Article

# *#London2012:* Towards Citizen-Contributed Urban Planning Through Sentiment Analysis of Twitter Data

Anna Kovacs-Gyori [1,*], Alina Ristea [1], Clemens Havas [1], Bernd Resch [1,2] and Pablo Cabrera-Barona [3,4]

[1] Department of Geoinformatics—Z_GIS, University of Salzburg, 5020 Salzburg, Austria; E-Mails: anna.gyori@sbg.ac.at (A.K.-G.), mihaela.ristea@sbg.ac.at (A.R.), clemensrudolf.havas@sbg.ac.at (C.H.), bernd.resch@sbg.ac.at (B.R.)
[2] Center for Geographic Analysis, Harvard University, Cambridge, MA 02138, USA; E-Mail: bresch@fas.harvard.edu
[3] Institute of Higher National Studies—IAEN, 170507 Quito, Ecuador; E-Mail: pablo.cabrera@iaen.edu.ec
[4] Latin American Social Sciences Institute—FLACSO Ecuador, 170201 Quito, Ecuador; E-Mail: pablo.cb@live.com

* Corresponding author

**Abstract**
The dynamic nature of cities, understood as complex systems with a variety of concurring factors, poses significant challenges to urban analysis for supporting planning processes. This particularly applies to large urban events because their characteristics often contradict daily planning routines. Due to the availability of large amounts of data, social media offer the possibility for fine-scale spatial and temporal analysis in this context, especially regarding public emotions related to varied topics. Thus, this article proposes a combined approach for analyzing large sports events considering event days vs comparison days (before or after the event) and different user groups (residents vs visitors), as well as integrating sentiment analysis and topic extraction. Our results based on various analyses of tweets demonstrate that different spatial and temporal patterns can be identified, clearly distinguishing both residents and visitors, along with positive or negative sentiment. Furthermore, we could assign tweets to specific urban events or extract topics related to the transportation infrastructure. Although the results are potentially able to support urban planning processes of large events, the approach still shows some limitations including well-known biases in social media or shortcomings in identifying the user groups and in the topic modeling approach.

## 1. Introduction

Cities are complex systems (Castells, 1996; Hall, 1966; Theodore, 2006), consisting of two main elements: the people as residents or visitors, and the infrastructure to fulfill their needs ranging from housing to recreation or even self-realization (Costanza et al., 2007; Maslow, 1943). Some of the infrastructure or related networks are static and mostly physical, such as buildings or the road and electricity networks, whereas others are more dynamic, like social, transportation, or financial networks. From an urban analysis viewpoint, the dynamic nature of these systems is challenging, especially in the case of large cities with millions of people constantly on the move and having different needs and preferences. These challenges do not only result from the sheer amount of

people, but also from the intense spatiotemporal variability originating from urban dynamism and from the constantly changing subjective needs of each person. Therefore, effective planning practice requires analysis at high spatial and temporal scales to understand this dynamism of urban life and processes.

Traditional methods, such as questionnaires or counting, are not capable of handling such fine temporal and spatial scales at all, or they are highly resource-consuming and, therefore, slow and costly, and thus not up-to-date. This is where the advantages of the data-driven era become relevant, most concretely with respect to the real-time availability of social media data. This data provides unseen contextual insights into spatiotemporal phenomena on a finer scale in cities through users' digital traces on different online platforms such as Twitter, Foursquare/Swarm and Flickr (Abbasi, Rashidi, Maghrebi, & Waller, 2015; Aubrecht, Ungar, & Freire, 2011; Crooks et al., 2015; Girardin, Vaccari, Gerber, Biderman, & Ratti, 2009). This is of central importance to urban planning dealing with the optimization of the above-mentioned networks.

Planners are responsible for land use strategies, the design of public places, or transportation planning, which constitute essential factors of urban life (McGill, 2017). An aspect of particular importance for urban planning is the investigation of the effects of a planned large event, considering residents and visitors. These events have a special role in planning because they are usually temporary and require completely different circumstances and conditions compared to the average daily routines of urban life. In contrast with other unplanned events such as emergencies (e.g., natural, industrial and manmade disasters), direct preparations can be made for planned events, not just precautionary measures (Getz & Page, 2016). As a consequence, they are preceded by extensive planning and preparation efforts; but such events frequently still face severe inconveniences or even disruptions, most strikingly with respect to the transportation of people, presumably in different ways for residents and visitors. Therefore, the distinction between these two groups is crucial in most of the analyses due to their different needs, behavioral patterns, and exposure to the effect of a large planned event.

Thus, analyses at fine temporal scales are inevitable for examining citizens' mobility, which can help in detecting patterns and anomalies through the understanding of underlying problems or phenomena in the context of urban transportation. For instance, citizens' trajectories and the number of people moving through the city vary over time during the day, but also between days depending on the weather, weekday, planned and sudden events, traffic density, and many other factors (Sagl, Resch, Hawelka, & Beinat, 2012). Therefore, traditional annual commuting statistics are not informative on such fine spatial and temporal scales because they are mostly produced only once or twice per year and aggregated to spatial planning units. First, this results in commuting data that do not actually reflect real travel directions due to their aggregated nature, and, second, everyday individual trajectory details are lost through the aggregation.

As a consequence, social media, providing digital spatiotemporal traces of individuals, grant valuable mobility information, particularly through their nature of a large and continuous source of data and their fine scale in space and time. Another advantage of using these sources is the potential for extracting direct feedback about city life-related topics, places or phenomena by revealing subjective aspects as well, such as public mood or emotions (Frank, Mitchell, Dodds, & Danforth, 2013; Quercia, Ellis, Capra, & Crowcroft, 2012; Resch, Summa, Zeile, & Strube, 2016). This provides an opportunity to investigate what people actually think about parts of the city and the direct or indirect effects of a large event. Thus, several methods and analyses have been developed in the area of opinion mining (e.g., Pak & Paroubek, 2010) and semantic topic extraction (e.g., Steiger, Westerholt, Resch, & Zipf, 2015) for use in urban planning.

However, to the best of our knowledge, limited research has been conducted to analyze social media data regarding planned large events, considering comparison days (before or after the event), different user groups (residents vs visitors) along with the linkage between sentiment analysis and topic extraction in one study. In our work, we intend to integrate all of these aspects to provide valuable knowledge about urban events, whereby our case study focuses on the 2012 Olympic Games in London. By exploring emotions and events in a city through social media analysis, we aspire to a better understanding of citizens' behaviors and needs in cities. Thereby, we can provide a basis to aid planners in identifying more specific urban planning issues for further in-depth analysis. In line with these goals, we intend to answer the following research questions in this article:

RQ1 → How can we identify distinctive characteristics of tweeting behavior in terms of spatiotemporal patterns and sentiments between "residents" and "visitors"?

RQ2 → Are there detectable changes in the spatial and temporal patterns, and sentiment of the tweets during the London Olympic Games compared to the days before and after it?

RQ3 → Which topics that are related to urban planning in the context of a large sports event can be identified through semantic analysis of social media posts?

## 2. Related Work

### 2.1. Citizen-Contributed Geographic Information to Describe Urban (and Spatial) Practices

Among the practical applications of geographic data extracted from social media, we can distinguish two main categories: quantitative and qualitative aspects.

*Quantitative* approaches describe spatiotemporal phenomena by leveraging the advantage of fine spatial and temporal scales of the data, such as crowdsourcing urban form and function (Crooks et al., 2015), or characterizing and classifying urban areas and location types (Noulas, Scellato, Mascolo, & Pontil, 2011).

Among quantitative analyses, mobility forms its own well-defined category. The applications range from the description of general spatiotemporal dynamics to revealing anomalies on urban (Aubrecht et al., 2011; Fujisaka, Lee, & Sumiya, 2010; Hasan & Ukkusuri, 2014), or global scales (Hawelka et al., 2014). Traffic and mobility modeling purposes are also present, such as validating travel demand models (Lee, Gao, & Goulias, 2016), and analyzing origin-destination flows (Cebelak, 2013).

There are also a few applications where researchers assess urban life from a *qualitative* point of view using different social media sources. Girardin et al. (2009) evaluated urban attractiveness by analyzing images from Flickr and mobile phone usage data, while Sun, Fan, Bakillah and Zipf (2015) used geo-tagged images for road-based travel recommendations. As another approach, several researchers improved and refined various methodologies for extracting emotions (Resch et al., 2016), transit rider satisfaction (Collins, Hasan, & Ukkusuri, 2013), and community happiness (Quercia et al., 2012) from Twitter data, also combined with demographics and other objective characteristics of a place such as education or obesity (Mitchell, Frank, Harris, Dodds, & Danforth, 2013), or even defined sentiment as a function of movement (Frank et al., 2013). The advantages of utilizing available additional datasets such as demographics, mobile phone data or mobility trajectories are twofold; they can help the interpretation of the primary results extracted from social media, and, on the other hand, they are also appropriate for validation purposes.

Further related approaches vary in their purpose, i.e., they are not specifically tailored to urban planning, but they can still be used in this context. They include classifying Twitter users (Pennacchiotti & Popescu, 2011), or describing geotemporal demographics (Longley, Adnan, & Lansley, 2015). Furthermore, the analysis of textual content to extract spatial information is becoming increasingly important (Cheng, Caverlee, & Lee, 2010; Dalvi, Kumar, & Pang, 2012; Kinsella, Murdock, & Hare, 2011; Li, Serdyukov, de Vries, Eickhoff, & Larson, 2011), including the generation of ambient geographic information (Stefanidis, Crooks, & Radzikowski, 2013), or the definition of geotag gazetteers (Keßler, Maué, Heuer, & Bartoschek, 2009).

### 2.2. Urban Planning, Social Media and Planned Large Events

Previous studies have shown that social media usage is generally more intensive during large events, and a concentration around the venue and impact on transportation is also identifiable (Gupta & Kumaraguru, 2012;

Zhang, Ni, He, & Gao, 2016). The Olympic Games are considered one of the world's largest events, involving a lot of organizational tasks from social, technical, environmental, economic, demographic and transportation-related perspectives (Chen, 2012; Cook & Ward, 2011; Malfas, Houlihan, & Theodoraki, 2004).

Furthermore, a growing body of literature is dealing with the use of Twitter data for investigating the characteristics of different types of events, which is of importance also to the field of urban planning. Recently, a number of machine-learning approaches have been used to investigate electoral predictions (Gayo-Avello, 2013), stock market flows (Zhang, Fuehres, & Gloor, 2011), flu trends (Culotta, 2010; Ritterman, Osborne, & Klein, 2009), natural disasters (Fraustino, Liu, & Jin, 2012; Resch, Usländer, & Havas, 2017), or to detect large events (Lee & Sumiya, 2010; Li, Lei, Khadiwala, & Chang, 2012; Weng & Lee, 2011), even in near real-time (Zhao, Zhong, Wickramasuriya, & Vasudevan, 2011) and with respect to their impacts (Panteras et al., 2015).

Large sports events like the FIFA World Cup can also be identified using the content of the tweets, hashtags and distribution of retweets. Kim et al. (2015) applied topic modeling before and during the event, while Corney, Martin and Göker (2014) identified phrases (word n-grams) that showed a sudden increase in frequency in the dataset and then selected co-occurring n-grams to identify topics. By using sentiment analysis, researchers identified relationships between the public mood and large socioeconomic events in the media (Bollen, Mao, & Pepe, 2011), together with trends and possible predictions of the disposition theory (Yu & Wang, 2015), such as fanship for sports. Clearly, these analyses identified changes in activity patterns (e.g., supporters induce a general increase in number of tweets), in topic diversity, and in the spatial distribution of topics related to the event.

### 3. Data

The study area for the present analysis is Greater London, which has an expansion of 3,458 km$^2$. The Twitter data was obtained using the Twitter Streaming Application Programming Interface (Twitter INC, 2017) for the year 2012, and consists of tweet content and attributes such as user name, user location, and message time. We only harvested geolocated tweets, as our study requires geospatial and temporal analysis. To the best of our knowledge, this database does not contain retweets. It shall be noted that due to user practice and the policy of Twitter, in general, the tweets containing coordinates represent only a smaller subset of all tweets posted in a given period, about 1–10% according to previous studies (Morstatter, Pfeffer, Liu, & Carley, 2013; Zhang et al., 2016). Moreover, they are not evenly spread in space and among user groups, as youngsters tend to use social media more actively (Li, Goodchild, & Xu, 2013; Resch et al., 2017). These issues have been thoroughly

discussed in existing literature (Steiger et al., 2015; Sui & Goodchild, 2011) and will be further detailed in the Discussion section. In addition, we want to point out that distinguishing personal and non-personal Twitter accounts was beyond the scope of this study. Although we are aware of the possible bias originating from it, we considered its effect on the final results marginal due to the large amount of data.

## 4. Methodology

As shown in Figure 1, our methodology comprises a sequential number of steps for pre-processing (defining temporal bins for before/during/after the Olympics and identifying residents vs visitors), textual analysis (sentiment analysis and automated semantic topic modeling), spatial hot spot detection, and finally evaluation and validation through a point pattern test of our results. The single steps are described in the following sub-sections.

### 4.1. Pre-Processing

We developed a two-step filtering procedure to prepare the raw data for the subsequent analysis:

*Temporal binning*: First, we created temporal bins from the raw data representing time periods before, during and after the Olympic Games (OG). This allows us to distinguish between "event days" and "comparison days". The reason for following this approach has been described in previous literature, as large-scale events such as the OG change the dynamics of a city for the time of the event. The temporal bins have been defined as follows: before: *June 27–July 13, 2012*; during: *July 27– August 12, 2012*; after: *August 27–September 12, 2012*.

*Spatiotemporal subsetting* (hypothesizing residents and visitors): The self-reported geolocation data from tweets and the frequency of their presence in the temporal subsets were used to identify presumable "residents" and "visitors" in London. Our approach is based on the work of Abbasi et al. (2015), who identified these user types in Sydney for city trip analysis. The rationale for identifying the two groups was the following: A person who tweeted at least once in each of the temporal subsets was considered a "resident", whereas a person who tweeted in just one of them was considered a "visitor" (non-resident). The remaining users of the dataset were not considered in the present study, as we could not

differentiate between less actively tweeting residents or those visitors who stayed longer than a month, without performing further extensive analysis. Although, it is possible to identify them based on their tweets' content, but that is a complex methodology on its own and, therefore, was beyond the scope of this study. We are aware of the limitations of our method and discuss them, along with the advantages in the Discussion section. Yet, our results underpin that the method is effective when there is no additional data available for classifying user types, and it sufficiently reflects the necessary differences between the two groups for the desired purposes in our case.

