Florida International University
FIU Digital Commons

6-19-2019

Comparing the Spatial and Temporal Activity Patterns between Snapchat, Twitter and Flickr in Florida

Levente Juhasz
Florida International University, ljuhasz@fiu.edu

Hartwig H. Hochmair
University of Florida

Follow this and additional works at: https://digitalcommons.fiu.edu/gis

Part of the Geographic Information Sciences Commons, and the Spatial Science Commons

Recommended Citation

Comparing the Spatial and Temporal Activity Patterns between Snapchat, Twitter and Flickr in Florida

Levente Juhász[1] and Hartwig Hochmair[2]

[1]Florida International University, Miami, USA
[2]University of Florida, Fort Lauderdale, USA

Abstract

Social media services generate enormous amounts of spatiotemporal data that can be used to characterize and analyse user activities and social behaviour. Although crowd-sourced data have the advantage of comprehensive spatial and temporal coverage compared to data collected in more traditional ways, the various social media platforms target different user groups, which leads to user selection bias. Since data from social media platforms are used for a variety of geospatial applications, understanding such differences and their implications for analysis results is important for geoscientists. Therefore, this research analyses differences in spatial and temporal contribution patterns to three online platforms, namely Flickr, Twitter and Snapchat, over a six-week period in Florida. For the comparison of spatial contribution patterns, a set of negative binomial regression models are estimated to identify which socio-economic factors and characteristics of the built and natural environments are associated with contribution activities. The contribution differences observed are discussed in light of the targeted user groups and different purposes of the three platforms.

Keywords:

social media, spatiotemporal data, Flickr, Twitter, Snapchat, spatial analysis

1   Introduction

The past decade has seen a rapid increase in the amount of shared social media content, leading to billions of georeferenced data points and a large collection of images. Data from social media and photo-sharing websites have been widely used for the study of human mobility, behavioural trends, and information flow within social networks (Girardin, Calabrese, Fiore, Ratti & Blat, 2008; Hawelka et al., 2014; Takhteyev, Gruzd & Wellman, 2012). Despite their wide range of geo-applications, social media platforms experience user selection and geographical bias, which affect data quality and validity. Understanding differences in user contribution behaviour is necessary for the assessment of data validity, accuracy and representativeness (Li, Goodchild & Xu, 2013). This paper analyses

spatiotemporal activity patterns of geographic data generated on Flickr, Snapchat and Twitter for Florida between December 2018 and January 2019, and has the following two objectives:

1. To compare the temporal distribution of Flickr images, snaps, tweets, and tweets with linked media

2. To estimate a set of negative binomial regression models to predict activity rates per areal unit for each of the four data collections, based on socio-economic factors and the built and natural environments.

Outcomes of these two objectives will help to better understand user contribution patterns and the nature of user bases for the different platforms. Although our study is concerned with the influence of the demographics of local residents on contribution patterns, it also aims to capture activities from any type of user, such as tourists, and not just local residents. Hence, no users are removed from the analysis other than for technical reasons (e.g. to eliminate bots).

A large body of previous research has already analysed user selection bias and the quality of social media data. One study, for example, identified tourist attractions as hotspots in Flickr usage patterns but not on Twitter (Li et al., 2013). Using partial least squares regression, the same study found that, after removing probable tourist raw data, well-educated people in management, business, science and the arts are more likely to be involved in the generation of georeferenced tweets and photos. Another study compared the mapping of photo-sharing services in Great Britain, including Flickr, Panoramio, Picasa Web and Geograph (Antoniou, Morley & Haklay, 2010), and related them to expected values based on population density. The paper suggested that the diversity of the contribution patterns between the platforms was a result of differences in nature of the Web applications analysed. Other studies were concerned with the spatial accuracy of images on photo-sharing platforms, including Flickr and Panoramio (Zielstra & Hochmair, 2013), and images obtained from social media platforms Twitter and Instagram (Cvetojevic, Juhász & Hochmair, 2016). The spatial aspect of Snapchat data has so far only been sparsely addressed in the research community, possibly due to the lack of an Application Programming Interface (API) for data download. One study, for example, visualized the spatiotemporal distribution of snaps in three US cities (Juhász & Hochmair, 2018). Other papers have examined the reasons that motivate users to share snaps (Habib, Shah & Vaish, 2019), and have explored the time spent snapping compared to other social media platforms (Billings, Qiao, Conlin & Nie, 2017). The research presented here extends the literature by identifying socio-economic factors associated with the use of three popular social media platforms, and by providing a spatial and temporal description of Snapchat activity patterns. A better understanding of how Snapchat activities relate to those on other platforms will be useful for the future integration of this data source into geo-applications.

