Applied Geography 123 (2020) 102306

Contents lists available at ScienceDirect

Applied Geography

journal homepage: http://www.elsevier.com/locate/apgeog

Mining public sentiments and perspectives from geotagged social media data for appraising the post-earthquake recovery of tourism destinations

Yingwei Yan [a,b,c], Jingfu Chen [d,*], Zhiyong Wang [e]

[a] Key Lab of Guangdong for Utilization of Remote Sensing and Geographical Information System, Guangdong Open Laboratory of Geospatial Information Technology and Application, Guangzhou Institute of Geography, Guangzhou, China
[b] Department of Geography, National University of Singapore, Singapore
[c] Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou, China
[d] School of Tourism Management, Sun Yat-Sen University, China
[e] Department of Human Geography and Spatial Planning, Utrecht University, The Netherlands

ARTICLE INFO

Keywords:
Post-disaster recovery
Sentiments and perspectives
Geotagged social media data
Volunteered geographic information
Sentiment analysis
Topic modeling

ABSTRACT

Post-disaster recovery involves interdependent processes of physical and psychological rehabilitations. Over the past few years, researchers have explored geotagged social media data to assist the planning, monitoring, and assessment of the post-disaster recovery of tourism destinations, given its advantages over traditional approaches. Nonetheless, recent studies have mostly focused on quantitatively accessing the physical elements of post-disaster recovery (e.g., infrastructure reconstruction and re-influx of tourists). Few studies have explored people's sentiments and perspectives over the process of post-disaster recovery. In this study, a mixed methods approach involving sentiment analysis and Latent Dirichlet allocation (LDA) topic modeling is designed for mining sheer volume of tweets about Lombok and Bali, generated by nonlocal Twitter users after a series of earthquakes in the two places in August 2018. The findings mainly suggest that people have generally become less negative about Lombok and Bali over time, despite fluctuations in their sentiment polarities' central tendencies. In addition, dissatisfactions about the housing reconstruction progress, tourism recovery status, and living conditions in the affected areas of Lombok still existed in 2019; contestations have been found with regard to the huge funds for hosting the 2018 Bali IMF-World Bank meeting after the earthquakes. The overall results of this study have proved that the adopted approach can effectively reveal the variations of people's sentiments and perspectives of general and specific issues regarding post-disaster tourism recovery over time.

1. Introduction

The reputation and images of tourism destinations are highly vulnerable to catastrophic events such as earthquake and nuclear accidents (Yan et al. 2018). Natural and man-made disasters often damage and even destroy tourist attractions, infrastructures, and visitors' good perception of destinations, arousing tourists' worries about instability and safety and weakening their desires for holiday in affected sites. Tourists' behavioural changes, including trip cancelation and last-minute booking, as a result of tarnished images of tourism destinations, often lead to loss of tourism revenue and damage local tourism industry at different degrees (Mair et al. 2016; Yan et al. 2018). It has been argued that tourism destinations which manage to minimize negative impacts of reputational damage are likely to show more resilience than destinations failing to do so (Mair et al. 2016).

The repairment of place images and reputation is closely linked to both the reconstruction of physical landscapes and the development of positive attitudes of the public, especially potential visitors, towards the post-disaster recovery of tourism destinations. It thus involves interdependent processes of physical and psychological rehabilitation. The physical (tangible) elements such as infrastructure reconstruction, re-influx of tourists, and economic rebound ((Yan et al., 2017a), (Yan et al., 2018)) can be readily assessed using secondary data, while the intertwined psychological elements (Schumann, 2018) are difficult to quantify, standardize, and measure. Qualitative approaches such as interview, focus group, photovoice, and participatory mapping have been mostly adopted to assess the non-physical elements of post-disaster recovery (Schumann, 2018). However, these approaches are to some

* Corresponding author.
E-mail addresses: yanyingwei@u.nus.edu (Y. Yan), chenjf75@mail.sysu.edu.cn (J. Chen), z.wang2@uu.nl (Z. Wang).

https://doi.org/10.1016/j.apgeog.2020.102306
Received 21 February 2020; Received in revised form 29 July 2020; Accepted 15 August 2020
Available online 1 September 2020

extent costly because they rely on researchers' personal and embodied linkages with the field and are confined to a relatively small proportion of stakeholders. Social media, as a source of big data (Schroeder, 2014), has the potential to compensate for the abovementioned shortcomings of traditional qualitative methods, as it is ubiquitous, rich in content, and big in volume, velocity, and variety (Gandomi & Haider, 2015). Despite the value of social media data, it has not been well included in the analysis of the recovery of tourism destinations after catastrophes.

