# A global comparison of bicycle sharing systems

James Todd [*], Oliver O'Brien, James Cheshire

*UCL Department of Geography, Gower Street, London WC1E 6BT, United Kingdom*

## ARTICLE INFO

## ABSTRACT

Increasing urban populations have created pressures on the transportation networks that serve them. Bicycle sharing systems (BSS) have seen a dramatic increase in popularity as cities around the world begin to implement and see significant use and benefit from this growing mode of urban micro-mobility. As a result, the research surrounding bicycle sharing systems has also increased, although this has been primarily focused on the analysis of individual systems. There is therefore a need for a global comparison of systems, particularly given that prior research often omits China, which is currently the largest BSS market in the world. This paper therefore marks a major step forward through its analysis of data from 322 schemes situated on all major continents. Conducting such analysis, there appear to be 5 main types of BSS: *very large, high use BSS, large BSS in major cities, medium BSS with extensive cycling infrastructure, small to medium efficient BSS* and *small to medium inefficient BSS*. From these major cluster groups, we are able to group schemes by usage, contextual indicators and the behavioural characterics of their users. This not only facilitates a global comparison of scheme performance, but also offers a basis to new schemes to identify established BSS with similar characteristics that can be used as a template for anticipating the likely demand from users.

## 1. Introduction

The number of Bicycle Sharing Systems (BSS) has seen a dramatic increase in recent years rising from 13 systems in 2004 to over 2000 in 2019 (Bicycle Sharing Blog, 2019), making it one of the fastest growing modes of transportation. The docked BSS we find in many cities – such as Santander Cycles in London and Paris's Vélib' System – are known as "third generation" schemes (Corcoran et al., 2014; Médard de Chardon and Caruso, 2015; Médard de Chardon et al., 2017; DeMaio and Gifford, 2004; ter Beek et al., 2014) and can be characterised as having customised bicycles with automated docking stations secured by user payment card details. They are typically accompanied by websites that provide data on the occupancy rates of each docking station (Médard de Chardon and Caruso, 2015; Davis, 2014).

In the decade since 2010, when the London BSS commenced operation, the global urban population increased from 3.6 billion to 4.3 billion (Worldbank, 2021). With urban growth comes increased levels of congestion (Çolak et al., 2016) that cities need to manage across both public and private transportation modes. As technologies around BSS have improved and become cheaper, they have become an attractive solution to help improve both congestion levels and also health as cycling is a form of active transport that reduces emissions (Fishman et al., 2013; O'Brien et al., 2014; Pucher and Buehler, 2012; Shaheen et al., 2010; Woodcock et al., 2014; Zhang et al., 2019). In a study conducted by Yang et al. (2019), they find evidence to show a change in travel behaviour of BSS users near metro stations, which exemplifies the uptake of BSS as a mode to help overcome issues around the 'last mile problem'. This refers to the difficulty that commuters encounter when travelling from a transportation hub, such as a train or bus station, to their final destination.

User-focused research has shown that weekday journeys during peak hours may actually be faster using BSS services compared to taxi journeys (Faghih-Imani et al., 2017). Similarly, with a careful integration with existing public transportation networks, BSS has been found to reduce overall travel times as well as increase use of these more environmentally friendly modes of public transportation as a whole (Jäppinen et al., 2013; Zhang et al., 2015). In addition, the benefits of cycling can outweigh disbenefits of air pollution exposure (Mueller et al., 2015) as well as provide a safer means to travel, in comparison to private cyclists (Fishman and Schepers, 2016).

Whilst the literature on individual BSS is extensive there have been few comparisons of schemes globally. This can be attributed to the lack of routine data releases from BSS operators (Mátrai and Tóth, 2016) and has meant there is currently no established standard for the comparison

---

**Table 1**

A summary of some of variables which influence the use of BSS, extracted from the papers discussed.

| Variable | | Relationship | Paper |
|---|---|---|---|
| BSS Factors | Number of Stations | Positive | Faghih-Imani et al. (2014), Médard de Chardon et al. (2017) |
| | Station Density | Positive | Médard de Chardon et al. (2017) |
| | Station Capacity | Positive | Tran et al. (2015), O'Brien et al. (2014) |
| | Distance to Station | Negative | Tran et al. (2015), Bachand-Marleau et al. (2012) |
| Socio-Demographic | Population | Positive | Médard de Chardon et al. (2017), Faghih-Imani et al. (2014), Tran et al. (2015) |
| | Income | Positive Negative | Fishman et al. (2014), Roy et al. (2019), Woodcock et al. (2014), Bachand-Marleau et al. (2012) |
| | Age | Positive | Fishman et al. (2014), Zhang et al. (2016) |
| | Gender | Male | Fishman et al. (2014), Zhang et al. (2016), Goodman and Cheshire (2014), Murphy and Usher (2015) |
| | Jobs | Positive | Tran et al. (2015), Woodcock et al. (2014) |
| | Education | Positive | Fishman et al. (2014), Shaheen et al. (2013) |
| | Ethnicity (White) | Positive | Buck et al. (2013) |
| Weather/Climate | Rainfall | Negative | Corcoran et al. (2014), Miranda-Moreno and Nosal (2011) |
| | Windspeed | Negative Insignificant | Corcoran et al. (2014), Miranda-Moreno and Nosal (2011) |
| | Air Pollution | Negative | Campbell et al. (2016) |
| | Temperature | Insignificant Positive | Corcoran et al. (2014), Faghih-Imani et al. (2014) |
| | Humidity | Negative | Faghih-Imani et al. (2014) |
| Topography | Slope | Negative | Frade and Ribeiro (2014), Mateo-Babiano et al. (2016) |
| | Altitude | Negative | Tran et al. (2015) |
| Cycling Infrastructure | Cycling Infrastructure | Positive | Fishman et al. (2014), Faghih-Imani et al. (2014), Buck and Buehler (2012), Mateo-Babiano et al. (2016) |
| Other | Public/School Holidays | Insignificant | Corcoran et al. (2014), Brandenburg et al. (2007), Borgnat et al. (2011) |
| | Helmet Requirement | Negative | Médard de Chardon et al. (2017), Fishman et al. (2014), O'Brien et al. (2014) |
| Temporality | Weekday/Weekend | – | Faghih-Imani et al. (2014), Faghih-Imani et al. (2017), O'Brien et al. (2014), Zaltz Austwick et al. (2013) |
| | Season | – | Ahmed et al. (2010), Faghih-Imani et al. (2017) |
| | Time of day | – | Faghih-Imani et al. (2017), O'Brien et al. (2014) |

of BSS (Médard de Chardon and Caruso, 2015). One important consequence of this is that it allows operators to overstate the success of their BSS in comparison to others in search of benefits such as increased investment (Médard de Chardon and Caruso, 2015). This paper therefore provides a timely update and extension of the global analysis conducted in O'Brien et al. (2014) and delivers a comprehensive comparison of an unprecedented number of docked BSS around the world.

Utilising a very large collection of BSS data, this research conducts a two-staged clustering analysis to help gain insights into the global landscape of docked BSS, their evolution, and individual system successes and failures. To achieve this, it was first necessary to establish comparative metrics across all systems in the dataset through the manipulation of dock capacity data. These metrics – detailed in Section 4.1 – provide the foundations to the global classification created through a combination of K-means and Dynamic Time Warping (DTW) clustering.

## 2. Literature review

The body of BSS literature has been comprehensively detailed in Fishman (2016) and Fishman (2019), so our focus in this literature review is on those studies that offer comparisons between BSS (Médard de Chardon et al., 2017; O'Brien et al., 2014; Kou and Cai, 2019; Sarkar et al., 2015; Zaltz Austwick et al., 2013; Zhang et al., 2015), as well as variables that influence usage across systems. The variables that are found to be influential (detailed in Section 2.1) have helped to inform decisions on the counterfactual variables to underpin the classification process detailed in Section 4.2.

### 2.1. Variables of influence

When assessing the factors that can influence the use of BSS, it makes sense to start with the attributes of BSS themselves. As Table 1 shows, the size of a system appears to play a large role in influencing users' propensity to use it as well as influence the absolute number of riders (O'Brien et al., 2014). We also find that the larger the system and higher the density of docking stations, the more a BSS is used (Médard de Chardon et al., 2017; O'Brien et al., 2014). There is also a growing body

of literature surrounding users' propensity to use systems based on factors such as their distance from the nearest docking station. Such insights are usually garnered through survey methods and can shed light on how far users are willing to travel in order to use BSS (Guerra et al., 2012; Bachand-Marleau et al., 2012). For example, Bachand-Marleau et al. (2012) found that the maximum distance a user is willing to walk to the nearest BSS station is around 500 m. Similarly, Gu et al. (2019) used 500 m as the 'acceptable walking distance' for commuters using BSS in combination with public transportation. Therefore, 500 m has been adopted as the standardised distance for the circle of operation around each docking station, which has been used to calculate the operational area for each BSS in this paper. A fixed distance is used in order to maintain metric homogeneity across BSS.

Another important determinant in the use of BSS are the characteristics of the population in close proximity to them. Fishman (2016) shows that higher population density and higher income areas have more active BSS use (Faghih-Imani et al., 2014; Fishman et al., 2014; Woodcock et al., 2014). Research by Bachand-Marleau et al. (2012) and Buck et al. (2013) have found that users' income levels have a mixed level of influence both within BSS as well as between BSS, with evidence showing that in Montreal, Canada, those users earning less than Can $40,000 were 32% less likely to use BSS, whereas those users in Washington D.C., United States, were typically of lower income in comparison to private cyclists. BSS is more likely have a closer gender balance although still with a male bias (Pucher et al., 2011). Research on the ethnicity and education of BSS users in western cities shows higher use rates in urban areas with a higher proportion of white people as well as more educated populations (Fishman et al., 2014; Martin and Shaheen, 2014; Buck et al., 2013). The schemes have often favoured more affluent areas in the city but Goodman and Cheshire (2014) found that residents from more deprived areas use BSS if these are built in their local areas and remain affordable relative to other modes.

Weather and topography have been found to play a significant role in the use patterns within BSS and seasonal impacts are clear from the number of trips (Ahmed et al., 2010), with areas/times with high levels of average rainfall seeing a decrease in the use of BSS (Campbell et al., 2016; Corcoran et al., 2014; Miranda-Moreno and Nosal, 2011; Nankervis, 1999). Temperature also appears to play a significant role in the
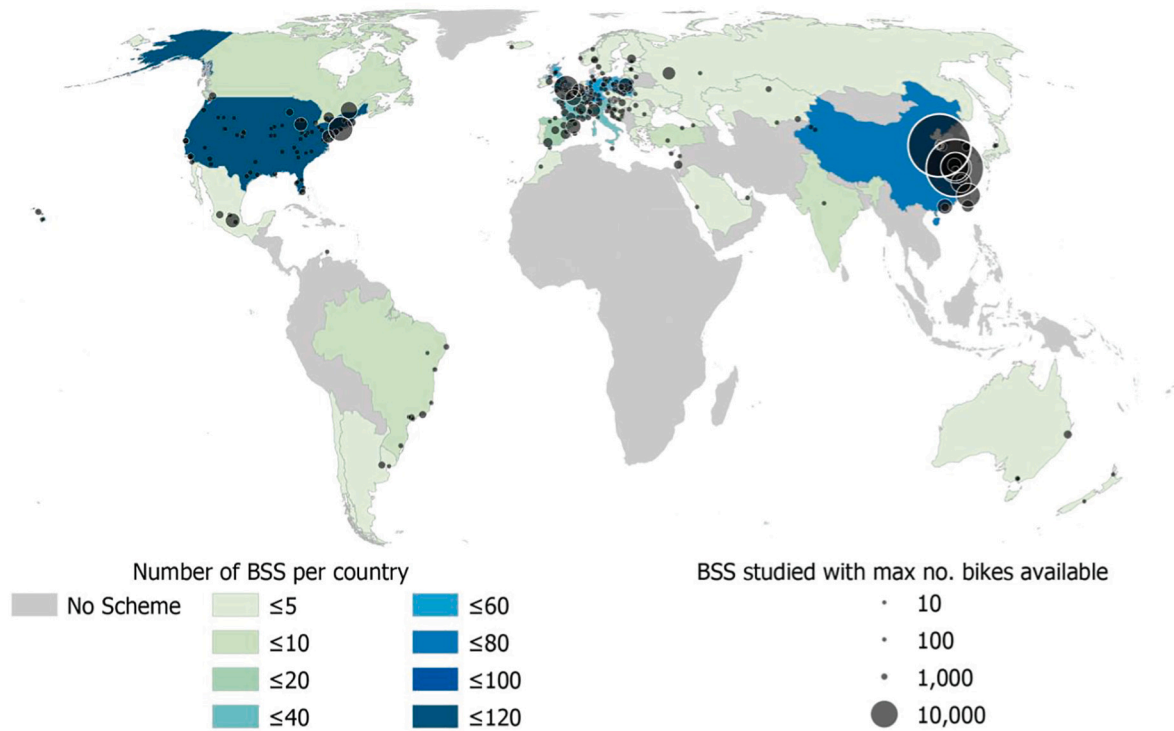
**Fig. 1.** The size and location of the 322 BSS used within the final analysis.

uptake of BSS. Generally, an increase in temperature has been associated with an increase in BSS activity (Faghih-Imani et al., 2014), up to a limit (Miranda-Moreno and Nosal, 2011). Other research shows a negative correlation between topographical relief and BSS use (Midgley, 2011), with Mateo-Babiano et al. (2016) finding that there was nearly 2 times more downhill than uphill journeys in the Brisbane BSS.

