

Osorio Arjona, J. (2022). Big Data y matrices Origen-Destino: cartografía de flujos de movilidad en España a partir de datos de Twitter y comparación con datos de telefonía móvil, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 29, 115–130. <https://dx.doi.org/10.21138/GF.776>

## **BIG DATA Y MATRICES ORIGEN-DESTINO: CARTOGRAFÍA DE FLUJOS DE MOVILIDAD EN ESPAÑA A PARTIR DE DATOS DE TWITTER Y COMPARACIÓN CON DATOS DE TELEFONÍA MÓVIL**

**Joaquín Osorio Arjona**  

Departamento de Población. Centro de Ciencias Humanas y Sociales CSIC  
C/ Albasanz 26, 28037, Madrid, España  
joaquin.osorio@cchs.csic.es

### **RESUMEN**

Este trabajo analiza el valor de nuevas fuentes de datos basados en Big Data para estudiar la movilidad en España y comparar diferentes patrones de movilidad observados según los datos usados. Para ello, este trabajo emplea datos de Twitter geolocalizados publicados en España durante un periodo de 30 meses, analiza la distribución espaciotemporal de los usuarios de Twitter según la provincia o mes en el que se ha publicado el *tweet*, y diseña una serie de matrices Origen-Destino con el objetivo de visualizar diferentes patrones en los flujos de movilidad en días laborales o en el periodo vacacional estival. Finalmente, se comparan los resultados obtenidos con las matrices Origen-Destino publicadas por el Ministerio de Fomento de España a partir de datos de telefonía móvil. Los resultados obtenidos indican una distribución espacial de usuarios de Twitter cercana a la realidad, un mayor número de usuarios en el mes de agosto, y el papel de la Comunidad de Madrid como provincia núcleo de atracción de viajes a nivel nacional. Respecto a los datos de telefonía móvil se ha observado una mayor concentración de viajes con origen o destino en la provincia de Madrid a partir de datos de Twitter.

Palabras clave: Twitter; matrices Origen-Destino; telefonía móvil; movilidad laboral; movilidad estival

### **BIG DATA AND ORIGIN-DESTINATION MATRICES: MAPPING OF MOBILITY FLOWS IN SPAIN BASED ON TWITTER DATA AND COMPARISON WITH CELL PHONE DATA**

### **ABSTRACT**

This work analyses the value of new data sources based on Big Data to study mobility in Spain and compares different mobility patterns observed according to the data used. To do this, this work uses geolocated Twitter data published in Spain over a period of 30 months, analyses the spatio-temporal

Osorio Arjona, J. (2022). Big Data y matrices Origen-Destino: cartografía de flujos de movilidad en España a partir de datos de Twitter y comparación con datos de telefonía móvil, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 29, 115–130. <https://dx.doi.org/10.21138/GF.776>

---

distribution of Twitter users according to the province or month in which the tweet was published, and designs a series of Origin-Destination matrices with the aim of visualizing different patterns in mobility flows on working days or during the summer vacation period. Finally, the results obtained are compared with the Origin-Destination matrices published by the Spanish Ministry of Development based on mobile phone data. The results obtained indicate a spatial distribution of Twitter users close to reality, a greater number of users in the month of August, and the role of the Community of Madrid as a core province that attracts travel at a national level. Regarding mobile phone data, a greater concentration of trips with origin or destination in the province of Madrid has been observed based on Twitter data.

Keywords: Twitter; Origin-Destination flows; mobile phones; weekdays mobility; summer mobility

## 1. Introducción

La movilidad es un proceso muy dinámico cada vez más complejo y menos sostenible. Este fenómeno está caracterizado por un aumento del número de viajes, una mayor diversidad de motivos para viajar, y un uso más intensivo del transporte motorizado (Gutiérrez & García-Palomares, 2007). Los modelos de demanda de transporte involucran análisis del número de viajes, destinos, rutas, y modos de transporte. Para poder entender la demanda de viajes, los planes de transporte tienen que tener en cuenta las distribuciones espaciotemporales de la localización de las personas en diferentes periodos de tiempo (Demissie *et al.*, 2019). Ante la complejidad que presenta la movilidad urbana, las fuentes tradicionales necesitan ser complementadas con nuevos datos que puedan actualizarse constantemente y que cuenten con una mayor resolución tanto espacial como temporal (Osorio-Arjona & García-Palomares, 2019a).

En la actualidad, vivimos en una sociedad hiperconectada en la que las personas usan y consumen internet de forma masiva y diaria y en la que casi todos los ciudadanos llevan consigo un teléfono móvil que permite la monitorización de su actividad humana diaria. El aumento del uso de los teléfonos móviles y la continua evolución tecnológica que estos dispositivos sufren han conllevado un importante aumento en la cantidad y calidad de los datos obtenidos con los que se puede analizar la movilidad diaria. A través de sus dispositivos móviles, los usuarios generan una huella digital de sus actividades y movimientos, una traza digital que puede ser seguida (Blanford *et al.*, 2015). Los datos geográficos que son creados por los dispositivos móviles pueden ser adquiridos fácilmente, a bajo coste, y pueden ser analizados en Sistemas de Información Geográfica, posibilitando la monitorización, cartografía y actualización constante de la información acerca de la movilidad en ciudades y regiones (Gutiérrez-Puebla *et al.*, 2016; Osorio-Arjona & García-Palomares, 2017).

Desde hace unos años, las empresas públicas y privadas han empezado a utilizar datos basados en el Big Data frente a la metodología clásica basada en encuestas domiciliarias y telefónicas debido a la ventaja de obtener grandes muestras de datos con un alto detalle espacial y temporal en poco tiempo y a bajo coste (Gutiérrez Puebla, 2018; Osorio-Arjona & García-Palomares, 2019a). Los datos de telefonía móvil son la nueva fuente de datos más utilizada, debido al colosal tamaño de sus muestras y la resolución temporal muy alta que poseen, con un registro producido en un intervalo de

Osorio Arjona, J. (2022). Big Data y matrices Origen-Destino: cartografía de flujos de movilidad en España a partir de datos de Twitter y comparación con datos de telefonía móvil, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 29, 115–130. <https://dx.doi.org/10.21138/GF.776>

---

pocos segundos (Gutiérrez-Puebla *et al.*, 2019). Sin embargo, la resolución espacial de los datos de telefonía móvil no depende de la ubicación del teléfono del usuario, sino de las antenas de telefonía móvil que recogen las llamadas o el uso de la red 4G o 5G (Chen, Ma, Susilo, Liu, & Wang, 2016).