### 4.2. Semantic Analysis

The semantic text analysis was performed in two consecutive steps: sentiment analysis, followed by automated topic extraction using the unsupervised machine-learning method Latent Dirichlet Allocation (LDA).

#### 4.2.1. Sentiment Analysis

The sentiment analysis algorithm used in our approach is based on the work of Breen (2012). Sentiment scores were calculated for each tweet to automatically define to what degree it contains positive or negative sentiments, by calculating the difference between the number of positive words and the number of negative words.

This approach requires a dictionary with positive and negative words, for which we selected the Hu Liu lexicon (Hu & Liu, 2004), which is the most acknowledged dictionary in recent literature. Generally, if the score value is higher than zero, the sentence is assumed to contain an overall "positive sentiment", whereas it is considered containing a "negative sentiment" if the value is below zero. If the score equals zero, then the sentence is considered "neutral".

The main disadvantage of this approach is that the algorithm has limitations in defining unambiguous negative or positive scores for sentiment values around zero because they are indeed either neutral or they are misclassified with a comparatively high probability. Thus, we categorize positive tweets with the score equal or higher than 2 and negative tweets with the score equal or lower than −2. This does not mean that all the tweets with the sentiment value of 1 and −1 are neutral; rather they have a lower accuracy of being identified as positive
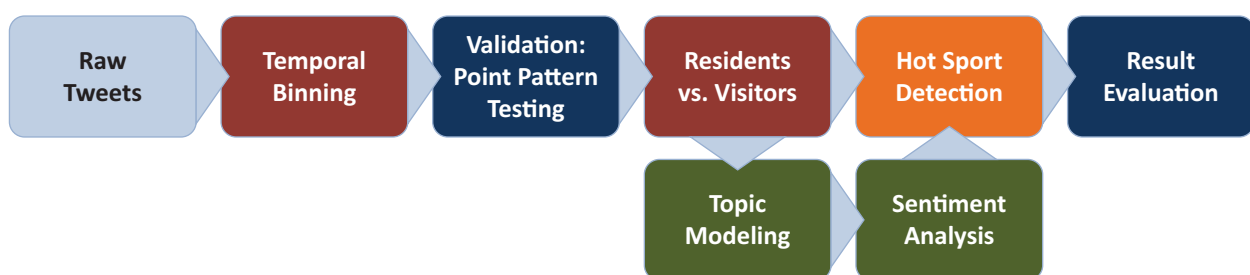


**Figure 1.** Workflow overview.

or negative and, therefore, we do not consider them in our analysis.

The terms "positive" and "negative" will be used throughout the article as defined above.

### 4.2.2. Machine-Learning Topic Modeling

As keyword-based approaches have limitation for social media data (Eisenstein, 2013), we used a machine-learning algorithm that extracts the latent structure of a dataset. This topic modeling approach clusters the data stream and filters the relevant tweets for further subsequent spatial analysis. Concretely, we used LDA, which is a probabilistic topic modeling algorithm that clusters semantic topics in a dataset. LDA is an unsupervised generative model that produces a document-topic distribution and a topic-word distribution. More information about the model and the hyperparameters of LDA can be found in (Blei, Ng, & Jordan, 2003).

Before the actual topic modeling procedure, social media posts need to be pre-processed, thus significantly improving the performance of LDA. We followed the steps defined by Resch et al. (2017), where every pre-processing step is explained in more detail. In the first step, every tweet is split at blank spaces so that every single character or sequence of characters can be treated individually (*tokenization*). Then all the words are set to lowercase to account for spelling mistakes and differences. In our experiment, URLs, special characters [e.g., ":" or ")"], short words (less than three characters), stop words (identified by a manual list and the list from Natural Language Toolkit (Manning et al., 2014, for English), and unique words that appear only once in the corpus, as well as numbers, are considered noise and are deleted. The remaining words are then reduced to their word stem using the Porter Stemmer (Porter, 1980).

In the next step, we applied LDA on the pre-processed data. We used the implementation of the Gensim library (Gensim, 2017) in Python and processed all the experiments with the following parameter values, which have been empirically derived, as no generically proven formal a-priori parameter estimation method exists so far: $\alpha = 0.0001$, $\beta = 1/\text{number\_of\_topics}$ and number_of_topics = 30. We set $\alpha$ to a value that is close

to zero because short documents such as tweets usually only contain a single topic (Zhao et al., 2011). The other two variables were chosen according to experimental evidence. In the final step, we classified the tweets in accordance with the topic with the highest probability. The extracted topics were then manually interpreted, focusing on *Olympics*-related and *transportation*-related topics. From our perspective, a topic is related to transportation when words like London, station, railway, underground, etc. have a high probability in a topic, whereas a topic is considered *Olympics*-related if the stem "olymp" has the highest probability, and other words like stadium, ticket, wembley, athlete, etc. also have a high probability. Examples of Olympics- and transportation-related topics can be found in the Results section.

### 4.3. Spatiotemporal Data Processing

In order to study the spatiotemporal behavior of residents and visitors in the three temporal bins (before, during and after the OG), we analyzed daily and hourly tweet intensities for the subsets of positive and negative tweets and the main semantic topics (LDA output), as well as the similarity patterns for spatial point distribution (Figure 2). In the last step, we investigated spatial hot spots using Kernel Density Estimation (KDE). The maps illustrating the results of the KDE can be found in the supplementary file.

To quantify spatial similarity between tweets before, during and after the OG, a nonparametric and area-based spatial point pattern test was used (Andresen, 2009; Andresen & Malleson, 2013). The test requires the following datasets: base points and test points for comparing spatial patterns and base polygons representing the areal units. We had 5,888 polygons as areal units using the administrative dataset of the Greater London Lower Super Output Area (LSOA) from 2011 (Greater London Authority's DataStore, 2017). The LSOA areas are only used for the similarity test in our study, to define a general pattern in tweeting behavior. The base points are the tweets from the "during OG" bin, both for residents and for visitors in two consecutive analyses. Whereas the *test points* are the tweets posted before and after the OG, first for residents then for visitors. The entire analy-
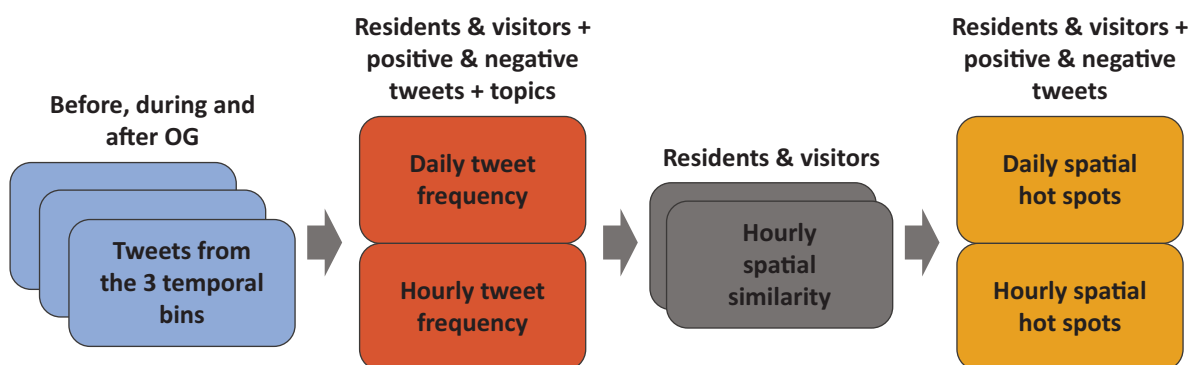


**Figure 2.** Spatiotemporal data processing overview.

sis was performed on hourly subsets. Regarding the base point dataset, the next step is to assign the points to the areal units (the LSOA) and then to calculate the percentage of points within each LSOA. For the dataset containing the test points, after assigning them to LSOA polygons as well, they should be randomly sampled, selecting 85% of the points and then calculate percentages (use Monte Carlo simulation to repeat this action 200 times). After that, we can create confidence intervals for each areal unit. Following these separate steps, the base percentage and test confidence for test interval are compared, and the result is the global index of similarity (for all the data) and the local one (for each areal unit). For this case study, the outcome of the test is a global index of similarity, where values range from 0 (no similarity) to 1 (identical). If the index is higher than 0.80, the two datasets are considered to be highly similar (Andresen, 2016). The index shows the level of similarity in the respective LSOA areas between the two analysis periods (Equation 1):

$$\frac{\sum_{i=1}^{n} s_i}{n} \quad \text{(Andresen, 2009)} \quad (1)$$

where $s_i$ is equal to one if two tweet datasets (in our case, similarity between the three temporal subsets, considered two by two) are similar in spatial unit $i$, and zero if the two are not similar at all. Further, $n$ is the total number of spatial units (the LSOA polygons).

To visually analyze the spatiotemporal characteristics of our findings, we generated hourly and daily density maps for positive and negative tweets during the three time bins according to the user groups (residents vs visitors). There are many spatial tools used to understand changes in geographical patterns (Chainey & Ratcliffe, 2005). For this case study, we chose the KDE method, which involves placing a kernel over each observation (tweet), and, by summing these kernels, showing a density estimation of the observations' distribution (Fotheringham, Brunsdon, & Charlton, 2000). We chose KDE be-

cause it belongs to a non-parametric class of density estimators, which has no fixed structure and depends on the point data to define an estimate; practically the form of the density is determined only from the data without any model. The parametric methods, such as Maximum Likelihood Estimation or Bayesian Estimation, assume to know the shape of the distribution. In addition, KDE is highly used for frequency distributions allowing a quick exploration of the dataset distribution. In the article the bandwidth selection was performed automatically by the software used, ArcGIS 10.4, where the kernel function is based on the quadratic kernel function. One of the main advantages of KDE is that it determines the spread of positivism and negativism in this case study, namely the area around a cluster where the likelihood for a positive or negative polarity is present based on spatial dependency. First, we split the data into hourly and daily segments and then ran the nonparametric KDE tool for each layer and temporal bin, which helped to illustrate spatial changes in residents' and visitors' tweeting behavior.

## 5. Results

Table 1 shows the summary of our main results that were generated through the methodology described above to provide an overview of the content and structure in this section.

### 5.1. RQ1 & RQ2: Geolocated Tweet Density and Sentiment Intensity for Temporal Subsets

We defined three temporal subsets for our analysis: the time period of the OG and the same number of comparison days before and after the Olympics to test the effect of the OG on spatiotemporal tweeting behavior and on the tweets' semantic content relating to RQ1 and RQ2.

One essential step for this study was to identify presumable residents and visitors. By applying the criteria

**Table 1.** Results summary.

|  | *Residents* | *Visitors* |
|---|---|---|
| **Positive** | • August 4: high positive peak for residents (gold medals for Great Britain), hot spots in the city center and at the Olympic Park; however, no increase in residents' raw tweets intensity.<br>• Opening Ceremony and Closing Ceremony clearly stand out in the number of positive tweets. | • August 4: positive sentiment peak in the daily temporal frame and slight increase in raw tweets intensity.<br>• Well-defined spatial hot spot at the Olympic Park. |
| **Negative** | • Mostly flat distribution on daily temporal patterns.<br>• More negative hot spots outside the city center during and after the OG. | • Low oscillations for all tweets and higher for the topics, e.g., the transportation topic.<br>• Before the OG, negative sentiment exceeds the positive for a few days. |
| **All tweets** | • Residents and visitors show different temporal and spatial patterns.<br>• Higher number of unique visitors tweeting during the OG.<br>• Tweets' spatial distribution per hour shows the highest similarity during the night (low number of tweets) and low similarity during the morning and evening (high number of tweets). | |

mentioned before (see Methodology section), we had to remove approximately 25% of the before OG tweets, 29% of the OG tweets and 24% of the after OG tweets. Practically speaking, we removed those users who tweeted exclusively in two temporal bins because it would have been difficult to distinguish whether they are just less actively tweeting residents or visitors who stayed a longer period than one month. The 11,571 London residents have the highest tweeting intensity during the OG compared to the visitors, who are texting more in the after OG period (Table 2). Regarding RQ1, we were able not only to distinguish residents and visitors in the dataset based on their temporal profile but also to identify clear and fundamental differences in the two groups' spatiotemporal behavior. Considering the different effect of planned events on residents and visitors, this finding has a key role in various planning-related social media analyses.

As for RQ2, large events tend to increase the social media participatory behavior (Wang, Can, Kazemzadeh, Bar, & Narayanan, 2012), which was also confirmed in this case by the highest density of tweets occurring during the OG (594,891 tweets). Another peak in tweeting intensity (545,693 tweets) was identified after the OG period (especially among visitors), which might be explained by the Paralympic games period and the London 2012 Festival as an accompanying event of the OG

to organize "the most culturally engaging" OG in history (Brown, 2012).

Further, one of our hypotheses was that positive sentiments in the text will occur more often during a large event compared to other usual days for the same locations. This assumption was confirmed by the obtained sentiment scores for the six datasets used in this study: 7.65% of the resident tweets and 6.02% of the visitor's tweets during OG are positive, while just 3.04% respectively 2.24% are negative (Figure 3). There was a noticeable decrease in negativity, while the positivity increased.

### 5.2. RQ3: Semantic Topic Extraction

In every sub-dataset (spatiotemporally divided, see Methodology section), we can identify one or more related topics for our target topics "Olympics" and "transportation". Table 3 shows the ten words with the highest probability in the topic. Due to the limited space, we visualize only some of the topics. The reason for the missing syllables of the words is the pre-processing step, stemming, which cuts the word to its root.

Table 3 shows that we can clearly identify topics related to "Olympic" and "transportation", distinguishing the periods *before*, *during* and *after* the OG, as well as between residents and visitors. In all of the "Olympic"-

**Table 2.** Residents and visitors for the three temporal subsets.

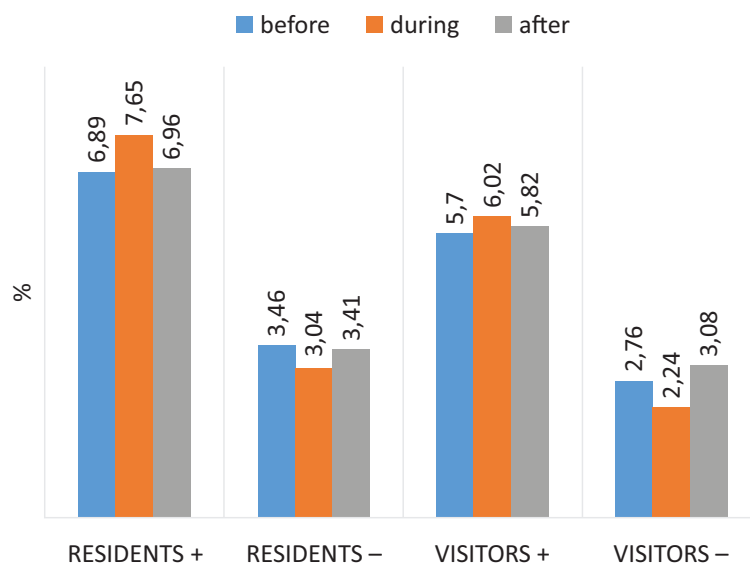| | STEP 1 | | STEP 2 | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Total geo-tagged tweets | | Residents | | Visitors | | Residents % | | Visitors % | |
| | *Tweets* | *Users** | *Tweets* | *Users** | *Tweets* | *Users** | *Tweets* | *Users** | *Tweets* | *Users** |
| *Before OG* | 478,551 | 46,357 | 195,319 | | 160,922 | 22,851 | 40.82 | 24.96 | 33.63 | 49.29 |
| *OG* | 594,891 | 59,248 | 210,024 | 11,571 | 190,770 | 30,443 | 35.30 | 19.53 | 32.07 | 51.38 |
| *After OG* | 545,693 | 54,956 | 178,100 | | 226,986 | 30,147 | 32.64 | 21.06 | 41.60 | 54.86 |

Note: * unique users.