With safety concerns found to be one of the largest barriers to BSS use (see Fishman et al., 2014) those urban areas with more dedicated cycling infrastructure benefit from greater BSS use (Faghih-Imani et al., 2014; Mateo-Babiano et al., 2016). However, research also suggests that regulations requiring users to wear helmets can limit the appeal of BSS (Médard de Chardon et al., 2017; Fishman et al., 2013; O'Brien et al., 2014), with BSS users significantly less likely to wear helmets as compared to private bicycle users in those cities which do not mandate cyclists to wear helmets (Fischer et al., 2012; LDA Consulting, 2013).

There are numerous other variables which have been assessed at the individual BSS level, but the impact of temporality is one which has been found to have the most consistent impacts across BSS. The number of journeys within a BSS conform to seasonal, weekly and daily trends (Ahmed et al., 2010; Faghih-Imani et al., 2017; Faghih-Imani et al., 2014; Zaltz Austwick et al., 2013). In particular we can see clear differences in the use patterns within a system, when comparing weekday and weekend journey patterns. As a result, the weekday-weekend divide has been a consistent feature of the majority of analysis which considers use patterns across BSS literature (Faghih-Imani et al., 2017; Faghih-Imani et al., 2014; Mateo-Babiano et al., 2016.

### 2.2. Bicycle sharing system comparisons

Comparisons between BSS provide a way of assessing the relationships between different attributes across systems. This enables us to develop a deeper understanding of the BSS landscape as a whole, as it allows us to see how the variables impact BSS performance globally. Previous work in this respect has been limited by poor data availability, therefore few regional comparisons have been attempted. There are exceptions, which include Zaltz Austwick et al. (2013) who compare 5

systems, and a number of studies that seek to investigate differences in usage patterns as well as clustering (see also Bieliński et al., 2019; Kou and Cai, 2019; Sarkar et al., 2015; Zhang et al., 2015).

The scope for substantive international comparisons across BSS is constrained by the inconsistency of BSS data collection practices and formats, and the limited availability of data for many systems. As a result, when making comparisons between BSS the "trips per day per bicycle" metric (TDB) is commonly used as it is straightforward to calculate and enables some understanding of the usage rates in a BSS. This metric is calculated using the dock capacity information, which is the most commonly available form of real-time BSS data. For example, it is the basis to Médard de Chardon et al.'s (2017) comparison of 75 systems from around the world that determined common variables that influence the use of each BSS. Similar to Table 1, the authors identified 5 categories of variable: BSS attributes, density and compactness, geography, weather and transportation infrastructure. Although the authors express TDB as a 'good standardised measure of success' (Médard de Chardon et al., 2017), in this paper TDB has been used primarily as a measure of system efficiency and rate of use.

There is a need for a large global comparison of systems, particularly given that prior research often omits China, which is currently the largest BSS market in the world with 1.9 million bicycles in over 430 cities in 2016 (Gu et al., 2019). The research in this paper therefore marks a major step forward given it benefits from data collected from schemes situated on all major continents. It can thus reveal new similarities/differences between systems and offer insights at a truly global level.

## 3. Data

### 3.1. Bicycle sharing system data

Dock capacity information is the most abundantly available data for BSS as it offers a real-time indication of the number of bicycles and spaces at each docking station within a particular BSS. These data are typically used to notify users and system operators of bike availability

**Table 2**

Confounding variable data sources, resolutions and buffer distances.

| Data source | Variables calculated | Raster/ Vector (resolution) | Time period | Buffer distance |
|---|---|---|---|---|
| Global Human Settlement Population Grid by the European Commission | Population | Raster (250 m cells) | 2015 | 1000 m |
| Mapzen Terrain Tiles accessed through *elevatr* R package | Topography | Raster (Variable [zoom = 9[a]]) | Accessed 11th July 2019 | 300 m |
| Average precipitation 1970–2000 from worldclim.org | Precipitation | Raster (1000 m cells) | 1970–2000 | 1000 m |
| OpenStreetMap | Cycling Infrastructure Length | Vector | Accessed 1st July 2020 | Convex Hull |

[a] https://github.com/tilezen/joerd/blob/master/docs/data-sources.md#what-is-the-ground-resolution
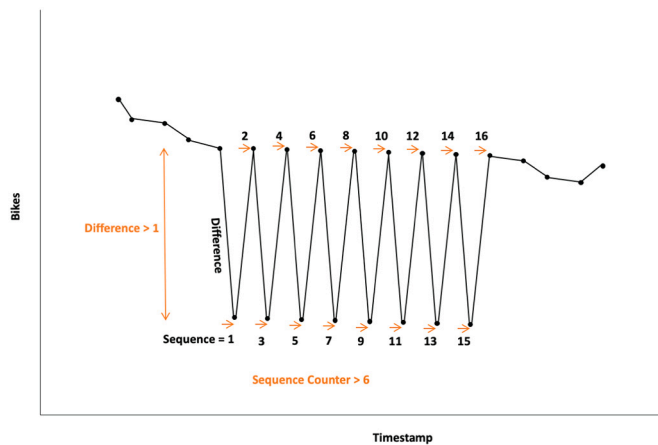


**Fig. 2.** Visual representation of the data noise detection algorithm.

across the system. But for the purposes of this research, these data were collected and stored within a database at 2-min intervals (for the majority of systems), enabling us to build up a detailed understanding of the size of the fleet within each system as well as the flow of bicycles in and out of docking stations. A subset of BSS (see Appendix) also release

Origin-Destination (O—D) journey data, which is used within this research for calibrating and validating journeys detected from the manipulation of dock capacity data (Section 4.1.1). Fig. 1 depicts the distribution and size of the 322 BSS for which we have complete data and used within the final analysis. The dataset provides the most holistic view of BSS which is currently available, with BSS data from all continents. The coordinates of docking stations within each BSS were also recorded and stored within a database. These were used to estimate the operational extent and to extract the confounding variables for each scheme (Sections 4.1.3 and 4.1.4).

### 3.2. Confounding variable data

In an effort to account for different contextual settings of each BSS additional confounding variables were calculated and clustered. These were chosen based on the literature in Section 2.1 and are detailed in Table 2. We have attempted to choose from each variable category, prioritising those which exhibited the most consistent impacts across BSS. Unfortunately, the choice of variables was limited to those with global coverage, as it was necessary to maintain metric homogeneity across the BSS within this analysis. After careful consideration, the metrics of population, topography, precipitation and cycling infrastructure were selected and used to calculate those additional confounding variables (detailed in Section 4.1.4).

Data on cycling infrastructure was derived from OpenStreetMap (OSM) data. OSM is widely used has been found to be a more updated and granular source of such data in comparison to proprietary alternatives like Google Maps (Natera et al., 2020).

## 4. Methodology

### 4.1. Metric creation

The careful manipulation of dock capacity data enables the creation of metrics indicating the size and use of each BSS, whilst dock location data can be used to determine the operational area of each BSS as well as enable the calculation of confounding variables within these areas. The metrics calculated provide the foundations for the creation of two datasets (Section 4.2) that split the time constant and time series variables and are used to conduct the two-staged classification process.

### 4.1.1. Journey estimation, validation and cleaning

In order to estimate the number of journeys within each system, the dock capacity records were examined sequentially. A fall in the number of bicycles between two consecutive collection periods (usually 2-min) was used as a proxy for the start of a journey from that docking
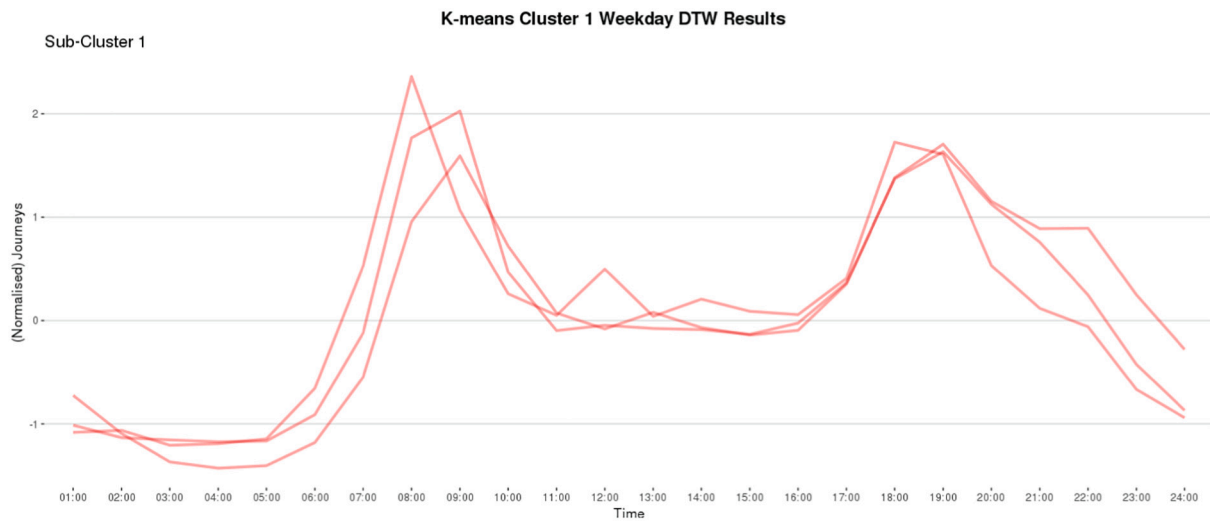
**Table 3**

K-means cluster characteristics.

| Cluster<br>Variable | Very large, high use BSS | Large BSS in major cities | Medium BSS with extensive cycling infrastructure | Small to medium efficient BSS | Small to medium inefficient BSS |
|---|---|---|---|---|---|
| No. of BSS | 3 | 15 | 5 | 66 | 233 |
| Population | 4,299,516.9 | 1,786,311.1 | 448,719.2 | 350,354.4 | 111,302.1 |
| Area (Km$^2$) | 441.3 | 167.5 | 73.6 | 36.1 | 16.7 |
| Slope (°) | 3.2 | 2.6 | 1.4 | 2.2 | 1.8 |
| Precipitation (mm) | 176.8 | 142.3 | 66.4 | 82.3 | 77.0 |
| Road Length (Km) | 5780.2 | 4741.4 | 14497.3 | 1294.1 | 748.1 |
| Cycle Length(Km) | 211.0 | 394.4 | 1533.5 | 181.5 | 80.4 |
| Weekday Journeys | 68659.5 | 24902.8 | 2708.1 | 4447.5 | 359.2 |
| Weekday Max Bikes | 18681.1 | 6001.5 | 1124.9 | 967.4 | 279.4 |
| Weekday TDB | 4.6 | 4.7 | 1.5 | 4.9 | 1.0 |
| Weekend Journeys | 65097.1 | 21728.7 | 2716.2 | 3384.3 | 305.0 |
| Weekend Max Bikes | 18839.5 | 6015.0 | 1121.7 | 968.7 | 277.5 |
| Weekend TDB | 4.2 | 3.9 | 1.4 | 3.8 | 0.9 |