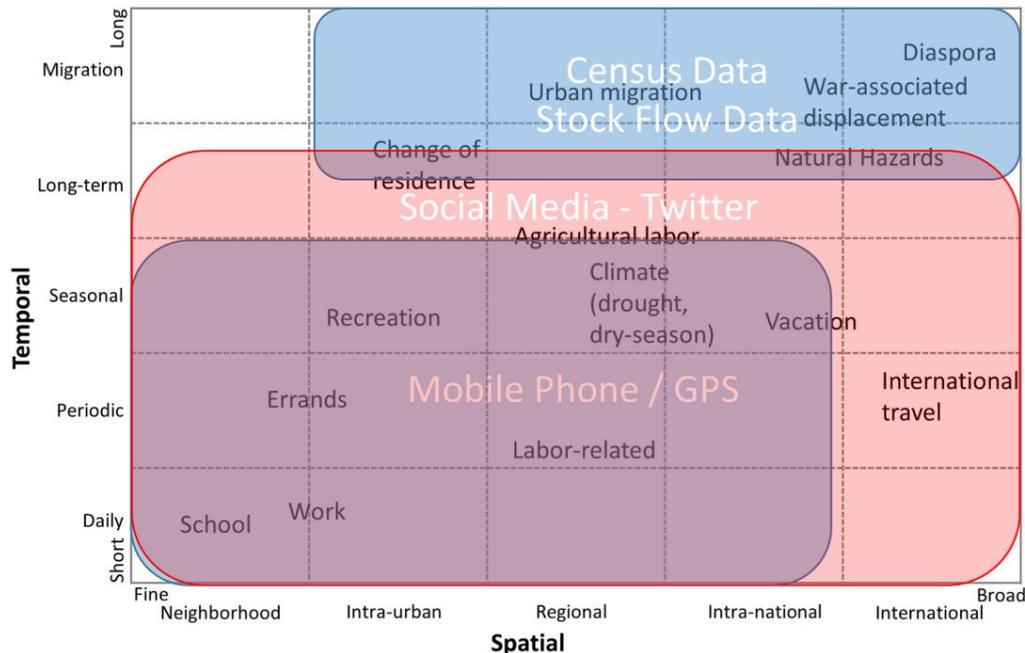
Otro tipo de fuentes de datos masivos utilizados en estudio de movilidad es la constituida por los sistemas de posicionamiento global o GPS. Los datos basados en GPS cuentan con una alta resolución espacial y temporal que permiten monitorizar y analizar diferentes procesos de movilidad (Gutiérrez-Puebla *et al.*, 2020). Entre las nuevas fuentes de datos que usan datos GPS destaca Twitter, red social en forma de microblogging que permite a los usuarios publicar un mensaje o un *tweet* con una estampa temporal completa, pero también con coordenadas espaciales xy. Por tanto, los datos de Twitter cuentan con una mayor resolución espacial que los datos de telefonía móvil además del valor semántico añadido que contiene un *tweet*. Otra ventaja del uso de datos de Twitter consiste en la posibilidad de descargar grandes cantidades de datos de forma gratuita, mientras que para obtener datos de telefonía móvil hay que depender de las compañías telefónicas (Hawelka *et al.*, 2014). Sin embargo, hay que tener en cuenta que la frecuencia de publicación de datos de Twitter es menor que los datos de telefonía móvil, por lo que la resolución temporal y el tamaño de las muestras son menores.

A la hora de estudiar la movilidad, las matrices Origen-Destino (OD) son una herramienta fundamental para visualizar los flujos de movilidad, predecir la demanda de transporte y diagnosticar mejoras en el uso de una red (Gao *et al.*, 2014). Las nuevas fuentes de datos masivos como los datos de telefonía móvil o las redes sociales como Twitter permiten trabajar con datos de alta resolución espacio-temporal en forma de punto, y usar la frecuencia espacial de estos puntos para poder definir un lugar de origen y un lugar de destino (Osorio-Arjona & García-Palomares, 2019b). Aunque los datos basados en Big Data permiten visualizar y analizar los flujos de movilidad con gran detalle en ciudades y áreas metropolitanas, su potencial para obtener información a mayores escalas como países está todavía poco explorado.

El objetivo de esta investigación es analizar el valor de las nuevas fuentes de datos basados en el Big Data para estudiar la movilidad a escala nacional a partir del diseño de matrices OD a escala nacional. Para ello, este trabajo usa datos de Twitter para visualizar flujos de movilidad entre las provincias que conforman España, teniendo en cuenta diferentes periodos de actividad como la movilidad en calendario laboral o en época estival. Los resultados obtenidos se comparan con matrices OD construidas por el Ministerio de Fomento de España a partir de datos de telefonía móvil, bajo la hipótesis de que los datos de Twitter tienen el potencial de capturar el movimiento humano a una mayor diversidad de escalas espaciales y temporales que los datos de telefonía móvil (Blanford *et al.*, 2015) (Figura 1).

Este artículo está dividido en cinco apartados. Después de esta introducción, el apartado 2 muestra las principales investigaciones que han usado datos de telefonía móvil o de Twitter para elaborar matrices OD a nivel nacional. El apartado 3 muestra el área de estudio, datos y metodología utilizados en esta investigación. Los resultados obtenidos son analizados en el apartado 4 para a continuación establecer una serie de conclusiones en el apartado 5.

Osorio Arjona, J. (2022). Big Data y matrices Origen-Destino: cartografía de flujos de movilidad en España a partir de datos de Twitter y comparación con datos de telefonía móvil, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 29, 115–130. <https://dx.doi.org/10.21138/GF.776>



**Figura 1. Escalas en las que actúan diferentes fuentes de datos para el estudio de la movilidad.**  
Fte. Blanford *et al.* (2015).

## 2. Estado del arte

En los últimos años se han publicado una serie de investigaciones que han empleado datos basados en Big Data para el diseño de matrices OD (Osorio-Arjona & García-Palomares, 2019b). La mayoría de estas investigaciones se han centrado en la escala urbana. Se pueden clasificar estos trabajos en función de la fuente de datos utilizada: datos de telefonía móvil (Alexander, Jiang, Murga, & González, 2015; Bonnel, Hombourger, Olteanu-Raimond, & Smoreda, 2015; Caceres, Wideberg, & Benitez, 2007; Louail *et al.*, 2015; Picornell *et al.*, 2015; Toole *et al.*, 2015) o datos de Twitter (Gao *et al.*, 2014; Lee, Goao, & Goulias, 2015; Osorio-Arjona & García-Palomares, 2019b; Perez, Dominguez, Rubiales, & Lotito, 2015; Salas-Olmedo & Rojas Quezada, 2017; Yin, Soliman, Yin, & Wang, 2017). Muchos trabajos siguen una metodología común de detección de lugares específicos de origen y destino para la construcción de matrices, expansión de los datos de las matrices mediante fuentes oficiales de datos como censos, y verificación de las matrices respecto a encuestas de transportes.

Es de especial interés el trabajo realizado por Lenormand *et al.* (2014), donde construyeron matrices OD en las áreas metropolitanas de Barcelona tanto con datos de telefonía móvil como con datos de Twitter, y compararon los resultados con una tercera matriz diseñada a partir de datos del Censo de 2011. Como resultados, obtuvieron que tanto los datos de telefonía móvil como los datos de Twitter producen patrones similares de densidad tanto en el espacio como en el tiempo, hallándose

Osorio Arjona, J. (2022). Big Data y matrices Origen-Destino: cartografía de flujos de movilidad en España a partir de datos de Twitter y comparación con datos de telefonía móvil, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 29, 115–130. <https://dx.doi.org/10.21138/GF.776>

---

las principales diferencias en la necesidad de obtener datos de Twitter en un periodo temporal mayor y en su mayor sesgo de edad al ser una red social usada principalmente por población joven.

Para esta investigación, son de interés los trabajos previos que han diseñado matrices OD a una escala regional o nacional. En primer lugar, se encuentran una serie de trabajos enfocados al estudio de la movilidad en Estonia mediante la construcción de matrices OD a partir de datos de telefonía móvil (Ahas, Silm, Järv, Saluveer, & Tiru, 2010; Puura, Silm, & Ahas, 2018; Silm & Ahas, 2010). El uso de datos de telefonía móvil permite además analizar diferentes flujos de movilidad dependiendo de si el idioma principal de la población de la muestra es el estonio o el ruso. Entre los resultados destaca la movilidad de la capital al campo o costa en verano, una mayor movilidad de la población joven o de género masculino, y una menor movilidad por parte de los usuarios cuyo idioma principal es el ruso. Kung *et al.* (2014) han usado datos de telefonía móvil para comparar los patrones de movilidad tanto a escala urbana (Boston en EE. UU.) como a escala nacional (Portugal, Costa de Marfil, Arabia Saudí), encontrando periodos diferentes de viaje trabajo-residencia en cada caso. Estas diferencias se deben a las diferencias en el uso de transporte, a la localización de los puestos de trabajo, y al periodo del día.

Por otra parte, Blanford *et al.* (2015) utilizaron datos de Twitter para estudiar la movilidad a diferentes escalas espaciales y temporales tanto dentro del país de Kenia como entre Kenia y los países colindantes. Los resultados validan Twitter como fuente alternativa para delinear patrones de movilidad humana dentro de un país y los flujos de movilidad en las fronteras. Yin *et al.* (2017) han buscado definir las fronteras urbanas en Reino Unido a partir de la movilidad de los usuarios de Twitter, encontrando como resultados que las distancias geográficas encontradas en los patrones de movilidad de los usuarios permiten establecer regiones geográficamente cohesivas.