Over the past few years, a few researchers have explored geotagged social media data to assist the planning, monitoring, and assessment of the post-disaster recovery of tourism destinations ((Yan et al., 2017a), (Yan et al., 2018)), given its advantages (i.e., richness in data coverage and volume, cost-effectiveness, and timeliness) over conventional approaches. Geotagged social media data is a type of volunteered geographic information (VGI) (Goodchild, 2007; Yan et al., 2017a), which is regarded by users as more interactive and credible in the recovery-marketing stage of tourism destinations (Schultz et al. 2011). It can provide fine granularity on-the-ground information (both quantitative and qualitative) for assessing the extent of damage (in the planning stage) and documenting the implementation status of reconstruction and rehabilitation (in the execution phase), complementing information sensed via traditional approaches such as remote sensing (Longueville, Luraschi, Smits, Peedell, & Groeve, 2010; Yan et al., 2017a). Despite the value of social media data, recent studies have mainly focused on quantitatively assessing the physical elements of post-disaster recovery. Scant studies have explored people's sentiments and perspectives over the course of the post-disaster recovery of tourism destinations.

Therefore, this study seeks to leverage on geotagged social media data to appraise the post-disaster recovery of tourism destinations via mining people's sentiments and perspectives regarding recovery status. Specifically, a case study is conducted in Lombok Island and Bali Island in Indonesia, which were affected by multiple earthquakes in August 2018. Geotagged Twitter data which is rich in text content is adopted to achieve the research objectives (Wang & Ye, 2018). To understand the impact of the earthquakes on the two international tourism destinations, it gathers data generated by international Twitter users who are non-locals of the two islands. A mixed methods approach involving sentiment analysis (Gauba et al. 2017) and Latent Dirichlet allocation (LDA) topic modeling (Blei et al. 2003) are adopted for text data mining to reveal (1) variations of people's general sentiments over time and (2) variations of people's sentiments and perspectives over time regarding specific aspects of post-disaster recovery.

2. Post-disaster recovery and social media

2.1. Approaches for post-disaster recovery

Post-disaster recovery refers to the rehabilitation of affected regions to a certain level of acceptability, and is mostly prompted by a series of damage rectification strategies and actions developed and implemented by individuals and (non)governmental agents (Yan et al., 2017a). Compared to other phases of disaster management (i.e., preparation, response, and mitigation), recovery is considered the least understood aspect of disaster management (Smith & Wenger, 2007, pp. 234–257).

Many approaches have been adopted to monitor and assess post-disaster recovery, such as remote sensing (Hoshi, Murao, Yoshino, Yamazaki, & Estrada, 2014), the combination of field surveys and remote sensing (Brown et al. 2012), resident interviews (Rathfon, Davidson, Bevington, Vicini, & Hill, 2013), social audits, household surveys, official building permit data, census, and damage assessment data analysis, outsourced data and insurance data analysis (Platt et al. 2016; Yan et al., 2017a). However, these approaches have different shortcomings (e.g., costly, labor-intensive, and time-consuming), as have been discussed by (Yan et al., 2017a).

Additionally, community-based disaster risk management (CBDRM) has become prevalent in major disaster management schemas across the world since the 1980s (Maskrey, 2011). CBDRM mobilizes community-level resources, capabilities, and knowledge. By enabling communication channels between the affected and policy-makers, CBDRM leads to a more appropriate and sustainable mitigation of risks that reduces the problem of disaster management measures (albeit technically sound) mismatching stakeholders' cultural and social experiences and expectations (Maskrey, 2011). CBDRM is particularly important in the recovery phase of disaster management, given communities' role in decision making and planning. However, from the angle of post-disaster image and reputation recovery of tourism destinations, perspectives outside the affected local community tend to be more important.