**K-means Cluster 1 Weekday DTW Results**

Sub-Cluster 1



**Fig. 4.** Results of DTW clustering on average weekday use for the very large, high use BSS cluster.

**K-means Cluster 1 Weekend DTW Results**
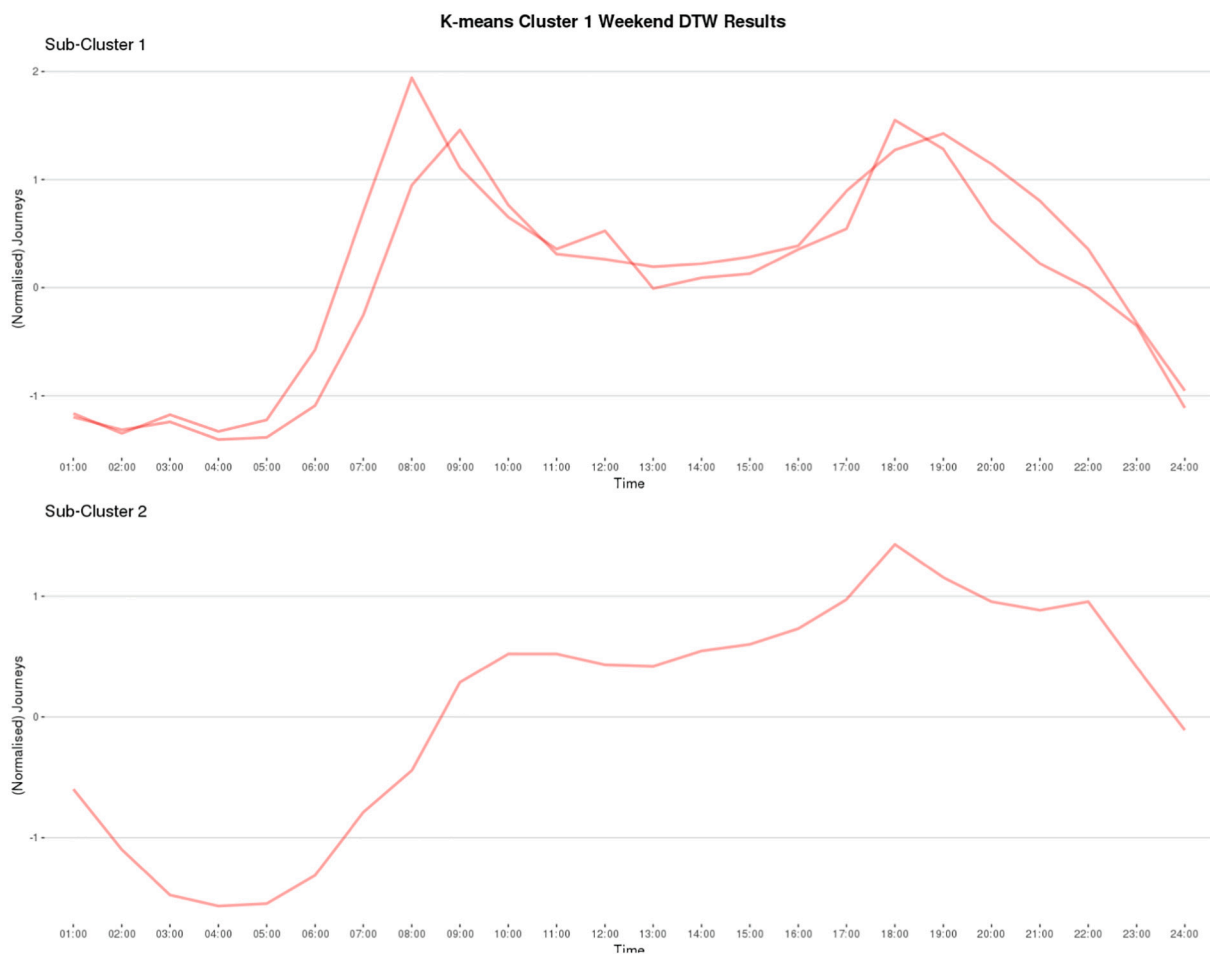
Sub-Cluster 1

Sub-Cluster 2



**Fig. 5.** Results of DTW clustering on average weekend use for the very large, high use BSS cluster.

station. This heuristic was applied to each docking station within a BSS, helping to build up a detailed picture of the frequency of journeys. Similar journey estimation methods have been used by Médard de Chardon and Caruso (2015) who found that aggregating journey estimations to the daily level proved to be the most accurate. Journeys were also aggregated to the hourly level, in order to be able to examine patterns of use throughout a day in more detail.

The validation of journey estimations was an essential step due to the importance of the variable in the creation of other metrics such as TDB as well as in the second-stage DTW cluster analysis of the use profile for a typical day. For these reasons, many precautions were taken in order to ensure its validity.
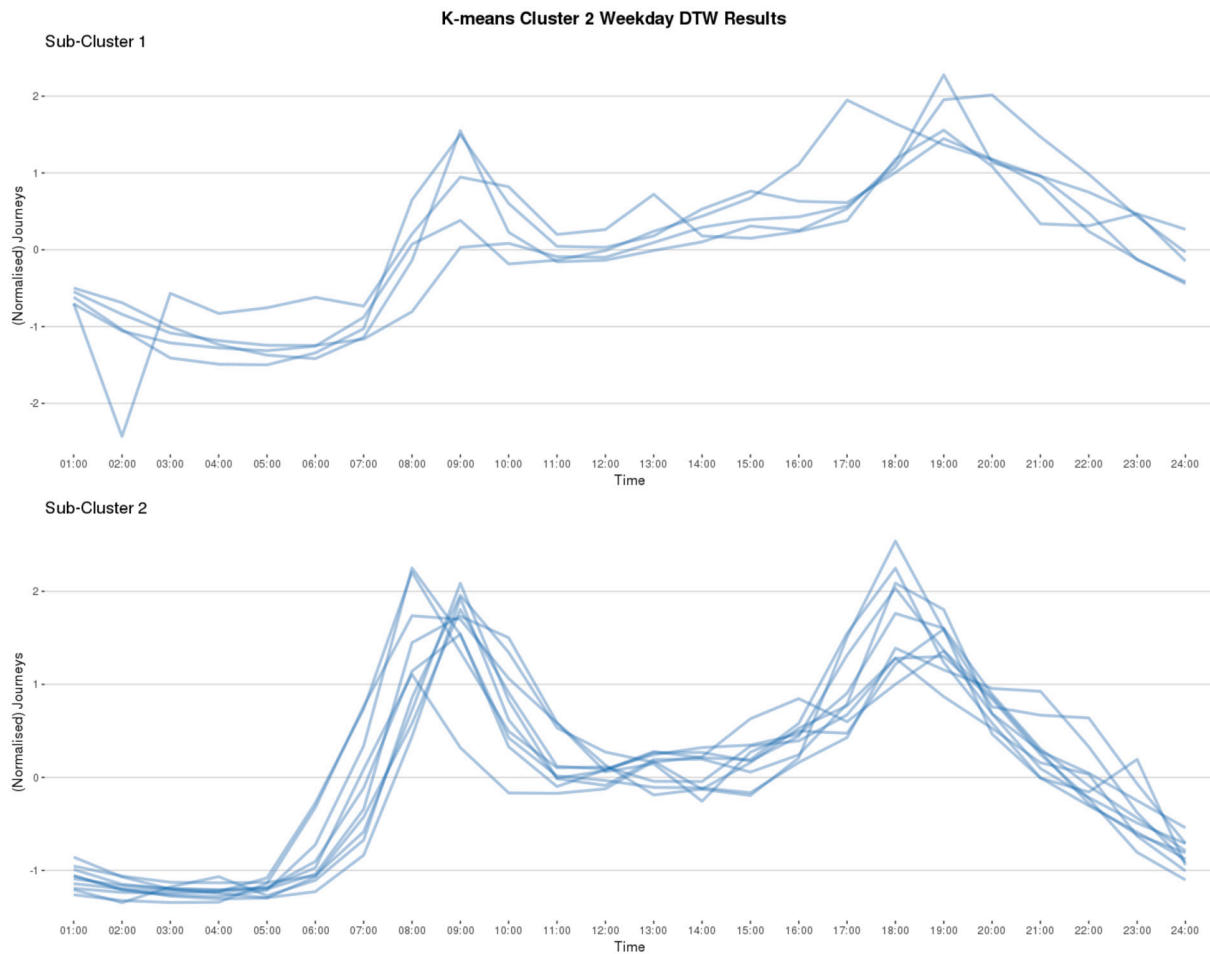
**Fig. 6.** Results of DTW clustering on average weekday use for the large BSS in major cities cluster.

For the systems for which we have journey data available (see Appendix), it was possible to compare the calculated journey estimations to the journey data. This contains the O—D flows of each journey taken within a BSS, as opposed to dock capacity data that is likely to contain operator redistributions as well as data errors. Therefore, the data is aggregated to the same 24-h period as the estimated journey figures and used to calculate the percentage difference. After initial investigations, journey estimations were found to both under and overestimate journey counts. On average, the estimated journey figures varied by up to 15% with respect to actual journey records from the journey data. Underestimations can arise from the frequency of observations not being sufficient to capture all dock changes. This effect can be reduced by using a shorter collection interval (Médard de Chardon and Caruso, 2015). Overestimations are typically due to operator actions such as rebalancing or bicycle removal for repairs, which are likely to get captured within dock capacity data. There were also periods of overestimations triggered largely by transmission errors from docking stations. These were identified within the dock capacity data where sequential records fluctuated between the same value across several collection periods (as seen in Fig. 2). For this method to be effective at identifying all periods of potential data collection errors across all BSS, it was necessary to test several threshold levels. Due to the high data volume and the large variability in BSS size, it was found that low threshold levels were necessary to capture all true positive cases of false data collection. This meant that differences in the number of bikes of more than 1 over a period greater than 6 sequential data collection instances were identified as periods of potential data error. Since the threshold values were very low, there were instances where the

algorithm captures some false positives. As a result, it was necessary to manually confirm each case identified as potential false data and remove those periods which were verified to be a false positive.

### 4.1.2. Maximum available bicycles and trips per day per bicycle

The number of operational bicycles (total number of bikes in a single collection period) was recorded and TDB calculated by dividing the estimated number of journeys within a particular day, by the maximum number of operational bicycles for the same day. TDB helps to account for differences in the size of the system and allows for cross comparison between systems in terms each bicycle's use per day.

### 4.1.3. Operational area

As mentioned in Section 2.1, the operational area of a BSS was defined by placing a 500 m radius buffer around each of the docking stations. These were dissolved into a single polygon and were used to calculate the average operational area. Similar to all the other metrics calculated within this analysis, it was important to ensure homogeneity in the way that the metrics were calculated, not only to ensure comparative rigor, but also to help automate the calculation process across large number of BSS. We therefore opted for this simpler rule of thumb, rather than attempting to determine operational areas – which are also subject to change – on an operator by operator basis. This approach has been taken in other studies, including Fuller et al. (2013), Hosford et al. (2019), Hosford et al. (2018) and Bachand-Marleau et al. (2012). A sensitivity analysis was conducted to determine the impacts of buffer size since smaller buffers have been used in other studies (see Kabra et al., 2019; Tran et al., 2015). The results suggest that using a
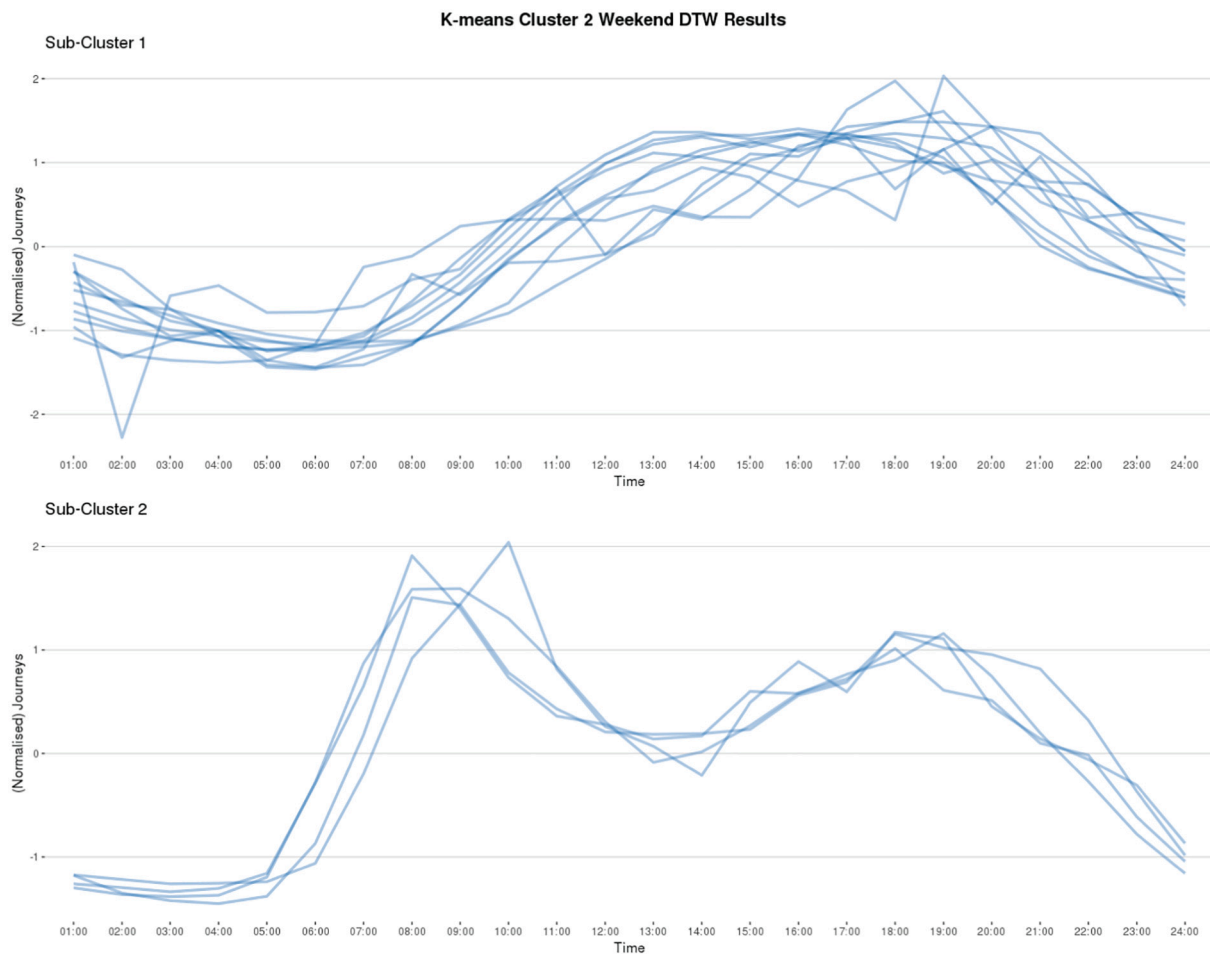
**Fig. 7.** Results of DTW clustering on average weekend use for the large BSS in major cities cluster.