Esta investigación se diferencia en trabajos previos en dos aspectos. Primero, este artículo busca comparar los resultados obtenidos por datos de Twitter y datos de telefonía móvil. Al contrario que artículos anteriores que verificaban las matrices OD construidas con datos de Twitter con fuente de datos tradicionales como encuestas de movilidad, este trabajo verifica las matrices diseñadas con datos de Twitter con las matrices que el Ministerio de Fomento ha construido a partir de datos de telefonía móvil, con el objetivo de ver las diferencias que ambos tipos de fuentes basadas en el Big Data recogen en la visualización de flujos de movilidad. Además, solo se consta de un trabajo previo que haya usado tanto datos de Twitter como de telefonía móvil (Lenormand *et al.*, 2014). El segundo aspecto radica en la escala de los datos. La mayor parte de los trabajos previos que ha analizado flujos de movilidad a partir de Big Data se ha centrado en las ciudades, por lo que hay muy pocas investigaciones que hayan usado Big Data para diseñar matrices OD a nivel de país. Combinando ambos aspectos, la originalidad de este artículo radica en la comparación a escala nacional de datos de Twitter y de telefonía móvil.

### 3. Área de estudio, datos, y metodología

#### 3.1. Área de estudio

Como se ha comentado previamente, el motivo por el que se ha seleccionado España como área de estudio parte de la originalidad que conlleva diseñar cartografía de matrices OD a escala

Osorio Arjona, J. (2022). Big Data y matrices Origen-Destino: cartografía de flujos de movilidad en España a partir de datos de Twitter y comparación con datos de telefonía móvil, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 29, 115–130. <https://dx.doi.org/10.21138/GF.776>

---

nacional. España está formada por 50 provincias (3 de ellas insulares) y 2 ciudades autónomas. Aproximadamente el 60 % de la población del país reside en provincias costeras, mientras que la provincia de Madrid acapara la mayor parte de la población del interior peninsular (conllevando como consecuencia un vacío demográfico en la mayoría de las provincias de las comunidades autónomas de Castilla-La Mancha, Castilla y León, Extremadura o Aragón). Esta distribución desigual de la población en el territorio le confiere de un interés añadido al estudio y visualización de los flujos de movilidad entre las diferentes provincias del país.

### 3.2. Datos utilizados

Para esta investigación se han empleado datos de Twitter publicados en España. Estos datos han sido descargados durante un periodo de 30 meses, desde marzo de 2016 hasta septiembre de 2018. Todos los *tweets* están geolocalizados y cuentan con una estampa temporal. En total se han descargado 12.691.289 *tweets* publicados por 640.234 usuarios.

Las matrices de viajes diseñadas con estos *tweets* fueron expandidas con datos de población a nivel provincial del año 2018 del Instituto Nacional de Estadística. Además, se han comparado los valores obtenidos en las matrices de viajes de Twitter con los datos de las dos matrices de viajes (correspondientes a los meses de julio y agosto, y al mes de octubre del año 2017) publicadas por el Ministerio de Fomento en el año 2018. Estas matrices fueron diseñadas a nivel provincial mediante datos de telefonía móvil procedentes de la red *Orange*. Los datos están disponibles de forma abierta en el apartado de estudio piloto de movilidad interprovincial de la página web *Observatorio del transporte y la logística en España* (<https://observatoriotransporte.mitma.es/>)

### 3.3. Metodología

La base de datos de *tweets* inicial fue incorporada a una base de datos de ArcGIS Pro, donde se realizaron una serie de procesos de filtrado de datos (Osorio-Arjona & García-Palomares, 2019b). Primero se eliminaron de la base de datos *tweets* cuyo identificador de usuario perteneciese a cuentas bot o robot. Estas cuentas pertenecen a usuarios compulsivos asociados a máquinas que generan mensajes automáticamente. Normalmente los bot suelen pertenecer a páginas de noticias o empresas que hacen publicidad de forma masiva y constante. Las cuentas bot fueron identificadas como cuentas con más de 1000 *tweets* publicados y con las mismas coordenadas en todos sus *tweets*, o cuentas con más de 10 *tweets* publicados con el mismo contenido semántico en el campo de texto de los *tweets*.

El siguiente proceso de limpieza consistió en eliminar usuarios cuyos *tweets* tienen siempre localizaciones muy similares. Es decir, se filtraron las cuentas de usuarios sin movilidad espacial, o con una movilidad espacial muy reducida. Para ello, se midieron las distancias entre los *tweets* de cada usuario, y se eliminaron los mensajes de aquellos usuarios con una distancia media menor a 50 metros en la localización de todos sus *tweets*. Después, se filtraron las cuentas de usuario con un periodo temporal reducido, eliminando aquellos usuarios que tienen todos sus mensajes concentrados en un periodo temporal de dos semanas seguidas. De esta forma, se filtran posibles visitantes o turistas, y se intenta asegurar trabajar con población residente. Finalmente, se filtraron los usuarios con muy baja actividad en Twitter, eliminando aquellos usuarios que hayan publicado menos de 5 *tweets* en total.

Osorio Arjona, J. (2022). Big Data y matrices Origen-Destino: cartografía de flujos de movilidad en España a partir de datos de Twitter y comparación con datos de telefonía móvil, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 29, 115–130. <https://dx.doi.org/10.21138/GF.776>

---

A continuación, se añadió información espacial a cada *tweet* de la provincia desde donde fue publicado. El siguiente paso fue asignar como provincia de residencia o de origen la provincia con un mayor número de *tweets* con *tweets* de cada usuario publicados en horario de 8 de la noche a 8 de la mañana. Esta metodología ha sido tradicionalmente usada en trabajos de movilidad usando datos de telefonía móvil, mediante los CDR o *call data records* (Ahas *et al.*, 2010; Alexander *et al.*, 2015; Picornell *et al.*, 2015).

En este estudio se busca visualizar todos los viajes interprovinciales detectados en la muestra de datos de Twitter, por lo que un usuario con una provincia de origen puede tener varias provincias de destino. Por lo tanto, los usuarios que han publicado todos sus mensajes en la misma provincia han sido eliminados de la muestra. El siguiente paso fue preparar las matrices OD utilizando los identificadores de usuario. Para este trabajo se han preparado tres matrices:

-Una matriz OD total con 572.757 *tweets* generados por 108.581 usuarios. En esta matriz se han utilizado todos los *tweets* de la muestra, independientemente del número de día o momento del año.

-Una matriz de *tweets* publicados en días laborables con 293.853 *tweets* de 90.948 usuarios. Aquí, se han seleccionado las provincias de destino a partir de *tweets* publicados en días laborables, filtrándose los publicados en fines de semana, días festivos, y meses de julio y agosto.

-Una matriz de *tweets* publicados en verano con 285.231 *tweets* de 82.974 usuarios. Para este caso, se han seleccionado los *tweets* publicados en los meses de julio y agosto, tanto en días laborables como fines de semana.

Se han escogido estos periodos de tiempo con el objeto de visualizar, por una parte, los principales patrones y flujos de movilidad en España a lo largo del periodo completo de la muestra, y, por otra parte, variaciones y principales lugares de destino durante periodo vacacional en contraposición con la matriz resultado de eliminar las anomalías que suele presentar la movilidad estival.

Las tres matrices se han diseñado a partir de relaciones entre las 52 provincias de España. Por último, estas tres matrices fueron expandidas a partir de datos de población a nivel provincial el Instituto Nacional de Estadística, empleando la siguiente fórmula:

$$T_{ij}^e = T_{ij} \cdot \frac{p_i}{\bar{p}_i}, \forall i \in N$$

donde  $\frac{p_i}{\bar{p}_i}$  son los pesos calculados para cada provincia  $N$ , basándose en el ratio entre el número total de habitantes de datos del padrón  $p_i$ , y las muestras de usuarios de Twitter  $\bar{p}_i$  obtenidos en la matriz. Estos pesos fueron multiplicados por el valor de los viajes de Twitter obtenidos en cada flujo  $T_{ij}$ , siendo el resultado los flujos expandidos  $T_{ij}^e$ .