**Figure 3.** Sentiment analysis distribution for resident and visitors during the three temporal subsets (%).

**Table 3.** Examples of words and their probabilities for the identified topics "Olympic" and "transportation".

| "transportation" visitors *before* OG | | "transportation" residents *during* OG | | "transportation" visitors *after* OG | | "Olympic" visitors *before* OG | | "Olympic" residents *during* OG | | "Olympic" visitors *after* OG | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| london | 0.3995 | london | 0.3435 | london | 0.3768 | olymp | 0.0426 | olymp | 0.2676 | paralymp | 0.1133 |
| greater | 0.0938 | greater | 0.0798 | greater | 0.0692 | loool | 0.0310 | excit | 0.0400 | olymp | 0.0958 |
| other | 0.0707 | other | 0.0670 | other | 0.0587 | year | 0.0280 | stadium | 0.0395 | stadium | 0.0526 |
| station | 0.0412 | station | 0.0465 | station | 0.0370 | walk | 0.0238 | final | 0.0395 | athlet | 0.0194 |
| hotel | 0.0207 | railway | 0.0228 | hotel | 0.0205 | point | 0.0228 | ticket | 0.0385 | found | 0.0186 |
| railway | 0.0191 | stratford | 0.0202 | bridg | 0.0167 | iphon | 0.0219 | photo | 0.0362 | serious | 0.0158 |
| underground | 0.0106 | street | 0.0173 | railway | 0.0166 | weird | 0.0144 | wembley | 0.0240 | watch | 0.0156 |
| victoria | 0.0098 | bridg | 0.0160 | tower | 0.0122 | shall | 0.0133 | game | 0.0237 | teamgb | 0.0149 |
| street | 0.0094 | venu | 0.0142 | arena | 0.0116 | togeth | 0.0132 | post | 0.0233 | paralympicsgb | 0.0147 |
| bridg | 0.0074 | underground | 0.0116 | cross | 0.0103 | restaur | 0.0127 | athlet | 0.0188 | problem | 0.0139 |

related topics, the word "olymp" has a high probability and a significantly high probability during OG compared to the other words in the topic. In the case of the "transportation"-related topics, multiple words show high probabilities, such as "station", "railway" or "underground". It is notable that the same words in the "transportation"-related topic show similar probabilities in the datasets for residents and visitors in different time periods. In the dataset after OG, when the Paralympics took place, "paralymp" is also the most probable word in the "Olympic"-related topic.

### 5.3. RQ2: Similarity Index

Figure 4 shows the similarity values in hourly bins as defined above (see Methodology section). The highest similarity values occur during the night when tweets are posted from the same LSOA areas, but they don't have a high density, according to the hourly intensity results. Starting at 5:00 a.m. the similarity curve decreases until
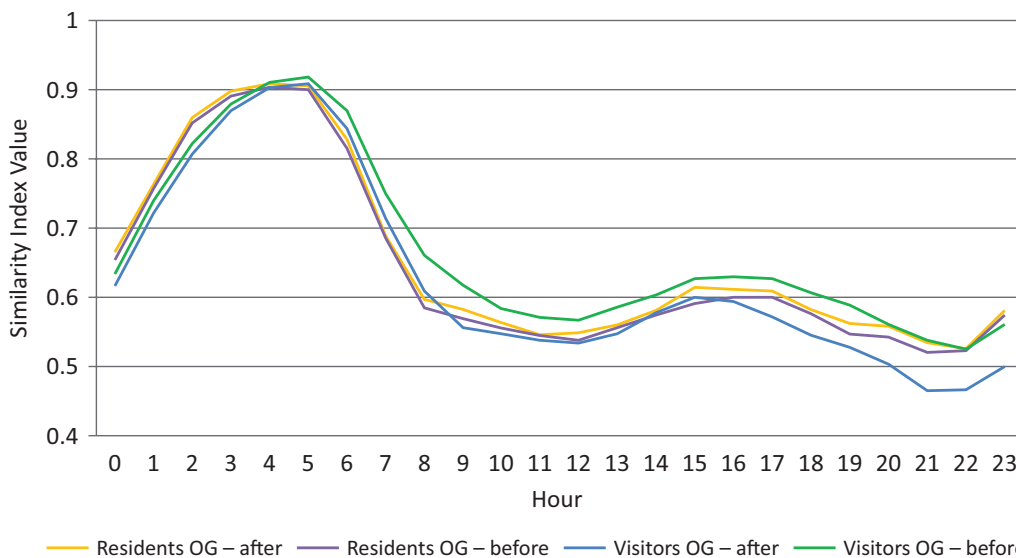
around 9:00 a.m. This shows that during the OG the spatiotemporal behavior of the users is different compared to before and after the OG. The more noticeable differences are at the end of the day, after 6:00 p.m., between residents and visitors after OG (ranging from 0.5814 to 0.5019), and between both visitor datasets (ranging from 0.5635 to 0.5019).

### 5.4. RQ1 & RQ2: Temporal Analysis

After extracting the topics and defining the sentiments of each tweet, we analyzed the temporal distribution of the negative and positive (see Methodology) tweets of residents and visitors, both on hourly and daily levels.

#### 5.4.1. Daily Patterns

The daily tweet intensity using the raw Twitter data for *residents* showed two temporal peaks during the OG, at the Opening Ceremony and at the Closing Ceremony. The

**Figure 4.** Similarity Index distribution for OG tweets as base points (range 0 to 1 where 1 means identical pattern for both analysis periods and 0 represents no similarity at all).
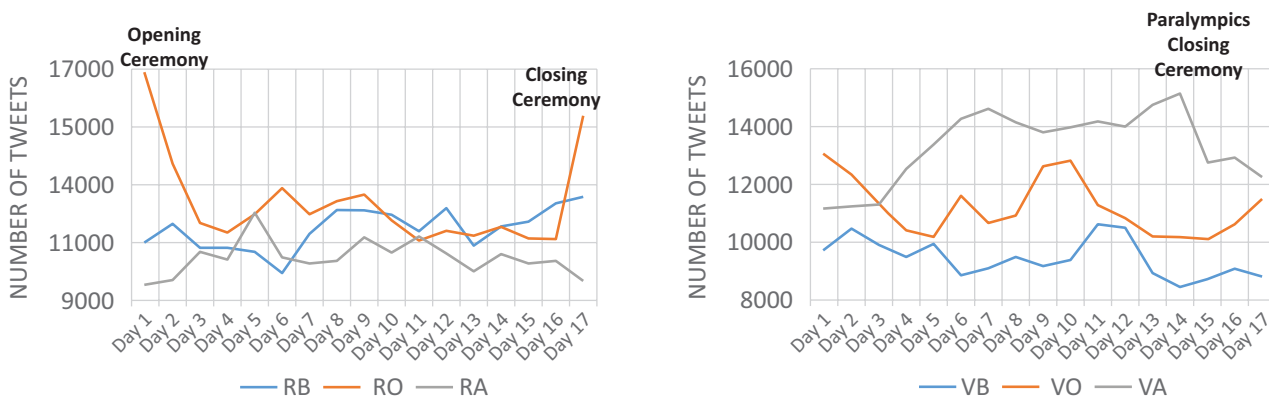
*visitors'* time series (unlike the *residents*) showed a peak in the OG period around August 4, when Great Britain won three gold medals in athletics. What is surprising is the higher volume of tweets after the OG for the visitors compared with the other time bins, including a peak during the Paralympics Closing Ceremony (Figure 5).

Next, we compared the daily patterns for sentiments in the data subsets. Figures 6–8 illustrate daily intensities in sentiment distribution for residents and visitors during the three temporal frames. It shall be mentioned that in a small number of days the tweets volume for the specific topics is low, especially for the "transportation" topic.
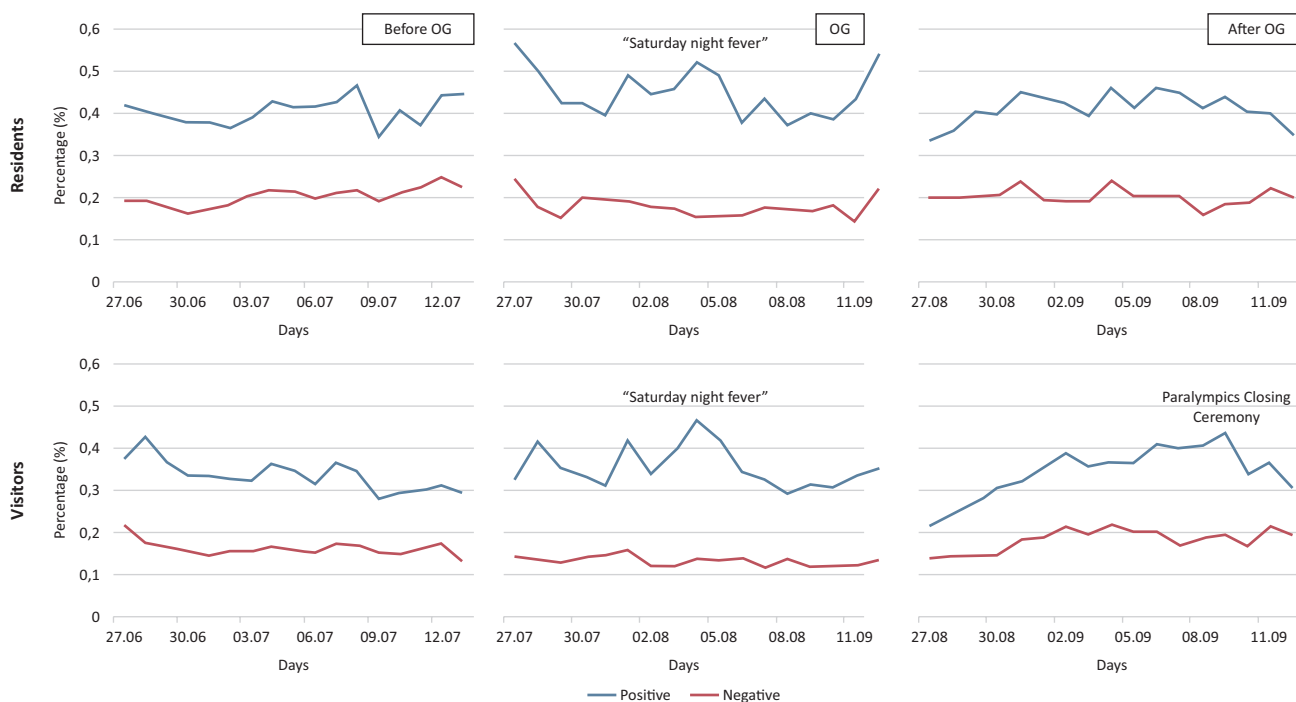
5.4.1.1. Positive vs Negative Tweet Trends

While analyzing the tweets for residents and visitors, the daily distribution of negative tweets was fairly equal

and smooth for "all tweets" (all six data frames, all topics). The negative tweets for residents included in the "olympic" topic have an almost flat trajectory, similar to the ones including "all tweets" (Figure 7 vs Figure 6), except July 11, when they showed an increase and the hot spot map showed higher intensity in the London Center areas, Lewisham and Morden, close to Wimbledon. The intensity of positive tweets was more predominant than the negative ones at any time (Figure 6 and Figure 8), with higher values during the OG and spatial concentration around the Olympic Park (Figure 9). For the "olympic" topic, the positivity curve reaches its maximum for residents on August 4 (from 0.15% negative tweets to 0.86% positive tweets), while for "all tweets" the peak is higher for visitors (Figure 6). In the newspapers, this day is referred to as "Saturday night fever", when Jessica Ennis, Greg Rutherford and Mo Farah all
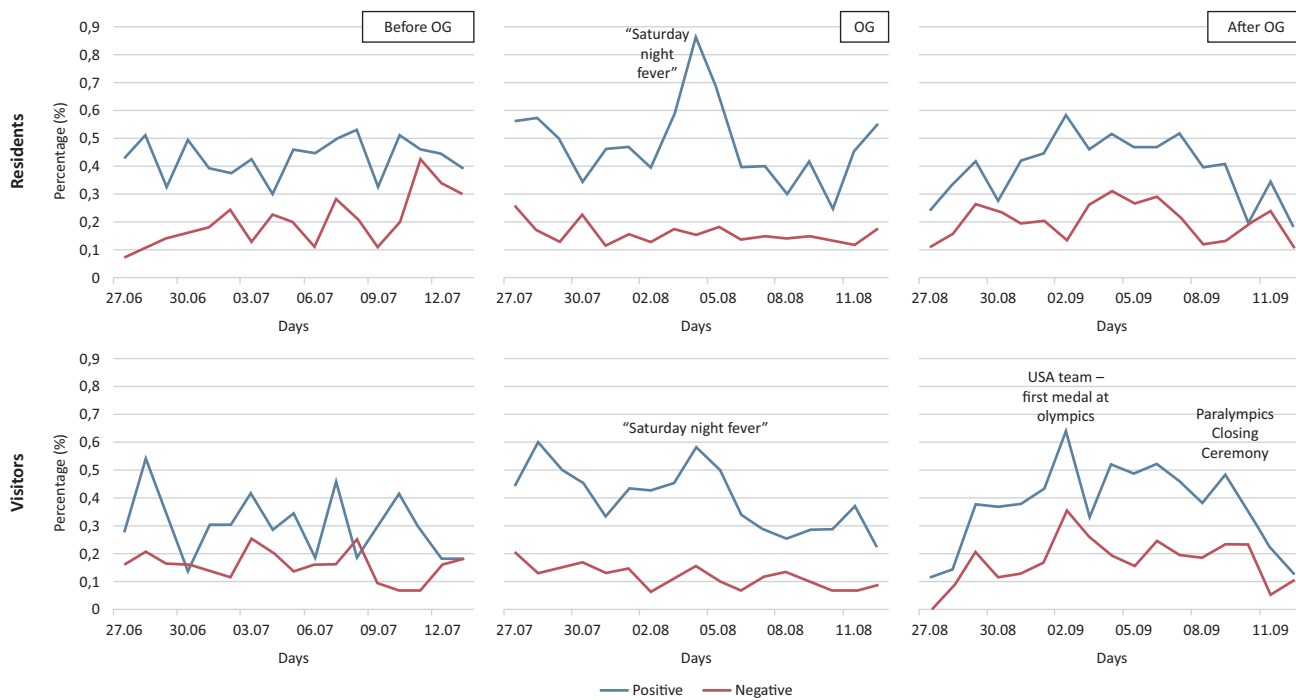