2 Platforms analysed

Flickr is a photo-sharing service intended to help people organize and share their images and videos. It is widely used by amateur and professional photographers. Twitter is a social media service that allows users to share short text messages (so-called tweets) of up to 280 characters. Users can also post media objects, such as images, videos and gifs (short animated images). Snapchat is a popular social media service that focuses on sending photos and videos. These images and short videos are called 'snaps', which can be shared publicly in a feature called 'Our Story'. The snaps are available for anyone to view and browse on a map interface called Snap Map, available at https://map.snapchat.com, and in smartphone applications. The snaps disappear after a certain length of time, which is currently 24 hours for publicly shared content.

All three online services presented in this study are popular among users worldwide. Snapchat has 188 million daily active users (Snapchat, 2018), and the total number of monthly active users is over 300 million[1]. Twitter has about 335 million monthly active users worldwide (Twitter, 2018). Flickr tends to have fewer active users than Snapchat and Twitter as it is highly specialized and focuses on photography, as opposed to instant user interaction. On its website, Flickr claims to have 90 million active monthly users[2]. A survey conducted in the United States found that Snapchat is used primarily by young people, while Twitter is used more by older generations (Smith & Anderson, 2018). The same survey reports that 78% of young adults (18- to 24-year-olds) use Snapchat, whereas within the same age group Twitter has only a 45% share. The popularity of Snapchat and Twitter can also be illustrated by the fact that among all adults who use them, 63% open Snapchat and 45% open Twitter multiple times a day. Another survey found that nearly half of US teenagers consider Snapchat their main social media platform (Piper Jaffray, 2017).

3 Study setup

3.1 Data collection and processing steps

This study considers data submitted to Snapchat, Flickr and Twitter between 14 December 2018 and 28 January 2019. For the regression analysis, however, the timeframe for Flickr data was extended to 1 September 2018 – 28 January 2019 in order to increase the sample size. Twitter and Flickr provide data access through standard APIs (Juhász, Rousell & Arsanjani, 2016). In this study, the Twitter streaming API was used to continuously collect geotagged tweets with exact locations. Flickr photo locations were harvested on 29 January 2019 through the Flickr API. Since Snapchat does not provide an open API, a self-developed tool was used to collect locations of public snaps submitted to the 'Our Story' feature (Juhász & Hochmair, 2018). Twitter and Flickr contain other metadata, such as user identifier, text and tags. However, Snapchat data includes only the location of the snap and

---

[1] https://www.omnicoreagency.com/snapchat-statistics/
[2] https://www.flickr.com/jobs/

the timestamp for the submission. Therefore, this study focuses solely on the spatial and temporal activity patterns on these three data sources.

Unlike Snapchat's 'Our Story' feature and Flickr's use specifically for sharing photos, Twitter users do not need to submit photos or videos in order to use the service. Therefore, a subset of tweets containing media objects (images, uploaded videos, or short animated pictures) was also extracted and used as another data source for analysis. Our data show that only 1.1% of geocoded tweets contain media objects. All data were stored in a spatially-enabled PostgreSQL database to allow for effective processing of spatiotemporal queries.

The final Florida dataset includes spatiotemporal information from 9,941 Flickr photos, 316,762 public Snapchat posts, 301,709 tweets, and 3,396 tweets with media objects. Figure 1 illustrates the spatial distribution within the data collection area (red outline) for the four datasets between 14 December 2018 and 28 January 2019 in and around Florida. Point sets are clustered around major metropolitan areas and college towns, including Gainesville (University of Florida) and Tallahassee (Florida State University).
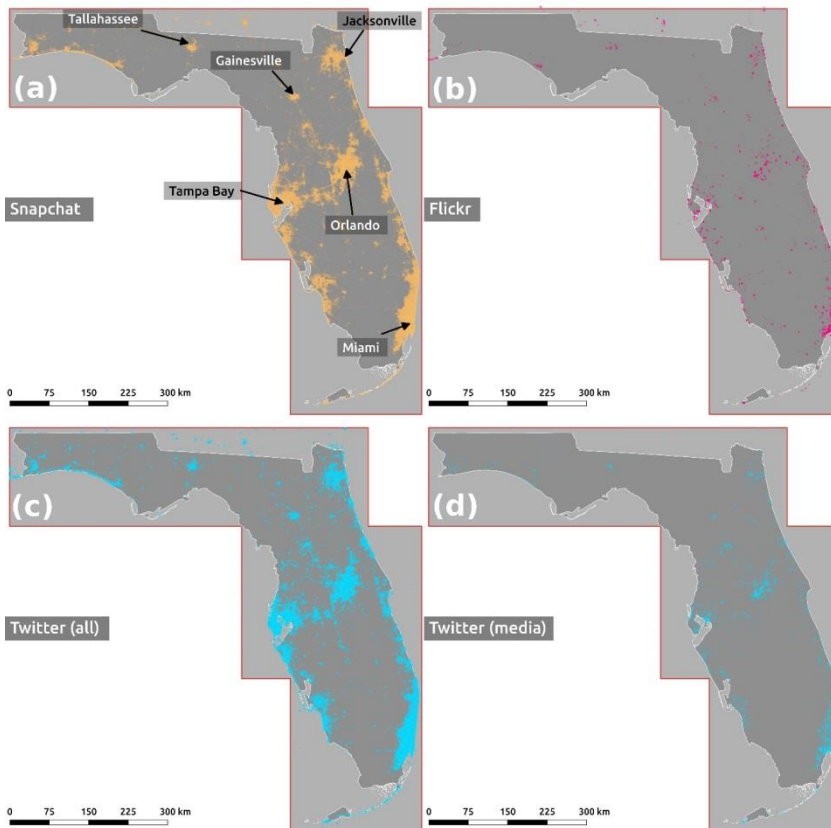

