Title
Functional data analysis approach for mapping change in time series: A case study using bicycle ridership patterns

Authors
Roy, Avipsa
Nelson, Trisalyn
Turaga, Pavan

Publication Date
2023

DOI
10.1016/j.trip.2022.100752

Transportation Research Interdisciplinary Perspectives 17 (2023) 100752

Contents lists available at ScienceDirect

Transportation Research Interdisciplinary Perspectives

journal homepage: www.sciencedirect.com/journal/transportation-research-interdisciplinary-perspectives

Functional data analysis approach for mapping change in time series: A case study using bicycle ridership patterns

Avipsa Roy [a,b,*], Trisalyn Nelson [c], Pavan Turaga [d,e]

[a] Department of Urban Planning and Public Policy, University of California, Irvine, USA
[b] Institute of Transportation Studies, University of California, Irvine, USA
[c] Department of Geography, University California Santa Barbara, Santa Barbara, USA
[d] School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, USA
[e] School of Arts, Media, and Engineering, Arizona State University, Tempe, USA

ARTICLE INFO

Keywords:
Change detection
Temporal alignment
Functional data analysis
Strava
Bicycling

ABSTRACT

Monitoring change is an important aspect of understanding variations in spatial–temporal processes. Recently, 'big data' on mobility, which are detailed across space and time, have become increasingly available from crowdsourced platforms. New methods are needed to best utilize the high spatial and temporal resolution of such data for monitoring purposes. These data can be considered mappable time series but are challenging to use owing to varying sampling rates and issues of temporal misalignment. We present a methodological framework for change detection from big data captured by crowdsourced fitness app Strava, which addresses misalignment issues in the underlying ridership patterns and maps temporal clusters of bicycling ridership change in the city of Phoenix, AZ between 2017 and 2018 at the street-segment level. Hourly and monthly changes were classified into four clusters for each time period - mapped along with crash density to highlight variations in bicycling ridership. Using spatially and temporally continuous data our study advances the existing approaches to mobility analysis, by using a functional data analysis approach. Our method is reproducible and can be used to expand studies in other cities for monitoring changes directly from crowdsourced ridership data thereby facilitating the decision-making process by practitioners to assess and plan safe bicycle infrastructure.

Introduction

Monitoring change from continuous time-series data is critical for city authorities to understand travel behavior and make targeted decisions related to transportation infrastructure changes that can improve the overall safety of its residents. Through monitoring, policymakers are more prepared to meet rising infrastructure demands (Miranda-Moreno et al. 2013) and ensure accessibility to existing infrastructure more expeditiously (Boss et al., 2018a,b). Change detection is essential to characterize the impacts of sudden fluctuations on overall spatio-temporal processes (Alaya et al. 2020), particularly where we observe changes in the frequency and/or in the intensity across multiple scales. Detection of changes is thus an essential step before performing any descriptive or predictive analysis.

Growth in the availability of crowdsourced GPS-data from smartphones has created an alternative source of high-resolution spatial–temporal data to enable researchers to understand mobility patterns. Although these datasets are biased towards a specific demographic (males between 25 and 45 years of age), they can be used as an indicator of ridership once the bias is accounted for by including geographic covariates (Roy et al. 2019). Crowdsourced fitness apps like Strava (Strava Metro, 2017) have been collecting anonymized bicycling trip data at the minute level, and such data can be used for monitoring change. Such fine-grained trip data can be represented using a functional form, as they are collected continuously over time. However, the major problem with using such data directly for change analysis is that they are often misaligned in time. These alignment issues, if not accounted for, may introduce errors in the analysis for getting accurate change estimates and introduce bias in policy decisions. Several studies in the remote sensing literature have highlighted that a direct pixel-to-pixel comparison from raster data for detecting changes can be challenging (White, 2006) often due to data misregistration issues arising

Abbreviations: DTW, Dynamic Time Warping; FDA, Functional Data Analysis; GPS, Global Positioning Systems; SRVF, Square Root Velocity Function.

* Corresponding author at: Department of Urban Planning and Public Policy, 300 Social Ecology I, Irvine, CA 92697, USA.
E-mail address: avipsar@uci.edu (A. Roy).

https://doi.org/10.1016/j.trip.2022.100752
Received 29 June 2022; Received in revised form 23 December 2022; Accepted 30 December 2022
Available online 10 January 2023

due to multiple platforms capturing the data (Knudsen and Olsen, 2003; Olsen, 2004). Previous studies (Townshend et al., 1992) have pointed out that the registration of data sets to a common framework is an essential precursor to monitoring change and can lead to reducing noise (Stow, 1999) caused due to misalignment.

Initial research (Boss et al., 2018a,b; Nelson and Boots, 2008) has shown that it is possible to detect changes in bicycling ridership patterns between two time periods – and that the changes are representative of actual changes in infrastructure. Such studies quantify the spatial variations of change in ridership using two snapshots in time. Some studies (Yang et al. 2018) have looked at spatial change detection from GPS trajectories however, they ignored the temporal component, whereas other studies (Kang and Aldstadt, 2019) have addressed the scale issue in land-cover change and activity zone detection from social-media platforms (Liu et al. 2019).

To utilize large volumes of raw time-series data, we must identify analytical methods that also account for patterns in changes across multiple scales, as the underlying data generating processes vary both in space and time. There is a lack of a well-defined framework for extracting actionable insights from such big time-series data for change monitoring across scales. To bridge this gap we propose a functional data analysis (FDA) (Ramsay and Silverman, 2005a,b) technique for mapping changes in bicycle ridership patterns.

Functional data analysis (FDA) (Ramsay and Silverman, 2005a,b) is an approach towards modeling time-series data that has very recently started to gain attention. FDA has already found applications in several areas of research including ecology (Bourbonnais et al. 2017, Gurarie et al., 2009), epidemiology (Aston and Kirch, 2012), remote sensing (Bourbonnais et al. 2017), physical activity recognition (Choi et al. 2018)), outlier detection in environmental applications (Torres et al., 2011) and traffic volume forecasting (Wagner-Muns et al. 2017). The basic idea behind FDA is to quantify discrete observations arising from time series in the form of a function that represents the entire measured function as a single observation, and then to draw modeling and/or prediction information from a collection of functional data by applying statistical concepts from multivariate data analysis. Such an approach has the advantage of generating models that can be described by continuous smooth functions, which enables getting accurate estimates of bicycling ridership volumes for change analysis by tackling effective noise reduction in the data through curve smoothing, and irregularities in sampling schedules introduced during the data collection by several users via the Strava app.

We hypothesize that changes in bicycle ridership patterns manifest a complex spatial–temporal phenomenon often emerging as a combined effect of several underlying factors involving route choice, built environment characteristics, infrastructure availability and safety, which are best viewed as observations from a dynamical process. Hence, small changes in any of these factors that might lead to significant changes in ridership volumes due to inherent non-linearity of the observations. Consequently, causing non-linear effects (Anirudh et al., 2015) on the observed time-series of such big data (i.e. bicycling ridership volumes of all street segments in a city) often manifested as misaligned functions which can severely distort the distance metric between time series. Failure to account for these effects, increases the observation variance (Marron et al., 2015), thereby leading to erroneous results for variance-based analysis (like clustering). By removing the phase variations in ridership data through temporal alignment technique like the square root velocity function (SRVF) which we introduced in this study, we make the results of subsequent cluster-based analysis more stable. All cluster analysis is based on variance-based computation. For example, even the popular k-means clustering algorithm is defined based on minimizing variance within each cluster. For semantically similar time-series, if the time-series are affected by time-warp, then the standard Euclidean distance between them increases, causing a larger variance even between otherwise similar time-series. Any variance-based cluster analysis would necessarily be affected by factors that increases the underlying variance between samples. The SRVF framework reduces the component of variance (Grohs et al., 2020; Guo and Srivastava, 2020; Srivastava et al., 2005) that is attributable to time-warping. The novelty of our approach lies in developing a methodological framework for mapping street-network level change in bicycling ridership patterns categorically from crowdsourced timeseries data using functional data analysis tools.

In this work, our goal is to demonstrate a method for detecting and mapping change in data collected continuously through time. To meet this goal we (1) quantify the change in bicycle ridership using a special case of elastic FDA known as the square-root velocity function (SRVF) representation and (2) visualize the temporal changes across hourly and monthly scales. We generate street-segment level maps for different time scales that enable practitioners to make targeted decisions regarding bicycle infrastructure planning.

Data & study area

Our study area is the City of Phoenix (Fig. 1), which lies within the state of Arizona in the USA. Phoenix is the largest metropolitan city in the county of Maricopa in Arizona with a population of 1,563,001 (US Census Bureau, 2015, City of Phoenix). Approximately 1.12 % of the population who commute to their workplace use bicycles as their preferred mode to work with the highest weekday ridership exceeding 270 bicyclists per day (City of Phoenix, Bicycle Master Plan report, 2015). Bias-corrected Strava ridership in Phoenix is representative of nearly 76 % of overall bicycling activities with bicyclist safety along with income and gender being the strongest indicators of overall ridership (Nelson et al., 2021). The city has an entire street network spanning 8,000 km, with approximately 1,140 km of total bicycle lanes that include 960 km of on-street bicycle facility and 190 km of off-street bicycle paths, 42 bicycle and pedestrian bridges/tunnels spanning the entire city (City of Phoenix, Bicycle Master Plan, 2015). The total number of bicycle trips increased from 52,976 to 74,191 between 2017 and 2018 (Strava Metro, 2017).

The City of Phoenix has gathered bicycle ridership from Strava Metro as part of a data acquisition effort by the Maricopa Association of Governments for estimating ridership estimates annually. Strava Metro provides information about anonymized bicycle trips recorded through the Strava fitness app. The data consists of activity counts (i.e. bicycle trips) per segment of transportation infrastructure in the Phoenix region, recorded every minute of the day.