2.2. Social media for disaster management

In the Web 2.0 era, social media can be a highly effective channel of disaster communication beyond local communities (Yates & Paquette, 2010, pp. 1–9). It facilitates international participation and connectedness (Haworth, 2018), driving disaster communication towards a paradigm of greater collaboration and transparency that facilitates a ubiquitous, timely, diverse, and constant flow of geospatial information ((Yan et al., 2017a), (Yan, Feng, & Chang, 2017b)). Although traditional media can also be a channel for disaster communication beyond local communities, The media has been criticized for its sensationalism (Mair et al. 2016). Public media reports after a disaster strike may be biased or misleading in order for attracting the public where sensationalism can emerge (Frisby, 2002). On the contrary, social media offers first-hand information about user experience, observations, sentiments, and perspectives. Although social media data may also be biased or misleading, it has been argued that "given enough eyes all bugs are shallow" according the Linus' Law (Goodchild & Li, 2012; Yan, Feng, & Wang, 2017c).

As such, social media data has been widely explored for disaster management. For example, based on various social media data (e.g., Twitter, Instagram, and Flickr), existing studies have investigated the characteristics of typhoon response communication (Takahashi et al. 2015), the hurricane evacuation responses of residents (Martín et al. 2017), the response and preventive monitoring of floods (de Albuquerque, Herfort, Brenning, & Zipf, 2015; Du, Cai, Sun, & Minsker, 2017; Fohringer, Dransch, Kreibich, & Schröter, 2015; Huang, Li, Wang, & Ning, 2019; Huang, Wang, & Li, 2018), and the responsive monitoring and detection of wildfires (Boulton, Shotton, & Williams, 2016, pp. 178–186; Goodchild & Glennon, 2010).

The use of social media for post-disaster recovery strategies, however, has not yet received any substantial empirical studies. Among the few works that have explored social media data for monitoring and assessing post-disaster recovery, a workflow and methods based on Flickr photos have recently been proposed by (Yan et al., 2017a) for assessing the quantitative (i.e., tourist quantity and distribution) and physical features (i.e., infrastructures) of the recovery of tourism destinations; and a method based on maximum entropy modeling and Flickr photos has been proposed by (Yan et al., 2018) to model the spatial distribution of tourists in a probabilistic way. These two works, however, neglect people's sentiments and perspectives over the process of the post-disaster recovery of tourism destinations.

3. Study area

The case study of this research is conducted in two popular international tourism destinations in Indonesia – Lombok Island and Bali Island (Fig. 1). The two islands were hit by a magnitude 6.9 earthquake on August 5, 2018 (Beech & LeBien, 2018). The two places were hit by two more earthquakes (magnitudes 6.3 and 7.0) on August 29, 2018 (BBC, 2018). As a result, a number of embassies issued "do not travel" warnings for the affected regions (Walden, 2019). According to the