Finalmente, las matrices diseñadas a partir de *tweets* publicados en días laborales y a partir de *tweets* publicados en verano fueron comparadas con las dos matrices OD diseñadas mediante datos

Osorio Arjona, J. (2022). Big Data y matrices Origen-Destino: cartografía de flujos de movilidad en España a partir de datos de Twitter y comparación con datos de telefonía móvil, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 29, 115–130. <https://dx.doi.org/10.21138/GF.776>

de telefonía móvil por el Ministerio de Fomento a partir del coeficiente de determinación  $R^2$  con el objeto de discernir el grado de aproximación que presentan ambas fuentes de datos.

#### 4. Descripción y análisis de resultados

##### 4.1. Distribución espacial y temporal de los usuarios de Twitter de la muestra

Los primeros resultados muestran como las provincias de Madrid y Barcelona (donde se encuentran las dos grandes aglomeraciones de España) son las dos provincias con mayor número de usuarios de Twitter residentes detectados en la muestra. A una segunda escala se encuentran las provincias más pobladas de Andalucía (Sevilla, Málaga, Cádiz y Granada), la Comunidad Valenciana, y las Islas Baleares como provincias con fuerte carácter residencial. Se puede destacar un mayor número de usuarios de Twitter en las provincias con una fuerte actividad turística o costera, mientras que las provincias de interior (salvo la excepción de la capital Madrid) cuentan con un número bajo de usuarios (Figura 2). El coeficiente  $R^2$  de correlación de usuarios de Twitter por provincia respecto a datos oficiales del padrón de población de 2018 del Instituto Nacional de Estadística presenta un valor de 0,87, indicando un ajuste muy próximo de la distribución de usuarios de Twitter a la realidad.



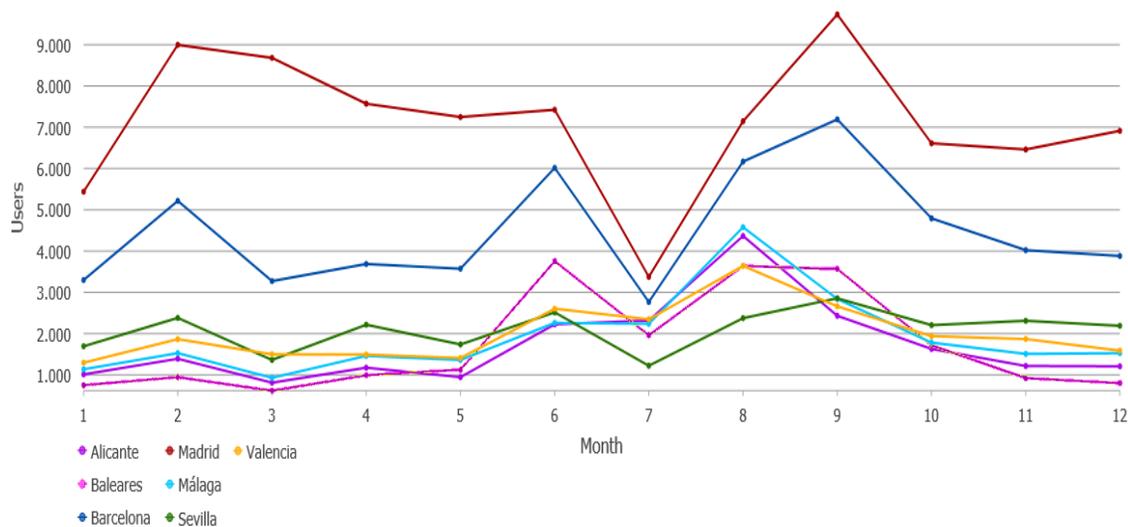
**Figura 2. Distribución espacial de los usuarios de Twitter de la muestra.**

Fte: Elaboración propia a partir de datos de Twitter.

Osorio Arjona, J. (2022). Big Data y matrices Origen-Destino: cartografía de flujos de movilidad en España a partir de datos de Twitter y comparación con datos de telefonía móvil, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 29, 115–130. <https://dx.doi.org/10.21138/GF.776>

Al elaborar un perfil temporal de la cantidad de usuarios en las provincias con mayor número de usuarios a lo largo del año, se puede visualizar un patrón muy similar en Madrid y Barcelona, con un número continuo de usuarios en primavera y en otoño, un decrecimiento destacado en los meses de enero y julio, y un crecimiento fuere en febrero y julio. Sin embargo, se pueden apreciar diferencias. Mientras que el número de usuarios en Madrid se mantiene continuo en junio, este número experimenta un importante crecimiento en Barcelona en junio. Otra diferencia radica en que el número de usuarios en Barcelona disminuye suavemente durante otoño, mientras que en Madrid crece ligeramente en el mes de diciembre.

En cuanto a las otras provincias destacadas, podemos visualizar como Baleares es la provincia con un menor número de usuarios de Twitter durante la mayor parte del año, pero cuenta con un fuerte crecimiento durante el verano que la convierte durante algunos meses en la tercera provincia de España con el mayor número de usuarios. En el caso contrario se encuentra Sevilla, la tercera provincia de España con más usuarios de Twitter durante casi todo el año, pero la provincia con menos usuarios durante el periodo estival. Málaga, Alicante y Valencia, ciudades pobladas y costeras, presentan una dinámica muy similar, con un número estable de usuarios a lo largo del año y un pico cumbre durante el verano (Figura 3).



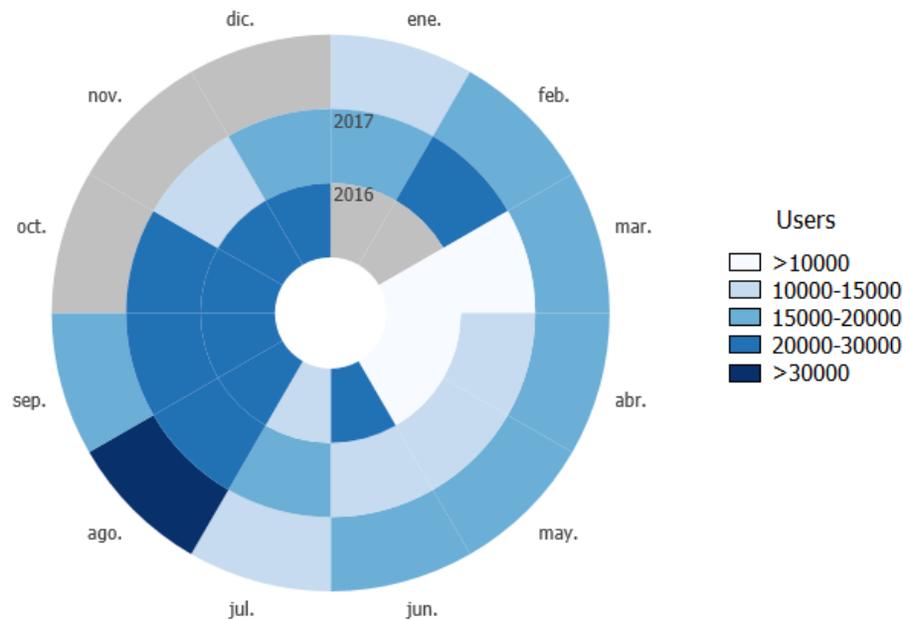
**Figura 3. Distribución mensual de usuarios de Twitter en las siete provincias con mayor número de usuarios de la muestra**

Fte. Elaboración propia a partir de datos de Twitter.

