÍO TIBER	1333	43
27 REFORMA-HAVRE	1287	41.5	14 REFORMA-RIO DE PLATA	1291	41.6
17 REFORMA-RIO TIBER	1157	37.3	267 AV. JESÚS GARCÍA-CARLOS J. MENESES	1228	39.6

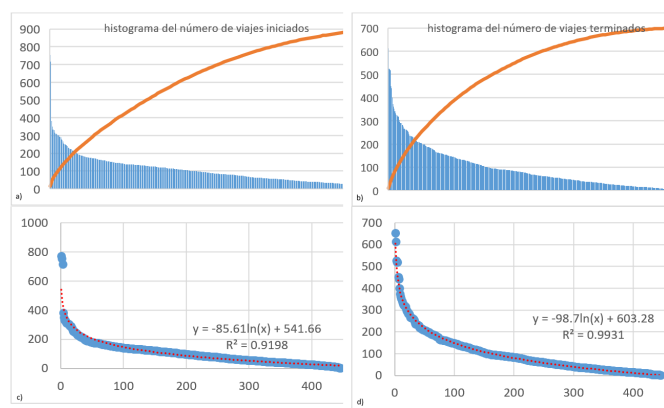


Figura 12: Cicloestaciones ordenadas según el número de viajes a) iniciados, y b) terminados en ella, durante la hora pico matutina en días laborables; los diagramas c) y d) muestran las curvas de regresión para la distribución de inicios y arribos respectivamente, durante el mismo lapso. / Figure 12. bike-sharing stations ordered according to a) the number of trips started, and b) finished in it, during the morning peak hour on weekdays; the diagrams c and d show the regression curves for the distribution of the beginnings and arrivals, respectively, during the same period.

Buenavista, y terminó en la vecindad de la glorieta de la Palma, se desconocen el origen y destinos específicos. Esta desventaja puede aminorarse mostrando la probabilidad de cada estación para el punto de origen o término de cada conexión, tomando

en cuenta que la probabilidad varía con la hora del día. Nótese la dinámica extraordinariamente asimétrica de estaciones ubicadas a unos cuantos metros entre sí (figura 14).

En el caso de que no sea factible o deseable representar todas las conexiones, el análisis de Pareto permite medir y explicitar qué proporción de rasgos está representada en el mapa y cuál es la proporción del fenómeno cubierta por este conjunto de rasgos. Dado el hecho de que en nuestro caso la altura representa la intensidad del fenómeno, podríamos usar el plano de corte que se muestra en la figura 8 para mostrar interactivamente el significado espacial de estos niveles de corte. Las curvas de regresión (figuras 12b y 12c) pueden utilizarse como un recurso eficiente para calcular estas proporciones.

A partir del monitoreo del estatus de las cicloestaciones en lapsos de tiempo muy cortos puede construirse un registro exhaustivo de la dinámica espaciotemporal de arribos y retiros lo que permitiría reproducir a diferentes velocidades la dinámica en tiempo real del sistema.

6. Conclusiones

En este trabajo revisamos la literatura existente sobre visualización en general y de manera específica sobre visualización de movilidad ciclista. Presentamos un esquema metodológico y resultados que ilustran una forma innovadora, hasta donde alcanza nuestro conocimiento, de explorar en línea estos patrones