We chose the period between 2017 and 2018 as several minor/major changes took place in the bicycling infrastructure during this period, which enables monitoring how ridership patterns varied before and after the changes were put into effect. Table 1 shows the trip information and the number of total activities recorded in each year. Strava is commonly used by recreational bicyclists which introduces a bias in the overall sampling of bicycle counts, which can be adjusted using additional geographical covariates (Roy et al., 2019) but in dense urban areas correlates with all bicyclists (Boss et al., 2018a,b). The demographics of the Strava users in Phoenix are not representative of the general bicycling population, there are differences in both gender (Table 1) and age. The trends in the Strava data used in this study are similar to age and gender trends of crowdsourced data used in other bicycling studies (Griffin and Jiao, 2015, Romanillos et al. 2016). We also use additional data from the City of Phoenix showing the bicycle crashes representative of the time period of study.

Methods

Our study can be broken into three main objectives – first, we convert Strava ridership volumes to time-series representing it as functional data, second, we then use a temporal alignment technique using square-root-velocity-function (SRVF) (Srivastava et al. 2011) to account for temporal variability and quantify change. Finally, a functional K-means

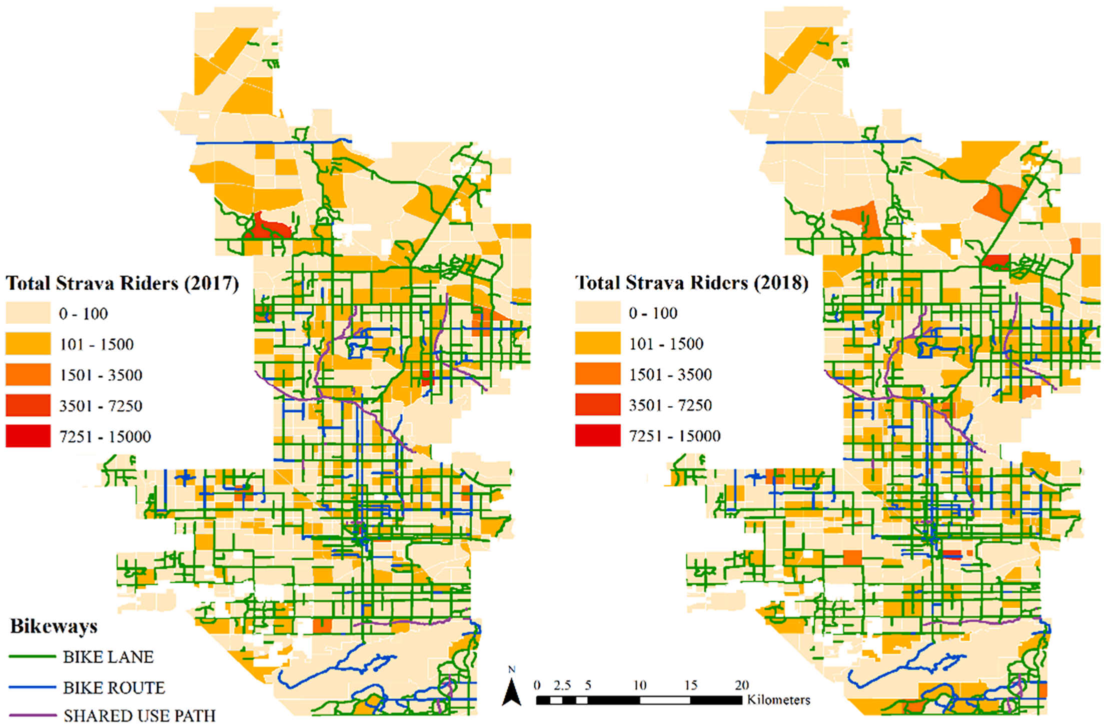

Fig. 1. Map showing the spatial distribution of Strava riders across all traffic analysis zones along with bikeways in the City of Phoenix (2017–2018).

Table 1
Summary of Strava ridership in Phoenix from 2017 to 2018.

Year	No. of commute trips	Total no. of activities	No. of street segments	% Male Strava riders	%Female Strava riders
2017	131,081	1.74 million	78,174	76.9 %	18.4 %
2018	138,714	1.78 million	74,191	76.5 %	19.7 %

Table 2
Features for functional data analysis on Strava ridership between 2017 and 2018.

Name	Operationalization	Time Period	Relevance
Mean ridership	The average number of bicycle trips at each temporal unit	Daily, Monthly	Understand hourly variability in ridership volumes (Brum-Bastos et al, 2019)
Mean Weekday Ridership	The average number of bicycle trips on weekdays at each temporal unit	Daily, Monthly	Weekday peak-period ridership helps identify commute patterns among riders & scale Strava data (Dadashova and Griffin, 2020)
Normalized Total Ridership	The ratio of the total number of bicycle trips in individual temporal unit and sum of riders across all temporal units	Daily, Monthly	Represents the proportion of all activity counts that occurred within that period on each segment. (Boss et al., 2018a,b)

clustering of the change in ridership is used to group street segments into different clusters based on functional means of the change clusters. We further elaborate each step in the following sub sections 3.1 through 3.5.

Converting raw Strava ridership to functional representation

Before detecting changes, it is essential to construct a continuous timeseries from the discrete ridership data collected from Strava fitness app. We convert the raw counts of trips into hourly and monthly timeseries representations, which are referred to as the functional data for our analysis.

The unique identifier of the street segments used in our analysis are typically obtained from OpenStreetMaps and are updated annually. Strava has internally matched these segments using their own map-matching algorithms and provided them for further analysis. We were able to run spatial joins and retain a total of 33,101matching street segments where data from both years 2017 and 2018 were available – missing data were removed prior to fitting SRVF algorithms to the data.

We aggregated the Strava Metro data available at one-minute temporal resolution for the years 2017 and 2018 to compute timeseries of average hourly and monthly trip counts (using weekday ridership data) for each of the 33,101 individual street segments spread throughout the city of Phoenix, which are referred to as "functional curves". The spatial patterns of the total annual ridership (Fig. 1) appear similar across both years, hence we focus on a finer temporal resolution like hourly/monthly to identify noticeable changes in ridership trends by transforming them into timeseries in the functional space.

Next, we normalized the mean of all the ridership variables to represent proportions of ridership in each time unit ranging between 0 and 1 for both hourly and monthly period from the functional curves. The scaled ridership data were then used to for temporal alignment in the selected periods Table 2.

Temporal alignment of Strava ridership volumes using SRVF

The basic idea behind temporal alignment is to transpose the actual count data from the Euclidean space to a functional space such that the phase and amplitude of the functions (Tucker et al., 2012; Lee and Jung, 2016), which are somewhat similar to signals from a sensor, are kept intact but are easier to compare as difference between curves. Ideally raw differences will only account for a discrete change at a single time point and not the entire time-series. The functions therefore need to be smoothed out prior to modeling. Although the changes of interest in Strava ridership may show similarities in feature-space (e.g. increase in ridership by 10 bicyclists versus increase by 100 bicyclists), signals may be temporally misaligned in the functional space due to variability in sampling rates, and temporal windowing operations as well. Due to these issues, we needed to more explicitly model temporal variability, before leveraging popular unsupervised statistical learning frameworks like K-means. We compare the results of quantifying change in ridership before and after temporal alignment to visualize the necessity of this approach.

For temporal alignment, we adopted a square root velocity function (SRVF) representation of the normalized hourly and monthly Strava ridership counts that would rectify temporal misalignment in the ridership data (which is now considered a signal) by separating its phase and amplitude components. SRVF is a pre-processing method that takes a time-series as input and outputs another time-series, with some rate variations factored out. Unlike traditional feature extraction methods like wavelets, SRVF preserves all the information in the original time-series, merely decoupling the observed variation into an amplitude signal and a phase signal. Any further feature extraction method applied to these signals, thus not precluding the use of other methods. SRVF has been used as pre-processing step in previous studies in computer vision with wearable time-series data (Choi et al., 2018) but used Fourier features and machine learning methods like support vector machines for classification problems. The use of SRVF as a pre-processing step was shown to improve end-classifier performance. SRVF inspired operations have also been integrated as differentiable layers in end-to-end deep learning methods, where also we show that addition of SRVF-layers improves the performance of deep networks for human activity classification problems (Lohit et al., 2019) and in another related work SRVF was found to improve the estimation of clinically relevant features related to gait and movement (Amor et al., 2019).

The SRVF method allows the development of proper Riemannian metrics (Srivastava et al. 2011) which are used to achieve a smoothing effect over time-series. It overcomes some limitations of Dynamic Time Warping such as the 'pinching' effect (Marron et al. 2015) which aligns completely different signals to each other by applying a warping function even though their phase and amplitude are not completely in synchronization. We first computed hourly ridership volumes (i.e. the average number of bicycling trips along each street segment) from raw Strava Metro data. For ease of analysis, we defined these hourly ridership volumes across a street segment as our function x. Then, x was converted into a corresponding SRVF representation to compute the Fréchet mean (Srivastava et al., 2011) for each street segment. We illustrate the temporal alignment procedure and change computation using in Algorithm 1. For each street segment, the original ridership function x was aligned using the estimated warping functions. The detailed warping functions are based on the Fischer-Rao metric and Fréchet means illustrated in details in Supplementary Materials S1.