300 m radius buffer to determine operational area makes no difference to the K-means cluster assignments in Section 5.1, whilst a 250 m radius buffer causes a change in the classification of just 1 system out of the 322 BSS within this analysis. This is a very minor change, reinforcing the strong cluster identities found within the results (Section 5).

### 4.1.4. Confounding variables

Time invariant confounding variables (population, precipitation, topography and cycle paths) were estimated for each BSS to add additional contextual information to our clusters. Active docking stations, in the period between April and September 2018, were buffered to generate polygons for overlay onto the confounding variable datasets (Section 3.2). The buffer distances used for each variable are shown in Table 2. For both the population and precipitation variables we decided to use a 1 km buffer to incorporate some of the urban characteristics beyond the immediate limits of the outer docking stations in a scheme. Although literature suggests that willingness to walk to a docking station is around 500 m for some systems (Gu et al., 2019; Bachand-Marleau et al., 2012), we know that users of BSS are likely to use BSS as a mode to solve issues around the "last mile problem" (Fan et al., 2019; Yang et al., 2019), which implies users travelling from outside that 500 m walking boundary. Similarly, when considering the topography and length of cycling infrastructure of a docked BSS, it is important to consider that journeys will be taken between docking stations, therefore calculations for these variables were limited to 300 m buffers and a convex hull. Much like the determination of the use of 500 m radius buffers for calculating operational area, the buffer sizes/methods were chosen to best capture the nature of each confounding variable in the context of

each BSS in a homogenous manner.

### 4.2. Panel style data curation and BSS clustering

Following the creation of the metrics detailed across Section 4.1, the data was split into time constant and time series data to allow for a two-staged clustering analysis to be conducted. This method was chosen to take advantage of the multiple metric types that were created as well as current limitations with the simultaneous clustering of time series and time constant variables.

For the first stage of clustering, a time constant dataset was created for analysis using a K-means technique (MacQueen, 1967). K-means clustering is a standard method for exploratory clustering where the number of clusters are specified by the user (Wagstaff et al., 2001). Here, it is employed to gain a detailed, static understanding of the global BSS landscape. The variables included within this time constant dataset (see Appendix) were a combination of BSS variables (Sections 4.1.1 to 4.1.3), as well as the confounding variables (Section 4.1.4). Since these metrics were calculated for each day that we have operational data, it was necessary to aggregate metrics to the same time period for comparative purposes. As a result, the metrics were aggregated for the 6-month period between April and September 2018, since this was the period for which the largest group of BSS had operational data. Each metric was aggregated between this period, and BSS variables split between weekday and weekend measures due to the large amount of variability found between patterns of use (Faghih-Imani et al., 2014; Faghih-Imani et al., 2017; O'Brien et al., 2014). In addition to the time aggregation, variables were normalised in order to remove scale and weighting issues
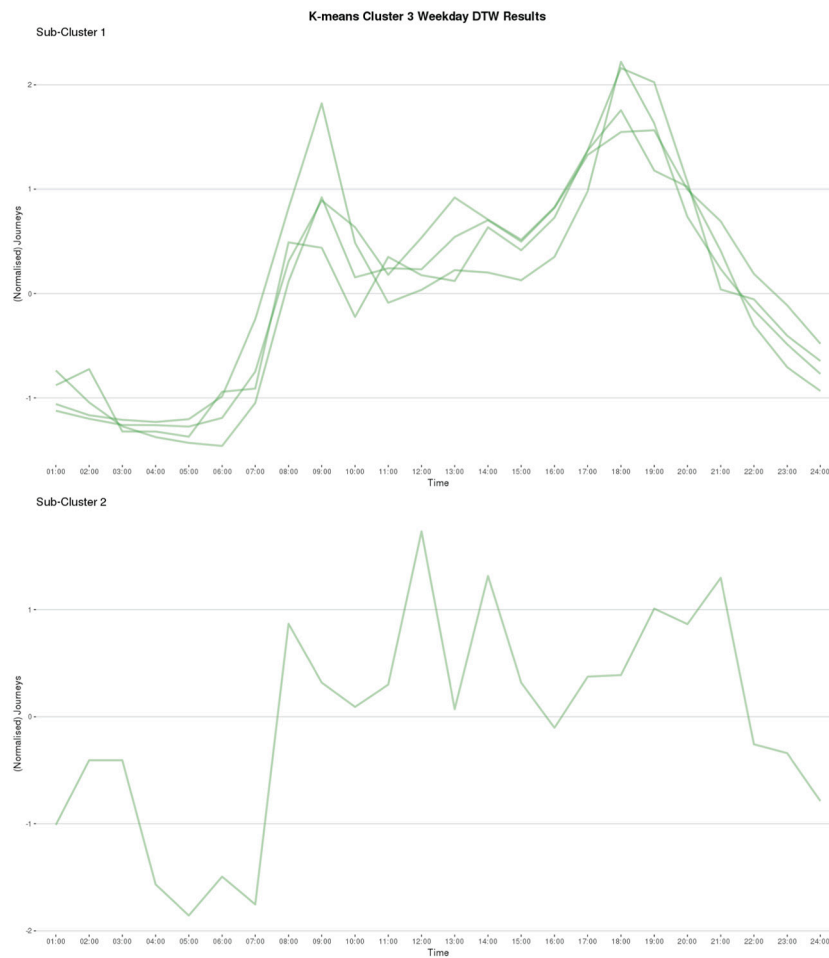
**Fig. 8.** Results of DTW clustering on average weekday use for the medium BSS with extensive cycling infrastructure cluster.

in the K-means clustering. Conducting this first-stage of clustering analysis allows us to gain an understanding of the type of system, in terms of its size, environment and general use as well as where each of these systems are situated in comparison to other BSS around the world.

In the second stage of clustering, the hourly journey estimations were used to create a time series dataset on each BSS. Similar to the creation of the time constant dataset for the first stage of clustering, the hourly journey estimations (Section 4.1.1) were aggregated for the same 6-month period between April and September 2018 and split between weekday and weekend (Faghih-Imani et al., 2014; Faghih-Imani et al., 2017; O'Brien et al., 2014). The time series data was also normalised, again to counter issues around scales. By aggregating the data in this way, the resulting dataset depicts the average pattern of use within each BSS for a typical weekday or weekend. For each of the clusters created in the first stage of clustering, a DTW cluster analysis was conducted. DTW is a shape-based clustering method for time series data, computing the smallest Euclidean distances between each time series based on moving window (Berndt and Clifford, 1994). These distances were then clustered using a hierarchical clustering method based on the Ward's method (Mojena, 1977). This was chosen after assessing the 3 alternatives – namely the single, complete and average methods (Edelbrock, 1979). It was found that the Ward's method provided the cleanest separation of clusters whilst minimising the within-cluster sum of squares variance. This second stage of clustering was employed to help give additional detail into each first stage cluster by investigating how systems within the first stage cluster are further grouped, based on their patterns of use.

## 5. Results

### 5.1. Bicycle sharing system clustering

The first-stage of the clustering was designed to create an understanding of the general groups that exist in the current BSS landscape. Although direct and statistical methods of deriving the number of clusters largely suggested 2 to be optimal, namely the elbow method, the Calinsky Criterion and average silhouette method, the decision was taken to group BSS into 5 clusters using a K-means clustering algorithm. Dividing 322 BSS into 2 unequal groups would not be beneficial in helping to uncover a deeper understanding of the global BSS landscape, whilst 5 clusters allowed for a large enough differentiation between clusters with clear cluster identities and kept the number of clusters at a minimum. The centres of each of these clusters are detailed in Table 3. The clusters have been named, based on these characteristics, as: '*very large, high use BSS*', '*large BSS in major cities*', '*medium BSS with extensive cycling infrastructure*', '*small to medium efficient BSS*', and '*small to medium inefficient BSS*'. Within this section, we refer to the 'efficiency' of a BSS using the TDB metric. Those systems that demonstrate higher TDB figures are taken to be more efficient, since this means that each bike is used on multiple occasions throughout an average day.

### 5.1.1. Very large, high use BSS

The first cluster, '*very large, high use BSS*', contains the largest and most used BSS in terms of the absolute number of journeys. This cluster only contains three systems, all of which are located in East Asia: Taipei, Suzhou and Weifang. These systems do not appear to have extensive
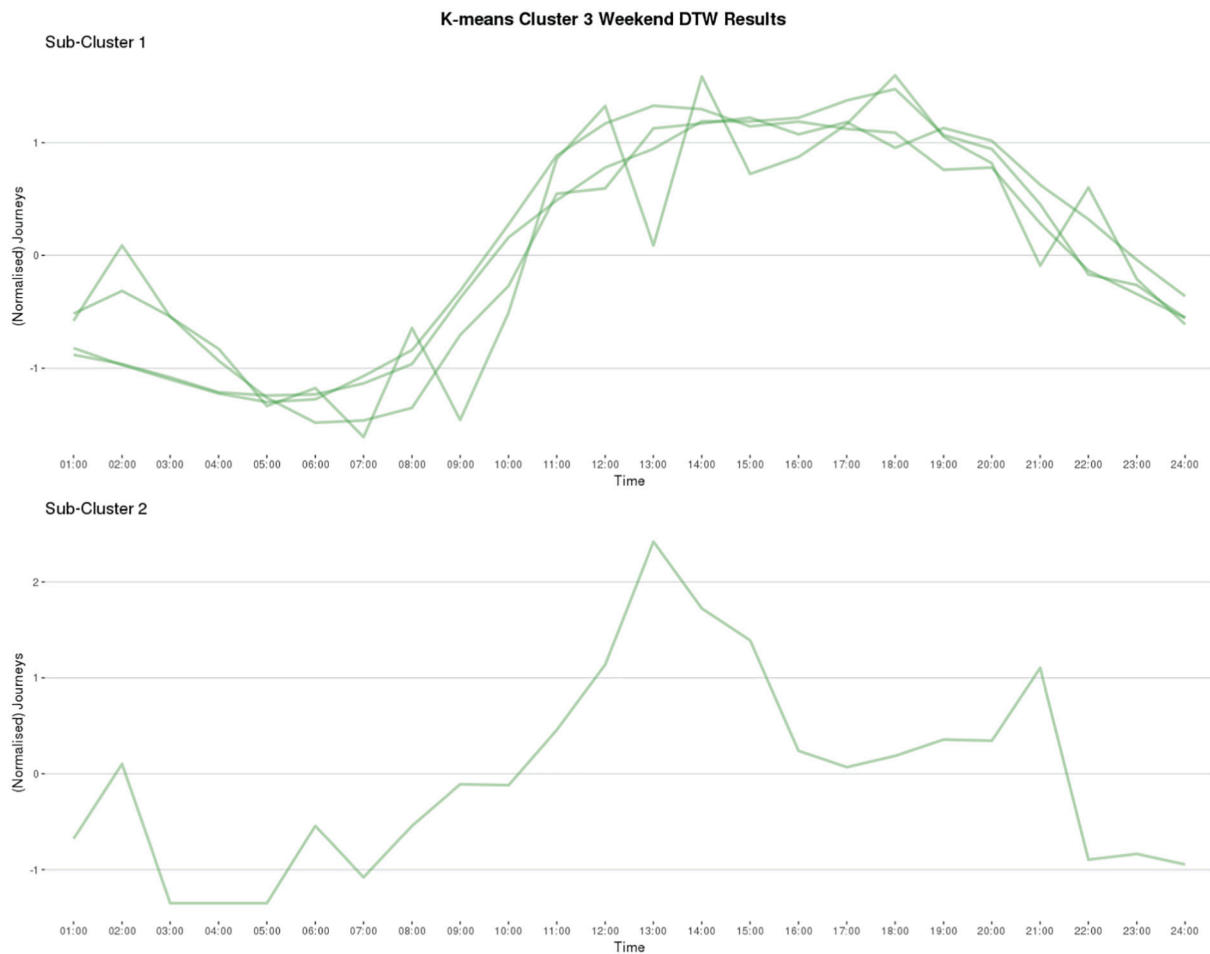
**Fig. 9.** Results of DTW clustering on average weekend use for the medium BSS with extensive cycling infrastructure cluster.

road or cycle infrastructures. Research suggests that there is a positive correlation between the level of cycling infrastructure and the number of journeys, so there may be further unmet demand in these cities.