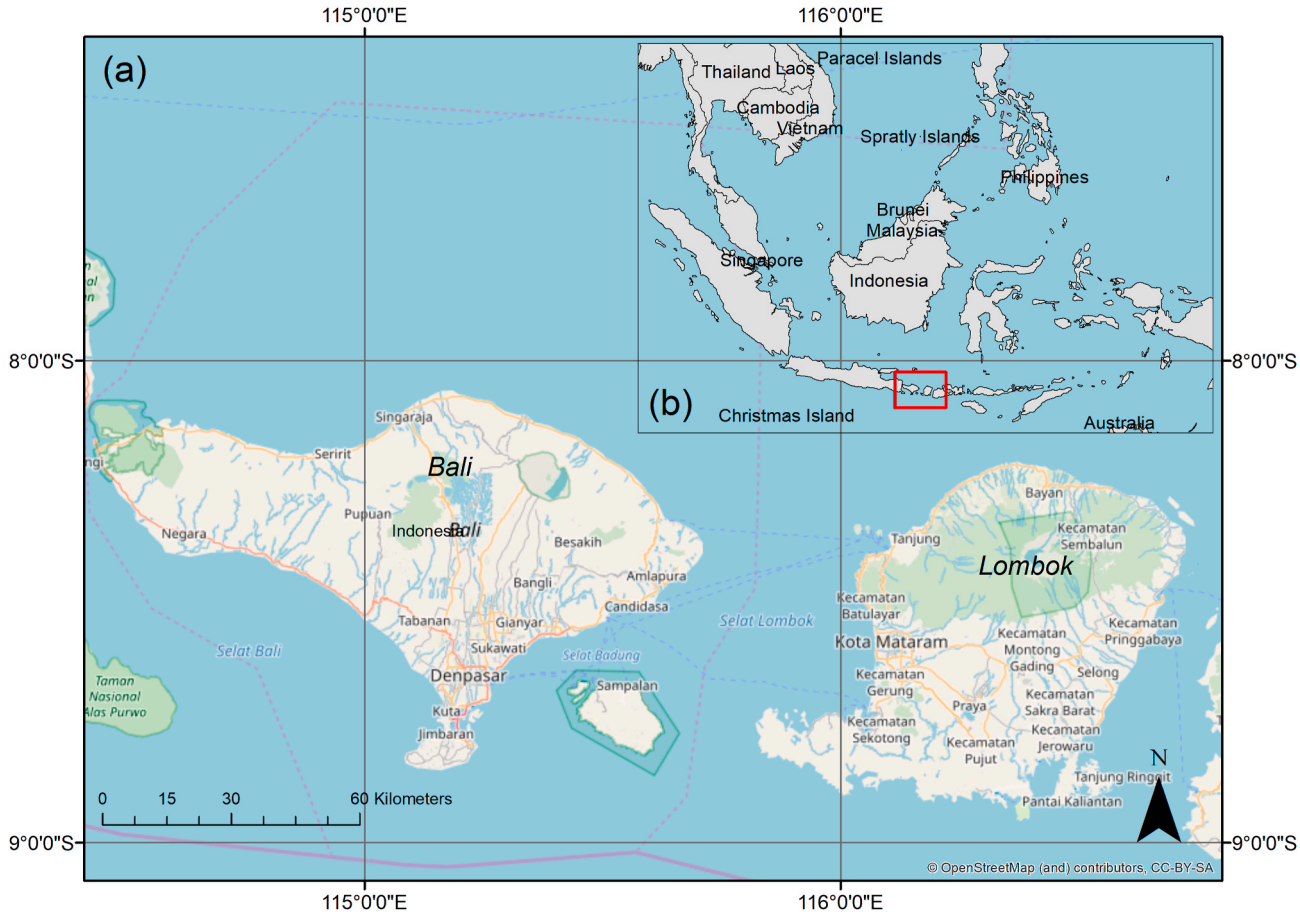

Fig. 1. Study area shown in OpenStreetMap. (a) Lombok Island and Bali Island. (b) Inset indicating the relative location of Lombok Island and Bali Island in Southeast Asia using a red rectangle. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

official statistics from the local government (Indonesia, 2020), the growth of the number of international visitor arrivals in Indonesia has slowed down strongly in 2019 (Fig. 2), suggesting that local tourism industry has been affected.

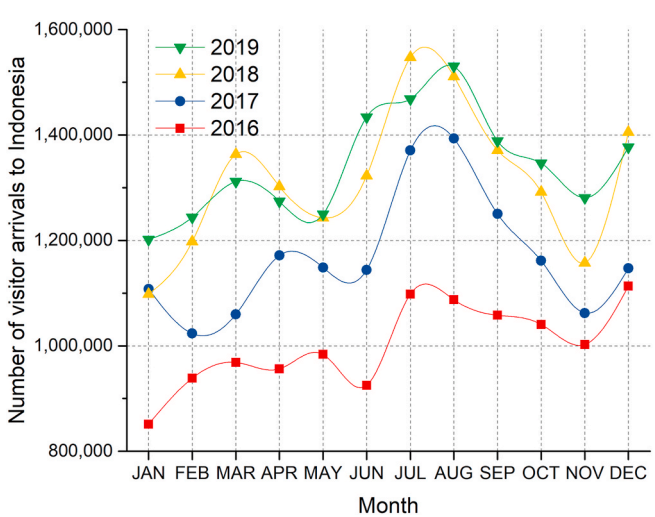

Fig. 2. Statistics of international visitor arrivals in Indonesia from 2016 to 2019.

4. Materials and methods

The workflow of the case study starts with Twitter data collection, data cleansing and pre-processing. Subsequently, sentiment analysis and topic-based sentiment analysis via LDA are deployed for data analysis. The following sections detail the methods.