Algorithm 1. *Steps for computing warped signals from Strava time-series for consecutive years.*

Input: Bicycling ridership time-series $x_1, x_2, \ldots, x_n$ for 'n' street segments and their Frèchet means for hourly and monthly scales
Output: Change in warped signals $\widetilde{x_1}, \widetilde{x_2}, , \ldots, \widetilde{x_n}$, for years y_{2017} *and* y_{2018} with varying elasticity coefficient i
for i = 0,0.1,0.2,….0.1.0 do
  for j = 1, 2, …, n do
    $\gamma2017_{i,j}$ = SRVF warp ($x2017_{i,j}$)
    $\gamma2018_{i,j}$ = SRVF warp ($x2018_{i,j}$)
    $[\widetilde{x}\ 2017_{i,j}] = [x2017_{i,j}] \circ \gamma2017_{i,j}$
    $[\widetilde{x}\ 2018_{i,j}] = [x2018_{i,j}] \circ \gamma2018_{i,j}$
    dist2017[i,j] = Euclidean distance($[x\_2017_{i,j}]$, $[\widetilde{x}\ 2017_{i,j}]$])
    dist2018[i,j] = Euclidean distance($[x\_2018_{i,j}]$, $[\widetilde{x}\ 2018_{i,j}]$])
    Temp_data2017[i,j] = $[\widetilde{x}\ 2017_{i,j}]$
    Temp_data2018[i,j] = $[\widetilde{x}\ 2018_{i,j}]$
  end
  Min_index2017[i] = min(dist2017[i,j])
  Min_index2018[i] = min(dist2018[i,j])
  Warped signals2017[i] = Temp_data2017[Min_index2017]
  Warped signals2018[i] = Temp_data2018[Min_index2018]
  Change[i] = Warped signals2018[i] – Warped signals2017[i-1]
end

For each street segment, the original ridership function x was aligned by composition with the estimated warping functions as shown in Equation (1). The detailed warping functions are based on the Fischer-Rao metric and Fréchet means discussed in the paper by Srivastava et al. (2010).

$$[\widetilde{x}] = [x]^{\circ}\gamma \tag{1}$$

Consequently, a new data set was created from which features could be extracted in a sliding window procedure (non-overlapping) with varying window lengths. We chose the warping function that resulted in the best fit. The warping function was computed by solving Equation (2)

$$\gamma^{*} = \underset{\gamma\in\Gamma}{\operatorname{argmin}}\|u - (q^{\circ}\gamma)\sqrt{\dot{\gamma}}\| \tag{2}$$

The γ is the warping function which is solved for using dynamic programming to get the optimal alignment of the curves (Srivastava et al., 2011), u is the Fréchet mean (Srivastava et al., 2011) obtained from the training phase, Γ is referred to as the warping group, and q is the SRVF representations of given functions defined as $q(t) = sign(\dot{f}(t))\sqrt{|\dot{f}(t)|}$, where f is the original timeseries function. The warped function $\widetilde{q}_t$ is given by equation (3).

$$\widetilde{q}_t = \frac{\frac{d}{dt}(f^{\circ}\gamma)(t)}{\sqrt{\left|\frac{d}{dt}(f^{\circ}\gamma)(t)\right|}} = (q^{\circ}\gamma)(t)\sqrt{\dot{\gamma}(t)} = (q,\gamma) \tag{3}$$

We use the nonparametric form of the Fisher-Rao metric (Srivastava et al., 2011) for analyzing SRVFs. In order to align the functions, we define an elastic distance d between two curves representing the bicycle ridership for a street segment on the functional space S given by equation (4). The solution to the optimization over Γ can be solved using dynamic programming.

$$d([q_1],[q_2]) = inf_{\gamma\in\Gamma}\|q_1 - (q_2,\gamma)\| \tag{4}$$

The functional curves were aligned in order to remove unnecessary noise from the raw ridership data so the difference between the curves and the respective mean curve in each year could be compared without altering the phase and amplitude of the functional representations.

We chose the warping function that resulted in the best fit using dynamic programming (Supplementary Materials S1) that minimized an elastic distance d between two curves representing the bicycle ridership for a street segment on the functional space S (Supplementary Materials S1). Finally, the functional curves were aligned in order to remove unnecessary noise from the raw ridership data so the difference between the curves in consecutive years could be compared without altering the phase and amplitude of the functional representations. We repeated the alignment process by fitting functions using different values of the elasticity coefficient 'λ' ranging from 0 to 1 represented as 'i' in Algorithm 1 which controls the amount of warping (Wu & Srivastava, 2011).

Calculating the change in Strava ridership from aligned functions

Once the alignment for both the years was completed using the Fréchet means of each curve, we calculated the functional change (C_i) in ridership patterns for consecutive years as shown in Equation (5) by calculating the difference of the aligned function of each street segment in a specific year from that of the previous year using different values of 'λ' as described in Algorithm 1. The process was repeated for all street segments in the study area to generate an $N \times M$ matrix where N represents each hour of the day and M represents the number of street segments.

Once the alignment for both the years was completed using the Fréchet means of each curve, we generated a mean signature from the aligned data corresponding to the overall hourly and monthly ridership across all street segments for 2017. Next, we quantified the functional change in ridership patterns in consecutive years by calculating the difference of the aligned function of each street segment in a specific year from the functional mean curve for all street segments in the previous year as shown in Equation (5).

$$C_i = \gamma_i^* - \mu_{i-1} \tag{5}$$

The design of the SRVF method in the '*fdasrvf*' R package allows us only to get mean signatures from the group of curves in a single time period. The mean curve of the previous year was used as a baseline of the average ridership trend of all street segments with which to compare the functional difference in ridership volumes for the following year. If the functional shift was higher from the mean in the previous year, it highlighted a high amount of change and no deviation from the functional mean indicated no significant change was observed.

K-Means clustering of change in ridership to categorize street segments

With the change (C_i) computed for each street segment, we ran a functional K-means clustering to group similar and dissimilar streets into 'k' groups. We determined the optimal 'k' for grouping the street segments using the gap statistic method (Tibshirani et al., 2001), based on the within-cluster sum of squares that measures the variability of the observations in each cluster. Once the desired number of 'k' clusters were determined, we visualized these groupings to identify and categorize ridership patterns. The street segments were grouped based on the similarity in functional change 'C_i'. The similarity is determined using a similarity index 'ρ' between two curves c_1 and c_2 (Supplementary Materials S1) which minimizes the within-cluster variance based on the functional change. We also calculated a silhouette score (ranging from -1 to $+1$) for different values of the elasticity coefficient 'λ' using the 'k' chosen clusters and retained the changes obtained with the optimal 'λ' that maximized the silhouette score (Algorithm 2). The silhouette score ensures that the SRVF alignment do not lead to overfitting of the functions and only slight shifts in peaks are accounted for prior to change computation.

Algorithm 2. *Steps for clustering change in ridership using optimal 'λ'*

Input: Changes in SRVF aligned ridership functions X= {x_1,x_2,….,x_n}for 'n' street segments
Output: Silhouette score 's' for 'k' clusters for varying *elasticity coefficient 'i'*
1: *for* i = 0,0.1,0.2,….0.1.0 do
2: k[i] = Kmeans_align (X,k, $\lambda = i$)
3: s[i] = Silhouette(k[*i*]$_{\text{dist}}$, k[*i*]$_{\text{labels}}$)
4: *end*
5: S = Max (s[i])

Mapping clusters to visualize changes

The mean functional change was finally used to categorize street segments into '*k*' groups. We calculated the mean and coefficient of variation of hourly, weekday, and total ridership along with the root mean squared of the functional change of the streets within each cluster. We then generated named categories based on the summarized cluster statistics and visualized the results of the K-Means clustering by color-coding each street segment by a unique color scheme corresponding to the category to which it belongs. Finally, we created a map for the entire city of Phoenix highlighting changes between 2017 and 2018 both at the hourly and monthly scale. To identify the potential causes for the change in ridership in each consecutive year we also incorporated an additional map layer indicating bicycle crash density in the City of Phoenix and infer the reason for changes by overlaying the results. The bicycle crash density was computed using a kernel density estimation technique in ArcGIS with a radius of 5 miles using data related to bicycling crash incidents from the Arizona Department of Transportation's open data portal. We also quantify the bicyclist exposure per cluster as the ratio of the total number of crashes across street segments in each cluster and the average length of all the street segments in the same cluster. The bicyclist exposure further indicates the safety aspects associated with ridership change patterns at the street segment level.

Results

Temporal alignment of hourly Strava ridership with SRVF

The functional curves of the hourly and monthly patterns of the raw ridership volumes are shown in Fig. 2. These curves were then aligned using SRVF to remove inconsistencies and mismatches in phase and amplitude of the functions. Fig. 2 shows the original temporal profiles of ridership for all street segments in our study area for 2017–2018 along with their aligned temporal profiles. Similar profiles have been generated from single bicycling counters (Miranda-Moreno et al. 2013), but using Strava allows much higher spatial resolution (every single segment within the city's street network) along with the temporal richness (every hour of a day during the entire year).

While Strava provides data based on only a sample of riders and there are demographic biases in the app users, research has shown the spatial patterns in this ridership data correlate with bicycle ridership volumes, especially in dense urban areas (Jestico et al. 2016, Boss et al., 2018a,b). However, post SRVF alignment (Fig. 2) when the spikes are realigned in time and the noise is removed, we observe that there is an increase in peak-period trips between 6 and 8 am and then there is a slight decrease in trips between 8 and 9 am – which is more reasonable given the street has attracted more commuters in 2018 than in 2017. We also varied the elasticity coefficient to visualize the effects of warping on original ridership functions and list three scenarios in Fig. 3. As the value of 'λ' varies, the warping functions change -we repeated the process for ten different values of 'λ' from 0 to 1.

The first plot shows the original average hourly ridership volumes represented as functional curves along all street segments. The warping functions (Fig. 3b) are generated by the SRVF algorithm (Algorithm 1) to realign the curves and remove additional noise using the Fisher-Rao metric. Finally, the temporally aligned warped data is generated (Fig. 3c) followed by the mean and standard deviations of both the original (Fig. 3d) and warped (Fig. 3e) functional curves. The final plot (Fig. 3f) shows the overall mean signature of all warped functional curves representing hourly ridership data across all street segments in a year in the city of Phoenix.

Determining functional change in hourly and monthly ridership

The changes in ridership were computed with varying the 'λ' coefficient and calculating year wise functional differences from the aligned curves between 2017 and 2018 for all segments in the study area. Fig. 4 shows the changes in ridership with $\lambda = 0.2$ before and after temporal alignment based on optimal choice of λ (Fig. 6). We chose four different street segments to show the individual functional changes along each of the segments before temporal alignment (Fig. 4a) as well as the warped functional curves (Fig. 4b) after SRVF alignment.