**Figure 5.** Daily tweets density for residents and visitors for the three temporal bins (R = residents, V = visitors; B = before OG, O = during OG, A = after OG).
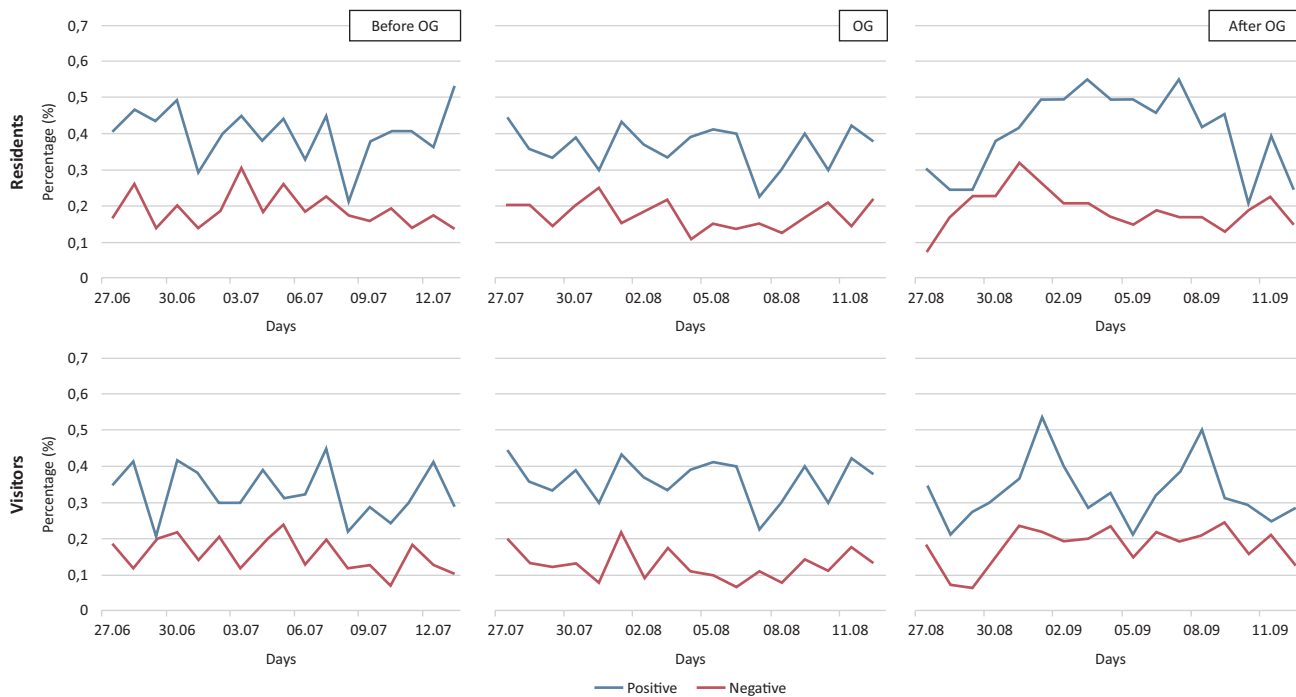


**Figure 6.** Daily sentiment analysis distribution for residents and visitors for "all tweets" (percentage of total number of tweets in the given temporal bin for the respective categories).

**Figure 7.** Daily sentiment analysis distribution for residents and visitors for the "olympic" topic (percentage of total number of tweets in the given temporal bin for the respective categories).



**Figure 8.** Daily sentiment analysis distribution for residents and visitors for the "transportation" topic (percentage of total number of tweets in the given temporal bin for the respective categories).

won gold medals for the host nation. This shows higher public engagement when an action such as winning a prize by co-nationals takes place.

Regarding the "after OG" period, on September 2, when the USA team won the first medal in the trunk and arms mixed at the Paralympics, a decrease of negativity happened for residents, while during the same day an increase occurred for the visitors, together with an in-

crease in the positive tweets (Figure 7). A common positive peak can be observed for "all tweets" and "olympic" on September 9 (Figure 6 and Figure 7), during the Paralympics Closing Ceremony. In comparison, the sentiments distribution of the "transportation" topic contains a smoothed zig-zag line, and an increase followed by a decrease in the positive tweets after the OG. Interestingly, on August 4 there was no peak in the results for ei-

ther residents or visitors, compared to the other subsets. The tweeting behavior after the OG for "transportation" shows a higher difference between positive and negative tweets for residents, mostly from September 1 to September 9 (Figure 8). However, the maximum tweet volume is 30 per day.
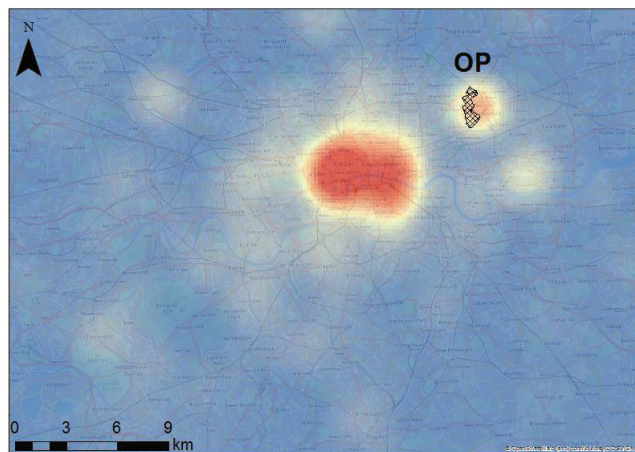
### 5.4.1.2. Daily Trends of Residents vs Visitors

*Before the OG*, residents and visitors for "all tweets" showed slight changes in the trend line (Figure 6), while for the topic "olympic" the visitors' tweets tended to form a zig-zag-like time series (Figure 7), similar to the "transportation" topic (Figure 8). No daily spatial hot spots were found in the Queen Elizabeth Olympic Park for this period.

*During the OG*, residents and visitors for "all tweets" showed a higher volume of positive tweets, on August 1 and August 4 (Figure 6). On August 1, the spatial hot spots are distributed between London's central area and the Olympic Park area, and on August 4 a high density is located specifically around the Olympic venues: the Olympic Park zone, the River zone (including Greenwich park), and the Central zone (including Hyde Park and Re-
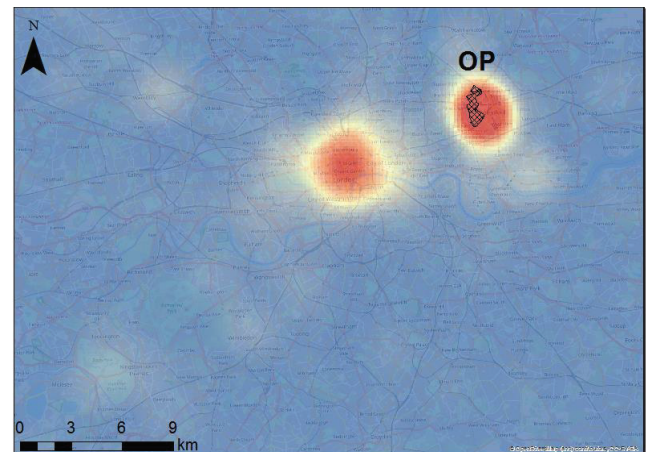
gent's Park). For the residents, we also notice smaller hot spots in many parts of the city, which suggest an increased interest in people's tweeting behavior for a special occasion (Figure 9). August 4 is also a common peak for "olympic" tweets, mostly for residents, with a hot spot location around the Olympic Park, the city center and another one between these two as well, almost continuously. In the same time, the visitors show the hot spot only around the Olympic Park and with much lower intensity in the city center (Figure 7, 9). Interestingly, August 4 showed a positive peak that is not connected with the increase in the intensity of the raw tweets for residents. Another dissimilarity arose for the "transportation" topic, where the graphic of sentiments distribution showed a different pattern (Figure 8). For example, the highest positive peaks for the OG period occur on August 1 for the visitors and August 2 for the residents. The visitors' tweeting hot spots are in the city center and at the Olympic Park, while the residents' tweets are clustered in an elongated hot spot with median values around the Olympic Park.

For the *After OG* period, September 2 was a peak of positive emotions for residents and visitors for the
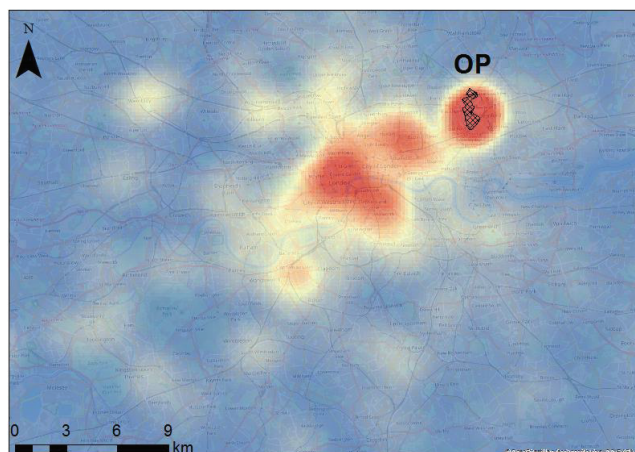
August 1 — residents (n = 1,029 mean = 0.31)
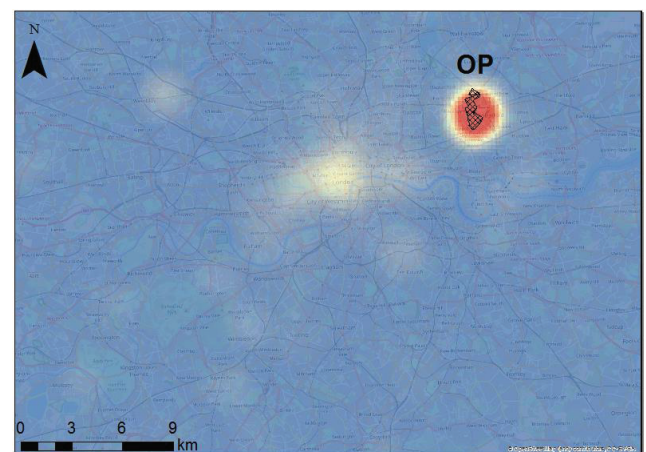


August 1 — visitors (n = 796 mean = 0.23)



August 4 — residents (n = 1,083 mean = 0.32)



August 4 — visitors (n = 878 mean = 0.26)



**Figure 9.** Hot spots of positive tweets (OP = Queen Elizabeth Olympic Park; red = high density, yellow = average density, blue = no density).

"olympic" topic, showing an intense hot spot for the visitors at the Olympic Park, while the residents' hot spot included the park, but the center was shifted towards the western part of the park. On September 9 residents from the "olympic" topic (Figure 7) and visitors for this topic and "all tweets" (Figure 6) showed an increased positive feeling, possibly caused by the Paralympics Closing Ceremony. Also, on the spatial density for this day the hot spots are located in the approximate city center and in the Olympic Park zone. A different temporal pattern occurred for the "transportation" topic of the visitors' tweets (Figure 8). Two predominant days showed positive peaks, including September 1, when the majority of tweets are from the Olympic Park, and September 8 when all the active tweeting happened in the London center.
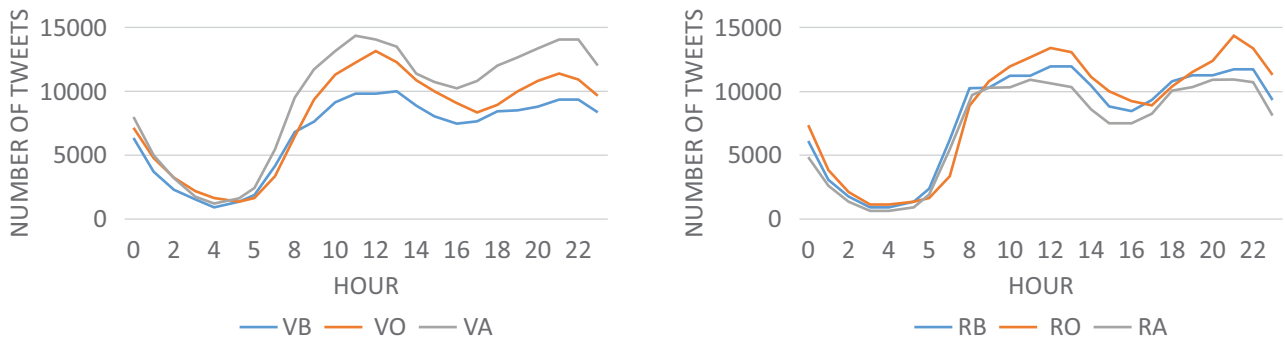
### 5.4.2. Hourly Patterns

After identifying peaks and patterns in the data on a daily level, we also analyzed the tweets' hourly spatial and temporal distribution. Figure 10 shows a general overview of the hourly distribution for the raw number of tweets per user groups in the two weeks temporal frames (a total of six weeks). Residents before OG showed a cyclical daily circular pattern, with low intensity overnight then a rapid increase in the morning and a steep decrease a few hours after midnight. Interestingly, we can define peaks for residents and visitors during the OG for the Opening Ceremony and also for the Closing Ceremony, together with the "Saturday night fever" mentioned in the daily patterns on August 4 for residents. Another peak in tweeting intensity occurs on September 9 for the residents, possibly due to the Paralympics Closing Ceremony.

After summing up the number of tweets per hour, we noticed only a slight change in the temporal tweeting behavior of the residents between the OG period and comparison days. However, after hour 20, a small increase occurs when compared to the "before" and "after" data (Figure 11). For the visitors, the tweets volume is surprisingly higher during the afternoon and the evening considering the after OG period. The number of unique users after the OG was lower than during the OG, which shows that there were fewer users after the OG, but they were more active than the ones who tweeted during the



**Figure 10.** Residents and visitors tweeting behavior per hour for the three temporal bins.

**Figure 11.** Hourly distribution of the tweets during the three analysis periods (R = residents, V = visitors; B = before OG, O = during OG, A = after OG).