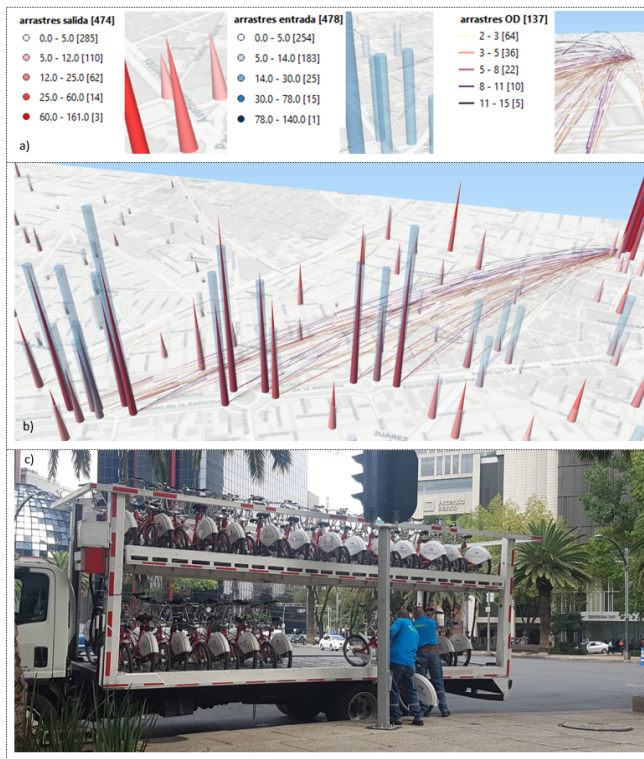


Figura 13: Arrastres y conexiones origen-destino dominantes. Se muestran promedios diarios para arrastres de salida (conos rojos), arrastres de entrada (cilindros azules) y arrastres entre cicloestaciones (parábolas clasificadas según el número de arrastres). La altura es proporcional al número. Se omiten todas las conexiones dirigidas con un promedio diario menor a 2 / Figure 13. Bike trawl and dominant origin-destination connections. Daily averages are shown for exit trawls (red cones), entry trawls (blue cylinders) and trawls between specific stations (classified parabolas, according to the number of trawls). Height is proportional to the number. All the directed connections with a daily average of less than 2 are omitted

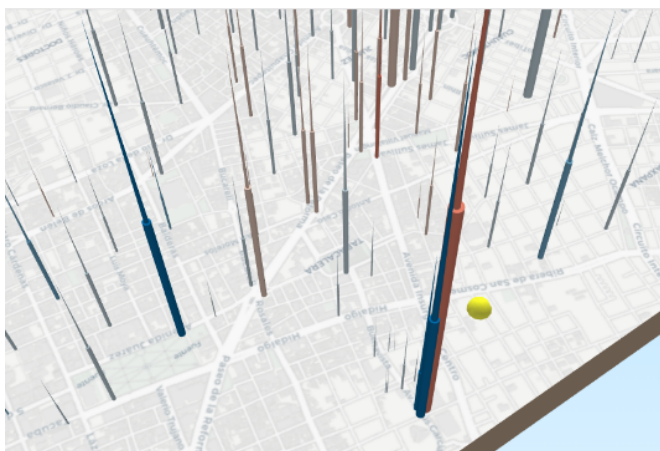


Figura 14: Ejemplo de estaciones separadas por unos cuantos metros, que presentan una fuerte asimetría en el balance de flujo./ Figure 14. Example of stations separated by a few meters, which have a strong asymmetry in the flow balance.

a través de escenas tridimensionales, espaciotemporales, interactivas.

Se realizó un análisis a diferentes escalas espaciales y temporales, de la dinámica del sistema Ecobici de la ciudad de México, que se presenta a través de gráficas y cuatro mapas referentes a 1) la estructura del sistema, 2) la dinámica horaria de las cicloestaciones a lo largo de un día, 3) las conexiones dominantes OD entre clúster de cicloestaciones agregadas por proximidad, 4) la distribución espacial de los arrastres realizados para compensar excesos espaciotemporales de demanda y oferta.

Se concluye que:

- Las escenas referidas son un recurso eficaz para comunicar visualmente 1) el nivel de actividad de las cicloestaciones y sus variaciones en el espacio y en el tiempo, 2) el grado de (a)simetría de los flujos entre zonas y sus variaciones durante las horas pico, 3) el nivel de actividad y balance acumulado de las cicloestaciones.
- El examen visual facilita la formulación de hipótesis a ser validadas mediante análisis estadísticos finos.
- La arquitectura propuesta, permite utilizar R Studio para realizar de manera remota, análisis sobre los datos almacenados en PostgreSQL.
- La publicación del tipo de escenas que se presentan aquí –dinámicas, interactivas y en línea, basada en datos abiertos y software libre, ofrece a la comunidad científica un recurso escalable para difundir estudios sobre fenómenos de movilidad, a través de mapas Web interactivos.