Generating k-means clusters from aligned Strava ridership functions

The changes in ridership curves after alignment were used to generate clusters that could be used as named categories. We varied the number of clusters as shown in Fig. 5 to choose 4 as the optimal value for

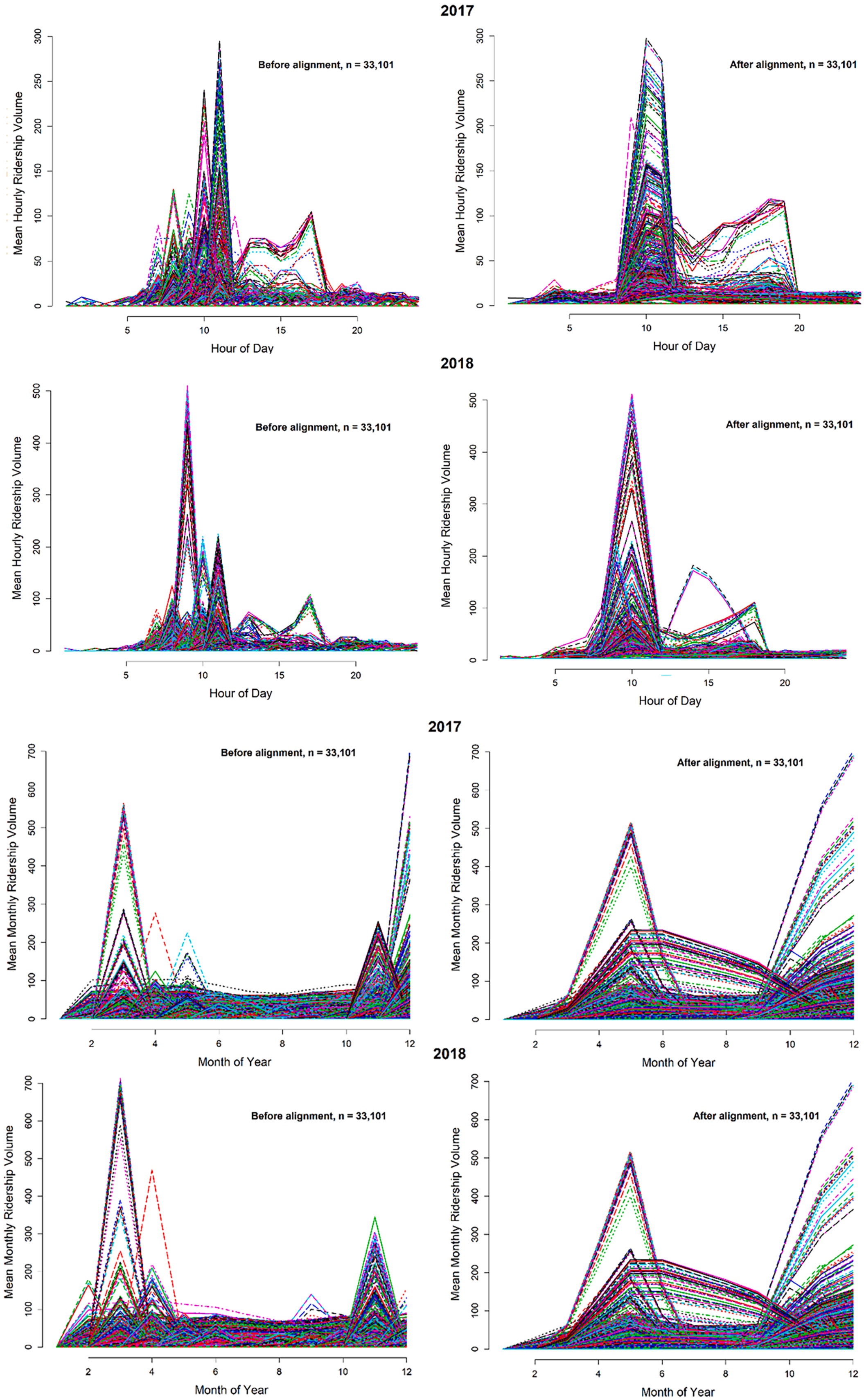

Fig. 2. Functional curves of actual Strava ridership in 2017 and 2018 at the (a) hourly and (b) monthly scales before and after SRVF alignment.

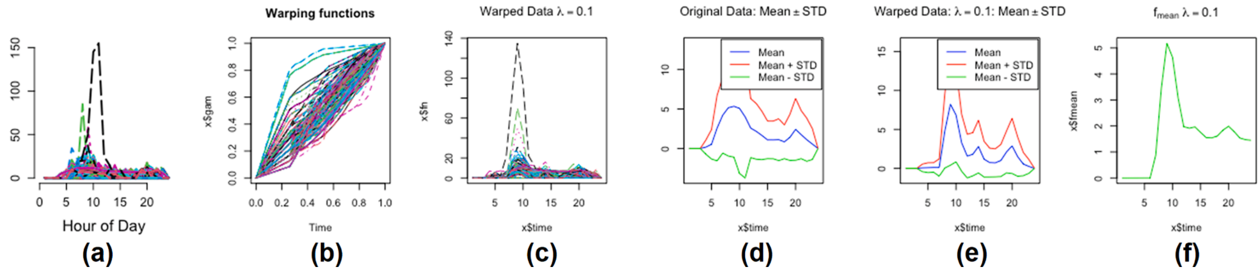

Fig. 3. Functional curves of original and warped Strava ridership using optimal elasticity coefficient 'λ'.

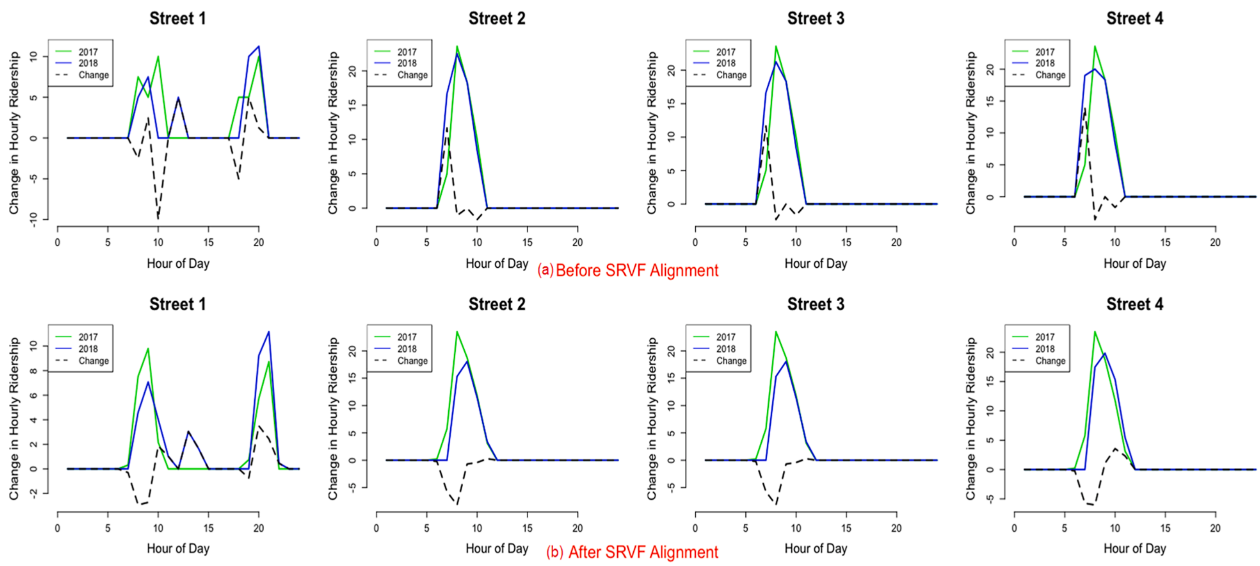

Fig. 4. Functional changes in hourly Strava ridership between 2017 and 2018 using the elasticity coefficient λ = 0.2. Four different street segments are shown to highlight how the temporal alignment warps the peaks to remove noise and helps in calculating changes in hourly ridership.

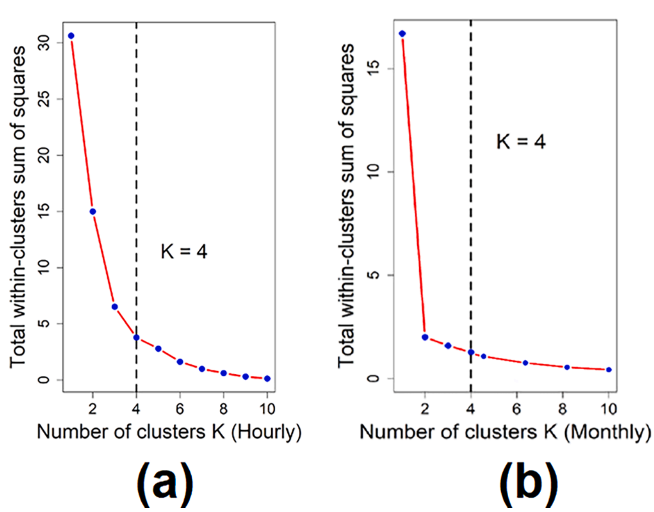

Fig. 5. Determining the optimal number of clusters using for different values of 'k'.

'*k*' based on the within cluster sum of squares using the 'elbow' method i. e. the point after which the within cluster sum of squares start decreasing in a linear fashion. Thus for the given data, we conclude that the optimal number of clusters for the data is 4. After finalizing the number of clusters we varied the elasticity coefficient as shown in Fig. 6 to determine the maximum silhouette score to find the optimal elasticity coefficient.