Assessing the results of the sub-clusters we can see that all of the systems have a very similar daily use pattern (Table 4). During an average weekday all of the systems have two major peaks at the start and end of the working day (Fig. 4), which informs us that these systems are typically used by commuters. Suzhou and Weifang appear to exhibit similar use patterns throughout the weekend, although the Taipei City BSS demonstrates a single, large peak throughout the weekend (Fig. 5). This suggests a dominance of leisure and tourism activity on non-working days (O'Brien et al., 2014).

### 5.1.2. Large BSS in major cities

The second cluster, '*large BSS in major cities*', contains 15 systems that include New York, Paris, London, Barcelona, and Seoul (Table 5). These BSS are typically located in sprawling conurbations with large and densely populated cores. These systems hold similar characteristics to those *very large, high use BSS*, the main difference being their absolute size. *Large BSS in major cities* appear to have a greater proportion of roads with cycling infrastructure compared to *very large, high use BSS*. The success of BSS within this cluster may help to reinforce previous literature which finds that greater investment in cycle safety helps encourage greater BSS use (Faghih-Imani et al., 2014; Fishman et al., 2014; Mateo-Babiano et al., 2016).

Assessing the sub-clusters, we can see that there are two distinct patterns of use that we can identify during the average weekday (Fig. 6) and weekend (Fig. 7) within these *large systems in major cities*. During the

weekdays we find that the majority of the systems have a typical two commuter peaks. We also find that five of the systems appear to have a more gradual use patterns, with less defined commuter peaks, although they are still identifiable. During the weekend, we find that the majority of systems appear to have one large peak, which is likely driven by leisure and tourism use. There are four systems that appear to have two commuter peaks at the weekend, displaying similar patterns to the weekday. These systems are all in Chinese cities. This gives us an indication that the users within these systems continue to work throughout the weekends.

### 5.1.3. Medium BSS with extensive cycling infrastructure

The third cluster, *medium BSS with extensive cycling infrastructure* is a small cluster of five systems (Table 6), which are characterised by an unusually large road and cycling infrastructure network for their size. Although these systems benefit from good levels of cycling infrastructure, they appear to have relatively low levels of use. The cities with these systems also are relatively flat and dry. Both these aspects contradict the earlier literature (Campbell et al., 2016; Corcoran et al., 2014; Midgley, 2011; Miranda-Moreno and Nosal, 2011; Nankervis, 1999), which goes to show that the confounding variables within this analysis do not have consistent implications on the activity and use within BSS. Further investigation into these systems would provide new understandings of how these variables can impact the use of BSS in different ways.

Looking into the aggregated daily travel patterns shows that for the majority of systems, there appears to be a clear two-commuter peak trend during the weekdays (Fig. 8) and a large extended peak during the
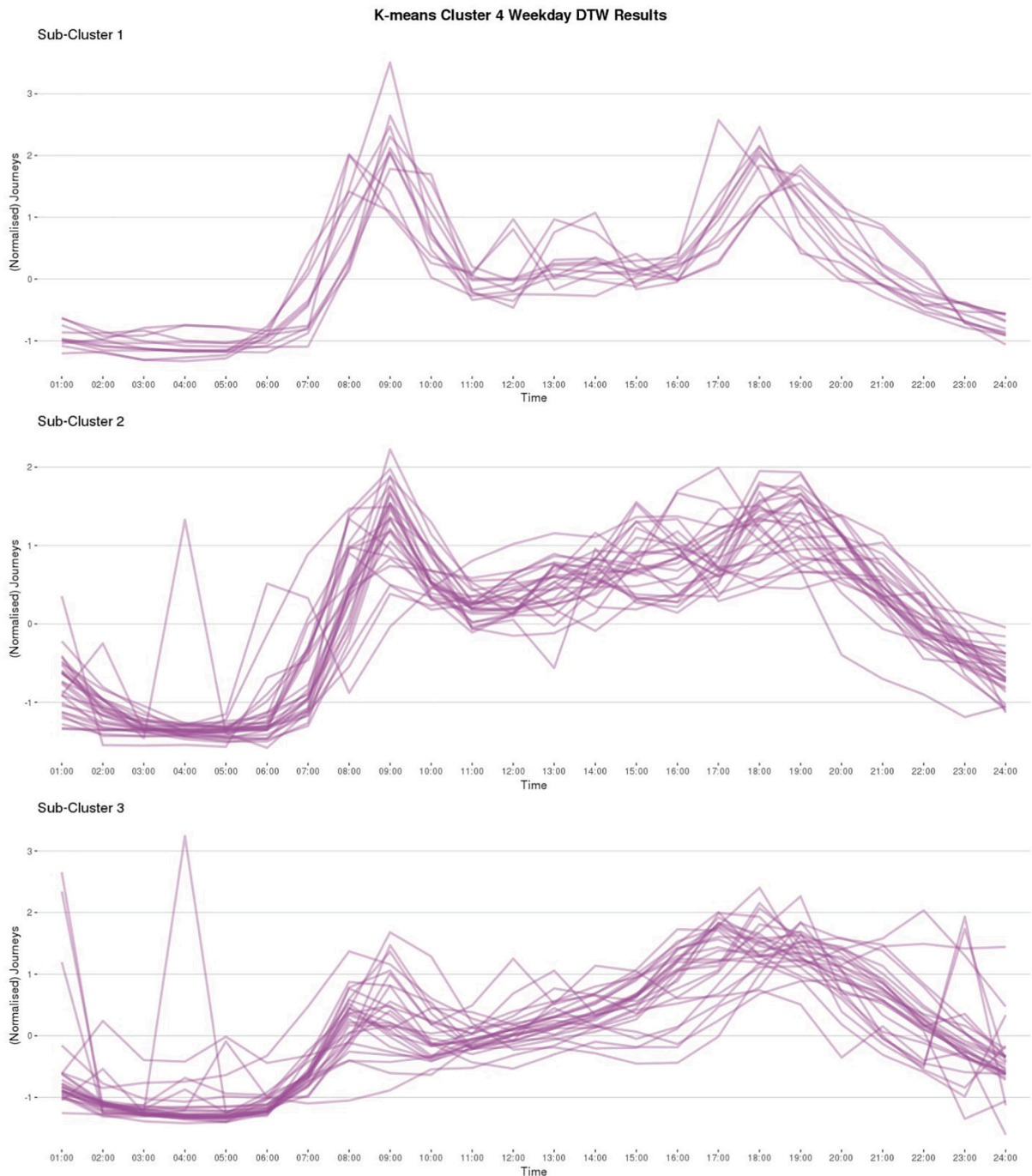
**Fig. 10.** Results of DTW clustering on average weekday use for the small to medium efficient BSS cluster.

weekends (Fig. 9). The BSS in Stuttgart (Germany) appears to be an anomaly in terms of its typical use on both the weekday and weekend. This is primarily due to its small size in comparison to the other systems in this cluster, as well as the fact that this system was used electric bicycles. Electric BSS were much more uncommon at this time, which may help to explain its differing journey patterns. At the end of October 2018, the system was replaced by another operator after losing its contract in a competitive retendering process, for whom we do not have data.

*5.1.4. Small to medium efficient BSS*

The fourth cluster, *small to medium efficient BSS,* is a large group of 66 systems (Table 7). Although much smaller, these systems appear to have a similar density of cycling infrastructure as those *very large, high use*

*BSS,* which follow the trends suggested in the current literature. For their size, these systems achieve high numbers both in terms of the number of journeys and TDB, during both the weekends and weekday. This suggests that these systems are well managed, and ideally set up to cater to the surrounding population.

An investigation into the sub-clusters shows that there are three main types of journey pattern during the weekday and weekend. During the weekday (Fig. 10), the first two sub-clusters depict the typical two commuter peak pattern, with the first sub-cluster depicting two defined peaks whilst the second sub-cluster has much smaller peaks. The third sub-cluster of journey patterns during the weekday depict two very small peaks during commuting hours, although there appears to be a gradual increase in journeys throughout the day until around 6 or 7 pm
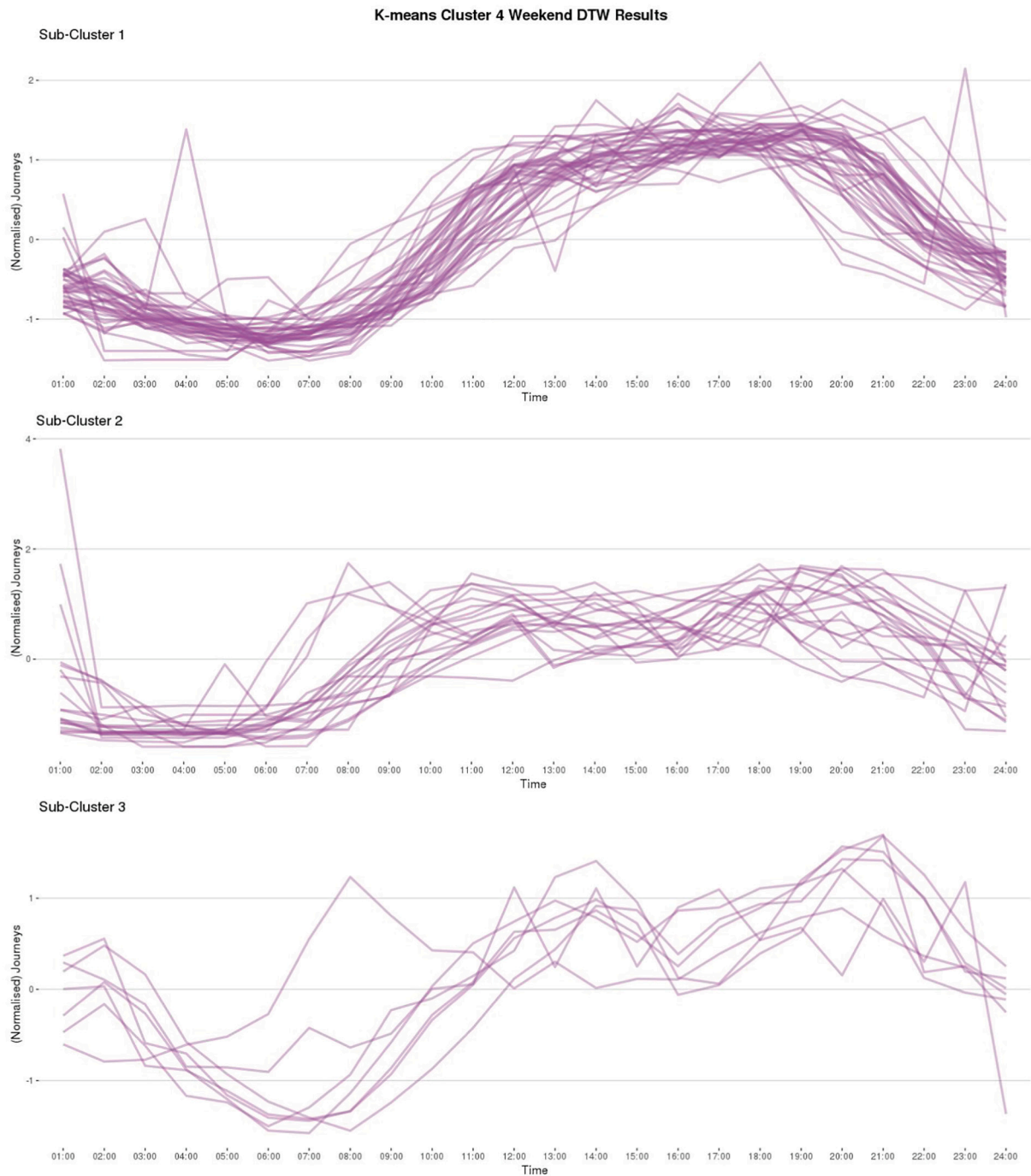
**Fig. 11.** Results of DTW clustering on average weekend use for the small to medium efficient BSS cluster.

when the number of journeys start to fall. These systems therefore appear exhibit a more varied user base, with general trends similar to those of leisure and tourism use, with small surges in demand during peak hours by commuters. During the weekends (Fig. 11), again we see that two of the clusters follow the typical trend of having one large peak across the weekend. Both of these sub-clusters show a steep incline throughout the morning hours between 8 am and noon, with a steep decline in use in the late evening around 9 pm. The third sub-cluster appears to have a much more even distribution of journeys throughout the day, with journeys appearing to gradually increase throughout the morning hours accompanied by a gradual decrease in the number of journeys into the evening and night.

### 5.1.5. Small to medium inefficient BSS

The final cluster, *small to medium inefficient BSS,* contain the largest number of systems (233) (Table 8 and 9) but are, on average, the smallest BSS within the dataset and experience the lowest levels of use. These systems appear to be inefficient, with bicycles used less than once a day. Although they have the lowest average measures for most metrics, they appear to have a comparatively similar levels of average precipitation and total relief to those *medium BSS with extensive infrastructure* and *small to medium efficient BSS* schemes.