Referencias

Aigner, W., Miksch, S., Schumann, H., Tominski, C., 2011. Visualization of Time-Oriented Data. Springer, London, <https://doi.org/10.1007/978-0-85729-079-3>.

Andrienko, N., Andrienko, G., 2013. Visual analytics of movement: An overview of methods, tools and procedures. Information Visualization 12 (1), 3–24, <https://doi.org/10.1177/1473871612457601>.

Basaraner, M., 2016. Revisiting cartography: towards identifying and developing a modern and comprehensive framework. Geocarto International 31 (1), 71–91, <https://doi.org/10.1080/10106049.2015.1041560>.

Batty, M., 2013. Big data, smart cities and city planning. Dialogues in Human Geography 3 (3), 274–279, <https://doi.org/10.1177/2043820613513390>.

Ferreira, N., Poco, J., Vo, H. T., Freire, J., Silva, C. T., 2013. Visual Exploration of Big Spatio-Temporal Urban Data: A Study of New York City Taxi Trips. IEEE Transactions on Visualization and Computer Graphics 19 (12), 2149–2158, <https://doi.org/10.1109/TVCG.2013.226>.

Gobierno de la Ciudad de México, 2014. Encuesta ECOBICI 2014 Retrieved from https://www.ecobici.cdmx.gob.mx/sites/default/files/pdf/ecobici_2014_encuesta.pdf.

Graser, A., Schmidt, J., Roth, F., Brändle, N., 2019. Untangling origin-destination flows in geographic information systems. Information Visualization 18 (1), 153–172, <https://doi.org/10.1177/1473871617738122>.

Harder, C., Brown, C., 2017. The ArcGIS Book: 10 Big Ideas About Applying The Science of Where. Esri Press, retrieved from <http://go.esri.com/1/82202/2017-05-24/fdd14n>.