Figure 1: Spatial distribution of data sources: (a) Snapchat posts submitted to 'Our Story'; (b) Flickr photos; (c) geotagged tweets; (d) geotagged tweets containing media objects. Individual point locations were rendered with 2% opacity.

3.2 Analysis methods

Temporal analysis

For tweets and snaps, timestamps correspond to the moment the data was posted to the service by users. For Flickr photos, this study utilizes the timestamp for when the photo was taken. As a first step, UTC timestamps were converted to match the local time in Florida (Eastern Standard Time).

Twitter and Snapchat downloaders require the continuous running of data collection processes. They are therefore vulnerable to network errors and other glitches. During this 1.5-month data collection campaign, we experienced six days with only partial Twitter data collection (between 23 December and 28 December 2018), and two days with only partial Snapchat data collection (31 December 2018 and 1 January 2019). Data from these days were removed from further analysis.

Bots are Twitter profiles that share automated messages. These user profiles can manipulate online conversations; they also affect the spatial pattern of tweeting activity by adding artificial noise to the dataset. Complete automation probability (CAP) is the probability that a Twitter account is completely automated (Yang et al., 2019). CAP scores were calculated for all users, and tweets from users with CAP scores greater than 0.5 were assumed to be bots and therefore excluded from further analysis. This affected 2.7% of users, equating to 3.4% of all tweets.

Data was aggregated by different temporal units to facilitate the analysis of temporal activity patterns. That is, we utilized daily aggregated numbers to describe activity patterns over the course of the study timeframe, and also aggregated data hourly in order to provide insights into the daily dynamics of photo-sharing and social media sites.

Regression analysis

The purpose of the regression analysis was to explore the relationship between counts of social media activities (A), sociodemographic variables (S), and characteristics of the built (B) and natural (N) environments at census-tract level. A stylized representation of this relationship in functional form can be given as $A = f(S, B, N)$. Since the observed variables are count data, a negative binomial regression model was developed and estimated for each of the four social media types. To express the left-hand side of the equation as a rate of events per areal unit exposure, an offset variable was introduced to the right-hand side of the equation and set to the natural logarithm of the tract area in m². The four models were developed in a manual stepwise approach by adding and removing variables in an exploratory manner to improve model fit while at the same avoiding significant spatial autocorrelation among residuals. Ignoring spatial dependence in spatial data can lead to coefficient estimation bias and biased standard errors (Anselin, 1988). In this case study, eigenvector spatial filtering (ESF) was applied to model spatial autocorrelation. ESF utilizes eigenvector decomposition to extract a set of eigenvectors from the spatial weight matrix that is incorporated in the numerator of the Moran's I coefficient (Griffith, 2000). A spatial filter that comprises all relevant eigenvectors can then be used as an additional predictor variable in standard statistical techniques (Helbich & Arsanjani, 2015). This spatial filter is

used as a predictor in each of the four regression models and explains a considerable part of the variance in the distribution of social media activities.