4.1. Data collection, cleansing and, pre-processing

Tweets are collected via TAGS (https://tags.hawksey.info/), a free Google Sheet template allowing us to set up and run automated collection of search results from Twitter without digging into the details of the Twitter's search Application Programming Interface (API). Specifically, tweets' texts and hashtags containing the keywords of "Lombok" and "Bali" posted worldwide between August 6, 2018 and July 31, 2019 are collected (a one-year Tweet dataset). The number of Tweets collected (i.e., maximum varies) per hour is set as 3000 (the default value, this optional input must be a number between 1 and 18,000). Duplicate data are disallowed. Non-English Tweets are translated into English using the Google Translate formula in Google Sheets.

The data collection results in 2,007,802 Tweets mentioning Lombok and 7,643,592 Tweets mentioning Bali after 49,397 and 193,762 Tweets posted by news channels are removed from the initial Tweets database, i.e., Tweets from user names involving the keyword "news" are removed. Fig. 3 shows the time series of the Tweets collected in this study. The number of Tweets mentioning Lombok is big in August 2018 but levels off afterwards. By contrast, Tweets mentioning Bali fluctuate in quantity throughout the year, in which Tweets containing the

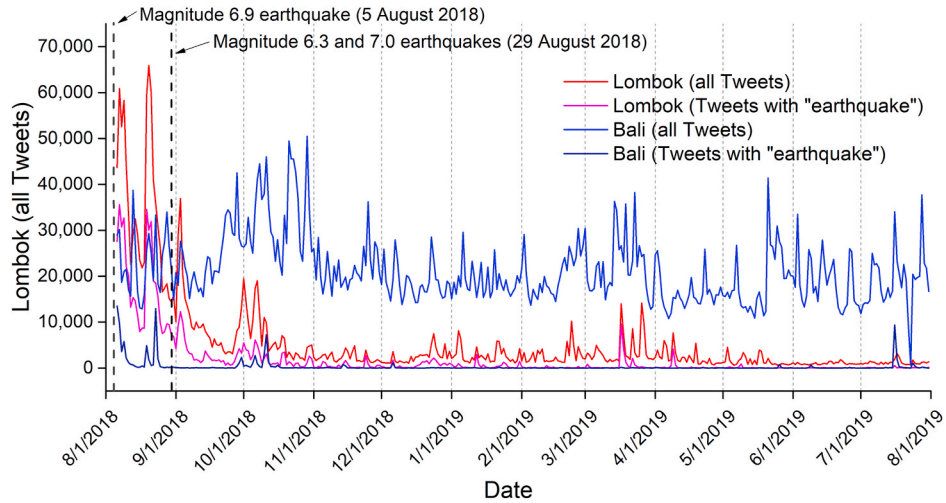

Fig. 3. Time series of Tweets mentioning (a) Lombok and (b) Bali.

keyword "earthquake" make some quantity peaks in August 2018.

Big data text mining is computationally intensive, therefore, for analyzing the spatiotemporal trends and patterns of public sentiments and perspectives over the nine months after the earthquake occurred on August 5, 2018, tweets in four selected months are extracted, including tweets timestamped in August 2018, October 2018, January 2019, and April 2019. For observing whether recovery measures were put in place rapidly and the related public attitudes, the two selected months in 2018 are closer to one another than the two selected months in 2019 are. The user locations of Tweets posted in the four selected months are geocoded based on users' profile locations (Kumar & Singh, 2019) using Geocoder which is a geocoding library written in Python (Carriere, 2019). The geocoded Tweets across the globe are visualized in Appendix 1. Lombok Island and Bali Island are renowned international tourism destinations, in order to reflect on their international images and reputation, Tweets generated by international Twitter users who are nonlocals of the two islands are subsequently extracted based on users' profile locations. Table 1 presents the statistics of the Tweets involved in the analysis to be detailed in the following sections.

4.2. Sentiment analysis

Sentiment analysis is conducted using Geocoded Tweets mentioning Lombok and Bali. The sentiment computation is performed using TextBlob (Loria, 2014) which is an open source library for performing natural language processing functions. TextBlob has been extensively used and proven effective in the sentiment analysis of Tweets (Budiharto & Meiliana, 2018, Al Walid et al., 2019; Yaqub et al. 2020). The default sentiment analysis implementation, PatternAnalyzer (based on the pattern library (Clips, 2018)), is used to generate a sentiment score (called polarity score) and a subjectivity score (for measuring the level subjectivity of the text content) for each Tweet. The range of the sentiment score varies from −1 to 1. Here, a value close to −1 stands for a highly negative text while a value close to 1 represents a highly positive text. The range of the subjectivity score varies from 0 to 1. A value close to 0 stands for an objective text while a value close to 1 indicates a highly subjective text. The subjectivity score enables us to identify high and low subjectivity Tweets based on a subjectivity threshold of 0.5 according to (Sahni, Chandak, Chedeti, & Singh, 2017).