- Hengl, T., Roudier, P., Beaudette, D., Pebesma, E., 2015. plotKML: Scientific visualization of spatio-temporal data. *Journal of Statistical Software* 63 (5), 1–25.
- INEGI, Gobierno del Distrito Federal, Gobierno del Estado de México, 2007. Encuesta Origen Destino 2007. México, Distrito Federal.
- Kapler, T., Wright, W., 2005. GeoTime Information Visualization. *Information Visualization* 4 (2), 136–146, <https://doi.org/10.1057/palgrave.ivs.9500097>.
- Keim, D., Andrienko, G., Fekete, J.-D., Görg, C., Kohlhammer, J., Melançon, G., 2008. Visual Analytics: Definition, Process, and Challenges. En: Kerren, A., Stasko, J. T., Fekete, J.-D., North, C. (Eds.), *Information Visualization*. Springer, Berlin, Heidelberg, pp. 154–175, https://doi.org/10.1007/978-3-540-70956-5_7.
- Ourednik, A., 2017. The third dimension of political mapping: exploiting map interactivity for a better understanding of election and referendum results. *Cartography and Geographic Information Science* 44 (4), 284–295, <https://doi.org/10.1080/15230406.2017.1295820>.
- O'Brien, O., Cheshire, J., Batty, M., 2014. Mining bicycle sharing data for generating insights into sustainable transport systems. *Journal of Transport Geography* 34, 262–273, <https://doi.org/10.1016/j.jtrangeo.2013.06.007>.
- Ren, L., Du, Y., Ma, S., Zhang, X. L., Dai, G. Z., 2014. Visual analytics towards big data. *Ruan Jian Xue Bao/Journal of Software* 25 (9), 1909–1936, <https://doi.org/10.13328/j.cnki.jos.004645>.
- Romanillos, G., Austwick, M. Z., Ettema, D., De Kruijf, J., 2016. Big Data and Cycling. *Transport Reviews* 36 (1), 114–133, <https://doi.org/10.1080/01441647.2015.1084067>.
- Romanillos, G., Moya-Gómez, B., Zaltz-Austwick, M., Lamíquiz-Daudén, P. J., 2018. The pulse of the cycling city: visualising Madrid bike share system GPS routes and cycling flow. *Journal of Maps* 14 (1), 34–43, <https://doi.org/10.1080/17445647.2018.1438932>.
- Sandvik, B., 2008. Using KML for thematic mapping. MSc in Geographical Information Science thesis. Institute of Geography School of GeoSciences. Edinburgh, University of Edinburgh, retrieved from: https://thematicmapping.org/~thematic/downloads/Using_KML_for_Thematic_Mapping.pdf.
- Steg, L., Gifford, R., 2005. Sustainable transportation and quality of life. *Journal of Transport Geography* 13 (1), 59–69, <https://doi.org/10.1016/j.jtrangeo.2004.11.003>.
- Suárez-Lastra, M., 2018. Plan Bici CDMX. Ciudad de México: Gobierno de Ciudad de México, Banco Interamericano de Desarrollo, Universidad Nacional Autónoma de México.
- Thakur, S., Hanson, A. J., 2010. A 3d Visualization of Multiple Time Series on Maps. En: 2010 14th International Conference Information Visualisation. pp. 336–343, <https://doi.org/10.1109/IV.2010.54>.
- Thakuriah, P. V., Tilahun, N., Zellner, M. (Eds.), 2017a. Seeing Cities Through Big Data: Research, Methods and Applications in Urban Informatics. Springer, <https://doi.org/10.1007/978-3-319-40902-3>.
- Thakuriah, P. V., Tilahun, N. Y., Zellner, M., 2017b. Big Data and Urban Informatics: Innovations and Challenges to Urban Planning and Knowledge Discovery. En: Thakuriah, P. V., Tilahun, N. Y., Zellner, M. (Eds.), *Seeing Cities Through Big Data*. Springer, pp. 11–45, https://doi.org/10.1007/978-3-319-40902-3_2.
- Thakuriah, P. V., Tilahun, N. Y., Zellner, M., 2017c. Introduction to seeing cities through big data: Research, methods and applications in urban informatics. En: Thakuriah, P. V., Tilahun, N. Y., Zellner, M. (Eds.), *Seeing Cities Through Big Data*. Springer, pp. 1–9, doi: https://doi.org/10.1007/978-3-319-40902-3_1.
- Tominski, C., Schulze-Wollgast, P., Schumann, H., 2005. 3d information visualization for time dependent data on maps. En: Ninth International Conference on Information Visualisation (IV'05). pp. 175–181, <https://doi.org/10.1109/IV.2005.3>.
- Tominski, C., Schumann, H., Andrienko, G., Andrienko, N., Dec. 2012. Stacking-Based Visualization of Trajectory Attribute Data. *IEEE Transactions on Visualization and Computer Graphics* 18 (12), 2565–2574, <https://doi.org/10.1109/TVCG.2012.265>.
- Wikipedia, 2019. List of bicycle-sharing systems. Accessed: 29/03/2019. Retrieved from https://en.wikipedia.org/wiki/List_of_bicycle-sharing_systems.
- Zhu, L., Li, Z., Wang, Z., Mar. 2018. Representing Time-Dynamic Geospatial Objects on Virtual Globes Using CZML—Part II: Impact, Comparison, and Future Developments. *ISPRS International Journal of Geo-Information* 7 (3), 102, <https://doi.org/10.3390/ijgi7030102>.

This article accompanies the following material:

- HTML: DOI: 10.22201/igg.25940694.2019.2.61.129
- (2D) Base map: DOI: 10.22201/igg.25940694.2019.2.61.130
- (3D) Hourly dynamics: DOI: 10.22201/igg.25940694.2019.2.61.131
- (3D) Bike trawl: DOI: 10.22201/igg.25940694.2019.2.61.133
- (3D) Dominant connections: DOI: 10.22201/igg.25940694.2019.2.61.132