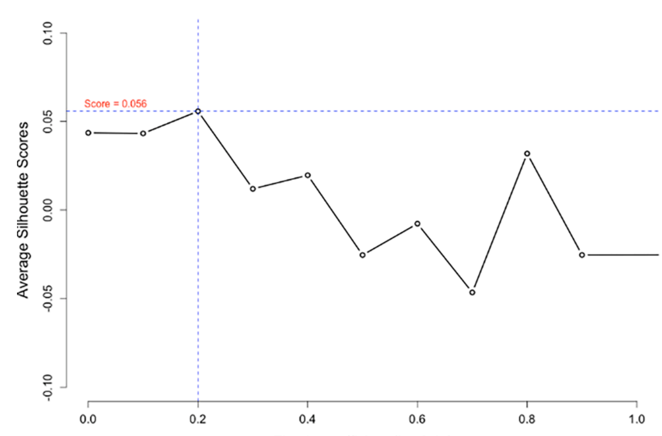

Fig. 6. Determining the optimal value of elasticity coefficient by varying the silhouette score for 'k = 4′ clusters.

Using the maximum silhouette score of 0.056 that was obtained for the elasticity coefficient of 0.2 (Fig. 6), we generated change maps as well as cluster summaries reported in Fig. 7 and Table 3 respectively. The clusters with optimal elastic coefficient 0.2 are grouped into four categories of 'high peak', 'low peak', 'high off-peak', and 'low off-peak' ridership for the hourly changes.

Similarly, for the monthly changes, they are grouped into 'high summer', 'low summer', 'high winter' and 'low winter' ridership categories based on seasonality. In Fig. 7, the grey lines indicate the

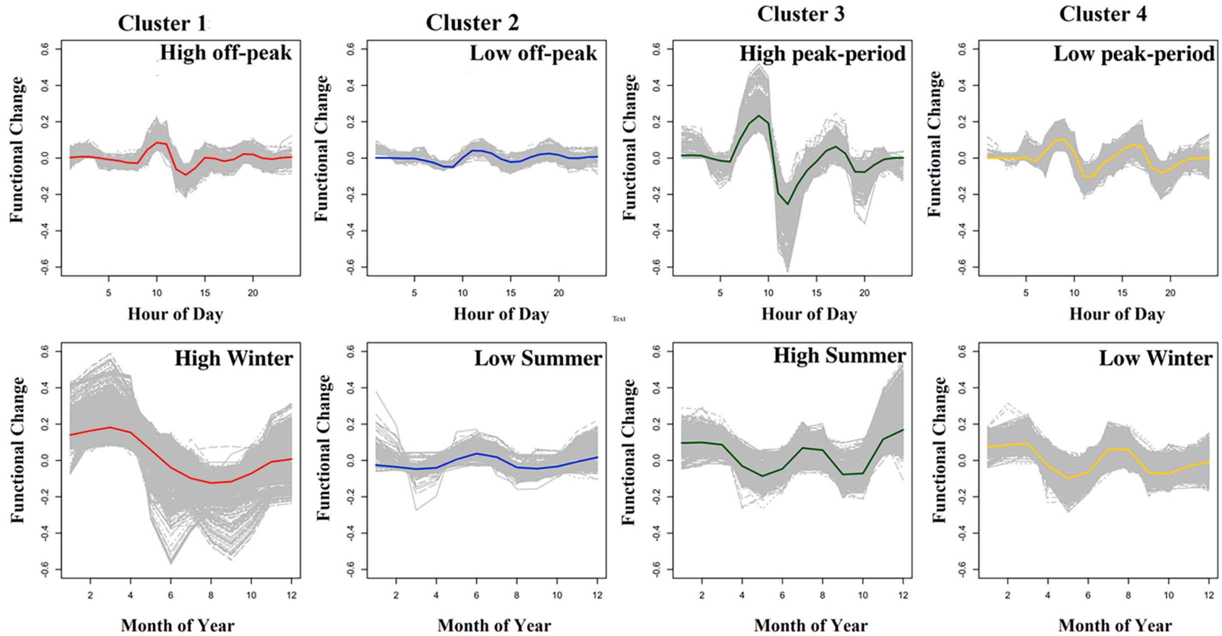

Fig. 7. Clusters showing streets grouped by the functional change in ridership for hourly and monthly changes.

Table 3
Summary of Strava ridership in each of the 4 clusters shown in Fig. 7 based on the functional change in Strava ridership from 2017 to 2018.

Time Period	Cluster	% of Segments	Mean Functional Change	Weekday ridership		Daily/Annual Ridership		Bicyclist Exposure (No. of crashes/ Road length)	Category
				Mean	C.V.	Mean	C.V.		
Hourly	1	31.15	0.04	3.59	1.13	42.16	2.24	0.51	High off-peak
	2	33.64	0.02	2.20	1.22	8.64	1.14	0.59	Low off-peak
	3	6.61	0.11	7.21	1.76	105.23	1.75	0.60	High peak-period
	4	28.60	0.05	4.66	0.92	82.95	1.80	0.35	Low peak-period
Monthly	1	11.33	0.08	5.01	0.94	123.61	0.97	0.41	High Winter
	2	42.53	0.02	4.32	1.12	67.90	2.98	0.60	Low Summer
	3	19.87	0.06	5.03	1.03	105.43	2.02	0.40	High Summer
	4	26.27	0.05	4.88	0.99	118.60	2.11	0.49	Low Winter

functional of the aligned hourly ridership in 2017 and the colored line indicates the overall cluster center of the functional curves in that cluster.

The summary statistics of each cluster shown in Fig. 7 are listed in Table 3, which identifies the percentage of streets (n) in each cluster (k) and the highest, lowest and average daily ridership along with the mean change of ridership (c_k) in each cluster 'k' calculated as the root mean square across all hours of the day and months of the year in each group. We also calculated the bicyclist exposure per cluster as the ratio of the number of crashes that occurred in the streets network and the total length of streets per cluster.

We have also summarized the total hourly trips in a day and the total trips in a month for each cluster in Fig. 8. The peak period clusters show more variability in ridership volumes compared to the off-peak clusters for the hourly trips whereas for the monthly trips the high summer and low winter clusters show more variability. A possible explanation could be that more number of bicyclists ride during peak periods and the overall increase/decrease in ridership is comparatively higher across multiple years compared to off-peak ridership. For the seasonal scenario, the low change in winter and high change in summer.

The categories were assigned based on the mean functional change in each class with the lowest 2 values indicating off-peak hourly ridership and highest 2 values indicating peak-period ridership. Similarly, for monthly ridership we used the lowest two values for low seasonal change and the highest two values for high seasonal change. The classes are subjective and we mapped the change clusters to further visually detect these categories. Existing literature (Brum-Bastos et al., 2019; Nelson et al., 2022) on classifying bicycling regions based on ridership patterns using clustering techniques are used to further support the categorization schemes based on temporal patterns. Further, differentiating classes by magnitude of change was useful for an end goal of stratified sampling and identifying areas for future bike infrastructure planning.

Mapping the change clusters in the Phoenix metropolitan area

We show change maps that highlight the spatial distribution of the clusters based on the functional change in ridership at both hourly and monthly scales (Fig. 9). The different categories of ridership changes (Table 3) indicate higher changes during peak periods at the hourly scale were more prominent near the downtown area. At the monthly scale, more changes were observed in the summer months compared to the winter months. In Fig. 9, average crash density calculated as a kernel density estimate from the total bicyclist collisions data between 2015 and 2017 provided by the City of Phoenix is overlain on maps to demonstrate the variations of ridership change in comparison with bicyclist exposure. The high crash density areas overlap with the high off-peak ridership changes at the hourly scale and high summer ridership changes at the monthly scale (Fig. 9).

To get a more detailed understanding of these changes at the individual street level we gathered infrastructure data from the City of Phoenix and mapped the hourly and monthly changes for 2 specific

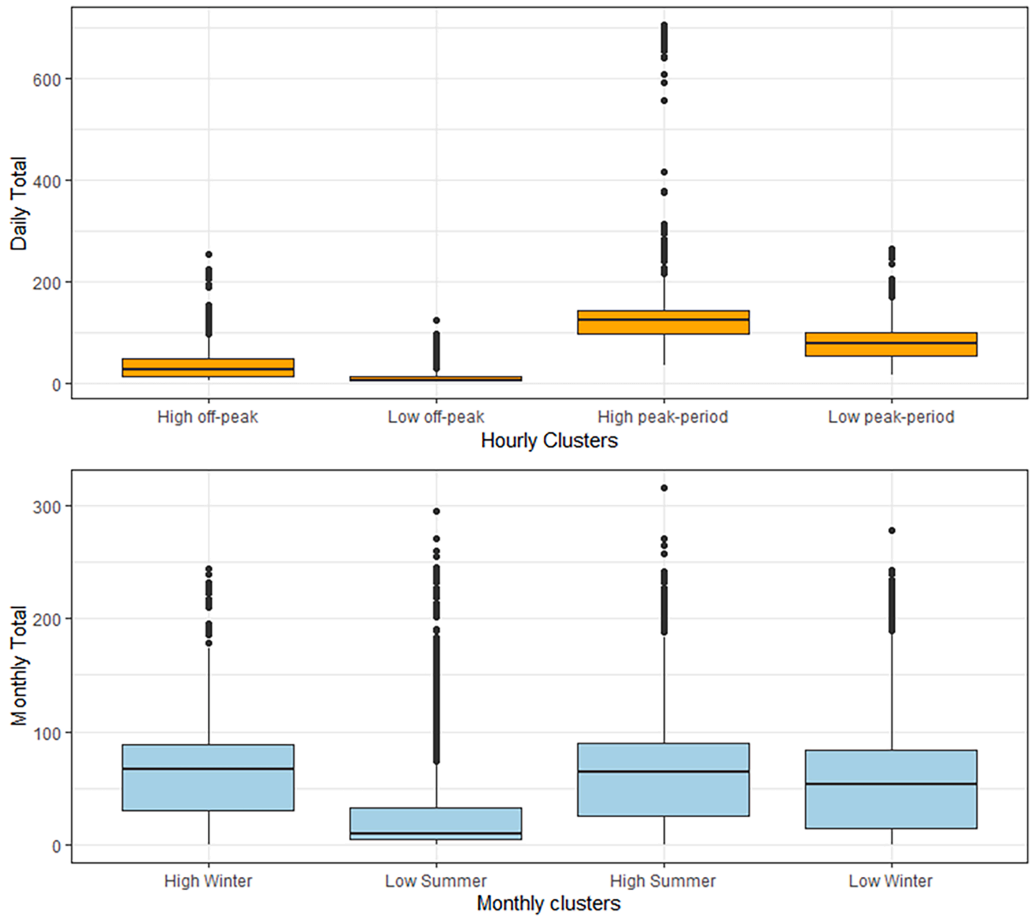

Fig. 8. Boxplots showing the variability of total ridership volumes across each cluster for the hourly and monthly time periods.

segment zones where new bike lanes were introduced in 2017 as part of the bicycle lane improvement plan. Fig. 9 shows the high peak-period changes that occur around Baseline Road between 7th Avenue and 19th Avenue (Fig. 10a) as well as along 24th Street between Baseline Road and Southern Avenue (Fig. 10b) where new bike lanes were introduced. Baseline Road experiences low change in ridership during the winter months (Fig. 10c) and high change during the summer months (Fig. 10d).