OG. This observation may be explained by various factors, such as the start of school when students are more active, who may potentially be misclassified as visitors because of their limited Twitter activity over the summer from the school/university location (in London).
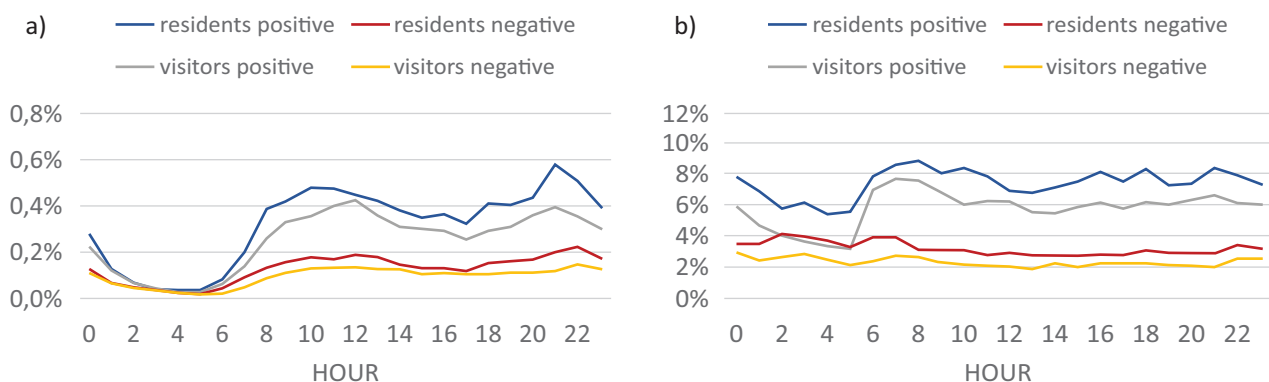
Furthermore, the hourly patterns were analyzed considering the sentiment score and also the topic allocation. Figure 12 shows the amount of positive and negative tweets for both groups per hour. It is evident that the number of negative tweets never exceeds the number of positive ones in either of the user groups. The positive peaks differ between residents and visitors: Residents tended to tweet more in the evening (around 9 p.m.), whereas visitors tweeted slightly more around 12:00 noon. If we only consider the tweets where the topic "olympic" was identified by the algorithm (Figure 13), these characteristics become even more striking. However, if we compare the relative tendencies (all positive/negative tweets in the given hour for each period) these peaks are smoother, which means, there are generally higher number of positive tweets in the evening.

The following videos[1] show the spatiotemporal pattern of positive and negative tweets (aggregated to 24h hours such as for Figure 12, in 10-minutes timeframes)

in both user groups for all the three analysis periods. (A static version containing four different hours during the day can be found in the supplementary file.) Blue points represent negative tweets, whereas the positive ones are visualized in red. The semi-transparent points representing each tweet stay there for two hours to illustrate the density of tweets. Each video shows all the three temporal bins after each other (3 × 24 hours) for our two groups of users (residents and visitors), and a clock shows the current time in the lower right corner.

5.4.2.1. Changes in the Pattern Comparing the OG Period to Before and After

This section reflects to RQ2, as we compared the before and after OG periods to the patterns during the OG. For the residents, we can see that the core of the main hot spot is constant for each hour throughout all the analysis periods, and it is located mainly in the city center. For the negative tweets, the before and after periods are quite similar during the day, but for the positive tweets, there is still a hot spot around the Olympic Park. The reason for this is that our analysis period after the OG includes the days of the Paralympic Games as well. At the same



**Figure 12.** Positive and negative tweets per hour (%): a) as *absolute* values–every hour compared to the number of all tweets during the OG for residents and for visitors; b) as *relative* values–every hour compared to the number of all tweets in that hour during the OG for residents and for visitors).

---

[1] http://giscience.zgis.at/gisce/Videos_Towards_Citizen-Contributed_Urban_Planning_through_Opinion_Mining_of_Twitter_Data

time, for the visitors, the pattern during the OG and in the other two periods is not that different. Except during the morning hours, but the lower number of tweets can explain this because in this case even 2–3 point can result in relatively strong hot spots.

**Figure 13.** Positive and negative tweets per hour (%) for residents and visitors for the "olympic" topic.



### 5.4.2.2. Pattern of the Average Day: Before, During and After the OG

For residents during the OG, there are a few extra hot spots reflecting the venues of the OG. In the periods before and after the OG, the smaller hot spots in the outer parts of the city occur mostly in the morning and the evening, probably due to commuting, but only for the negative tweets. Interestingly, for positive tweets from visitors, the morning hot spot is more concentrated (they do not commute) but only before and during the Olympic Games. During the rest of the day, the patterns do not change significantly in any of the analysis periods, both for positive and negative tweets.

### 5.4.2.3. Positive vs Negative Tweet Trends

In general, the positive tweets of the residents tend to be more concentrated with one main hot spot, except the morning after the OG where the negative tweets are conglomerated. However, during the OG in the evening, the size of the positive tweet hot spot is much larger. Probably that was the time when most of the residents were tweeting about the Olympics. For visitors, this concentration of one large hot spot for the positive tweets is not significant; there are smaller hot spots around the city in both cases, and, in general, the hot spot for the negative tweets is larger in extent.

### 5.4.2.4. Residents vs Visitors

The most significant difference in patterns between residents and visitors is the distribution of negative tweets in the morning, but again this could be a result of the low number of tweets. Also, the smaller hot spots of the visitors' tweets tend to be more on the Eastern side of the study area, especially in the period after the OG. In general, it can be significant for planners to further investigate the trends and possible causes for negative tweets,

as partially it might be connected to planning-related issues such as low satisfaction with infrastructure or poor quality of services.

## 6. Discussion

### 6.1. Integration of the Results into Planning Processes

The major objective of our case study was to illustrate the general potential of Twitter data analysis for urban planning purposes in the case of large planned events. Thus, we identified and addressed research gaps, such as the distinction between residents and visitors regarding the Olympics and comparing event days and non-event days along with both spatiotemporal and content analysis in one study. Consequently, our results serve as a basis for further, more in-depth analyses.

In general, both previous research and the work presented in this article have shown that results from social media analysis are directly usable in urban planning processes, including the general ability to detect sentiments that are associated with places (Resch et al., 2016). In this regard, social media provide people (local citizens and visitors) with a simple and powerful instrument to share their opinions and subjective impressions. This is particularly relevant with respect to connecting social media posts to specific urban events such as Olympic Games or other large sports events, for gaining insight into the perceptions of the urban population regarding these events. In fact, social media are a valuable, open source of information for urban planning.

This openness is of particular importance because urban planning processes are oftentimes still characterized by closed communication between local and official actors, lacking open discussion and transparent procedures (Resch et al., 2016). Moreover, openness and transparency are increasingly a key factor for successful urban planning, allowing for an efficient weighing process that considers the opinions and sentiments of different stakeholders. Current planning processes, however, are mostly shaped by deductive processes, which are typically introduced and controlled by urban governments, oftentimes neglecting or not sufficiently integrating the needs of the citizens. In this context, social media play a key role because they provide an instrument for organizing public participation activities and citizen initiatives. On the positive side, the integration of public discussions on social media and other digital platforms also increase the validity and acceptance of governmental decision-making because traditional planning methods are complemented by new "human sensor" data that reflect the citizens' wishes and needs (Zeile, Resch, Exner, & Sagl, 2015).

This is in clear contrast to top-down approaches that follow different decision-making principles. Integrating social media into urban planning may be able to provide unseen insights into citizens' thoughts, perceptions and expectations concerning urban events in an induc-

tive bottom-up approach. In this sense, urban planning-related discussions are, to some degree, self-organizing, giving citizens the chance to discuss planning issues in a peer-to-peer process, rather than in a government-driven one. However, the issues of the digital divide, that mostly younger, better educated, and more technologically savvy people participate in social media networks, should be addressed in social media analysis (Czepkiewicz, Jankowski, & Młodkowski, 2017). Due to this digital divide social media platforms are currently not representing the entire society or population appropriately (Diaz, Gamon, Hofman, Kiciman, & Rothschild, 2016; Mellon & Prosser, 2017), therefore conclusions drawn from the analysis depending on the phenomenon should be handled accordingly. The extremes are especially underrepresented in terms of age (very young, and older generations), economic situation (those who cannot afford access through internet or gadget), etc. Consequently, we are aware that the social media-based approach shows a number of limitations; still it may complement current urban planning procedures through an improved understanding of the city as a living organism through proactively engaging citizens into urban planning (Resch, 2013).

Based on the methods we used and their outcomes, we can identify two main types of further planning-related investigations (the list of the examples is not complete):

a) *Macro-scale:*

- On a city level, it is possible to point out the differences in the general mobility patterns compared to non-event days (also at different times during the day) and use it for further transportation modeling. The different needs of residents and visitors should be considered;
- Planners can also further investigate the hot spots for negative tweets in both groups (residents and visitors); as they show different trends, there might be different reasons behind them. These hot spots can be compared with the extracted topics in these areas, whether they are related to transportation or other planning-related topics, the event itself, or something else entirely;
- Regarding the extracted topics, it is possible to search for more specific terms, if the planners provide expert knowledge. Furthermore, other terms can be identified that are connected to the planning-related topics and have not been considered by urban planners yet. The sentiment and spatial distribution of these topics all over the city can then be explored, both for residents and visitors, as these might also differ in this case.

b) *Micro-scale*:

- We could clearly identify activity patterns related to individual venues of the Olympic Games. An interesting example could be to focus on a venue and explore the behavioral patterns of residents and visitors and the effect of a given event at that venue. (Specifically, right before or after and during the analysis.) Do the residents tend to be more negative? Or maybe less active during that time?
- Additional datasets are definitely advantageous for the micro-scale analysis. Planners can explore deeper connections between the event and other urban processes. For example, the effect on the local economy of using bank cards can also be analyzed. (Habidatum, 2017);
- Extracting information on a user level is also an option. For visitors, planners can trace the intra-urban mobility patterns, if they tweet regularly during the day (between their accommodation and the venues). This analysis is even more accurate with additional mobile phone data analysis.

*6.2. Psychological Biases in Human Language and Social Media and Their Relevance for Urban Planning*

Most generally, there is a universal positive bias in human language (Dodds et al., 2015). The findings of the present study are congruent with this kind of bias: A higher percentage of positive tweets than negative tweets were identified. Moreover, the residents show a clear peak of positive sentiments during the OG, even though there have been several examples observed recently where local people opposed the organization of the Olympics in their cities (e.g., Kaufmann, 2015; Moore, 2015; Sims, 2017). It might suggest that, once underway, world-class events in a city boost self-respect and pride of the city residents, and the perceptions of their benefits are typically optimistic (Whitson & Macintosh, 1993), which might not be true in the planning phase (e.g., Dempsey, & Zimbalist, 2017). However, it is important to consider that the positive cognitive bias and the homeostatic happiness maintain satisfaction in life, and self-beliefs can act as reality buffers (Cummins & Nistico, 2002).

This, again, raises the question of whether the positive cognitive bias can be a buffer that masks some inconveniences in the city that large events such as the OG may cause. For example, Ritchie, Shipway and Cleeve (2009) identified that, in general, urban residents supported important events in their area, but were concerned with some issues such as traffic congestion and an increasing cost of services. Additionally, the benefits of large events in a city can differ between social groups.

For example, younger residents, residents that have a higher socioeconomic status, and residents that live farther away from the event's location, are more likely to perceive additional benefits from the event (Ritchie et al., 2009; Whitson & Macintosh, 1993). The results of our study suggest positivism related to the OG, but beyond this social media positivism, there are several considerations explained above that urban governments and urban planners need to study. In other words, the obtained results are good indicators of the importance of large sports events for residents' life satisfaction, but, in urban planning, these results cannot be isolated from the rest of the city dynamics.

These issues become critical if we consider that urban governments are usually open to investing in consumer and entertainment-oriented developments, such as sports events (Harvey, 1987). However, citizens in a city are more than consumers. Additionally, a large sports event causes changes in different dimensions in a city such as image, knowledge, and emotions, where the long-term effects of these changes are complex to understand (Preuss, 2007). We believe that long-term social media analysis can be considered a necessary instrument to monitor these effects in a city and to offer more pluralistic information to urban planners. Urban planners can use this information to evaluate different qualitative and quantitative costs and benefits of a large sports event. Further research needs to develop new approaches to study large urban events' legacies using social media. At the same time, these approaches need to be enriched with robust epistemologies to understand the complex and dynamic human behavior in the virtual world (social media), without disconnection of the human behavior in the real world.

### 6.3. The Effect of the Paralympic Games on the Selection of the Temporal Bins

We were aware of the fact that there were days related to the Paralympics in the third temporal bin and, thereby, our comparison might show less significant differences. However, we also tested a fourth temporal bin (September 27–October 13, 2012), and the patterns in the original temporal bin (August 27–September 12, 2012) were not biased, except the day of the closing event. Therefore, we decided to keep the original after OG temporal bin because selecting days so much later can also have an effect on the final results, and those would be more difficult to interpret, such as different seasonal effects, or extraordinary events.

### 6.4. Identifying Residents vs Visitors

The process of categorizing Twitter users into "residents" and "visitors" is challenging, and, to our best knowledge, there is no "ground truth" methodology, providing unquestionable results, in the related literature. Also, Abbasi et al. (2015) stated that dividing social media users

into residents and visitors is not an easy task. They categorized these types of people in Sydney for city trips supporting urban planning. Our study follows an adapted method from the original when the residents are defined as users tweeting at least ten times in at least $n-1$ phases of the temporal data analysis. One reason for using this adapted approach is that our datasets had specific time frames related to the OG event, and we hypothesized that for finding so-called active residents, they would need to tweet in all n temporal bins, while the visitors had to tweet just in one of the temporal bins.

The spatiotemporal patterns identified and described in this study show a relative verification of this approach. For example, August 4 was an important day for Great Britain and it was undoubtedly reflected in the spatial and temporal patterns for both presumable residents and visitors: as a monocentric, well-defined daily hot spot around Olympic Park for the visitors, and as polycentric hot spots for residents in many parts of the city, with particularly high density around the city center and also around the park. The daily temporal graphic for positive tweets supports the hot spot map for residents by highlighting a larger increase in positive sentiments compared to visitors. This may be because visitors are generally more excited and tweeting positively for all the OG results, while residents are more interested in Great Britain's performance. August 4 is important in the analysis because it emphasizes people's behavior and how the positive event of winning three gold medals changes the spatial distribution of tweets. Also, for the hourly hot spot detection, we notice more intense hot spots around the Olympic Park for visitors than for residents, i.e. at 8:00 a.m. or 12:00 noon there are only mild or nonexistent hot spots for residents at the park.