Subsequently, Tweets are further classified into four clusters as follows based on whether the keyword "earthquake" is involved in individual Tweets or not.

- HE (high subjectivity and specific Tweets with the keyword "earthquake").
  - Example: "All our support and affection to the people of Indonesia after the terrible earthquake on the island of Lombok. Our heartfelt condolences to the families of the deceased and our desire for a speedy recovery of the wounded so they can return to their homes soon."
- LE (low subjectivity and specific Tweets with the keyword "earthquake").
  - Example: "Indonesian students organized a sausage sizzle fundraiser to raise funds for the victims of the Lombok earthquakes."
- HG (high subjectivity and general Tweets without the keyword "earthquake")
  - Example: "Beautiful Places Honeymoon in Lombok."
- LG (low subjectivity and general Tweets without the keyword "earthquake")
  - Example: "Travelling to Lombok and exploring Kuta Bali!"

A scatter plot of the four clusters of Tweets is created for each of the four observation months based on the great circle distance between the geolocations of individual Tweets and the geographic centers of tourism destinations. The sentiment polarities' central tendencies of the four clusters of Tweets are visualized for the four observation months using box plot.

Table 1
Statistics of the Tweets involved in the analysis. Note that "No." is the abbreviation of "number".

		No. of all Tweets (No. of all geocoded Tweets)	No. of all nonlocal Tweets (No. of users)	No. of nonlocal Tweets with "earthquake" (No. of users)
Lombok	August 2018	892,614 (521,874)	513,074 (154,091)	264,904 (103,491)
	October 2018	211,818 (120,253)	117,819 (44,894)	35,058 (17,373)
	January 2019	87,244 (51,349)	48,834 (18,228)	7,542 (4,653)
	April 2019	68,646 (39,492)	34,568 (15,937)	3,431 (3,088)
Bali	August 2018	593,767 (306,348)	264,593 (157,119)	34,050 (27,596)
	October 2018	1,040,585 (578191)	506,486 (293,042)	11,193 (8,922)
	January 2019	579,806 (324,693)	241,071 (146,240)	229 (125)
	April 2019	488,683 (264,108)	226,483 (138,528)	259 (199)