Discussion

The increase in popularity of health and fitness apps, such as Strava, has provided a novel source of high-resolution spatio-temporal big data. Strava data have been used to examine where cyclists ride and several studies have examined the use of Strava data as a proxy for ridership volumes (Griffin and Jiao, 2015, Jestico et al. 2016). Heesch and Langdon (2017) used heatmaps and counts of cyclists from Strava data to assess the impact of infrastructure change on bicycling behavior and Boss et al. (2018a,b) use spatial autocorrelation techniques to monitor annual changes in spatial patterns of ridership.

In the current study, we advance change detection approaches by operationalizing an FDA approach which have been used in several other areas of research including ecology (Embling et al. 2012), environmental monitoring (Lee et al. 2015), and climate science (Ballari et al. 2018). The study highlights the use of temporal alignment before detecting changes to account for elastic variations in the functional data (Srivastava et al. 2011). The SRVF technique used in this study can overcome the challenges faced by other methods such as wavelet analysis (Antoniadis et al. 2013) which have not considered real-world scenarios and alignment issues. The SRVF method allows development of proper Riemannian metrics over time-series data, avoids some limitations of DTW-based methods such as the 'pinching' effect (Marron et al., 2015) and helps to tackle the issue of noisy signals in bicycling ridership volumes that helps to refine the spectra of similar signals which are more aligned to each other and lead to a better monitoring of changes in ridership patterns over time.

We generated change maps of bicycling ridership for all street segments in the city of Phoenix post-alignment of the functional curves into 4 categories for both hourly and monthly scales (Table 3). Our results indicate that nearly 32 % of the street segments in Phoenix show a high hourly change in ridership during the off-peak hours (Table 3). There are also 6.6 % of segments that account for high peak-period change in ridership (Table 3). These patterns indicated that improved infrastructure between 2017 and 2018 has led to a major increase in bicycling ridership which is probably due to increased sense of safety and comfort among bicyclists. However, further investigation is needed to better understand the specific infrastructure changes such as reduced speed limits, addition/removal of bike lanes and addition/removal of bicyclist signal at intersections that may lead to these changes. These results are consistent with previous studies (Akar and Clifton, 2009) which show that bicyclists tend to ride more in areas with a high density of bicycling infrastructure as they feel safe biking and have a higher sense of comfort (Teixeira et al. 2020) bicycling in these areas. The average number of bicyclists varies from 83 to 105 during the high and low hourly peak periods (Table 3) and from 9 to 42 bicyclists during off-peak hours (Table 3). The changes during peak-period hourly ridership occur mostly around downtown Phoenix. The reason being commuters use bike lanes and bike paths for their regular commutes around this area the most. The high change areas also overlap with regions of high crash density as more incidents occur owing to exposure to a high volume of motorized

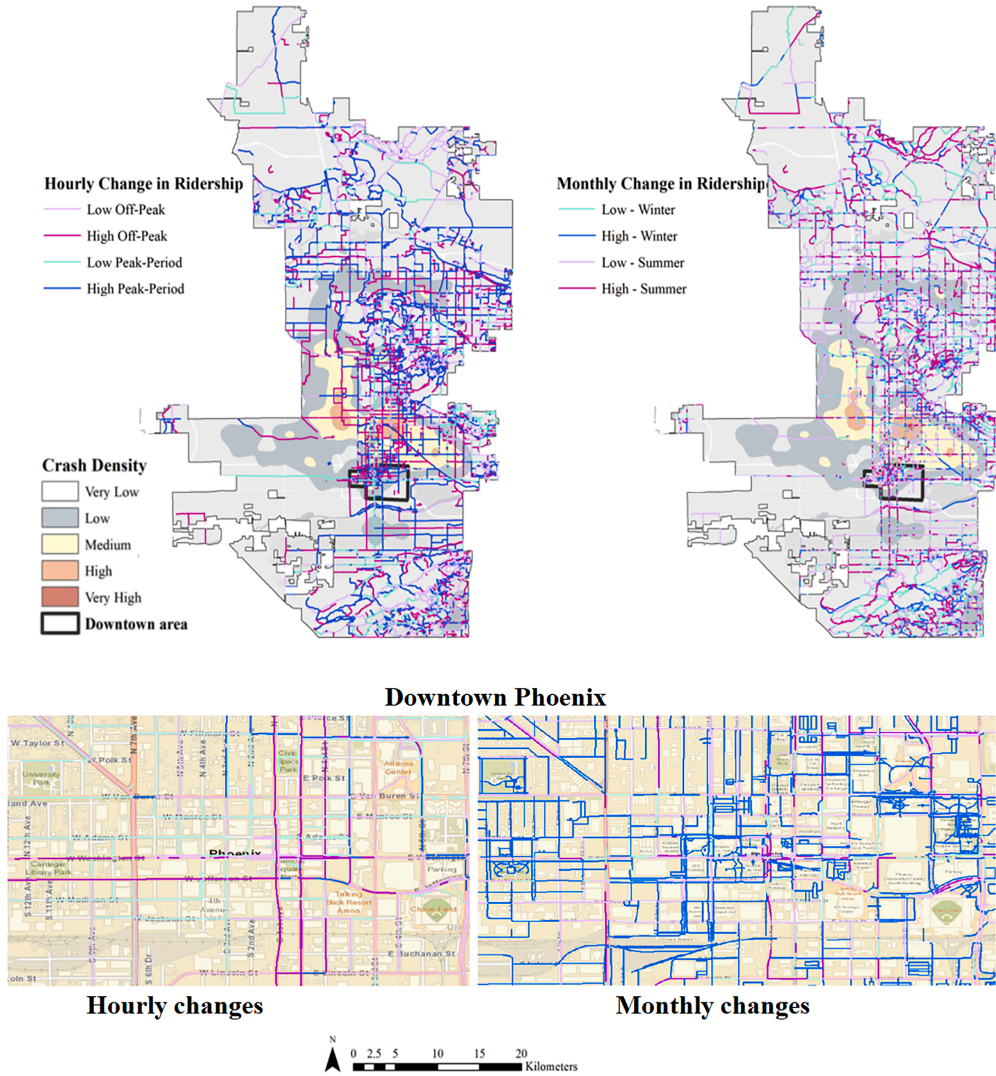

Fig. 9. Maps showing different clusters of the change in hourly and monthly ridership between 2017 and 2018 along with bicyclist crash density.

traffic interspersed with bicycle trips during peak periods in these areas which is consistent with the results of the study by Fournier et al. (2019) and Saha et al. (2018) which highlight that traffic volumes have a positive correlation with bicycle crashes. The bicyclist exposure for high change during peak periods is 0.60 and for off-peak hours is 0.51 (Table 3) indicating a sharp increase in hourly ridership with lack of suitable infrastructure might lead to more crashes and effect bicyclist safety. Previous studies (Pucher and Buehler, 2016, Vanparijs et al. 2020) have shown that North-American cities like Portland, Washington DC, and New York have already improved bicycling safety and increasing bicycling levels by greatly expanding their bicycling infrastructure. Therefore, to reduce exposure in areas with a high change in ridership authorities need to provide more bicycling infrastructure.

At the monthly scale, most of the changes that occur during the winter months which consist of 37.6 % of the street segments (Table 3) (including high and low changes in winter) located mostly on the outskirts of the city with recreational riders making more trips along trails and parks (Fig. 9). However, the changes in summer across the street segments in and around the downtown area (Fig. 9) are comparatively low as those areas are mostly used by commuters that have ridership patterns that are not as impacted by weather, a trend that is consistent with previous studies (Brandenburg et al. 2007, Miranda-Moreno and Nosal, 2011). The remainder of the streets which are used for recreational trips experienced a sharp dip in commutes owing to high temperatures in the summer season (Brandenburg et al., 2007). The high and low changes during summer overlap with high crash density areas as the crashes occur more frequently in and around the high traffic zones specifically near the city center. Surprisingly, 42.5 % of street segments with low change during summer months (Table 3) have a high exposure of 0.60 (Table 3) indicating that the overall risk of crashes in streets with bicycle commuters in and around the city center is typically high (Loidl et al. 2016) throughout the year. Hence, local authorities should invest more in introducing bicycle-friendly infrastructure that reduces exposure in those areas even with a low change in monthly ridership. The monthly change in ridership indicates a total increase or decrease in ridership spanning both years – it may be linked to temperatures but there are additional factors to consider at the ground level – such as street-closures, roadwork, construction of new bike lanes, removal of existing bike lanes that also govern these ridership volumes. Hence it is a combined effect of all these factors that lead to the ultimate manifestation of functional change in the summer/winter months. In the future work, we aim to integrate temperature data to further capture actual impacts of heat exposure, wind effects etc. on total ridership change. Previous studies (Prato, Kaplan, Rasmussen, & Hels, 2016) have found where there are more bicycling trips there are lower bicycle crash rates since bicyclists tend to choose safer routes and motor vehicle drivers are forced to pay attention to non-motorized travelers. Strava data have also been used in bicyclist safety studies to map exposure at finer spatial

Fig. 10. Street-level change maps of hourly and monthly ridership along segments where new bike infrastructure was introduced by the City of Phoenix(In the street segments along (a) Baseline Road: 7th Avenue to 19th Avenue and along (b) 24th Street: Baseline Road to Southern Avenue, the City of Phoenix introduced new bike lanes which led to high changes in hourly peak and off-peak period ridership along that stretch. The monthly ridership changes along (c) Baseline Road: 7th Avenue to 19th Avenue is mixed with high summer and low winter ridership change and along (d) 24th Street: Baseline Road to Southern Avenue, experienced low change in ridership during summer and winter seasons.).

scales (Ferster et al., 2021) which highlighted that largest incident hotspots occurred outside of the central downtown area and the high exposure areas typically occur on multi-use paths where incidents are under-reported (Jestico et al., 2016). It has also been suggested in prior studies (Branion-Calles et al., 2019) (that increasing the availability of bike infrastructure can increase perceptions of bicyclist safety in mid-sized North American cities. Fig. 9 shows high change clusters (which indicates a decrease/increase) in areas outside of downtown area that support previous results in this direction.