Assessing the travel patterns of these systems, the journey patterns are arranged into 4 sub-clusters for both the weekdays and weekends. For those weekday travel patterns (Fig. 12), we have 2 sub-clusters which appear to have the typical two-commuter peak, one containing
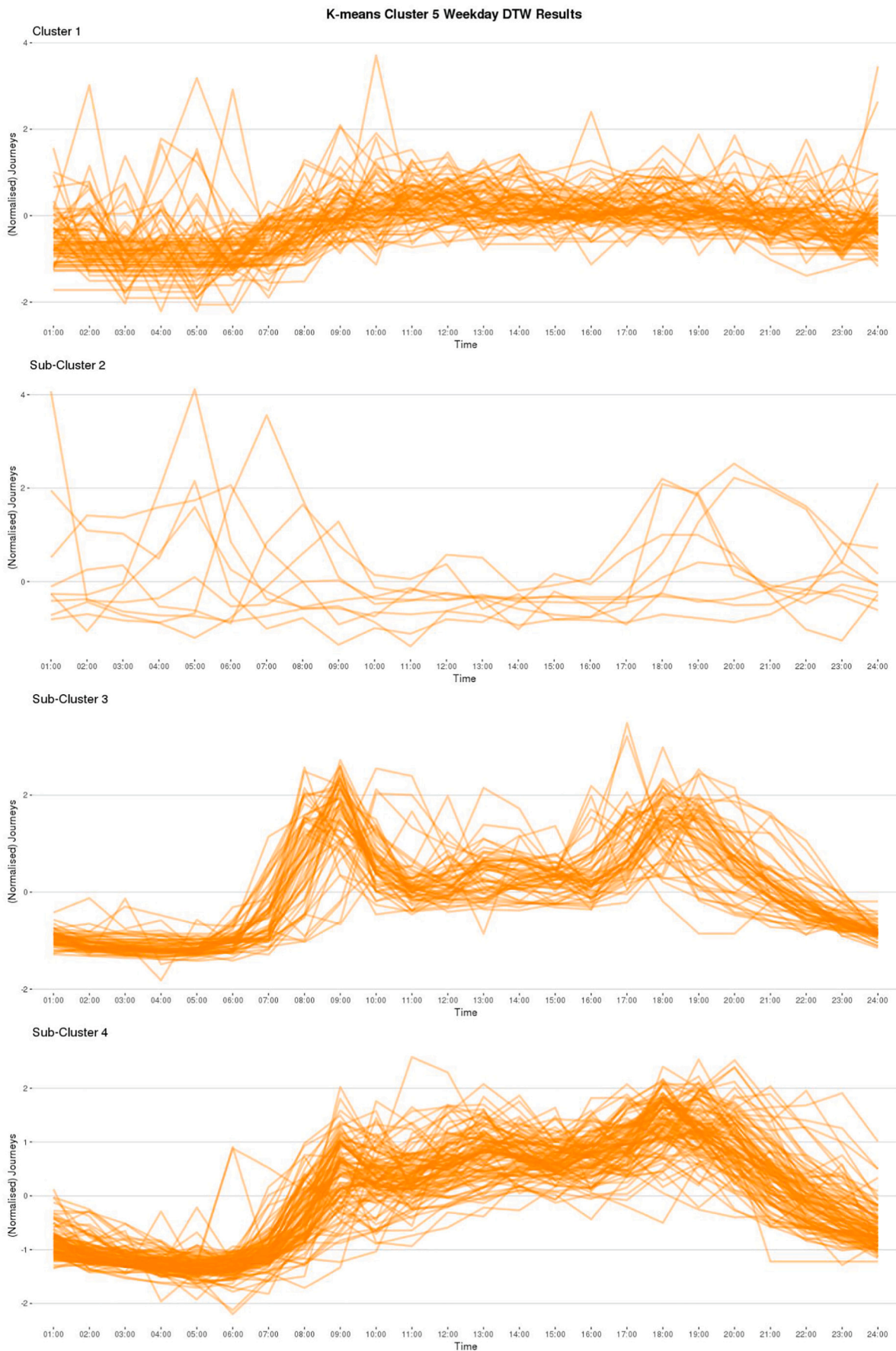
**Fig. 12.** Results of DTW clustering on average weekday use for the small to medium inefficient BSS cluster.
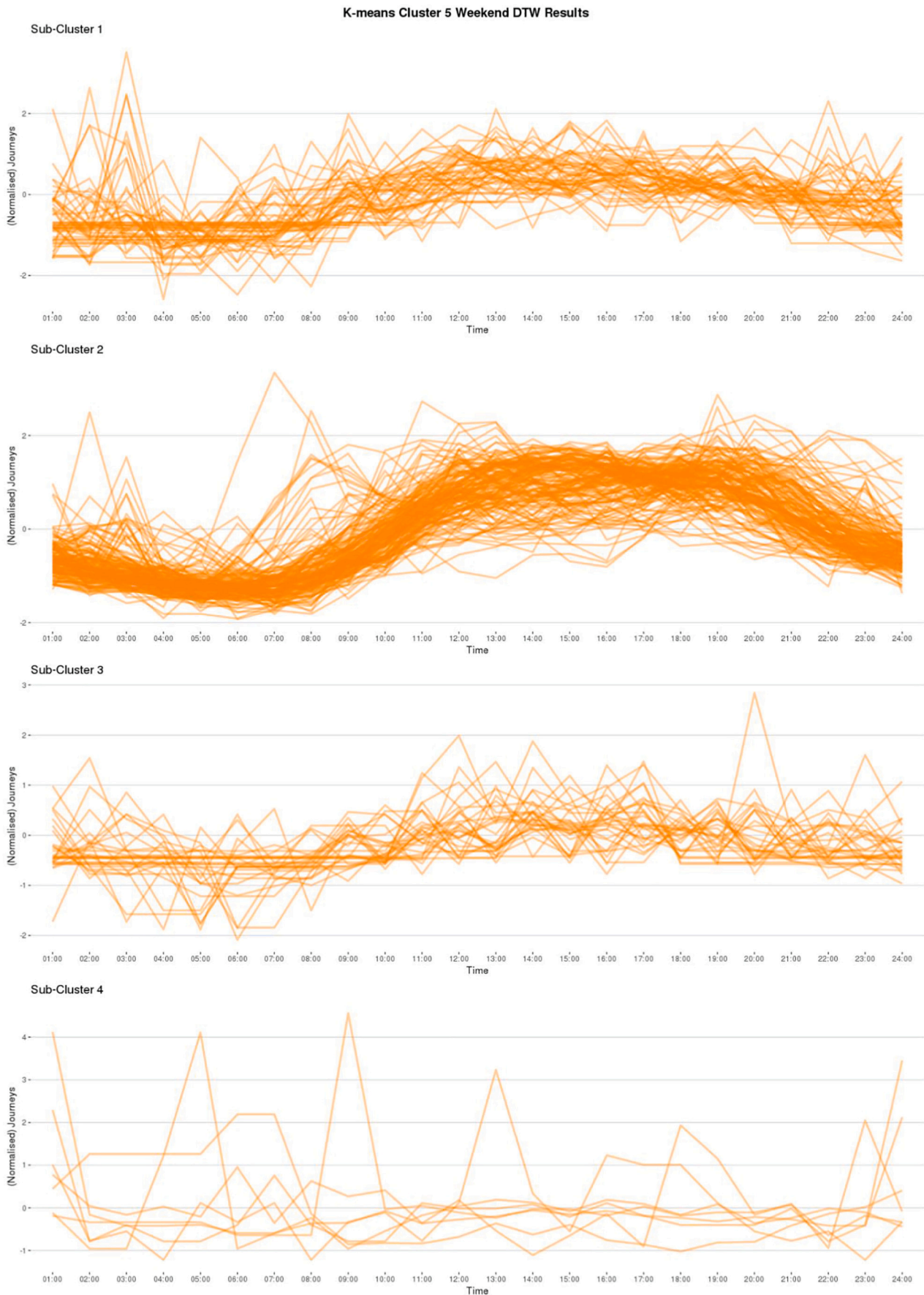
**Fig. 13.** Results of DTW clustering on average weekend use for the small to medium inefficient BSS cluster.

**Table 4**
K-means Cluster 1 DTW sub-cluster allocation.

| BSS sub-cluster allocation in K-means cluster 1 | | |
| --- | --- | --- |
| Weekday | Weekend | |
| 1 | 1 | 2 |
| - Weifang | - Weifang | - Taipei City |
| - Suzhou | - Suzhou | |
| - Taipei City | | |

**Table 5**
K-means Cluster 2 DTW sub-cluster allocation.

| BSS sub-cluster allocation in K-means cluster 2 | | | |
| --- | --- | --- | --- |
| Weekday | | Weekend | |
| 1 | 2 | 1 | 2 |
| - Barcelona | - London | - London | - Changshu |
| - Paris | - Chicago | - Barcelona | - Kunshan |
| - Moscow | - Montreal | - Chicago | - Wenzhou |
| - Seoul | - Changshu | - Kaohsiung | - Zhongshan |
| - Warsaw | - Kaohsiung | - Mexico City | |
| | - Kunshan | - Montreal | |
| | - Mexico City | - Moscow | |
| | - New York | - New York | |
| | - Wenzhou | - Paris | |
| | - Zhongshan | - Seoul | |
| | | - Warsaw | |

**Table 6**
K-means Cluster 3 DTW sub-cluster allocation.

| BSS sub-cluster allocation in K-means cluster 3 | | | |
| --- | --- | --- | --- |
| Weekday | | Weekend | |
| 1 | 2 | 1 | 2 |
| - Bremen | - Stuttgart | - Bremen | - Stuttgart |
| - Los Angeles | | - Los Angeles | |
| - Mannheim | | - Mannheim | |
| - Washington D.C. (CaBi) | | - Washington D.C. (CaBi) | |

large and distinct peaks (sub-cluster 3), with the other characterised by smaller peaks during commuting hours with a more even use throughout the day (sub-cluster 4). The first sub-cluster appears to have a fairly uniform use throughout the day, which may suggest systems characterised by a mix of leisure and commuter users, although large peaks and variations identify systems may help to identify systems with poor use or poor data quality. Similarly, the second sub-cluster has no prominent pattern due to infrequent use of the systems therefore with the potential for messy data. The patterns we find for the weekend use (Fig. 13) is similar with the second and third sub-clusters showing typical weekend behaviour with one large peak, with a steady increase throughout the morning and steady use throughout the afternoon and early evening. The first sub-cluster appears to depict a more of a uniform BSS use

throughout a typical weekend, although we find large spikes during the early hours of the morning at 3 am, which may be a result the 24 h operation of BSS in comparison to alternative public transportation modes or operational maintenance such as bicycle rebalancing. The final weekend sub-cluster appears to contain those systems with the lowest number of journeys and therefore appear to show sporadic use throughout the day. This occurs since these systems rarely get used and when there a few journeys that occur they cause a large spike the number of journeys.

## 6. Applications

The analysis here offers new insights and perspectives on the global BSS landscape, allowing for a greater understanding of the structure and relationships between BSS at the global level. A prominent issue with analysis of BSS is availability and the quality of the data (Médard de Chardon and Caruso, 2015; Mátrai and Tóth, 2016). This has allowed for unchallenged exaggerations in terms of the published statistics from BSS operators (Médard de Chardon and Caruso, 2015). The data and heuristics adopted here could therefore be used to verify the published statistics for those systems whose operators release highlight statistics, as well as helping to predict the use for those systems that do not release

**Table 8**
K-means Cluster 5 DTW weekday sub-cluster allocation.

| BSS sub-cluster allocation in K-means cluster 6 (Weekday) | | | |
| --- | --- | --- | --- |
| 1 | 2 | 3 | 4 |
| - Mödling | - Yangzi | - Nantong | - Tel Aviv |
| - Lake Neusiedl | - Dajiawa | - Brussels | - Minneapolis |
| - Hamburg | - Bhopal | - Brisbane | - Budapest |
| - Vannes | - Abu Dhabi | - Vancouver | - Miami Beach |
| - Wachau | - King Abdullah Economic City | - Anqiu | - Almaty |

**Table 9**
K-means Cluster 5 DTW weekend sub-cluster allocation.

| BSS sub-cluster allocation in K-means cluster 6 (Weekend) | | | |
| --- | --- | --- | --- |
| 1 | 2 | 3 | 4 |
| - Lake Neusiedl | - Nantong | - Mödling | - Dajiawa |
| - Hamburg | - Brussels | | |
| - Vannes | - Tel Aviv | | |
| | - Brisbane | | |
| | - Minneapolis | | |
| | - Yangzi | | |
| | - Wachau | | |
| | - Bhopal | | |
| | - Abu Dhabi | | |
| | - King Abdullah Economic City | | |
| | - Vancouver | | |
| | - Anqiu | | |
| | - Budapest | | |
| | - Miami Beach | | |
| | - Almaty | | |

**Table 7**
K-means Cluster 4 DTW sub-cluster allocation.

| BSS sub-cluster allocation in K-means cluster 4 | | | | | |
| --- | --- | --- | --- | --- | --- |
| Weekday | | | Weekend | | |
| 1 | 2 | 3 | 1 | 2 | 3 |
| - Milan | - Lyon | - Helsinki | - Toronto | - Milan | - Valencia |
| - Toronto | - Huangyan | - Sanxiang | - Boston | - Gaomi | - Seville |
| - Gaomi | - Valencia | - Daejeon | - Lyon | - Changyi | - Sanxiang |
| - Changyi | - Toulouse | - Oslo | - Toulouse | - Huangyan | |
| - Boston | - Seville | - Łódź | - Helsinki | - Daejeon | |
| | | | - Łódź | - Oslo | |

data at all. Since the metrics calculated within this research are independent of any operator motivations, they may be considered impartial. They also provide insights into systems that fail to share data.