However, this particular approach has limitations: We did not consider the declared language of the users (e.g., maybe non-English speakers are more likely to be visitors tweeting before and during OG); the tweets' "user location" field was not used considering the biased information introduced subjectively by the user; data availability–we only had access to 2012 London data, whereas having worldwide data would have been helpful for exploring user's activity status. An interesting future approach may be to adopt all location-related features and create an index of defining residents and visitors.

### 6.5. Topic Modeling

In our analysis we used the basic LDA model for topic modeling that follows a "bag of words" approach, meaning that it uses solely the frequency of terms in a document and does not take grammar or word order into account. A significant problem is names, which consist of two words like "Greater London", where each word is treated independently. As "Greater" and "London" are words that are commonly used in combination, analyzing biterms may increase the quality of our results. However, there are other names like "Olympic Games" where

"Games" is a common word in an English conversation. For our particular case, relevant biterms include "Greater London", "Olympic Games" and "Victoria Station", which ware included as single words in Table 3.

### 6.6. Influence of Topic Modeling on the Sentiment Score

The distribution of positive tweets during the OG for the "olympic" topic is different for both residents and visitors. While depicting the possible reasons, we noticed the different word probabilities resulting from the LDA algorithm, such as the word "olymp", with a probability of 0.2676 during the OG period for residents and 0.0426 after the OG for visitors. The inclusions of words not relevant to the topic (which is subjectively named after checking the highest words probabilities), such as "iphon" in the Olympic topic might lead to an unexpected temporal distribution. Also, the sentiment score function is limited in defining a high volume of positive and negative tweets, their majority being labeled as neutral, which may result in increased fuzziness in interpreting the results.

### 7. Conclusion

In conclusion, our findings validly answer our research questions: Through spatiotemporal and sentiment analysis of the tweets, we could identify significant patterns in terms of our two defined user groups as well as for the days before, during and after the event. Additionally, the uncertainty originating from the identification of the members of each group due to the lack of additional details can be reduced by integrating further datasets (e.g., cab rides, bicycle network and public transport usage, mobile phone data).

Regarding the utilization for planning purposes, we can state that despite the limitations described above, by applying our workflow to the sample dataset we can provide valuable information about the spatiotemporal behavior and sentiment of residents or visitors concerning large planned events. By comparing our results to important dates of the event (e.g., the Closing Ceremony, "Saturday night fever") or location of the venues, we could validate our results both content-wise and for spatiotemporal patterns, even on finer spatial and temporal scales. Last but not least, topics that are directly related to planning and transportation could be extracted and can be further analyzed for specific urban planning purposes in the future.

Concluding, the case study was also appropriate to illustrate the potential of utilizing social media data for sentiment analysis and topic modeling in order to provide general feedback regarding large planned events. Nevertheless, there are possible ways for improvement beyond the scope of the current study that can also aid to overcome some of the already mentioned limitations. One such option is to design a geovisual analytical tool to interpret the large amounts of data (e.g., maps, graphs, tweets, time periods), also supporting users who are less familiar with GIS concepts and methods. Furthermore, as an outlook to participatory planning, the acquired knowledge could be presented in a Volunteered Geographic Information platform, which is directly connected to the event and where people can provide feedback with location data.

### Conflict of Interests

The authors declare no conflict of interests.

### References

Abbasi, A., Rashidi, T. H., Maghrebi, M., & Waller, S. T. (2015). Utilising location based social media in travel survey methods: Bringing Twitter data into the play. In *Proceedings of the 8th ACM SIGSPATIAL International Workshop on Location-Based Social Networks* (pp. 1–9). Redondo Beach, USA: ACM.

Andresen, M. A. (2009). Testing for similarity in area-based spatial patterns: A nonparametric Monte Carlo approach. *Applied Geography*, *29*(3), 333–345.

Andresen, M. A., & Malleson, N. (2013). Spatial heterogeneity in crime analysis. In M. Leitner (Ed.), *Crime modeling and mapping using geospatial technologies* (pp. 3–23). Dordrecht, Netherlands: Springer.

Andresen, M. A. (2016). An area-based nonparametric spatial point pattern test: The test, its applications, and the future. *Methodological Innovations*, *9*, 1–11.

Aubrecht, C., Ungar, J., & Freire, S. (2011). Exploring the potential of volunteered geographic information for modeling spatio-temporal characteristics of urban population A case study for Lisbon Metro using Foursquare check-in data. In *Proceedings of the 7th International Conference on Virtual Cities and Territory* (pp. 57–60). Lisbon: Nova University of Lisbon.

Blei, D., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022.

Bollen, J., Mao, H., & Pepe, A. (2011). Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In *Proceedings of the Fifth International Conference on Weblogs and Social Media* (pp. 450–453) Barcelona, Spain: AAAI Press.

Breen, J. O. (2012). Mining Twitter for airline consumer sentiment. In G. Miner, J. Elder, A. Fast, T. Hill, R. Nisbet, & D. Delen (Eds.), *Practical text mining and statistical analysis for non-structured text data applications* (pp. 133–149). Waltham, MA: Elsevier Academic Press.

Brown, M. (2012). Cultural Olympiad 2012 reaches the critical masses. *The Guardian*. Retrieved from https://www.theguardian.com/culture/2012/mar/12/cultural-olympiad-2012-critical-masses

Castells, M. (1996). *The rise of the network society: Volume I: The information age: Economy, society, and culture* (1st ed.). Oxford, UK: Blackwell Publishing.

Cebelak, M. (2013). *Location-based social networking data: Doubly-constrained gravity model origin-destination estimation of the urban travel demand for Austin, TX* (Master Thesis). Faculty of Graduate School, University of Texas, Austin.

Chainey, S., & Ratcliffe, J. (2005). *GIS and crime mapping* (1st ed.). Chichester, UK: John Wiley & Sons.

Chen, Y. (2012). *Urban strategies and post-event legacy: The case of summer Olympic cities*. Paper presented at the AESOP (Association of European Schools of Planning) 26th Annual Congress, Ankara, Turkey, 11–15 July 2012.

Cheng, Z., Caverlee, J., & Lee, K. (2010). You are where you tweet: A content-based approach to geo-locating Twitter users. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management* (pp. 759–768). Toronto, Canada: ACM.

Collins, C., Hasan, S., & Ukkusuri, S. V. (2013). A novel transit rider satisfaction metric: Rider sentiments measured from online social media data. *Journal of Public Transportation*, *16*(2), 21–45.

Cook, I. R., & Ward, K. (2011). Trans-urban networks of learning, mega events and policy tourism: The case of Manchester's Commonwealth and Olympic Games projects. *Urban Studies*, *48*(12), 2519–2535.

Corney, D., Martin, C., & Göker, A. (2014). Spot the ball: Detecting sports events on Twitter. In *Proceedings of Advances in Information Retrieval* (pp. 449–454). Amsterdam, Netherlands: Springer.

Costanza, R., Fisher, B., Ali, S., Beer, C., Bond, L., Boumans, R., . . . Snapp, R. (2007). Quality of life: An approach integrating opportunities, human needs, and subjective well-being. *Ecological Economics*, *61*(2/3), 267–276.

Crooks, A., Pfoser, D., Jenkins, A., Croitoru, A., Stefanidis, A., Smith, D., . . . Lamprianidis, G. (2015). Crowdsourcing urban form and function. *International Journal of Geographical Information Science*, *29*(5), 720–741.

Culotta, A. (2010). Towards detecting influenza epidemics by analyzing Twitter messages. In *Proceedings of the First Workshop on Social Media Analytics*. Washington DC, USA: ACM.

Cummins, R. A., & Nistico, H. (2002). Maintaining life satisfaction: The role of positive cognitive bias. *Journal of Happiness Studies*, *3*, 37–69.

Czepkiewicz, M., Jankowski, P., & Młodkowski, M. (2017). Geo-questionnaires in urban planning: Recruitment methods, participant engagement, and data quality. *Cartography and Geographic Information Science*, *44*(6), 551–567.

Dalvi, N., Kumar, R., & Pang, B. (2012). Object matching in tweets with spatial models. In *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining* (pp. 43–53). Seattle, USA: ACM.

Dempsey, C., & Zimbalist, A. (2017). *Boston Olympics: How and why smart cities are passing on the torch* (1st ed.). Lebanon, USA: University Press of New England.

Diaz, F., Gamon, M., Hofman, J. M., Kiciman, E., & Rothschild, D. (2016). Online and social media data as an imperfect continuous panel survey. *PloS one*, *11*(1).

Dodds, P. S., Clark, E. M., Desu, S., Frank, M. R., Reagan, A. J., Williams, J. R., . . . & Danforth, C. M. (2015). Human language reveals a universal positivity bias. *Proceedings of the National Academy of Sciences*, *112*(8), 2389–2394.

Eisenstein, J. (2013). What to do about bad language on the internet. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human language technologies* (pp. 359–369). Atlanta, USA: The Association for Computational Linguistics.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2000). *Quantitative geography: Perspectives on spatial data analysis* (1st ed.). London, UK: Sage.

Frank, M. R., Mitchell, L., Dodds, P. S., & Danforth, C. M. (2013). Happiness and the patterns of life: A study of geolocated tweets. *Scientific Reports*, *3*(1), 2625.

Fraustino, J. D., Liu, B., & Jin, Y. (2012). *Social media use during disasters: A review of the knowledge base and gaps*. National Consortium for the Study of Terrorism and Responses to Terrorism, University of Maryland.

Fujisaka, T., Lee, R., & Sumiya, K. (2010). Exploring urban characteristics using movement history of mass mobile microbloggers. In *Proceedings of the Eleventh Workshop on Mobile Computing Systems & Applications* (pp. 13–18). Annapolis, USA: ACM.

Gayo-Avello, D. (2013). A meta-analysis of state-of-the-art electoral prediction from Twitter data. *Social Science Computer Review*, *31*(6), 649–679.

Gensim. (2017). *Gensim Python package*. Retrieved from https://radimrehurek.com/gensim

Getz, D., & Page, S. J. (2016). *Event studies: Theory, research and policy for planned events* (3rd ed). New York, USA: Routledge.

Girardin, F., Vaccari, A., Gerber, A., Biderman, A., & Ratti, C. (2009). Quantifying urban attractiveness from the distribution and density of digital footprints. *International Journal of Spatial Data Infrastructures Research*, *4*, 175–200.

Greater London Authority's DataStore. (2017). Statistical GIS boundary files for London. *London Datastore*. Retrieved from https://data.london.gov.uk/dataset/statistical-gis-boundary-files-london

Gupta, A., & Kumaraguru, P. (2012). Credibility ranking of tweets during high impact events. in *Proceedings of the 1st Workshop on Privacy and Security in Online Social Media*. (Article No. 2) Lyon, France: ACM.

Habidatum. (2017). Sport events & local economy.

*Habidatum*. Retrieved from https://habidatum.com/projects/sport-events-and-local-economy

Hall, P. (1966). *The world cities* (2nd ed.). London, UK: Weidenfeld and Nicolson.

Harvey, D. (1987). Flexible accumulation through urbanization: Reflections on "postmodernism" in the American city. *Antipode*, *19*(3), 260–286.

Hasan, S., & Ukkusuri, S. V. (2014). Urban activity pattern classification using topic models from online geolocation data. *Transportation Research Part C: Emerging Technologies*, *44*, 363–381.

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

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 168–177). Seattle, USA: ACM.

Kaufmann, B. (2015). If you want Olympics, do not ask the people. *Swissinfo.ch*. Retrieved from https://www.swissinfo.ch/eng/directdemocracy/democracy-and-olympic-spirit_if-you-want-the-olympics–do-not-ask-the-people/41595460

Keßler, C., Maué, P., Heuer, J. T., & Bartoschek, T. (2009). Bottom-up gazetteers: Learning from the implicit semantics of geotags. In K. Janowicz, M. Raubal, & S. Levashkin (Eds.), *GeoSpatial Semantics. GeoS 2009. Lecture Notes in Computer Science* (Vol. 5892). Berlin and Heidelberg: Springer.

Kim, J. W., Kim, D., Keegan, B., Kim, J. H., Kim, S., & Oh, A. (2015). Social media dynamics of global co-presence during the 2014 FIFA World Cup. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 2623–2632). Seoul: ACM.

Kinsella, S., Murdock, V., & Hare, N. O. (2011). "I'm Eating a Sandwich in Glasgow": Modeling locations with Tweets. In *Proceedings of the 3rd International Workshop on Search and Mining User-Generated Contents* (pp. 61–68). Glasgow, UK: ACM.

Lee, J. H., Gao, S., & Goulias, K. G. (2016). Can Twitter data be used to validate travel demand models? In *95th Annual Transportation Research Board Meeting* (pp. 1–27). Washington, DC, USA: TRB.

Lee, R., & Sumiya, K. (2010). Measuring geographical regularities of crowd behaviors for Twitter-based geo-social event detection. In *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location-Based Social Networks* (pp. 1–10). San Jose, USA: ACM.

Li, R., Lei, K. H., Khadiwala, R., & Chang, K. C.-C. (2012). TEDAS: A Twitter-based event detection and analysis system. In *Proceedings of the 28th International Conference on Data Engineering* (pp. 1273–1276). Washington, DC, USA: IEEE.

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.

Li, W., Serdyukov, P., de Vries, A. P., Eickhoff, C., & Larson, M. (2011). The where in the tweet. *Proceedings of the 20th ACM International Conference on Information and Knowledge Management* (pp. 2473–2476). Glasgow, UK: ACM.

Longley, P. A., Adnan, M., & Lansley, G. (2015). The geotemporal demographics of Twitter usage. *Environment and Planning A*, *47*(2), 465–484.

Malfas, M., Houlihan, B., & Theodoraki, E. (2004). Impacts of the Olympic Games as mega-events. *Proceedings of the Institution of Civil Engineers—Municipal Engineer*, *157*(3), 209–220.

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J., & McClosky, D. (2014). The Stanford CoreNLP Natural Language Processing Toolkit. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System demonstrations* (pp. 55–60). Baltimore, USA: Association for Computational Linguistics.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, *50*(4), 370–396.

McGill. (2017). About urban planning. *McGill, School of Urban Planning*. Retrieved from https://mcgill.ca/urbanplanning/planning

Mellon, J., & Prosser, C. (2017). Twitter and Facebook are not representative of the general population: Political attitudes and demographics of British social media users. *Research & Politics*, *4*(3), 1–9.