The change maps shown in Fig. 6 are categorized based on the continuous temporal changes derived from our functional change (Table 3) technique that captures fine-grained changes during all hours of a day and each month of the year. Our results also highlight the importance of considering multi-scale temporal changes in bicycling when infrastructure changes in a city. Hourly changes can be useful to detect commute patterns (e.g., Heinen et al. 2011) throughout the day whereas monthly changes give a summary of seasonal ridership patterns (e.g., Jestico et al. 2016) in the city. Practitioners can use maps at both scales (Fig. 9) to identify regions that need improved infrastructure for tackling daily bicycle traffic as well as make long term plans for future interventions that assure bicyclist safety within a region to quantify the cumulative impact over time.

Previous studies that have evaluated the mapped change in ridership focused on the comparison of two discrete snapshots in time (Boss et al., 2018a,b). Snapshot approaches remove much of the temporal detail that could be valuable for understanding more nuanced changes. As well, the snapshot selected for evaluating change is often subjective. Our study overcomes the challenge of possible information loss by discrete snapshot approaches by combining the changes at hourly and monthly scales from continuous time-series data. The increasing availability of big data from urban sensing technologies such as Strava has enabled monitoring change from spatial-temporal processes such as bicycle ridership continuously across multiple scales utilizing the functional data analysis framework. We have developed the technique as a way for policymakers to visually represent change through maps and identify infrastructure needs of a city. Our method is a step forward towards easing the process of detecting changes from big data by policymakers using visual approaches using an FDA framework. The clusters shown in Fig. 8 are specific to the study area and will change in different geographic contexts but our method is broadly applicable and can be nuanced for any city using the change detection framework developed in this study based upon availability of good quality timeseries data. The results from this study would be a good starting point for planners to make informed decisions on investments for modifying an existing infrastructure or installing new infrastructure such as paved bike lanes, adding a new bike path, increasing width of lanes as well as reducing the number of motor vehicle lanes.

Conclusion

Big crowdsourced data pose numerous challenges ranging from the extraction of actionable information (Yang et al., 2017) to temporal misalignment (Choi et al. 2018) and bias on app usage (Roy et al. 2019). The need for more accurate and reliable understanding and predictions requires improvements to algorithms that can recognize data inaccuracies, sampling errors. Efficiently integrating big data from different spatial–temporal scales is critical for many research fields beyond transportation including earth system sciences (Deser et al., 2020). Our research opens a new avenue for using functional approaches to data preprocessing and analysis across multiple scales from big spatio-temporal data. Functional approaches help in identifying the latent spatial–temporal patterns, which cannot be observed directly, through a data-driven perspective. Inferring such pattern changes from a raw noisy stream of individual trips is a rather non-trivial task and an ongoing area of geographic information science research. Developing generalized techniques as outlined in our study, to automatically detect pattern changes from individual-level longitudinal spatial–temporal data, is therefore critical to developing behavior models that are adaptive over time.

While using big spatio-temporal data it is essential to account for nonlinear warpings for proper alignment and co-registration of functional curves. Our method highlights the use of square-root velocity functions to overcome such challenges and detect changes in hourly and monthly scales from functional data. From a broader perspective, this paper contributes to debates in time geography based on the theoretical foundation on how time and space constitute social life from the scale of individuals (Hägerstrand, 1985). Considering previous research (Kwan, 2002, Miller, 2005, Long and Nelson, 2013, Kwan and Neutens, 2014) that highlight the role of underlying time and scale issues in geography, this paper builds a framework for analyzing change from real-world data at fine-grained scales and contributes to the field of urban analytics from a methodological perspective which can help policymakers.

Data and codes availability statement

A sample dataset and reproducible code that support the findings of this study are openly available in Figshare (Roy et al., 2020) at https://doi.org/10.6084/m9.figshare.13171862.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

A code repository is made available through Figshare

Acknowledgments

Pavan Turaga was supported by an NSF Career grant 1452163 at Arizona State University. The authors would like to thank the University of California for supporting the research as well as Strava Metro and the City of Phoenix for their assistance in data collection, sharing and support throughout the project.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.trip.2022.100752.

References

Akar, G., Clifton, K.J., 2009. Influence of individual perceptions and bicycle infrastructure on decision to bike. Transp. Res. Rec. 2140 (1), 165–172.
Alaya, M.A., Ben, C.T., Dabo-Niang, S., Chebana, F., Ouarda, T.h.B.M.J., 2020. Change point detection of flood events using a functional data framework. Adv. Water Resour. 137, 103522.
Amor, B.B., Srivastava, A., Turaga, P., Coleman, G., 2019. A framework for interpretable full-body kinematic description using geometric and functional analysis. IEEE Trans. Biomed. Eng. 67 (6), 1761–1774.
Anirudh, R., Turaga, P., Su, J., & Srivastava, A., 2015. Elastic functional coding of human actions: From vector-fields to latent variables. in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 3147-3155).
Antoniadis, A., Brossat, X., Cugliari, J., Poggi, J.-M., 2013. Clustering functional data using wavelets. Int. J. Wavelets Multiresolut. Inf. Process. 11 (01), 1350003.
Aston, John AD, and Claudia Kirch. "Detecting and estimating changes in dependent functional data." Journal of Multivariate Analysis 109 (2012): 204-220.In-person at Einstein Hi MDea.
Ballari, D., Giraldo, R., Campozano, L., Samaniego, E., 2018. Spatial functional data analysis for regionalizing precipitation seasonality and intensity in a sparsely monitored region: Unveiling the spatio-temporal dependencies of precipitation in Ecuador. Int. J. Climatol. 38 (8), 3337–3354.
Boss, D., Nelson, T., Winters, M., 2018a. Monitoring city wide patterns of cycling safety. Accid. Anal. Prev. 111, 101–108.
Boss, D., Nelson, T., Winters, M., Ferster, C.J., 2018b. Using crowdsourced data to monitor change in spatial patterns of bicycle ridership. J. Transp. Health 226–233.
Bourbonnais, M.L., Nelson, T.A., Stenhouse, G.B., Wulder, M.A., White, J.C., Hobart, G. W., Hermosilla, T., Coops, N.C., Nathoo, F., Darimont, C., 2017. Characterizing spatial-temporal patterns of landscape disturbance and recovery in western Alberta, Canada using a functional data analysis approach and remotely sensed data. Eco. Inform. 39, 140–150.
Brandenburg, C., Matzarakis, A., Arnberger, A., 2007. Weather and cycling—a first approach to the effects of weather conditions on cycling. Meteorol. Appl. 14 (1), 61–67.
Branion-Calles, M., Nelson, T., Fuller, D., Gauvin, L., Winters, M., 2019. Associations between individual characteristics, availability of bicycle infrastructure, and city-wide safety perceptions of bicycling: A cross-sectional survey of bicyclists in 6 Canadian and US cities. Transp. Res. Part A: policy and practice 123, 229–239.
Brum-Bastos, Vanessa, Ferster, Colin J., Nelson, Trisalyn, Winters, Meghan, 2019. Where to put bike counters? Stratifying bicycling patterns in the city using crowdsourced data. Transp. Findings.
Choi, Hongjun, Wang, Qiao, Toledo, Meynard, Turaga, Pavan, Buman, Matthew, Srivastava, Anuj, 2018. Temporal alignment improves feature quality: an experiment on activity recognition with accelerometer data. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 349–357.
City of Phoenix, Bicycle Master Plan report, 2015 https://www.phoenix.gov/streetssite/Documents/Bicycle%20Master%20Plan/2014bikePHX_Final_web.pdf. Accessed October 7, 2020.
Dadashova, B., Griffin, G.P., 2020. Random parameter models for estimating statewide daily bicycle counts using crowdsourced data. Transp. Res. Part D: Transp. Environ. 84, 102368.
Deser, C., Lehner, F., Rodgers, K.B., Ault, T., Delworth, T.L., DiNezio, P.N., Ting, M., 2020. Insights from Earth system model initial-condition large ensembles and future prospects. Nat. Clim. Chang. 10 (4), 277–286.
Embling, C.B., Illian, J., Armstrong, E., van der Kooij, J., Sharples, J., Camphuysen, K. CJ., Scott, B.E., 2012. Investigating fine-scale spatio-temporal predator–prey patterns in dynamic marine ecosystems: a functional data analysis approach. J. Appl. Ecol. 49 (2), 481–492.
Ferster, C., Nelson, T., Laberee, K., Winters, M., 2021. Mapping bicycling exposure and safety risk using Strava Metro. Applied geography 127, 102388.
Fournier, N., Christofa, E., Knodler Jr, M.A., 2019. A mixed methods investigation of bicycle exposure in crash rates. Accid. Anal. Prev. 130, 54–61.
Griffin, G.P., Jiao, J., 2015. Where does bicycling for health happen? Analysing volunteered geographic information through place and plexus. J. Transp. Health 2 (2), 238–247.
Grohs, P., Holler, M., Weinmann, A. (Eds.), 2020. Handbook of Variational Methods for Nonlinear Geometric Data. Springer Nature.
Guo, X., Srivastava, A., 2020. Shape analysis of functional data. In: Handbook of Variational Methods for Nonlinear Geometric Data. Springer, Cham, pp. 379–394.