In addition, this research offers a tool to help those cities better plan BSS, should they wish to implement new or improve existing BSS. By comparing the characteristics of the city to those within the analysis, city officials can learn from those more efficient systems. This is timely due to the fast-expanding nature of this mode of micro-mobility. Alternative modes of micro-mobility, such as scooters and dockless BSS, have a similar data format and are similarly limited in terms of data availability. The methods outlined in this research can easily be extended to other new modes of urban micro-mobility and help to provide a framework for comparison between system metrics.

Finally, by using the metric creation methodology outlined in this paper it is possible to gain a detailed understanding into the operation of a BSS using dock capacity data alone. This means that it is straightforward to scale the analysis by running the heuristics on data inputs from additional schemes as their data become available.

### 6.1. Summary

The analysis detailed here offers new insights into the global landscape of BSS, outlining a robust method that manipulates a voluminous dataset in a manageable way to create a comprehensive global comparison of BSS.

Dock capacity data provides the foundations for the majority of the comparative metrics created within this research, which has been supplemented with additional confounding variables to enable further contextual understanding of each BSS. Using a two-staged clustering process on the calculated metrics provides a framework for comparison between BSS and helps to situate individual systems within this rapidly growing mode of urban micro-mobility.

Conducting such analysis, there appear to be 5 main types of BSS: *very large, high use BSS, large BSS in major cities, medium BSS with extensive cycling infrastructure, small to medium efficient BSS* and *small to medium inefficient BSS*. From these major cluster groups, we are able to group schemes by usage, contextual indicators and the behavioural characteristics of their users. This not only facilitates a global comparison of scheme performance, but also offers a basis to new schemes to identify established BSS with similar characteristics that can be used as a template for anticipating the likely demand from users.

### Declaration of conflicting interests

### Acknowledgements

### Appendix

*Fixed variable panel style dataset (K-means clustering)*

| System | Population | Area | Slope | Road Length | Cycle Length | Weekday Journeys | Weekday Max Bikes | Weekday TDB | Weekend Journeys | Weekend Max Bikes | Weekend TDB |
|--------|-----------|------|-------|-------------|--------------|------------------|-------------------|-------------|------------------|-------------------|-------------|
|        |           |      |       |             |              |                  |                   |             |                  |                   |             |

*BSS with published or inferable journey data*

| | | |
|---|---|---|
| Albany | Glasgow | Ottawa |
| Atlanta | Hailey/Ketchum | Oxford (Hourbike) |
| Berlin | Hamilton | Oxford (oBike) |
| Beverly Hills | Hoboken | Phoenix |
| Bialystok | Karlsruhe | Portland |
| Bielsko-Biała | Kraków | Reading |
| Boise | Lincoln (UK) | San Mateo |
| Brighton | Liverpool | Santa Monica |
| Brno | Long Beach (California, USA) | Sheffield |
| Budapest | Long Beach (New York, USA) | Tampa |
| Buffalo | Monash University | Topeka |
| Bydgoszcz | New Orleans | UCLA |
| Cardiff | Northampton | Warsaw |
| Charleston | Nottingham | Washington D.C. (JUMP) |
| Cologne | Orlando | West Hollywood |
| Derby | | |
| Gibraltar | | |

*BSS cluster allocation*

*K-Means Cluster 1 (Very large, High use (3)).*

- DTW Weekday Sub-Cluster 1

Suzhou, Taipei City, Weifang.
*K-Means Cluster 2 (Large systems in major cities (15))*

- DTW Weekday Sub-Cluster 1

Barcelona, Moscow, Paris, Seoul, Warsaw.

- DTW Weekday Sub-Cluster 2

Changshu, Chicago, Kaohsiung, Kunshan, London, Mexico City, Montreal, New York, Wenzhou, Zhongshan.
*K-Means Cluster 3 (Medium systems with extensive cycling infrastructure (5))*

- DTW Weekday Sub-Cluster 1

Bremen, Los Angeles, Mannheim, Washington D.C. (CaBi).

- DTW Weekday Sub-Cluster 2

Stuttgart
*K-Means Cluster 4 (Small to medium efficient systems (66)).*

- DTW Weekday Sub-Cluster 1

Belfast, Boston, Changyi, Dublin, Gaomi, Milan, Nantes, San Francisco, Saragossa, Sarajevo, Toronto, Washington D.C. (JUMP).

- DTW Weekday Sub-Cluster 2

Bergen, Bordeaux, Buenos Aires, Bydgoszcz, Castellon, Girona, Guadalajara, Honolulu, Huangyan, Kanazawa, Lille, Ljubljana, Lyon, Madrid, Marseille, Nice, Palma, Pisa, Querétaro, Rio, Santos, Seville, Toulouse, Turin, Valencia, Vienna, Vila Velha, Vilnius.

- DTW Weekday Sub-Cluster 3

Białystok, Brighton, Cardiff, Daejeon, Gliwice, Helsinki, Kalisz, Katowice, Kołobrzeg, Kraków, La Rochelle, León, Łódź, Lublin, Ordu, Oslo, Porto Alegre, Poznan, Radom, Recife, Salvador, Sanxiang, São Paulo, Sorocaba, Szczecin, Wrocław.
*K-Means Cluster 5 (Small to medium inefficient systems (233))*

- DTW Weekday Sub-Cluster 1

Vor Wien, Amstetten, Aruba, Auckland, Bath, Bottrop, Brinje, Byblos, Christchurch, Columbia, Exeter, Flensburg, Gelsenkirchen, Gibraltar, Greenville, Groß-Enzersdorf, Gütersloh, Hailey/Ketchum, Hamburg, Hamm, Herne, Hollabrunn, Ivanić-Grad, Jurmala, Karlovac, Kent State University, Kingston (UK), Laa an der Thaya, Lincoln (UK), Lunz am See, Makarska, Marchfeld, Metković, Mödling, Munich, Norderstedt, Northampton, Nottingham, Oberhausen, Offenbach am Main, Prague, Quickborn, Romerland, Rüsselsheim am Main, Saratoga Springs, Schenectady, Serfaus, Sisak, Slavonskibrod, Slough, Sursee, Traisen, Tubingen, Tulln, Tuzla, Vannes, Velikagorica, Victoria, Vukovar, Wachau, West Palm Beach, Wienerwald, Wiesbaden, Wiener Neustadt, Wurzburg, Zadar, Lower Traisental

- DTW Weekday Sub-Cluster 2

Abu Dhabi, Bhopal, Dajiawa, Jackson, King Abdullah Economic City, Konstancin-Jeziorna, Porec, Šibenik, Yangzi.

- DTW Weekday Sub-Cluster 3

Aksu, Amiens, Anqiu, Aral, Besançon, Brisbane, Brussels, Bucharest, Changle, Cluj, Cergy, Dayton, Denver, Dijon, Drammen, Fenhuzhen, Glasgow, Gospić, Gothenburg, Hamilton, Hoboken, Indianapolis, Innsbruck, Lillestrøm, Limassol, Lincoln (NE, USA), Linqu, Lund, Luxembourg City, Lucerne, Maastricht, Malta, Milton Keynes, Mulhouse, Namur, Nancy, Nantong, Offenburg, Omaha, Petrolina, Philadelphia, Puebla, Reading, Rennes, Rouen, Salt Lake City, Shengzezhen, St Polten, Taibaozhuang, Toyama, University of South Florida, Valence, Vancouver, Zagreb.

- DTW Weekday Sub-Cluster 4

Albany, Almaty, Aspen, Astana, Atlanta, Augsburg, Austin, Avignon, Batumi, Belfort, Berlin, Bochum, Boise, Boulder, Broward County, Budapest, Buffalo, Calais, Charlotte, Chattanooga, Cincinnati, Clermont-Ferrand, Cologne, Colorado Springs, Columbus, Créteil, Derby, Des Moines, Detroit, Dortmund, Dresden, Duisburg, Düsseldorf, Essen, Fort Worth, Frankfurt, Grodzisk Mazowiecki, Houston, Kaiserslautern, Kansas City, Karlsruhe, Kazan, Klagenfurt, Kona, Konya, Kranj, Lake Neusiedl, Las Vegas, Leipzig, Liverpool, Long Beach (CA, USA), Long Beach (NY, USA), Louisville, Lviv, Madison, Marburg, Marrakech, Melbourne, Memphis, Miami Beach, Milwaukee, Minneapolis, Monash University, Montevideo, Montpellier, Mülheim an der Ruhr, Nashville, New Orleans, Nitra, Nuremberg, Opole, Orlando, Park City, Phoenix, Pioneer Valley, Pittsburgh, Portland, Potsdam, Reykjavik, Richmond (CA, USA), Riga, San Antonio, San Diego, San Jose, Santa Monica, Santander, Shymkent, Stalowa Wola, Saint-Étienne, Stirling, St Petersburg (FL, USA), Szamotuły, Tallinn, Tampa, Tel Aviv, Timisoara, Topeka, Trondheim, Tucson, University of Warwick, Usedom, Virginia, Oakland/Berkeley.

# References

Ahmed, F., Rose, G., Jacob, C., 2010. Impact of weather on commuter cyclist behaviour and implications for climate change adaptation. In: Proceedings of 2010 ATRF Australasian Transport Research Forum. Australian National Audit Office, Canberra, Australia, pp. 1–19.

Bachand-Marleau, J., Lee, B.H.Y. El-Geneidy, AM., 2012. Better understanding of factors influencing likelihood of using shared bicycle systems and frequency of use. Transp. Res. Rec. 2314 (1), 66–71. https://doi.org/10.3141/2314-09.

Berndt, D.J., Clifford, J., 1994. Using Dynamic Time Warping to Find Patterns in Time Series, 12.

Bicycle Sharing Blog, 2019. The Bike-Sharing World Map [Online] Available at: https://www.google.com/maps/d/u/0/viewer?mid=1UxYw9YrwT_R3SGsktJU3D-2GpMU&hl=en&ll=28.192374373465793%2C15.232058864590385&z=4.

Bieliński, T., Kwapisz, A., Ważna, A., 2019. Bike-sharing systems in Poland. Sustainability 11, 2458. https://doi.org/10.3390/su11092458.

Borgnat, P., Abry, P., Flandrin, P., Robardet, C., Rouquier, J.-B., Fleury, E., 2011. Shared bicycles in a city: a signal processing and data analysis perspective. Adv. Complex Syst. 14, 415–438. https://doi.org/10.1142/S0219525911002950.

Brandenburg, C., Matzarakis, A., Arnberger, A., 2007. Weather and cycling – a first approach to the effects of weather conditions on cycling. Meteorol. Appl. 14, 61–67. https://doi.org/10.1002/met.6.

Buck, D., Buehler, R., 2012. Bike Lanes and Other Determinants of Capital Bikeshare Trips [WWW Document]. undefined. URL. paper/Bike-Lanes-and-Other-Determinants-of-Capital-Trips-Buck-Buehler/cbcf3e1361fb1870bdadf06623555f3a0fc39753. accessed 9.18.20.

Buck, D., Buehler, R., Happ, P., Rawls, B., Chung, P., Borecki, N., 2013. Are bikeshare users different from regular cyclists?: A first look at short-term users, annual members, and area cyclists in the Washington, D.C., Region. Transp. Res. Rec. 2387, 112–119. https://doi.org/10.3141/2387-13.

Campbell, A.A., Cherry, C.R., Ryerson, M.S., 2016. Factors influencing the choice of shared bicycles and shared electric bikes in Beijing. Transp. Res. C Emerg. Technol. 67, 399–414. https://doi.org/10.1016/j.trc.2016.03.004.

Çolak, S., Lima, A., González, M.C., 2016. Understanding congested travel in urban areas. Nat. Commun. 7, 10793. https://doi.org/10.1038/ncomms10793.