Agosto ha sido el mes con mayor número de usuarios de Twitter detectados, mientras que marzo ha sido el mes con un menor número de usuarios. Por tanto, se puede decir que hay un aumento de la actividad en Twitter en el periodo en el que los usuarios se desplazan a otros destinos del país de vacaciones. Además, se ha podido visualizar un mayor número de usuarios en 2016. Esto es debido a que las actualizaciones de las políticas de privacidad de Twitter llevadas a cabo en 2017 han

Osorio Arjona, J. (2022). Big Data y matrices Origen-Destino: cartografía de flujos de movilidad en España a partir de datos de Twitter y comparación con datos de telefonía móvil, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 29, 115–130. <https://dx.doi.org/10.21138/GF.776>

conllevado una disminución del número de *tweets* geolocalizados. Sin embargo, el mes de agosto del año 2018 fue el mes con mayor número de usuarios de toda la muestra (Figura 4).

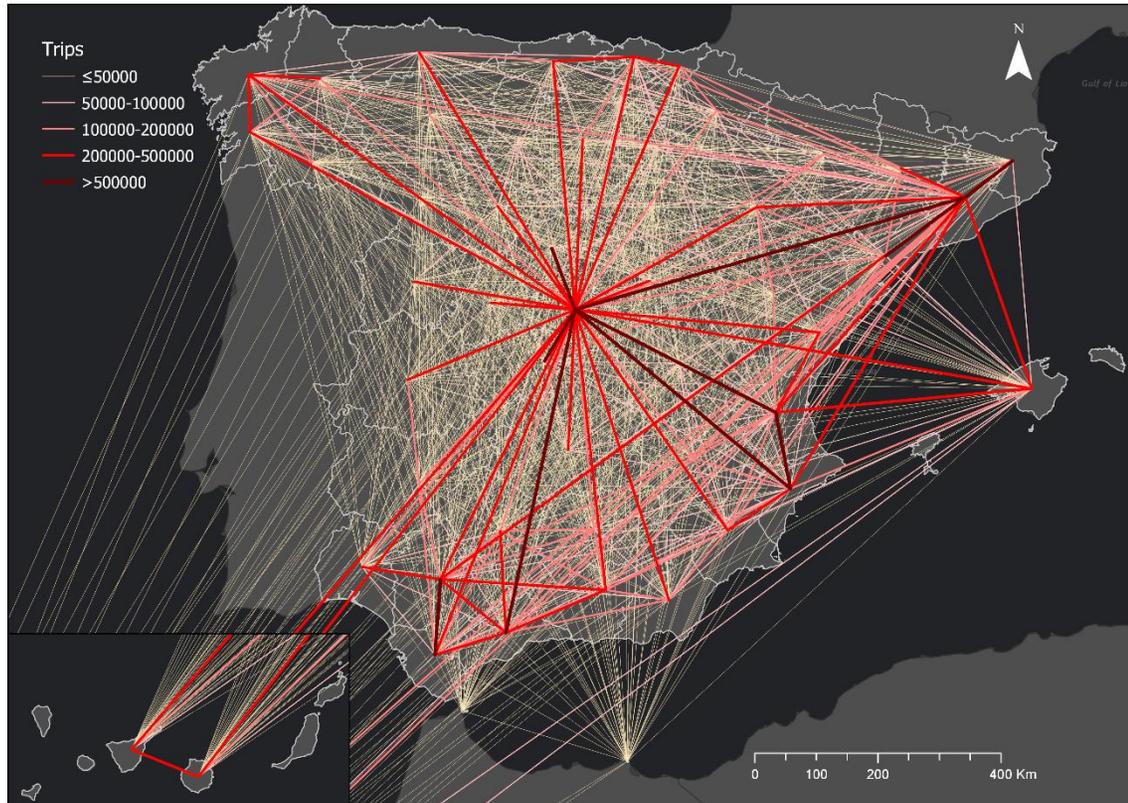


**Figura 4. Distribución temporal de los usuarios de Twitter de la muestra**  
Fte. Elaboración propia a partir de datos de Twitter.

#### 4.2. Visualización de matrices OD diseñadas a partir de usuarios de Twitter

La matriz OD de flujos de viajes obtenida durante todo el periodo temporal de la muestra indica como Madrid es la principal provincia de atracción de viajes. Los flujos con mayor número de usuarios se hallan entre Madrid y las provincias adyacentes de Toledo y Segovia, pero también entre Madrid y las provincias de Barcelona, Málaga, Valencia o Alicante. Mientras que se puede decir que Madrid es un polo de atracción de viajes a nivel nacional, Barcelona lo es a nivel regional, contando con altos flujos de movilidad con origen en las otras provincias de Cataluña. También destacan en cuanto a volumen los flujos de usuarios entre Valencia y Alicante, y entre Sevilla y Cádiz. A una segunda escala se pueden visualizar las conexiones de Madrid con las provincias costeras de casi toda España además de con las Islas Canarias, las conexiones triangulares entre Cataluña, comunidad Valenciana e Islas Baleares, y los flujos internos en Andalucía con Sevilla y Málaga como grandes focos (Figura 5).

Osorio Arjona, J. (2022). Big Data y matrices Origen-Destino: cartografía de flujos de movilidad en España a partir de datos de Twitter y comparación con datos de telefonía móvil, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 29, 115–130. <https://dx.doi.org/10.21138/GF.776>

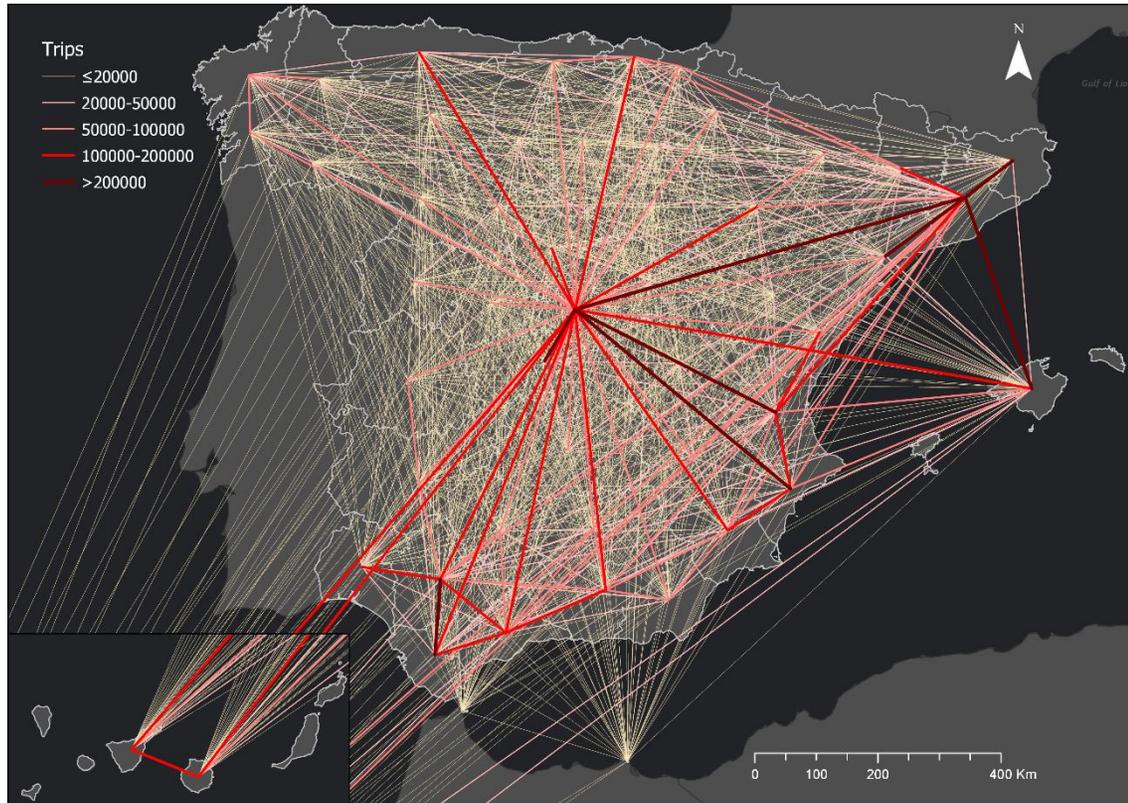


**Figura 5. Matriz de viajes OD a partir del total de tweets de la muestra**

Fte. Elaboración propia a partir de datos de Twitter.