Gurarie, E., Andrews, R.D., Laidre, K.L., 2009. A novel method for identifying behavioural changes in animal movement data. Ecol. Lett. 12 (5), 395–408.
Hägerstrand, T., 1985. Time geography: focus on the corporeality of man, society and environment. In: Aida, S. (Ed.), The Science and Praxis of Complexity. The United Nations University, Tokyo, pp. 193–216.
Heesch, K.C., Langdon, M., 2017. The usefulness of GPS bicycle tracking data for evaluating the impact of infrastructure change on bicycling behaviour. Health Promot. J. Austr. 27 (3), 222–229.
Heinen, E., Maat, K., Van Wee, B., 2011. Day-to-day choice to commute or not by bicycle. Transp. Res. Rec. 2230 (1), 9–18.
Jestico, B., Nelson, T., Winters, M., 2016. Mapping ridership using crowdsourced cycling data. J. Transp. Geogr. 52, 90–97.
Kang, J.-Y., Aldstadt, J., 2019. Using multiple scale spatio-temporal patterns for validating spatially explicit agent-based models. Int. J. Geogr. Inf. Sci. 33 (1), 193–213.
Knudsen, T., Olsen, B.P., 2003. Automated change detection for updates of digital map databases. Photogramm. Eng. Remote Sens. 69 (11), 1289–1296.
Kwan, M.-P., 2002. Time, information technologies, and the geographies of everyday life. Urban Geogr. 23 (5), 471–482. https://doi.org/10.2747/0272-3638.23.5.471.
Kwan, M.-P., Neutens, T., 2014. Space-time research in GIScience. Int. J. Geogr. Inf. Sci. 28 (5), 851–854. https://doi.org/10.1080/13658816.2014.889300.
Lee, S., & Jung, S., 2016. Combined analysis of amplitude and phase variations in functional data. arXiv preprint arXiv:1603.01775.
Lee, D.-J., Zhu, Z., Toscas, P., 2015. Spatio-temporal functional data analysis for wireless sensor networks data. Environmetrics 26 (5), 354–362.
Liu, X., Huang, Q., Gao, S., 2019. Exploring the uncertainty of activity zone detection using digital footprints with multi-scaled DBSCAN. Int. J. Geogr. Inf. Sci. 33 (6), 1196–1223.
Lohit, S., Wang, Q., Turaga, P., 2019. Temporal transformer networks: Joint learning of invariant and discriminative time warping. In: Proceedings - 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2019, pp. 12418–12427.
Loidl, M., Wallentin, G., Wendel, R., Zagel, B., 2016. Mapping bicycle crash risk patterns on the local scale. Safety 2 (3), 17.
Long, J.A., Nelson, T.A., 2013. A review of quantitative methods for movement data. Int. J. Geogr. Inf. Sci. 27 (2), 292–318. https://doi.org/10.1080/13658816.2012.682578.
Marron, J.S., Ramsay, J.O., Sangalli, L.M., Srivastava, A., 2015. Functional data analysis of amplitude and phase variation. Stat. Sci. 468–484.
Miller, H.J., 2005. A measurement theory for time geography. Geogr. Anal. 37 (1), 17–45. https://doi.org/10.1111/gean.2005.37.issue-1.
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.
Miranda-Moreno, L.F., Nosal, T., Schneider, R.J., Proulx, F., 2013. Classification of bicycle traffic patterns in five North American Cities. Transp. Res. Rec. 2339 (1), 68–79.
Nelson, T.A., Boots, B., 2008. Detecting spatial hot spots in landscape ecology. Ecography 31 (5), 556–566.
Nelson, T.A., Ferster, C., Roy, A., Winters, M., 2022. Bicycle streetscapes: a data driven approach to mapping streets based on bicycle usage. Int. J. Sust. Transport. 1–11.
Nelson, T., Roy, A., Ferster, C., Fischer, J., Brum-Bastos, V., Laberee, K., Winters, M., 2021. Generalized model for mapping bicycle ridership with crowdsourced data. Transp. Res. Part C: Emerg. Technol. 125, 102981.
Olsen, B.P., 2004. Automatic change detection for validation of digital map databases. Int. Arch. Photogrammetry Remote Sens. 30, 569–574.
Prato, C.G., Kaplan, S., Rasmussen, T.K., Hels, T., 2016. Infrastructure and spatial effects on the frequency of cyclist-motorist collisions in the Copenhagen Region. J. Transp. Saf. Security 8 (4), 346–360.
Pucher, John, and Ralph Buehler. "Safer cycling through improved infrastructure." (2016): 2089-2091.
Ramsay, J.O., Silverman, B.W., 2000. Functional data analysis. Springer.
Ramsay, J.O., Silverman, B.W., 2005b. Functional data analysis. Springer, New York, NY.
Romanillos, G., Austwick, M.Z., Ettema, D., De Kruijf, J., 2016. Big data and cycling. Transp. Rev. 36 (1), 114–133.
Roy, A., Nelson, T.A., Stewart Fotheringham, A., Winters, M., 2019. Correcting bias in crowdsourced data to map bicycle ridership of all bicyclists. Urban Sci. 3 (2), 62.
Roy, A., Turaga, P., Nelson, T., 2020. Data and code for FDA analysisFunctional Data Analysis for Spatial-Temporal Change Detection. figshare. Software. https://doi.org/10.6084/m9.figshare.13171862.
Saha, D., Alluri, P., Gan, A., Wanyang, W.u., 2018. Spatial analysis of macro-level bicycle crashes using the class of conditional autoregressive models. Accid. Anal. Prev. 118, 166–177.
Srivastava, Anuj, Wei Wu, Sebastian Kurtek, Eric Klassen, and James Stephen Marron, 2011. "Registration of functional data using Fisher-Rao metric." arXiv preprint arXiv:1103.3817.
Srivastava, A., Joshi, S.H., Mio, W., Liu, X., 2005. Statistical shape analysis: Clustering, learning, and testing. IEEE Transactions on pattern analysis and machine intelligence 27 (4), 590–602.
Srivastava, A., Klassen, E., Joshi, S.H., Jermyn, I.H., 2010. Shape analysis of elastic curves in euclidean spaces. IEEE transactions on pattern analysis and machine intelligence 33 (7), 1415–1428.
Stow, D.A., 1999. Reducing the effects of misregistration on pixel-level change detection. Int. J. Remote Sens. 20 (12), 2477–2483.
Strava Metro, 2017. Comprehensive User Guide: Version 5.01. Strava LLC. San Francisco, California. http://metro.strava.com/wp-content/uploads/2017/04/StravaMetro_5.01_User_Guide.pdf. Accessed October 7, 2020.
Teixeira, I.P., Rodrigues, A.N., da Silva, T., Schwanen, G.G., Manzato, L.D., Zeile, P., Dekoninck, L., Botteldooren, D., 2020. Does cycling infrastructure reduce stress biomarkers in commuting cyclists? A comparison of five European cities. J. Transp. Geogr. 88, 102830.
Tibshirani, R., Walther, G., Hastie, T., 2001. Estimating the number of clusters in a data set via the gap statistic. J. R. Stat. Soc. Ser. B (Stat Methodol.) 63 (2), 411–423.
Torres, J.M., Nieto, P.G., Alejano, L., Reyes, A.N., 2011. Detection of outliers in gas emissions from urban areas using functional data analysis. J. Hazard. Mater. 186 (1), 144–149.
Townshend, J.R., Justice, C.O., Gurney, C., McManus, J., 1992. The impact of misregistration on change detection. IEEE Trans. Geosci. Remote Sens. 30 (5), 1054–1060.
Tucker, J.D., Wu, W., Srivastava, A., 2012. Generative Models for Function Data using Phase and Amplitude Separation. Comput. Stat. Data Anal. https://doi.org/10.1016/j.csda.2012.12.001.
US Census Bureau, American Community Survey 2015, https://www.phoenixopendata.com/dataset/phoenix-az-demographic-data/resource/6f460cd1-d0aa-4005-aadb-c371772cbd7b. Accessed October 7, 2020.
Vanparijs, J., Jelle Van Cauwenberg, L., Panis, I., Van Hecke, E., Gillis, D., Gautama, Sidharta, Meeusen, R., de Geus, Bas, 2020. Cycling exposure and infrastructural correlates in a Flemish adolescent population. J. Transp. Health 16, 100812.
Wagner-Muns, I.M., Guardiola, I.G., Samaranayke, V.A., Kayani, W.I., 2017. A functional data analysis approach to traffic volume forecasting. IEEE Trans. Intell. Transp. Syst. 19 (3), 878–888.
White, R., 2006. Pattern based map comparisons. J. Geogr. Syst. 8 (2), 145–164.
Wu, W., Srivastava, A., 2011. An information-geometric framework for statistical inferences in the neural spike train space. J. Comput. Neurosci. 31 (3), 725–748.
Yang, C., et al., 2017b. Utilizing cloud computing to address big geospatial data challenges. Comput., Environ. Urban Syst. 61, 120–128. https://doi.org/10.1016/j.compenvurbsys.2016.10.010.
Yang, X., Tang, L., Stewart, K., Dong, Z., Zhang, X., Li, Q., 2018. Automatic change detection in lane-level road networks using GPS trajectories. Int. J. Geogr. Inf. Sci. 32 (3), 601–621.