Mitchell, L., Frank, M. R., Harris, K. D., Dodds, P. S., & Danforth, C. M. (2013). The geography of happiness: Connecting Twitter sentiment and expression, demographics, and objective characteristics of place. *PLoS ONE*, *8*(5).

Moore, J. (2015). When Denver rejected the Olympics in favour of the environment and economics. *The Guardian*. Retrieved from https://www.theguardian.com/sport/blog/2015/apr/07/when-denver-rejected-the-olympics-in-favour-of-the-environment-and-economics

Morstatter, F., Pfeffer, J., Liu, H., & Carley, K. M. (2013). Is the sample good enough? Comparing data from Twitter's streaming API with Twitter's firehose. In *Proceedings of the 7th International Conference on Weblogs and Social Media* (pp. 400–408). Cambridge, USA: AAAI Press.

Noulas, A., Scellato, S., Mascolo, C., & Pontil, M. (2011). Exploiting semantic annotations for clustering geographic areas and users in location-based social networks. In *Proceedings of the Social Mobile Web* (pp. 32–35). Barcelona, Spain: AAAI Press.

Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. In *Proceedings of the International Conference on Language Resources and Evaluation* (pp. 1320–1326). Valletta, Malta: European Language Resources Association.

Panteras, G., Wise, S., Lu, X., Croitoru, A., Crooks, A., &

Stefanidis, A. (2015). Triangulating social multimedia content for event localization using Flickr and Twitter. *Transactions in GIS*, *19*(5), 694–715.

Pennacchiotti, M., & Popescu, A. (2011). A machine learning approach to Twitter user classification. In *Proceedings of the Fifth International Conference on Weblogs and Social Media* (pp. 281–288). Barcelona, Spain: AAAI Press.

Porter, M. F. (1980). An algorithm for suffix stripping. *Program*, *14*(3), 130–137.

Preuss, H. (2007). The conceptualisation and measurement of mega sport event legacies. *Journal of Sport & Tourism*, *12*(3/4), 207–228.

Quercia, D., Ellis, J., Capra, L., & Crowcroft, J. (2012). Tracking "gross community happiness" from tweets. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (pp. 965–968). Seattle, USA: ACM.

Resch, B. (2013). People as sensors and collective sensing: Contextual observations complementing geo-sensor network measurements. In J. Krisp (Ed.), *Progress in location-based services* (pp. 391–406). Berlin and Heidelberg: Springer.

Resch, B., Summa, A., Zeile, P., & Strube, M. (2016). Citizen-centric urban planning through extracting emotion information from Twitter in an interdisciplinary space-time-linguistics algorithm. *Urban Planning*, *1*(2), 114–127.

Resch, B., Usländer, F., & Havas, C. (2017). Combining machine-learning topic models and spatiotemporal analysis of social media data for disaster footprint and damage assessment. *Cartography and Geographic Information Science*, 1–15.

Ritchie, B. W., Shipway, R., & Cleeve, B. (2009). Resident perceptions of mega-sporting events: A non-host city perspective of the 2012 London Olympic Games. *Journal of Sport & Tourism*, *14*(2/3), 143–167.

Ritterman, J., Osborne, M., & Klein, E. (2009). Using prediction markets and Twitter to predict a swine flu pandemic. In *Proceedings of the 1st International Workshop on Mining Social Media* (pp. 9–17).

Sagl, G., Resch, B., Hawelka, B., & Beinat, E. (2012). From social sensor data to collective human behaviour patterns–Analysing and visualising spatio-temporal dynamics in urban environments. In T. Jekel, A. Car, J. Strobl, & G. Griesebner (Eds.), *GI_Forum 2012: Geovizualisation, society and learning* (pp. 54–63). Berlin and Offenbach, Germany: Herbert Wichmann Verlag.

Sims, S. (2017). How Bostonians defeated the Olympics. *Citylab*. Retrieved from https://www.citylab.com/equity/2017/06/how-bostonians-defeated-the-olympics/529155

Stefanidis, A., Crooks, A., & Radzikowski, J. (2013). Harvesting ambient geospatial information from social media feeds. *GeoJournal*, *78*(2), 319–338.

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.

Sui, D., & Goodchild, M. (2011). The convergence of GIS and social media: Challenges for GIScience. *International Journal of Geographical Information Science*, *25*(11), 1737–1748.

Sun, Y., Fan, H., Bakillah, M., & Zipf, A. (2015). Road-based travel recommendation using geo-tagged images. *Computers, Environment and Urban Systems*, *53*, 110–122.

Theodore, D. (2006). Sense of the city: An alternative approach to urbanism. *Journal of Architectural Education*, *60*(2), 69–70.

Twitter INC. (2017). Retrieved from https://developer.twitter.com/en/docs

Wang, H., Can, D., Kazemzadeh, A., Bar, F., & Narayanan, S. (2012). A system for real-time Twitter sentiment analysis of 2012 US presidential election cycle. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics* (pp. 115–120). Jeju, Republic of Korea: Association for Computational Linguistics.

Weng, J., & Lee, B.-S. (2011). Event detection in Twitter. In *Proceedings of the Fifth International Conference on Weblogs and Social Media* (pp. 401–408). Barcelona, Spain: AAAI Press.

Whitson, D., & Macintosh, D. (1993). Becoming a world-class city: Hallmark events and sport franchises in the growth strategies of western Canadian cities. *Sociology of Sport Journal*, *10*(3), 221–240.

Yu, Y., & Wang, X. (2015). World Cup 2014 in the Twitter world: A big data analysis of sentiments in US sports fans' tweets. *Computers in Human Behavior*, *48*, 392–400.

Zeile, P., Resch, B., Exner, J.-P., & Sagl, G. (2015). Urban emotions: Benefits and risks in using human sensory assessment for the extraction of contextual emotion information in urban planning. In S. Geertman, J. Ferreira, R. Goodspeed, & J. Stillwell (Eds.), *Planning support systems and smart cities* (pp. 209–225). Switzerland: Springer International Publishing.

Zhang, X., Fuehres, H., & Gloor, P. A. (2011). Predicting stock market indicators through Twitter "I hope it is not as bad as I fear". *Procedia: Social and Behavioral Sciences*, *26*, 55–62.

Zhang, Z., Ni, M., He, Q., & Gao, J. (2016). *Mining transportation information from social media for planned and unplanned events*. Buffalo, NY, USA: University at Buffalo.

Zhao, W. X., Jiang, J., Weng, J., He, J., Lim, E.-P., Yan, H., & Li, X. (2011). Comparing Twitter and traditional media using topic models. In *Proceedings of 33rd European Conference on IR Research* (pp. 338–349). Dublin, Ireland: Springer Verlag.

Zhao, S., Zhong, L., Wickramasuriya, J., & Vasudevan, V. (2011). *Human as real-time sensors of social and physical events: A case study of Twitter and sports games* (Technical Report TR0620-2011). Rice University and Motorola Labs.

**About the Authors**

**Anna Kovács-Győri** has been a PhD candidate at the University of Salzburg, Department of Geoinformatics–Z_GIS since 2015. She studied Geography with a specialization in GIS at the Department of Physical Geography and Geoinformatics, University of Szeged, Hungary. During her PhD studies, she investigates the quality of the person-environment relationship in the urban context. Her research focuses on livability assessment and the role of spatiality, mobility, human scale, and health in it to promote urban planning and decision-making.
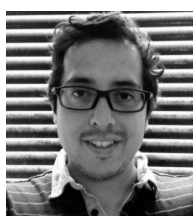
**Alina Ristea** is a fully funded PhD student at the Doctoral College GIScience, and her research topic is "Integration and Evaluation of Social Media Data into Crime Prediction Models", under the supervision of Professor Michael Leitner. Her background studies are in Geography–Cartography-Cadastral Elements (Bachelor) and GIS (Master) from the University of Bucharest, Romania.

**Clemens Havas** is a PhD student with a background in computer science at University of Salzburg's Department of Geoinformatics–Z_GIS since 2017. His research interests lie in the fields of machine learning, social media and geospatial analysis. Besides doing his PhD programme, Clemens Havas has been working for various projects funded by international organizations/programmes (e.g., Horizon 2020) by using new data sources to assess the impact of disasters.

**Bernd Resch** is an Assistant Professor at University of Salzburg's Department of Geoinformatics—Z_GIS and a Visiting Fellow at Harvard University (USA). His research interests revolve around fusing data from human and technical sensors, including the analysis of social media. Amongst a variety of other functions, Bernd Resch is Editorial Board Member of the *International Journal of Health Geographics*, Associated Faculty Member of the doctoral college "GIScience", and Executive Board member of Spatial Services GmbH.

**Pablo Cabrera-Barona** is Assistant Professor at the Institute of Higher National Studies (IAEN), and visitor Professor at the Latin American Social Sciences Institute (FLACSO) and San Francisco de Quito University (USFQ). His main interests lie in understanding different interactions between geography, economics, ecology, health and development processes. He is used to working at different scales, considering different social contexts.
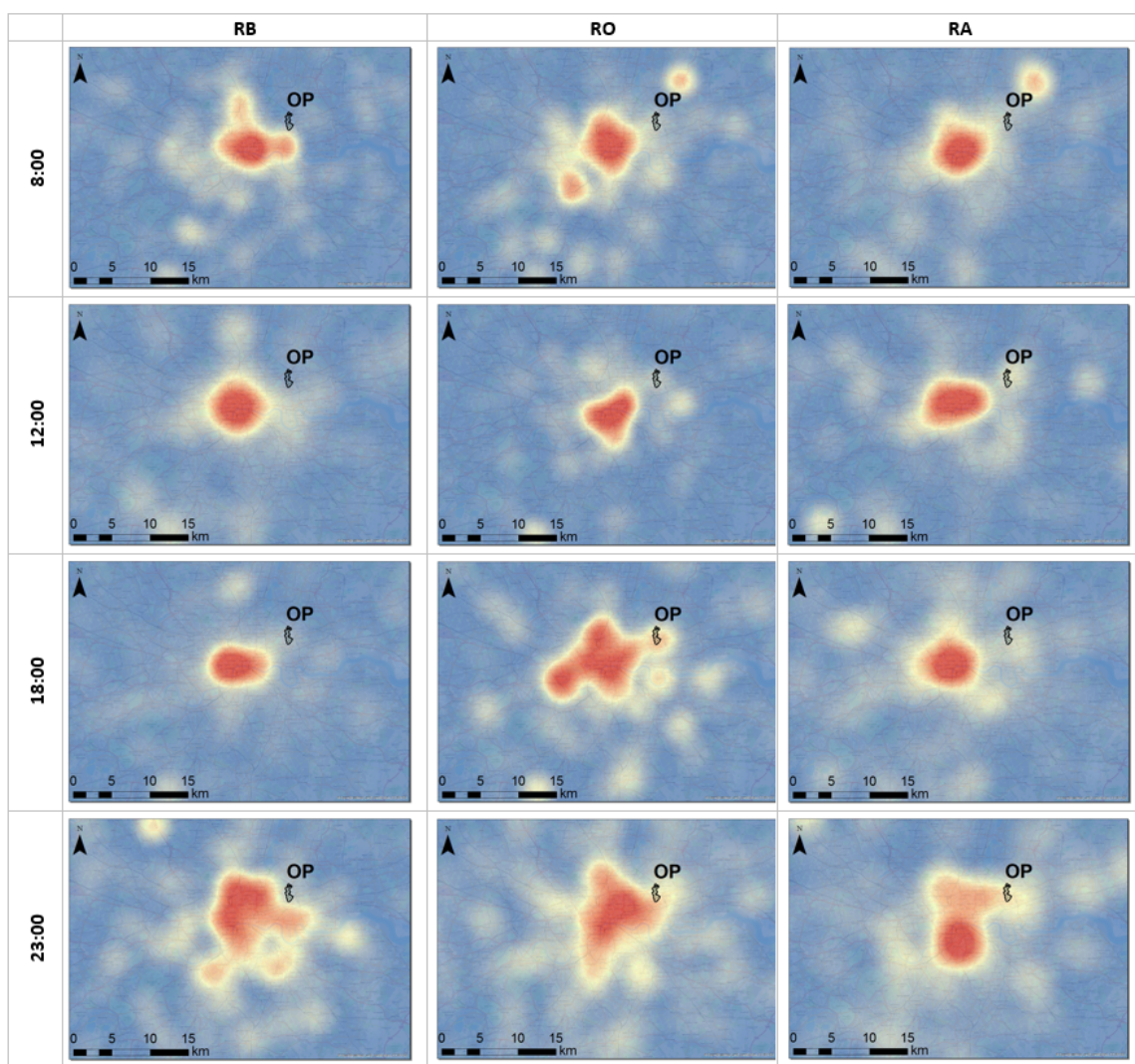
**Supplementary Files**

**Hourly spatiotemporal patterns (KDE)—for residents and visitors, positive and negative sentiments**

The first row in each figure shows morning patterns (8:00 a.m.), the second one represents one of the most active hours in the data (12:00 noon) while the third and fourth represent the early (6:00 p.m.) and late evening (11:00 p.m.). Red spots are the densest areas in the given hour and category, while yellow shows slightly lower density. Blue represents those areas where based on the KDE algorithm, the density in a cell is lower than the expected. The cell size was in each map 200 m$^2$. The tables provide general information on the cell values' statistics, and it can aid the interpretation of how high the differences were between blue and red areas. The higher the standard deviation is, the higher the difference in the density, while using the maximum and mean values, the different maps become comparable.