The study area for the regression analysis is a region of central Florida extending between the Ocala National Forest to the north and Jupiter to the south that comprises 882 census tracts (Figure 2).
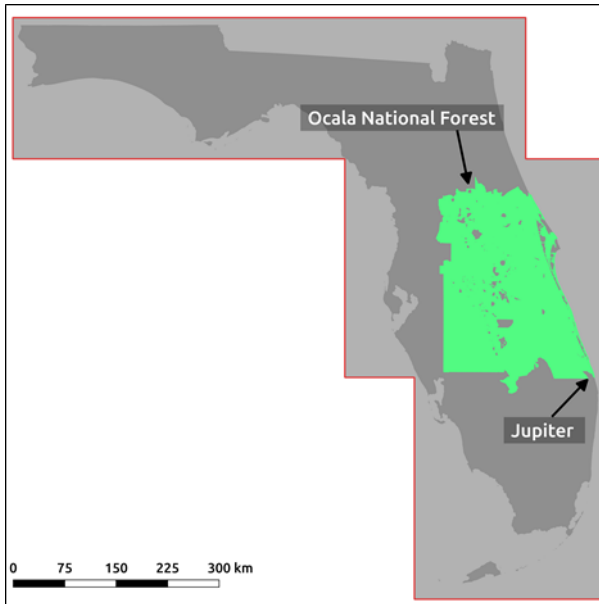

Figure 2: Study area for the spatial regression analysis

The list of explanatory variables considered for the regression models (Table 1) is subdivided into sociodemographic attributes and factors of the built and the natural environments. The list is partially drawn from related studies that model activity patterns on various social media and photo-sharing platforms (Antoniou et al., 2010; Lenormand et al., 2014; Li et al., 2013). Table 1 also describes the operationalization of the variables at census-tract level, as well as the sources from which these data were obtained.

As part of data preparation, Spearman's rho correlation coefficient was computed between all candidate explanatory variables. High correlations of $|r| > 0.7$ occurred only between % bachelor and % highschool ($r = -0.77$), and between % bachelor and % household income > 150 K ($r = 0.71$). To mitigate multicollinearity, these predictor combinations were avoided during the model-building process.

All snaps and tweets posted between 14 December 2018 and 28 January 2019 falling in the study region were used for analysis. The Flickr observation timeframe was a few months longer, as explained above. Flick revealed strong user participation inequality, with some individuals posting hundreds of photos during the analysis timeframe. To avoid a high level of geographic bias for pictures posted by these individuals, all photos from users who posted

more than 50 photos were removed. Most often these pictures were clustered in a small region, featuring a special event such as a politician's speech. In addition, images taken by users who posted between 11 and 49 photos were manually reviewed for content that was not directly associated with the surrounding outside environment. That is, images from users who shared primarily pictures of events (e.g. barbecue, soccer game), and museum, art and personal collections (e.g. cars, parked airplanes), or who promoted their business (e.g., posting detailed pictures of hotel rooms) were removed. The final dataset of activities in the Central Florida study region used in the regression models contained: 1,402 Flickr images from 294 users; 92,145 snaps (number of users unknown); 74,102 tweets from 14,841 users; and 894 tweets with media content from 294 users.

Table 1: Tract-based explanatory variables considered for the regression analysis

Variable	Operationalization	Data Source
Sociodemographic		
Population density	Population per km^2	
Median age	Median age in years	
% Black	% of black population	
% Hispanic	% of Hispanic population	
% Highschool	% of population aged over 25 with highschool diploma as highest qualification	2012–2016 American Community Survey 5-Year estimates
% Bachelor	% of population aged over 25 with bachelor's degree as highest qualification	
% Household income > 150 K	% of households with an annual income > $150,000	
% Household income < 25 K	% of households with an annual income < $25,000	
Job density	Number of jobs per km^2	U.S. Census Bureau - LEHD
Built environment		
Highway density	Length of Class 1 and 2 roads per km^2	HERE NAVSTREETS
Scenic highway density	Length of designated scenic highways per km^2	Florida Dept. of Transportation

% CBD	% of tract overlapping with the Orlando Central Business District	http://www.city-data.com
% University	% of tract overlapping with colleges and universities	TIGER areal landmarks
Hotel density	Hotels per km^2	OpenStreetMap
Restaurant density	Restaurants per km^2	
Museum density	Museums per km^2	
Amusement park density	Amusement parks per km^2	
Zoo (binary)	Presence of zoo	HERE NAVSTREETS
Performing arts (binary)	Presence of performing arts theater	
Movie theater (binary)	Presence of movie theater	
Airport (binary)	Presence of Orlando International Airport or Kennedy Space Center	Natural Earth
Natural environment		
Natural spring (binary)	Presence of natural spring	OpenStreetMap
% Forest	% of tract overlapping with state or national forest	Florida Forest Service, U.S. Dept. of Agriculture
% State park	% of tract overlapping with state park	Florida Dept. of Environmental Protection
% National park	% of tract overlapping with national park	National Park Service
Beach (binary)	Presence of ocean beach	TIGER areal landmarks
Bay/ocean (binary)	Adjacent to bay or ocean	