Corcoran, J., Li, T., Rohde, D., Charles-Edwards, E., Mateo-Babiano, D., 2014. Spatio-temporal patterns of a Public Bicycle Sharing Program: the effect of weather and calendar events. J. Transp. Geogr. 41, 292–305. https://doi.org/10.1016/j.jtrangeo.2014.09.003.

Davis, L.S., 2014. Rolling along the last mile: bike-sharing programs blossom nationwide. Planning 80, 10–16.

DeMaio, P., Gifford, J., 2004. Will smart bikes succeed as public transportation in the United States? J. Public Transp. 7, 1–15. https://doi.org/10.5038/2375-0901.7.2.1.

Edelbrock, C., 1979. Mixture model tests of hierarchical clustering algorithms: the problem of classifying everybody. Multivar. Behav. Res. 14 (3), 367–384. https://doi.org/10.1207/s15327906mbr1403_6. Routledge.

Faghih-Imani, A., Eluru, N., El-Geneidy, A.M., 2014. How land-use and urban form impact bicycle flows: evidence from the bicycle-sharing system (BIXI) in Montreal. J. Transp. Geogr. 41, 306–314. https://doi.org/10.1016/j.jtrangeo.2014.01.013.

Faghih-Imani, A., Anowar, S., Miller, E.J., 2017. Hail a cab or ride a bike? A travel time comparison of taxi and bicycle-sharing systems in New York City. Transp. Res. A Policy Pract. 101, 11–21. https://doi.org/10.1016/j.tra.2017.05.006.

Fan, A., Chen, X., Wan, T., 2019. How have travelers changed mode choices for first/last mile trips after the introduction of bicycle-sharing systems: an empirical study in Beijing, China. J. Adv. Transp. 2019, 5426080. https://doi.org/10.1155/2019/5426080.

Fischer, C.M., Sanchez, C.E., Pittman, M., Milzman, D., Volz, K.A., Huang, H., Gautam, S., Sanchez, L.D., 2012. Prevalence of bicycle helmet use by users of public bikeshare programs. Ann. Emerg. Med. 60, 228–231. https://doi.org/10.1016/j.annemergmed.2012.03.018.

Fishman, E., 2016. Bikeshare: a review of recent literature. Transp. Rev. 36, 92–113. https://doi.org/10.1080/01441647.2015.1033036.

Fishman, E., 2019. Bike Share. Routledge. https://doi.org/10.4324/9781315545127.

Fishman, E., Schepers, P., 2016. Global bike share: what the data tells us about road safety. J. Saf. Res. 56, 41–45. https://doi.org/10.1016/j.jsr.2015.11.007.

Fishman, E., Washington, S., Haworth, N., 2013. Bike share: a synthesis of the literature. Transp. Rev. 33, 148–165. https://doi.org/10.1080/01441647.2013.775612.

Fishman, E., Washington, S., Haworth, N., 2014. Barriers to bikesharing: an analysis from Melbourne and Brisbane. J. Transp. Geogr. 41, 325–337. https://doi.org/10.1016/j.jtrangeo.2014.08.005.

Frade, I., Ribeiro, A., 2014. Bicycle sharing systems demand. Procedia Soc. Behav. Sci. 111, 518–527. https://doi.org/10.1016/j.sbspro.2014.01.085.

Fuller, D., Gauvin, L., Kestens, Y., Daniel, M., Fournier, M., Morency, P., Drouin, L., 2013. Impact evaluation of a public bicycle share program on cycling: a case example of BIXI in Montreal, Quebec. Am. J. Public Health 103, e85–e92. https://doi.org/10.2105/AJPH.2012.300917.

Goodman, A., Cheshire, J., 2014. Inequalities in the London bicycle sharing system revisited: impacts of extending the scheme to poorer areas but then doubling prices. J. Transp. Geogr. 41, 272–279. https://doi.org/10.1016/j.jtrangeo.2014.04.004.

Gu, T., Kim, I., Currie, G., 2019. Measuring immediate impacts of a new mass transit system on an existing bike-share system in China. Transp. Res. A Policy Pract. 124, 20–39. https://doi.org/10.1016/j.tra.2019.03.003.

Guerra, E., Cervero, R., Tischler, D., 2012. Half-mile circle: does it best represent transit station catchments? Transp. Res. Rec. 2276, 101–109. https://doi.org/10.3141/2276-12.

Hosford, K., Fuller, D., Lear, S.A., Teschke, K., Gauvin, L., Brauer, M., Winters, M., 2018. Evaluation of the impact of a public bicycle share program on population bicycling in Vancouver, BC. Prev. Med. Rep. 12, 176–181. https://doi.org/10.1016/j.pmedr.2018.09.014.

Hosford, K., Winters, M., Gauvin, L., Camden, A., Dubé, A.-S., Friedman, S., Fuller, D., 2019. Evaluating the impact of implementing public bicycle share programs on cycling: the International Bikeshare Impacts on Cycling and Collisions Study (IBICCS). Int. J. Behav. Nutr. Phys. Act. 16 https://doi.org/10.1186/s12966-019-0871-9.

Jäppinen, S., Toivonen, T., Salonen, M., 2013. Modelling the potential effect of shared bicycles on public transport travel times in Greater Helsinki: an open data approach. Appl. Geogr. 43, 13–24. https://doi.org/10.1016/j.apgeog.2013.05.010.

Kabra, A., Belavina, E., Girotra, K., 2019. Bike-share systems: accessibility and availability. Manag. Sci. 66, 3803–3824. https://doi.org/10.1287/mnsc.2019.3407.

Kou, Z., Cai, H., 2019. Understanding bike sharing travel patterns: an analysis of trip data from eight cities. Phys. A Stat. Mech. Appl. 515, 785–797. https://doi.org/10.1016/j.physa.2018.09.123.

LDA Consulting, 2013. 2013 Capital Bikeshare Member Survey Report [Online] Available at: http://1105am3mju9f3st1xn20q6ek-wpengine.netdna-ssl.com/wp-content/uploads/2013/05/CapitalBikeshare_2013MemberSurveyReport.pdf.

MacQueen, J., 1967. Some methods for classification and analysis of multivariate observations. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1, pp. 281–297 (14).

Martin, E.W., Shaheen, S.A., 2014. Evaluating public transit modal shift dynamics in response to bikesharing: a tale of two U.S. cities. J. Transp. Geogr. 41, 315–324. https://doi.org/10.1016/j.jtrangeo.2014.06.026.

Mateo-Babiano, I., Bean, R., Corcoran, J., 2016. How does our natural and built environment affect the use of bicycle sharing? Transp. Res. A Policy Pract. 94, 295–307. https://doi.org/10.1016/j.tra.2016.09.015.

Mátrai, T., Tóth, J., 2016. Comparative assessment of public bike sharing systems. Transp. Res. Procedia 14, 2344–2351. https://doi.org/10.1016/j.trpro.2016.05.261.

Médard de Chardon, C., Caruso, G., 2015. Estimating bike-share trips using station level data. Transp. Res. B Methodol. 78, 260–279. https://doi.org/10.1016/j.trb.2015.05.003.

Médard de Chardon, C., Caruso, G., Thomas, I., 2017. Bicycle sharing system 'success' determinants. Transp. Res. A Policy Pract. 100, 202–214. https://doi.org/10.1016/j.tra.2017.04.020.

Midgley, P., 2011. Bicycle-sharing schemes: enhancing sustainable mobility in urban areas. U. N. Dep. Econ. Soc. Aff. 8, 1–12.

Miranda-Moreno, L.F., Nosal, T., 2011. Weather or not to cycle: temporal trends and impact of weather on cycling in an urban environment. Transp. Res. Rec. 2247 (1), 42–52. https://doi.org/10.3141/2247-06.

Mojena, R., 1977. Hierarchical grouping methods and stopping rules: an evaluation. Comput. J. 20, 359–363. https://doi.org/10.1093/comjnl/20.4.359.

Mueller, N., Rojas-Rueda, D., Cole-Hunter, T., de Nazelle, A., Dons, E., Gerike, R., Götschi, T., Int Panis, L., Kahlmeier, S., Nieuwenhuijsen, M., 2015. Health impact assessment of active transportation: a systematic review. Prev. Med. 76, 103–114. https://doi.org/10.1016/j.ypmed.2015.04.010.

Murphy, E., Usher, J., 2015. The role of bicycle-sharing in the city: analysis of the Irish experience. Int. J. Sustain. Transp. 9, 116–125. https://doi.org/10.1080/15568318.2012.748855.

Nankervis, M., 1999. The effect of weather and climate on bicycle commuting. Transp. Res. A Policy Pract. 33, 417–431. https://doi.org/10.1016/S0965-8564(98)00022-6.

Natera, L., Battiston, F., Iñiguez, G., Szell, M., 2020. Data-driven strategies for optimal bicycle network growth. Royal Society Open Science 7 (12), 201130. https://doi.org/10.1098/rsos.201130.

O'Brien, O., Cheshire, J., Batty, M., 2014. Mining bicycle sharing data for generating insights into sustainable transport systems. J. Transp. Geogr. 34, 262–273. https://doi.org/10.1016/j.jtrangeo.2013.06.007.

Pucher, J.R., Buehler, R. (Eds.), 2012. City Cycling 11. MIT Press, Cambridge, MA.

Pucher, J., Garrard, J., Greaves, S., 2011. Cycling down under: a comparative analysis of bicycling trends and policies in Sydney and Melbourne. J. Transp. Geogr. 19, 332–345. https://doi.org/10.1016/j.jtrangeo.2010.02.007.

Roy, A., Nelson, T.A., Fotheringham, A.S., Winters, M., 2019. Correcting bias in crowdsourced data to map bicycle ridership of all bicyclists. Urban Sci. 3, 62. https://doi.org/10.3390/urbansci3020062.

Sarkar, A., Lathia, N., Mascolo, C., 2015. Comparing cities' cycling patterns using online shared bicycle maps. Transportation 42 (4), 541–559. https://doi.org/10.1007/s11116-015-9599-9.

Shaheen, S.A., Guzman, S., Zhang, H., 2010. Bikesharing in Europe, the Americas, and Asia: past, present, and future. Transp. Res. Rec. 2143, 159–167. https://doi.org/10.3141/2143-20.

Shaheen, S., Martin, E., Cohen, A., 2013. Public Bikesharing and Modal Shift Behavior: A Comparative Study of Early Bikesharing Systems in North America.

ter Beek, M.H., Fantechi, A., Gnesi, S., 2014. Challenges in modelling and analyzing quantitative aspects of bike-sharing systems. In: Margaria, T., Steffen, B. (Eds.), Leveraging Applications of Formal Methods, Verification and Validation. Technologies for Mastering Change. ISoLA 2014, Lecture Notes in Computer Science, Vol. 8802. Springer, Berlin, Heidelberg.

Tran, T.D., Ovtracht, N., d'Arcier, B.F., 2015. Modeling Bike Sharing System using Built Environment Factors. In: Procedia CIRP, 7th Industrial Product-Service Systems Conference – PSS, Industry Transformation for Sustainability and Business, 30, pp. 293–298. https://doi.org/10.1016/j.procir.2015.02.156.

Wagstaff, K., Cardie, C., Rogers, S., Schroedl, S., 2001. Constrained K-means Clustering with Background Knowledge, p. Vol. 8.

Woodcock, J., Tainio, M., Cheshire, J., O'Brien, O., Goodman, A., 2014. Health effects of the London bicycle sharing system: health impact modelling study. BMJ 348, g425. https://doi.org/10.1136/bmj.g425.

Worldbank, World Development Indicators, 2021. Urban Population. Available at: https://data.worldbank.org/indicator/SP.URB.TOTL.

Yang, Y., Heppenstall, A., Turner, A., Comber, A., 2019. A spatiotemporal and graph-based analysis of dockless bike sharing patterns to understand urban flows over the last mile. Comput. Environ. Urban. Syst. 77, 101361. https://doi.org/10.1016/j.compenvurbsys.2019.101361.

Zaltz Austwick, M., O'Brien, O., Strano, E., Viana, M., 2013. The structure of spatial networks and communities in bicycle sharing systems. PLoS One 8, e74685. https://doi.org/10.1371/journal.pone.0074685.

Zhang, L., Zhang, J., Duan, Z., 2015. Sustainable bike-sharing systems: characteristics and commonalities across cases in urban China. J. Clean. Prod. 97, 124–133. https://doi.org/10.1016/j.jclepro.2014.04.006.

Zhang, J., Pan, X., Li, M., Yu, P.S., 2016. Bicycle-sharing system analysis and trip prediction. In: 2016 17th IEEE International Conference on Mobile Data Management (MDM). Presented at the 2016 17th IEEE International Conference on Mobile Data Management (MDM), pp. 174–179. https://doi.org/10.1109/MDM.2016.35.

Zhang, Y., Lin, D., Mi, Z., 2019. Electric fence planning for dockless bike-sharing services. J. Clean. Prod. 206, 383–393. https://doi.org/10.1016/j.jclepro.2018.09.215.