Al visualizar la matriz OD obtenida a partir de datos publicados en días laborables, prevalece la conexión entre Madrid y Barcelona, pero también entre Madrid y la Comunidad Valenciana, y entre Barcelona y las Islas Baleares. En cuanto a provincias adyacentes, Madrid solo presenta un elevado número de viajes de usuarios respecto a Toledo, mientras que Barcelona sigue contando con flujos de más de 200.000 viajes con las provincias de Tarragona y de Girona. También cabe destacar una vez más la alta movilidad entre las provincias de Sevilla y Cádiz (Figura 6).

Osorio Arjona, J. (2022). Big Data y matrices Origen-Destino: cartografía de flujos de movilidad en España a partir de datos de Twitter y comparación con datos de telefonía móvil, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 29, 115–130. <https://dx.doi.org/10.21138/GF.776>

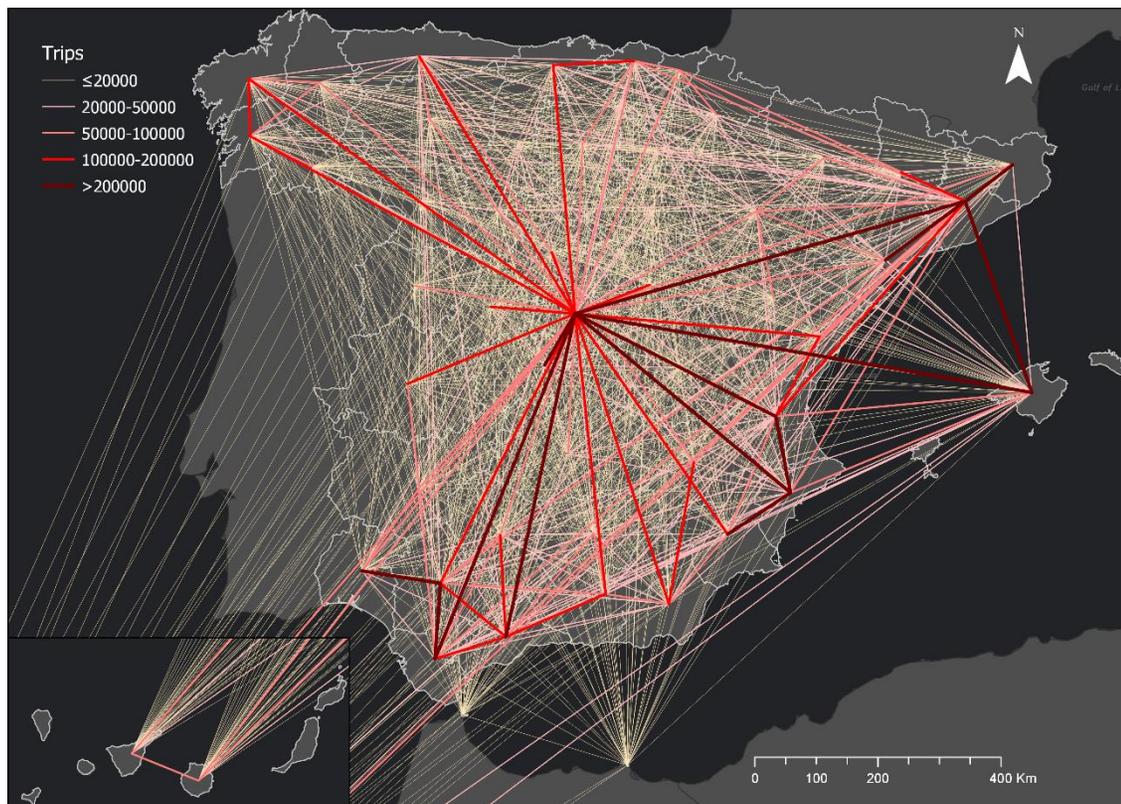


**Figura 6. Matriz de viajes OD a partir de los tweets publicados en días laborables**

Fte. Elaboración propia a partir de datos de Twitter.

Mientras que la matriz de viajes originados durante días laborables presentaba un bajo número de conexiones con más de 200.000 viajes, durante el verano el número de flujos con esa cifra de usuarios aumenta sustancialmente. La principal novedad respecto a la anterior matriz es el elevado número de viajes entre Madrid y las Islas Baleares o las provincias costeras de Andalucía (Cádiz y Málaga). También se puede visualizar un crecimiento del número de viajes entre Murcia y Alicante, o entre Sevilla y Huelva. Aunque por lo general el número de viajes en muchos flujos aumenta, este número sorprendentemente disminuye entre la Península Ibérica y las Islas Canarias (Figura 7).

Osorio Arjona, J. (2022). Big Data y matrices Origen-Destino: cartografía de flujos de movilidad en España a partir de datos de Twitter y comparación con datos de telefonía móvil, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 29, 115–130. <https://dx.doi.org/10.21138/GF.776>



**Figura 7. Matriz de viajes OD a partir de los tweets publicados en días de verano**

Fte. Elaboración propia a partir de datos de Twitter.

#### 4.3. Comparación con los datos de telefonía móvil del Ministerio de Fomento

Al validar los resultados obtenidos a partir de Twitter con las matrices OD publicadas por el Ministerio de Fomento a partir de datos de telefonía móvil, se obtienen bajos valores de autocorrelación. El coeficiente  $R^2$  entre la matriz de tweets publicados en días laborables y la matriz de movilidad en octubre del Ministerio de Fomento es de 0,36, mientras que el coeficiente entre la matriz de tweets generados en verano y la matriz de movilidad en julio y agosto del Ministerio de Fomento es de 0,4. Estos resultados se deben a la diferencia de la penetración entre datos de Twitter y datos de telefonía móvil causada por los sesgos geográficos y de edad (hay un mayor uso de Twitter en las grandes ciudades, donde se concentra la población entre 20 y 39 años).

Para observar cómo afecta esta diferencia a la distribución de flujos de viajes, se han registrado las 20 conexiones interprovinciales con un mayor número de viajes en cada matriz. Mientras que en las matrices generadas por datos de Twitter destacan los viajes entre Madrid y Barcelona, Valencia o Málaga, en las encuestas generadas a partir de datos de telefonía móvil priman los viajes entre provincias limítrofes, destacando las conexiones entre Madrid y Toledo o Guadalajara, Murcia y Alicante, o Vizcaya y Cantabria, pasando los flujos entre las grandes ciudades españolas a un segundo plano. Cabe destacar que tanto Twitter como los datos de telefonía móvil cuentan entre sus principales

Osorio Arjona, J. (2022). Big Data y matrices Origen-Destino: cartografía de flujos de movilidad en España a partir de datos de Twitter y comparación con datos de telefonía móvil, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 29, 115–130. <https://dx.doi.org/10.21138/GF.776>

rutas los flujos entre Barcelona y las otras provincias de Cataluña, y la conexión entre Sevilla y Cádiz (Tabla 1).

**Tabla 1. Diez rutas con mayor número de viajes observados en cada matriz**

Matriz a partir de todos los <i>tweets</i>	Matriz a partir de <i>tweets</i> en días laborables	Matriz del Ministerio de Fomento de octubre	Matriz a partir de <i>tweets</i> en verano	Matriz del Ministerio de Fomento de julio
Barcelona-Madrid	Barcelona-Madrid	Madrid-Toledo	Barcelona-Girona	Madrid-Toledo
Barcelona-Girona	Madrid-Barcelona	Toledo-Madrid	Madrid-Alicante	Toledo-Madrid
Madrid-Barcelona	Barcelona-Girona	Girona-Barcelona	Madrid-Valencia	Barcelona-Girona
Madrid-Valencia	Valencia-Madrid	Tarragona-Barcelona	Sevilla-Cádiz	Girona-Barcelona
Madrid-Toledo	Barcelona-Tarragona	Barcelona-Tarragona	Barcelona-Tarragona	Barcelona-Tarragona
Barcelona-Tarragona	Madrid-Valencia	Guadalajara-Madrid	Madrid-Barcelona	Tarragona-Barcelona
Madrid-Alicante	Madrid-Toledo	Murcia-Alicante	Madrid-Málaga	Alicante-Murcia
Valencia-Madrid	Girona-Barcelona	Alicante-Murcia	Valencia-Alicante	Murcia-Alicante
Sevilla-Cádiz	Alicante-Madrid	Madrid-Guadalajara	Madrid-Cádiz	Madrid-Guadalajara
Madrid-Segovia	Sevilla-Cádiz	Alicante-Valencia	Barcelona-Madrid	Guadalajara-Madrid

Fte. Elaboración propia a partir de datos de Twitter y del Ministerio de Fomento.