4   Analysis results

4.1 Temporal patterns

Figure 3 shows the temporal dynamics of the social media services analysed by plotting aggregated daily count values. On an average day, Twitter and Snapchat users share between 5,000 and 10,000 geolocated posts within the study area. Even though not conclusive due to gaps in data availability, the increased number of posts on public holidays, such as Christmas Day (25 December) on Snapchat and 1 January on Twitter, suggest increased social media and photo-sharing activities on these days. On the busiest day on Snapchat (Christmas Day), users posted 3.8 times as many snaps (14,272) as on the least busy day (23 January 2019). The daily data volumes of snaps submitted to 'Our Story' and public geotagged tweets are comparable with each other. However, the temporal characteristics of these two are

different. Figure 3a suggests that Snapchat users tend to be more active over the weekend (in grey), which is confirmed in a two-sample t-test, $t(42.4) = -7.2$, $p < 0.001$. Geotagged tweets, on the other hand, show a less differentiated temporal pattern, with no obvious activity difference between weekends and weekdays. This is somewhat different from an earlier study that reported an 11% decrease in tweeting activity during the weekends (Gao, Abel, Houben & Yu, 2012).

Figure 3b shows that the daily number of Flickr photos taken ranges between 29 (24 January) and 694 (14 December). Even though on some weekends Flickr users tend to take more photos than during the weekdays that preceded, the difference is not statistically significant. Tweets with media content have the lowest contribution volume among all the sources analysed.
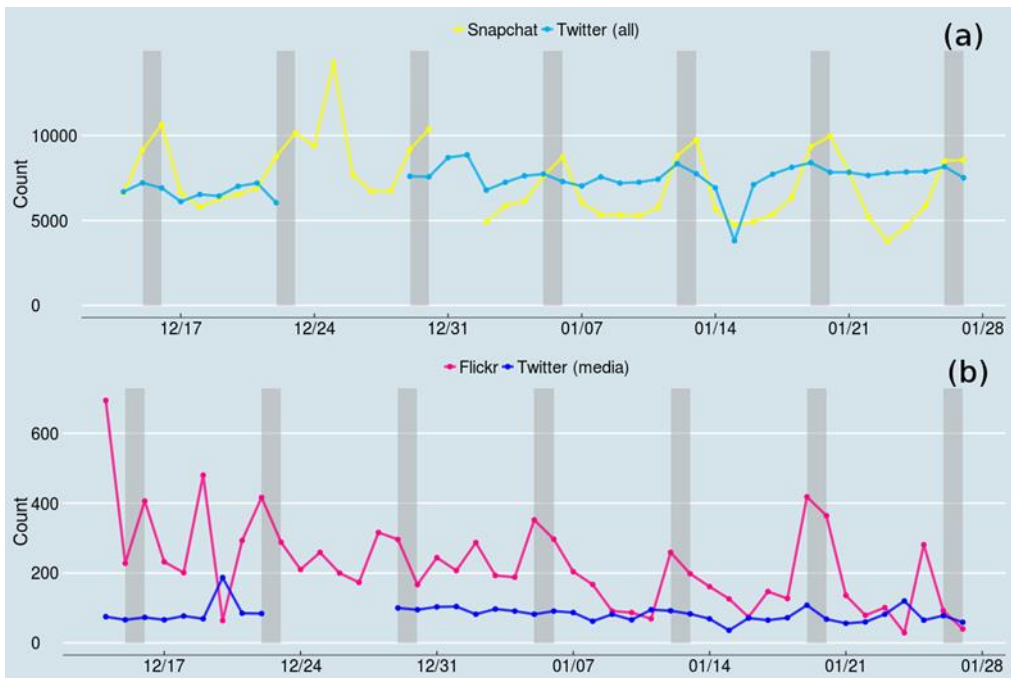

Figure 3: Daily activity count for (a) snaps and all geolocated Twitter, and (b) Flickr photos and Twitter content with media objects. Weekends (Saturday–Sunday) are highlighted as grey vertical bars

Figure 4 plots the average number of photos and posts per hour for weekdays and weekend days for the data sources analysed. Figure 4a shows that Snapchat users are more active on weekends, during the evening and in the early morning, suggesting that Snapchat is a social media tool used in the party scene. For Twitter (Figure 4b), on the other hand, weekend and weekday curves follow each other closely. Twitter activities tend to pick up early in the morning, increase until they reach a steady level in the afternoon, then drop in the late evening. This pattern mirrors a typical day for the general population, as most people get up in the morning and go to bed before midnight. Potentially, this can be explained by the fact

that Twitter is used by older people than Snapchat, who are more active during the day than at night. Twitter is also known to be used for work-related activities (Steiger, Westerholt, Resch & Zipf, 2015), which may explain why activity patterns tend to peak during the day. Media-sharing patterns on Twitter follow the same dynamics (Figure 3d).