**Table 1.** Values for the hot spot analysis using KDE for residents' negative tweets for specific hours.

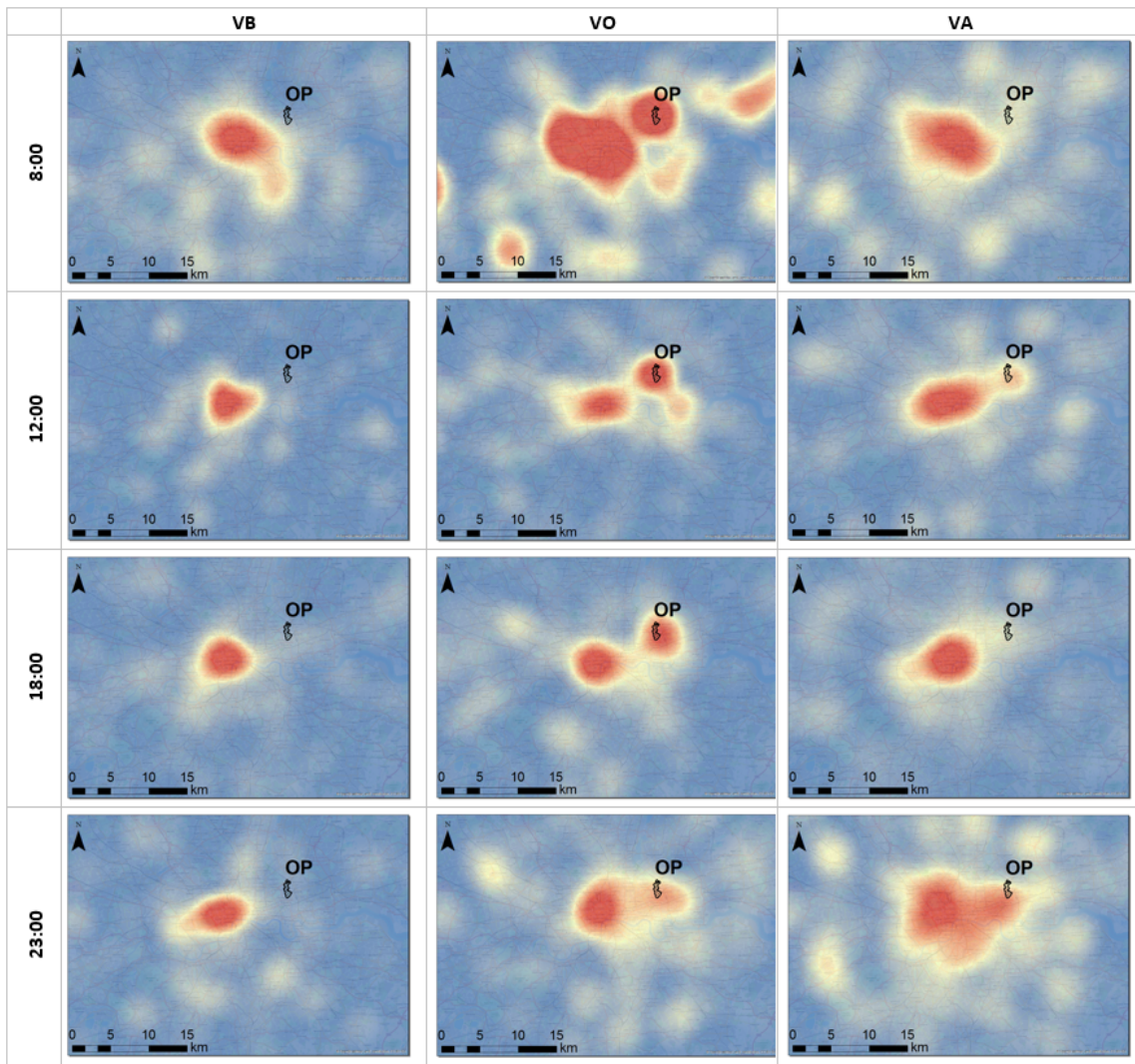|  | before | OG | after |  | before | OG | after |
|---|---|---|---|---|---|---|---|
| **8:00 a.m.** | n = 366 mean = 0.11 stdev = 0.25 | n = 287 mean = 0.10 stdev = 0.19 | n = 372 mean = 0.10 stdev = 0.21 | **6:00 p.m.** | n = 365 mean = 0.11 stdev = 0.23 | n = 323 mean = 0.10 stdev = 0.19 | n = 351 mean = 0.10 stdev = 0.16 |
| **12:00 noon** | n = 351 mean = 0.11 stdev = 0.18 | n = 400 mean = 0.12 stdev = 0.28 | n = 304 mean = 0.08 stdev = 0.17 | **11:00 p.m.** | n = 418 mean = 0.12 stdev = 0.21 | n = 364 mean = 0.11 stdev = 0.16 | n = 288 mean = 0.08 stdev = 0.13 |



**Figure 1.** Hot spots of residents' negative tweets for specific hours.

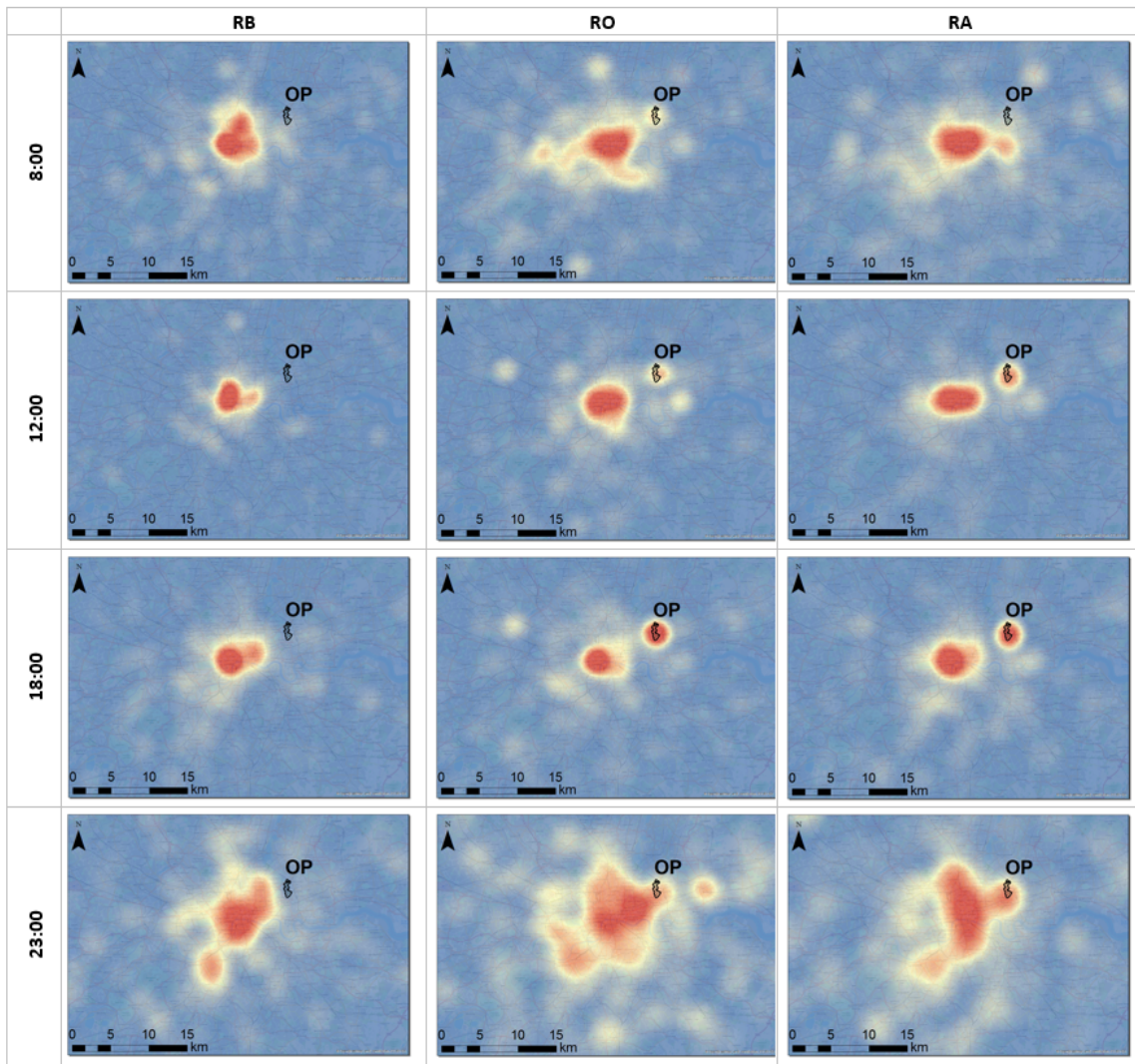**Table 2.** Values for the hot spot analysis using KDE for visitors' negative tweets for specific hours.

| | before | OG | after | | before | OG | after |
|---|---|---|---|---|---|---|---|
| **8:00 a.m.** | n = 217 mean = 0.06 stdev = 0.1 | n = 171 mean = 0.05 stdev = 0.09 | n = 317 mean = 0.09 stdev = 0.11 | **6:00 p.m.** | n = 228 mean = 0.06 stdev = 0.13 | n = 202 mean = 0.06 stdev = 0.11 | n = 330 mean = 0.09 stdev = 0.15 |
| **12:00 noon** | n = 236 mean = 0.07 stdev = 0.18 | n = 260 mean = 0.08 stdev = 0.15 | n = 369 mean = 0.11 stdev = 0.17 | **11:00 p.m.** | n = 248 mean = 0.08 stdev = 0.21 | n = 238 mean = 0.07 stdev = 0.11 | n = 445 mean = 0.12 stdev = 0.16 |



**Figure 2.** Hot spots of visitors' negative tweets for specific hours.

**Table 3.** Values for the hot spot analysis using KDE for residents' positive tweets for specific hours.

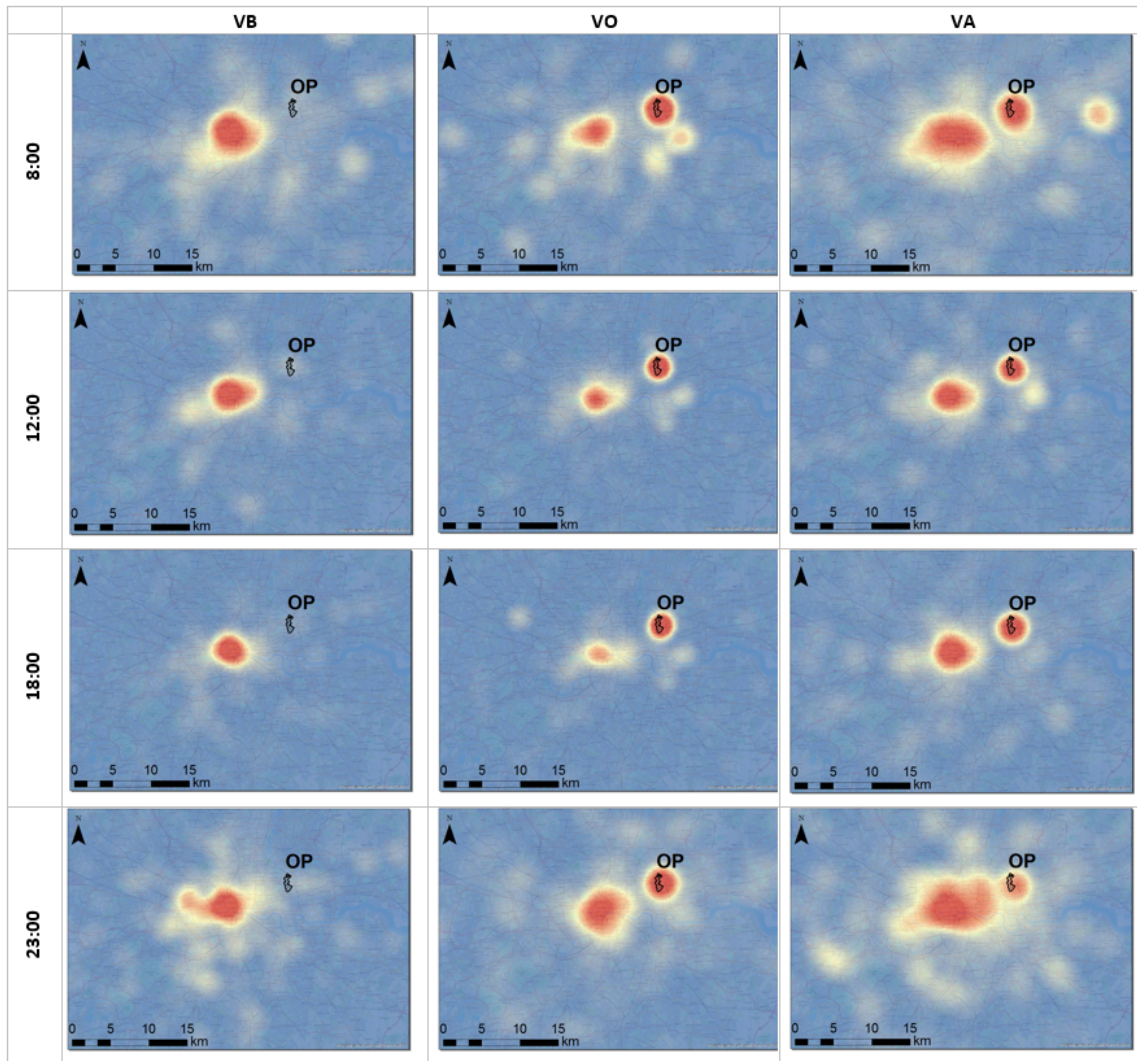|  | before | OG | after |  | before | OG | after |
|---|---|---|---|---|---|---|---|
| **8:00 a.m.** | n = 839 mean = 0.25 stdev = 0.7 | n = 819 mean = 0.24 stdev = 0.53 | n = 809 mean = 0.23 stdev = 0.47 | **6:00 p.m.** | n = 773 mean = 0.23 stdev = 0.62 | n = 868 mean = 0.25 stdev = 0.62 | n = 661 mean = 0.19 stdev = 0.47 |
| **12:00 noon** | n = 819 mean = 0.24 stdev = 0.83 | n = 939 mean = 0.27 stdev = 0.68 | n = 727 mean = 0.23 stdev = 0.54 | **11:00 p.m.** | n = 702 mean = 0.20 stdev = 0.41 | n = 828 mean = 0.24 stdev = 0.38 | n = 565 mean = 0.16 stdev = 0.27 |



**Figure 3.** Hot spots of residents' positive tweets for specific hours.

**Table 4.** Values for the hot spot analysis using KDE for visitors' positive tweets for specific hours.

| | before | OG | after | | before | OG | after |
|---|---|---|---|---|---|---|---|
| **8:00 a.m.** | n = 452<br>mean = 0.13<br>stdev = 0.28 | n = 498<br>mean = 0.15<br>stdev = 0.32 | n = 753<br>mean = 0.21<br>stdev = 0.36 | **6:00 p.m.** | n = 483<br>mean = 0.14<br>stdev = 0.46 | n = 562<br>mean = 0.16<br>stdev = 0.56 | n = 684<br>mean = 0.20<br>stdev = 0.48 |
| **12:00 noon** | n = 565<br>mean = 0.16<br>stdev = 0.48 | n = 821<br>mean = 0.24<br>stdev = 0.78 | n = 737<br>mean = 0.21<br>stdev = 0.55 | **11:00 p.m.** | n = 523<br>mean = 0.15<br>stdev = 0.38 | n = 577<br>mean = 0.16<br>stdev = 0.36 | n = 706<br>mean = 0.20<br>stdev = 0.34 |



**Figure 4.** Hot spots of visitors' positive tweets for specific hours.