## 5. Conclusiones

El Big Data geolocalizado está revolucionando los estudios de movilidad gracias a la facilidad de obtener enormes cantidades de datos actualizados, en muchos casos a bajo coste. Recientemente organismos públicos como el Ministerio de Fomento de España están empleando datos basados en las nuevas tecnologías para elaborar estudios de movilidad con los que poder ofrecer rápidamente resultados actualizados. Los datos de telefonía móvil son los datos masivos más utilizados en este tipo de estudios ya que permiten el acceso de una muestra gigante de datos en muy poco tiempo (Chen *et al.*, 2016). Sin embargo, los datos de telefonía móvil no suelen estar basados en coordenadas GPS, sino que su ubicación espacial depende de la ubicación de las antenas de telefonía móvil.

Osorio Arjona, J. (2022). Big Data y matrices Origen-Destino: cartografía de flujos de movilidad en España a partir de datos de Twitter y comparación con datos de telefonía móvil, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 29, 115–130. <https://dx.doi.org/10.21138/GF.776>

---

En este artículo se ha trabajado con datos de Twitter para obtener matrices de viajes como alternativa a los datos de telefonía móvil. Al estar basados en coordenadas GPS, los datos de Twitter presentan un mayor detalle espacial que los datos de telefonía móvil. Además, la descarga de estos datos es gratuita y facilita una actualización constante de los datos. Sin embargo, hay que tener en cuenta que el tamaño de las muestras y la resolución temporal es menor al que presentan los datos de telefonía móvil.

Bajo este panorama, en este artículo se han diseñado matrices OD a escala nacional. La mayoría de los estudios previos de movilidad basados en Big Data se enfocan en la escala urbana. En este artículo se ha querido aprovechar las matrices OD realizadas a nivel nacional por el Ministerio de Fomento a partir de datos de telefonía móvil, con el objetivo de comparar en esta escala los resultados obtenidos al usar dos fuentes de datos basadas en el Big Data, pero con propiedades diferentes. Como resultados, se ha visualizado que la distribución de usuarios residentes obtenida en la muestra de Twitter es similar a los datos oficiales de población, con una concentración mayor de usuarios en la Comunidad de Madrid y en las provincias costeras del Mar Mediterráneo, mientras que la distribución de usuarios es baja en las provincias de interior. Al añadir la variante temporal a la distribución de la muestra, se visualiza un mayor número de usuarios durante el verano, y un menor número en invierno.

Analizando las matrices OD diseñadas a partir de datos de Twitter, se ha visualizado la Comunidad de Madrid como polo principal de atracción de viajes a escala nacional, mientras que la provincia de Barcelona ha destacado como importante foco regional de atracción de viajes. También caben destacar los flujos de viajes entre Valencia y Alicante, y entre Sevilla y Cádiz. Se observa también el aumento del número de viajes y una mayor diversidad de rutas en la matriz elaborada en los meses de verano respecto a la matriz construida durante el periodo laboral, ratificando los resultados previos que mostraban el mes de agosto como principal periodo de actividad en Twitter.

Al comparar los resultados obtenidos con las matrices de viajes de datos de telefonía móvil obtenidos a partir del Ministerio de Fomento, se observa una baja autocorrelación y una diferencia en la distribución espacial de las rutas con mayor número de viajes. Mientras que con los datos de Twitter se pueden observar con más detalle los viajes entre las áreas metropolitanas y los principales lugares de turismo, los datos de telefonía móvil recopilan mejor los flujos de viaje entre provincias adyacentes. Hay que tener en cuenta que la geolocalización de los datos de telefonía móvil es más precisa en las áreas urbanas y que Twitter es una red social usada principalmente por población joven residente en ciudades, mientras que los datos de telefonía pueden recoger mejor la actividad en zonas rurales y pueblos con menor población, pero dependiendo igualmente de la ubicación de los datos de telefonía móvil. Aun así, los datos de telefonía móvil también presentan sesgos parecidos a los datos de Twitter en cuanto a edad y uso urbano.

La metodología empleada en esta investigación por tanto corrobora los datos de Twitter como una fuente de datos eficaz para el estudio de la movilidad a escala nacional debido a su fácil accesibilidad, su alta resolución tanto espacial y temporal, la posibilidad de obtener continuamente datos actualizados, y su aproximación a los patrones de movilidad que ocurren en la realidad. Sin embargo, esta metodología presenta limitaciones a tener en cuenta. Hay una serie de variables necesarias para el estudio del análisis de la movilidad y su aplicación en las políticas públicas que no se pueden obtener a partir del diseño de matrices OD mediante datos de Twitter, principalmente el modo de transporte utilizado para realizar el viaje, y características socioeconómicas como el género,

Osorio Arjona, J. (2022). Big Data y matrices Origen-Destino: cartografía de flujos de movilidad en España a partir de datos de Twitter y comparación con datos de telefonía móvil, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 29, 115–130. <https://dx.doi.org/10.21138/GF.776>

---

edad o nivel de renta que si es posible obtener a partir de datos de telefonía móvil. Por tanto, es aconsejable aprovechar las fortalezas tanto de los datos de Twitter como de los datos de telefonía móvil para obtener datos con una alta resolución espacial, temporal, y con la información necesaria para el diseño de planes de movilidad aprovechando el alto volumen y la alta velocidad a la que se obtienen estos datos. Aunque los datos de Twitter tampoco cuentan con el motivo del viaje, es posible discernir el motivo a partir de la combinación de la ubicación espacial de los *tweets* con información del uso del suelo de parcelas catastrales.

Este trabajo cuenta con una gama de interesantes futuras líneas de investigación. Por una parte, las matrices OD construidas con datos de telefonía móvil presentan información complementaria sobre el modo de viaje realizado, como el viaje en red ferroviaria o a través de la red de carreteras. Una futura línea de investigación consiste en entrar en más detalle acerca del modo de viaje realizado a partir de los datos de Twitter a partir de la estimación de rutas. Otra línea de trabajo consiste en combinar los datos de Twitter con datos de uso del suelo con el objeto de añadir información sobre el motivo de viaje. En este trabajo además se han tratado los viajes interprovinciales, descartando los flujos de movilidad que tienen origen y destino dentro de una misma provincia. Otra futura línea de investigación consiste en incluir estos viajes interprovinciales a la matriz OD de viajes diseñada a una escala nacional. Finalmente, en este artículo se han trabajado con escalas temporales agregadas con el objetivo de tener visualizar más claramente los volúmenes que presentan los flujos de movilidad y los patrones característicos de la movilidad laboral y la movilidad estacional. Sin embargo, es posible realizar análisis de movilidad a escala diaria y discernir flujos esporádicos.

## Referencias bibliográficas

- Ahas, R., Silm, S., Järv, O., Saluveer, E., & Tiru, M. (2010). Using mobile positioning data to model locations meaningful to users of mobile phones. *Journal of Urban Technology*, 17(1), 3–27. <https://doi.org/10.1080/10630731003597306>.
- Alexander, L., Jiang, S., Murga, M., & González, M. C. (2015). Origin-destination trips by purpose and time of day inferred from mobile phone data. *Transportation Research Part C: Emerging Technologies*, 58, 240–250. <https://doi.org/10.1016/j.trc.2015.02.018>.
- Blanford, J. I., Huang, Z., Savelyev, A., & MacEachren, A. M. (2015). Geo-Located Tweets. Enhancing Mobility Maps and Capturing Cross-Border Movement. *PLOS ONE*, 10(6), e0129202. <https://doi.org/10.1371/journal.pone.0129202>.
- Bonnel, P., Hombourger, E., Olteanu-Raimond, A.-M., & Smoreda, Z. (2015). Passive Mobile Phone Dataset to Construct Origin-destination Matrix: Potentials and Limitations. *Transportation Research Procedia*, 11, 381–398. <https://doi.org/10.1016/j.trpro.2015.12.032>.
- Caceres, N., Wideberg, J. P., & Benitez, F. G. (2007). Deriving origin–destination data from a mobile phone network. *IET Intelligent Transport Systems*, 1(1), 15. <https://doi.org/10.1049/iet-its:20060020>.