The hourly distribution of Flickr contributions follows daylight hours closely (Figure 4c), with more photos taken during the day. This can be expected, since light is essential for photography.
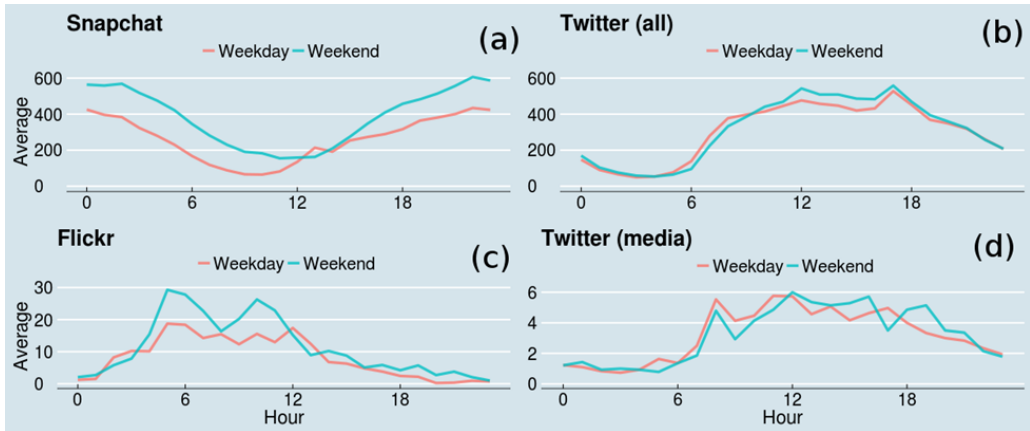

Figure 4: Hourly distribution of average photo and social media post counts for (a) weekdays and weekends for Snapchat; (b) geotagged tweets; (c) Flickr; (d) geotagged tweets containing media objects

4.2 Regression analysis

Table **2** presents the results of the spatially-filtered negative binomial regression models for activity counts related to Flickr images, snaps, tweets, and tweets with media content, based on the analysis of 882 census tracts. Only significant coefficients were retained for the final models. The Variance Inflation Factor (VIF) is below 3 for all models, which means that multi-collinearity is not a concern for the final set of variables used. The coefficient associated with the spatial filter is significant in all four models. Low Moran's I coefficients together with their p-values of above 0.05 indicate that spatial autocorrelation among residuals has been largely eliminated. Since in negative binomial regression models the link function is the natural log, the interpretation of β coefficients is that each one-unit increase in X_i increases the mean social activity per square km by a multiplication factor $\exp(\beta_i)$. This allows the computation of the percentage difference in social media activity for a one-unit change in each correlate, controlling for other correlates. For example, based on a mean density of 1.95 restaurants per km^2 in census tracts in the study area, an increase to 2.95 restaurants per km^2 is associated with an increase in Twitter activity by a factor of $e^{0.115} = 1.122$, which translates to a 12.2 % increase.

Comparison of regression coefficients between the four models reveals some commonalities but also some differences between the spatial activity patterns. One commonality is that in all four models user activity increases with population density. This means that the major generator of the data is the urban environment. Furthermore, model results show that all data sources, including Flickr, receive higher activity rates in areas of higher job density. Flickr, however, is the only platform with increased contribution rates in any of the outdoor recreational environments (e.g. state parks) that were considered. This could indicate that part of the Flickr user community aims for scenic outdoor images and therefore travels to natural environments more frequently. As opposed to this, natural areas such as forests, parks and natural springs are associated with significantly lower snap and tweeting (without media) rates. There are several possible reasons for this. First, the Snapchat community might be less focused on landscape and nature images than involved in events and happenings in society. This is supported by the positive coefficients associated with restaurants, amusement parks and movie theaters in the Snapchat model. Similarly, the Twitter platform is designed for immediate sharing of news and events, which are less likely to occur in remote areas, leading to reduced Twitter activity around natural springs and in forests. However, news in Twitter appears to be more frequently shared in proximity to restaurants and museums, and near amusement parks for tweets with media attached. Second, the lower activity rates of the Snapchat and Twitter users in natural areas could also be due to reduced cell phone and/or Wi-Fi coverage in remote areas (Cvetojevic et al., 2016). Flickr images can be taken outdoors even when there is no Internet availability, and uploaded later (e.g. from home) when the Internet is available again. Typically, the original geographic location of the picture (e.g. in a state park) is stored in EXIF headers and is retained even when the image is not uploaded immediately. Snaps and tweets, on the other hand, are designed to be posted instantaneously, a possibility which is limited in areas of low cell phone or Wi-Fi coverage.

Twitter usage has been found to be more frequent on roads with high traffic volume (Lenormand et al., 2014), which is confirmed in our analysis. Nevertheless, our models show that this is not true for tweets with media, which are more frequently found at social event locations (e.g. amusement parks) or in scenic areas (e.g. bay and ocean). The Florida Scenic Highways are 26 designated highways that pass through sites which the state of Florida determines to be historically, culturally, recreationally, naturally or archaeologically significant. Results show a positive correlation between Flickr activity rates and density of scenic highways, suggesting that the scenic highways indeed present scenic views that are worth being captured on images.

In terms of socio-economic variables, results show that Snapchat activities decrease with age, which confirms earlier surveys (Smith & Anderson, 2018). Educational and household income variables are not reflected in clearly different social media use, since areas with low income and ones characterized by higher educational attainment are associated with higher activity rates on the different platforms.

Table 2: Estimation results for the spatially-filtered negative binomial models

	Flickr	*Snapchat*	*Tweets*	*Tweets with Media*
Sociodemographic variables				
Population density	9.55E-04***	1.10E-03***	9.46E-04***	9.998E-04***
Median age		-1.16E-02***		
% Hispanic		-7.08E-03***		
% Bachelor	2.78E-02**		3.28E-02***	2.433E-02**
% Household income < 25 K		8.79E-03***	1.74E-02***	
Job density	5.05E-04***	2.39E-04***	3.63E-04***	
Built environment variables				
Highway density			2.60E-04**	
Scenic highway density	6.80E-04***		4.62E-04***	
% CBD			-0.186**	
Restaurant density		8.14E-02***	0.115***	0.141***
Museum density			0.732**	0.952*
Amusement park density		0.848**		1.142***
Performing arts	0.626*			
Movie theater		0.263**		
Natural environment variables				
Natural spring			-0.551*	
% Forest		-3.90E-02***	-2.96E-02***	
% State park	2.18E-02*	-1.06E-02**		
% National park		-2.72E-02*		
Bay/ocean	0.843***	0.232**		0.691***
Spatial filter	0.738***	0.680***	0.719***	0.740***
Constant	-19.42***	-12.75***	-14.71***	-19.19***
Number of observations	882	882	882	882
Residuals Moran's I	-2.14E-03	1.262E-03	1.10E-02	8.86E-03
Residuals Moran's I (p value)	0.885	0.261	0.058	0.075
McFadden's pseudo R^2	0.227	0.103	0.080	0.191
Adjusted R^2	0.220	0.100	0.078	0.185

5   Conclusions and future work

This study compared the activity patterns of three major social networking platforms, revealing differences both in temporal and in spatial aspects. Differences in daily activity patterns can be explained by the different user bases that the platforms try to attract. Snapchat, which is often used to share media related to social occasions, shows activity peaks late at night, in the early hours of the morning, and on weekends. Twitter, which is also commonly used for work-related purposes, typically shows an activity distribution pattern that starts in the early morning and slows down in the late evening. There is no increase in tweeting activity at the weekend. Flickr image-capturing activities follow a similar pattern to Twitter, with a faster decay, however, in the afternoon and evening hours, possibly due to less favorable lighting conditions later in the day.

The spatial regression estimations point towards differences in activity patterns between the four data sources analysed. Results show that natural areas, such as state parks, increase Flickr photo uploads, but that they reduce Snapchat and Twitter activities. Various possible explanations were put forward for the latter observation. Snapchat is more often used at party and event locations, such as restaurants, amusement parks or movie theaters, whereas Flickr users seem to be drawn more to scenic outdoor locations, such as along scenic highways or by the ocean. Tweets were found to be spatially related to typical locations of daily activities, such as restaurants and workplaces. Twitter also seems to be used for communication when travelling on highways.

For future work, we plan to extend this analysis in both the temporal and the spatial realms, and to correlate the different activity counts with ground truth data for human mobility and activity, such as visitor numbers to state parks.