Osorio Arjona, J. (2022). Big Data y matrices Origen-Destino: cartografía de flujos de movilidad en España a partir de datos de Twitter y comparación con datos de telefonía móvil, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 29, 115–130. <https://dx.doi.org/10.21138/GF.776>

---

Chen, C., Ma, J., Susilo, Y., Liu, Y., & Wang, M. (2016). The promises of big data and small data for travel behavior (aka human mobility) analysis. *Transportation Research Part C: Emerging Technologies*, 68, 285–299. <https://doi.org/10.1016/j.trc.2016.04.005>.

Demissie, M. G., Phithakkitnukoon, S., & Kattan, L. (2019). Trip Distribution Modeling Using Mobile Phone Data: Emphasis on Intra-Zonal Trips. *IEEE Transactions on Intelligent Transportation Systems*, 20(7). <https://doi.org/10.1109/TITS.2018.2868468>.

Gao, S., Yang, J., Yan, B., Hu, Y., Janowicz, K., & McKenzie, G. (2014). Detecting Origin-Destination Mobility Flows From Geotagged Tweets in Greater Los Angeles Area. *Eight International Conference on Geographic Information Science (GIScience'14)*.

Gutiérrez-Puebla, J., Benitez, C., García-Palomares, J. C., Romanillos, G., Rubinstein, E., Leño, J. M., ... Condeço-Melhorado, A. M. (2020). *Cómo aplicar Big Data en la planificación del transporte: el uso de datos de GPS en el análisis de la movilidad urbana*. (C. Benitez, Ed.). Banco Interamericano de Desarrollo.

Gutiérrez-Puebla, J., Benitez, C., Leño, J. M., García-Palomares, J. C., Condeço-Melhorado, A., Mojica, C., ... Romanillos, G. (2019). *Cómo aplicar Big Data en la planificación del transporte urbano. El uso de datos de telefonía móvil en el análisis de la movilidad*. (C. Benitez, Ed.). Banco Interamericano de Desarrollo.

Gutiérrez, J., & García-Palomares, J. C. (2007). New spatial patterns of mobility within the metropolitan area of Madrid: Towards more complex and dispersed flow networks. *Journal of Transport Geography*, 15(1), 18–30. <https://doi.org/10.1016/j.jtrangeo.2006.01.002>.

Gutiérrez-Puebla, J., García-Palomares, J. C., & Salas-Olmedo, M. H. (2016). Big (Geo) Data en Ciencias Sociales: Retos y Oportunidades. *Revista de Estudios Andaluces*, 33(331), 1–23. <https://doi.org/10.12795/rea.2016.i33.0>.

Gutiérrez Puebla, J. (2018). Big Data y nuevas geografías: la huella digital de las actividades humanas. *Documents d'Anàlisi Geogràfica*, 64(2), 195. <https://doi.org/10.5565/rev/dag.526>.

Hawelka, B., Sitko, I., Beinat, E., Sobolevsky, S., Kazakopoulos, P., & Ratti, C. (2014). Geo-located Twitter as proxy for global mobility patterns. *Cartography and Geographic Information Science*, 41(3), 260–271. <https://doi.org/10.1080/15230406.2014.890072>.

Kung, K. S., Greco, K., Sobolevsky, S., & Ratti, C. (2014). Exploring Universal Patterns in Human Home-Work Commuting from Mobile Phone Data. *PLoS ONE*, 9(6), 1–15. <https://doi.org/10.1371/journal.pone.0096180>.

Lee, J. H., Goao, S., & Goulias, K. G. (2015). Can Twitter data be used to validate travel demand models? *GEOTRANS Report 2015-5-03*, 1–27.

Lenormand, M., Picornell, M., Cantú-Ros, O. G., Tugores, A., Louail, T., Herranz, R., ... Ramasco, J. J. (2014). Cross-Checking Different Sources of Mobility Information. *PLoS ONE*, 9(8), 1–10. <https://doi.org/10.1371/journal.pone.0105184>.

Osorio Arjona, J. (2022). Big Data y matrices Origen-Destino: cartografía de flujos de movilidad en España a partir de datos de Twitter y comparación con datos de telefonía móvil, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 29, 115–130. <https://dx.doi.org/10.21138/GF.776>

---

Louail, T., Lenormand, M., Picornell, M., García Cantú, O., Herranz, R., Frias-Martinez, E., ... Barthelemy, M. (2015). Uncovering the spatial structure of mobility networks. *Nature Communications*, 6(1), 1–8. <https://doi.org/10.1038/ncomms7007>

Osorio-Arjona, J., & García-Palomares, J. C. (2017). Nuevas fuentes y retos para el estudio de la movilidad urbana. *Cuadernos Geográficos*, 56(3), 247–267. <https://doi.org/10.30827/CUADGEO.V56I3.5352>.

Osorio-Arjona, J., & García-Palomares, J. C. (2019a). Big Data y universidades: análisis de movilidad de los estudiantes universitarios a partir de datos de Twitter. *GeoFocus Revista Internacional de Ciencia y Tecnología de La Información Geográfica*, 24(9), 37–57. <https://doi.org/10.21138/GF.648>

Osorio-Arjona, J., & García-Palomares, J. C. (2019b). Social media and urban mobility: Using twitter to calculate home-work travel matrices. *Cities*, 89, 268–280. <https://doi.org/10.1016/j.cities.2019.03.006>.

Perez, A. J., Dominguez, L. D., Rubiales, A. J., & Lotito, P. A. (2015). Optimización de matrices origen-destino estimadas a partir de datos georeferenciados en redes sociales. *13º Simposio Argentino de Investigación Operativa*, 47–56.

Picornell, M., Ruiz, T., Lenormand, M., Ramasco, J. J., Dubernet, T., & Frías-Martínez, E. (2015). Exploring the potential of phone call data to characterize the relationship between social network and travel behavior. *Transportation*, 42(4), 647–668. <https://doi.org/10.1007/s11116-015-9594-1>.

Puura, A., Silm, S., & Ahas, R. (2018). The Relationship between Social Networks and Spatial Mobility: A Mobile-Phone-Based Study in Estonia. *Journal of Urban Technology*, 25(2), 7-25. <https://doi.org/10.1080/10630732.2017.1406253>

Salas-Olmedo, M. H., & Rojas Quezada, C. (2017). The use of public spaces in a medium-sized city: from Twitter data to mobility patterns. *Journal of Maps*, 13(1), 40–45. <https://doi.org/10.1080/17445647.2017.1305302>.

Silm, S., & Ahas, R. (2010). The seasonal variability of population in estonian municipalities. *Environment and Planning A*, 42(10), 2527–2546. <https://doi.org/10.1068/a43139>.

Toole, J. L., Colak, S., Sturt, B., Alexander, L. P., Evsukoff, A., & González, M. C. (2015). The path most traveled: Travel demand estimation using big data resources. *Transportation Research Part C: Emerging Technologies*, 58, 162–177. <https://doi.org/10.1016/j.trc.2015.04.022>.

Yin, J., Soliman, A., Yin, D., & Wang, S. (2017). Depicting urban boundaries from a mobility network of spatial interactions: a case study of Great Britain with geo-located Twitter data. *International Journal of Geographical Information Science*, 31(7), 1293–1313. <https://doi.org/10.1080/13658816.2017.1282615>.