References

Anselin, L. (1988). Spatial Econometrics: Methods and Models. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Antoniou, V., Morley, J., & Haklay, M. (2010). Web 2.0 geotagged photos: Assessing the spatial dimension of the phenomenon. Geomatica, 64(1), 99–110.

Billings, A. C., Qiao, F., Conlin, L., & Nie, T. (2017). Permanently desiring the temporary? Snapchat, social media, and the shifting motivations of sports fans. Communication & Sport, 5(1), 10–26. https://doi.org/10.1177/2167479515588760

Cvetojevic, S., Juhász, L., & Hochmair, H. (2016). Positional Accuracy of Twitter and Instagram Images in Urban Environments. GI_Forum 2016, 1, 191–203. https://doi.org/10.1553/giscience2016_01_s191

Gao, Q., Abel, F., Houben, G.-J., & Yu, Y. (2012). A comparative study of users' microblogging behavior on Sina Weibo and Twitter. 88–101. https://doi.org/10.1007/978-3-642-31454-4_8

Girardin, F., Calabrese, F., Fiore, F. D., Ratti, C., & Blat, J. (2008). Digital Footprinting: Uncovering Tourists with User-Generated Content. https://doi.org/10.1109/MPRV.2008.71

Griffith, D. A. (2000). A linear regression solution to the spatial autocorrelation problem. Journal of Geographical Systems, 2(2), 141–156. https://doi.org/10.1007/PL00011451

Habib, H., Shah, N., & Vaish, R. (2019). Impact of Contextual Factors on Snapchat Public Sharing. https://doi.org/10.1145/3290605.3300256

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

Helbich, M., & Arsanjani, J. J. (2015). Spatial eigenvector filtering for spatiotemporal crime mapping and spatial crime analysis. Cartography and Geographic Information Science, 42(2), 134–148. https://doi.org/10.1080/15230406.2014.893839

Juhász, L., & Hochmair, H. H. (2018). Analyzing the spatial and temporal dynamics of Snapchat. Presented at the VGI-Alive pre-conference workshop at AGILE 2018, Lund, Sweden.

Juhász, L., Rousell, A., & Arsanjani, J. J. (2016). Technical Guidelines to Extract and Analyze VGI from Different Platforms. Data, 1(3), 15. https://doi.org/10.3390/data1030015

Lenormand, M., Picornell, M., Cantú-Ros, O. G., Tugores, A., Louail, T., Herranz, R., … Ramasco, J. J. (2014). Cross-checking Different Sources of Mobility Information. PLOS ONE, 9(8). https://doi.org/10.1371/journal.pone.0105184

Li, L., Goodchild, M. F., & Xu, B. (2013). Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. Cartography and Geographic Information Science, 40(2), 61–77. https://doi.org/10.1080/15230406.2013.777139

Piper Jaffray. (2017). Taking Stock with Teens - Fall 2017. Retrieved from https://www.businesswire.com/news/home/20171011005804/en/

Smith, A., & Anderson, M. (2018). Social Media Use in 2018. Retrieved from Pew Research Center website: https://www.pewinternet.org/wp-content/uploads/sites/9/2018/02/PI_2018.03.01_Social-Media_FINAL.pdf

Snapchat. (2018). Snap Inc. Reports Second Quarter 2018 Results. Retrieved from https://investor.snap.com/news-releases/2018/08-07-2018-211104059

Steiger, E., Westerholt, R., Resch, B., & Zipf, A. (2015). Twitter as an indicator for whereabouts of people? Correlating Twitter with UK census data. Computers, Environment and Urban Systems, 54, 255–265. https://doi.org/10.1016/j.compenvurbsys.2015.09.007

Takhteyev, Y., Gruzd, A., & Wellman, B. (2012). Geography of Twitter networks. Social Networks, 34(1), 73–81. https://doi.org/10.1016/j.socnet.2011.05.006

Twitter. (2018). Q2 2018 Newsletter to Shareholders. Retrieved from https://s22.q4cdn.com/826641620/files/doc_news/archive/9f9933a2-4285-4912-8fac-7846916f1c49.pdf

Yang, K.-C., Varol, O., Davis, C. A., Ferrara, E., Flammini, A., & Menczer, F. (2019). Arming the public with artificial intelligence to counter social bots. Human Behavior and Emerging Technologies, 1(1), 48–61. https://doi.org/10.1002/hbe2.115

Zielstra, D., & Hochmair, H. H. (2013). Positional accuracy analysis of Flickr and Panoramio images for selected world regions. Journal of Spatial Science, 58(2), 251–273. https://doi.org/10.1080/14498596.2013.801331