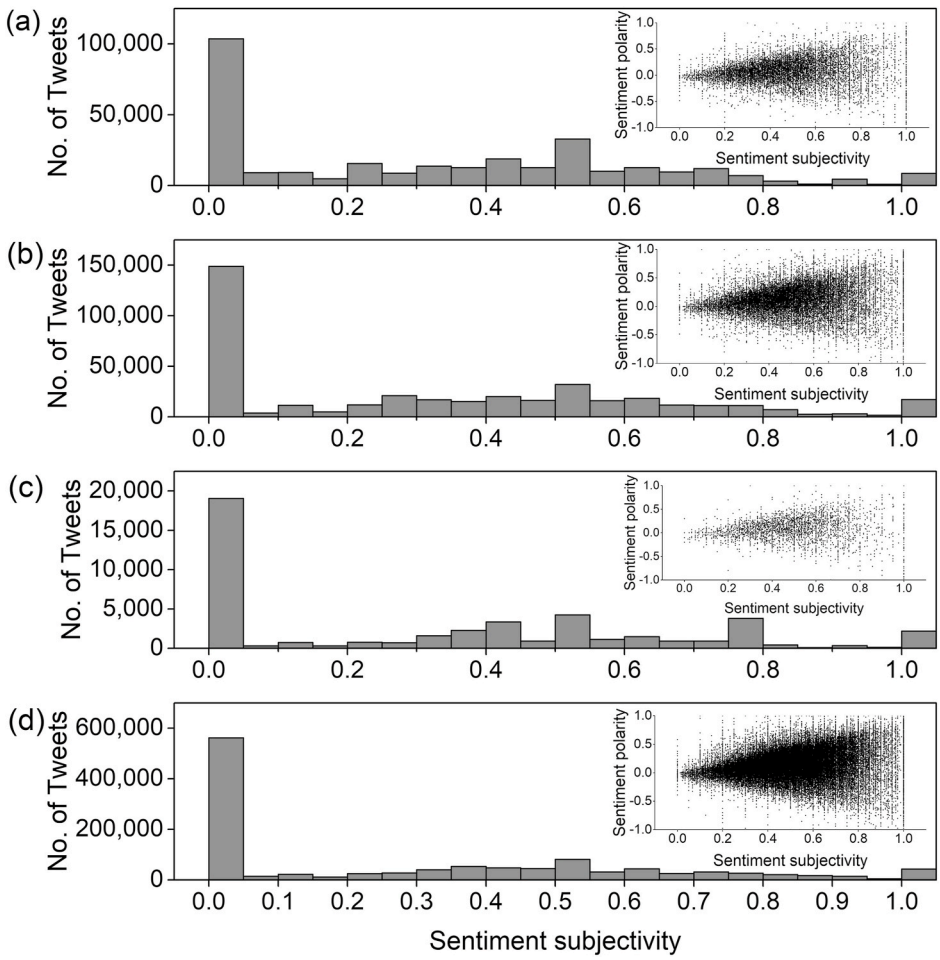

Fig. 4. Histograms showing the distribution of sentiment subjectivity and scatter plots (insets) showing the correlation between sentiment subjectivity and sentiment polarity for (a) Lombok-related Tweets with the keyword "earthquake", (b) Lombok-related Tweets without the keyword "earthquake", (c) Bali-related Tweets with the keyword "earthquake", (d) Bali-related Tweets without the keyword "earthquake".

4.3. Topic-based sentiment analysis

4.3.1. LDA topic modeling

The sentiment analysis detailed above does not consider the actual contents of Tweets. Therefore, topic-based sentiment analysis is included and performed via LDA for mining specific perspectives from the Tweets. Following (Yan et al., 2020), the LDA modeling is performed using the gensim Python wrapper for LDA from MALLET (McCallum, 2002; Řehůřek, 2019). LDA differentiates similar phrases with different contexts and separates them into different topics. A distribution over topics derived from Tweets is generated to represent the semantics in Tweets. The LDA modeling also generates a distribution over words to represent each topic.

The following natural language pre-processing steps are performed to reduce the semantic dimension of raw Tweets texts, creating word vectors required by the LDA topic modeling (Steiger et al. 2016).

- Tokenization: Cohesive strings from Tweets are spilt up into single words or "tokens" (Metke-Jimenez et al. 2011).
- Stop word removal: Stop words (e.g., "of" and "to") included in the NLTK natural language toolkit (Bird & Loper, 2004), namely the frequently occurring short-function words without valuable content, are removed.
- Lemmatization: Words are converted to their root forms to simplify further analyses using the Python library of spaCY (2018).
- Keyword removal: For Tweets falling in the clusters of LE and HE, the word "earthquake" is excluded from the word vectors, as earthquakes comprise a generic theme of the Tweet dataset, while what need to be derived and analyzed are subtopics pertinent to earthquakes.

Following Steyvers and Griffiths (2007), the two Dirichlet parameter priors α and β for the LDA modeling are computed by $50/K$ (K represents the number of topics to be derived) and assigned as 0.01, respectively. The K is optimized via the four stage topic coherence pipeline as presented in Röder et al. (2015), using Gensim (Řehůřek, 2019). The LDA posterior parameter inference and optimization are performed using Gibbs sampling which is a form of Markov chain Monte Carlo (Griffiths & Steyvers, 2004; Steiger et al. 2016). Lastly, for discussing the Tweets associated with the topic modeling results, "*****" is used to cover personal information (e.g., names mentioned in a Tweet).

4.3.2. Sentiment analysis of major topics

Since LDA derives multiple (K) topics from each cluster of Tweets, the top five topics ranked based on the number of Tweets associated with each topic, together with the corresponding sets of representative topic keywords and the most relevant Tweet in each topic (i.e., the Tweet that is most dominated by a given topic, identified based on the percentage contribution of the topic generated by the LDA algorithm) (Blei et al. 2003; Prabhakaran, 2020) are extracted. This narrows down the contents of Tweets for analysis to the major topics of Tweeting in the individual clusters. The major topics are interpreted to derive the patterns and trends of the Tweets over time. Specifically, the same as (Yan

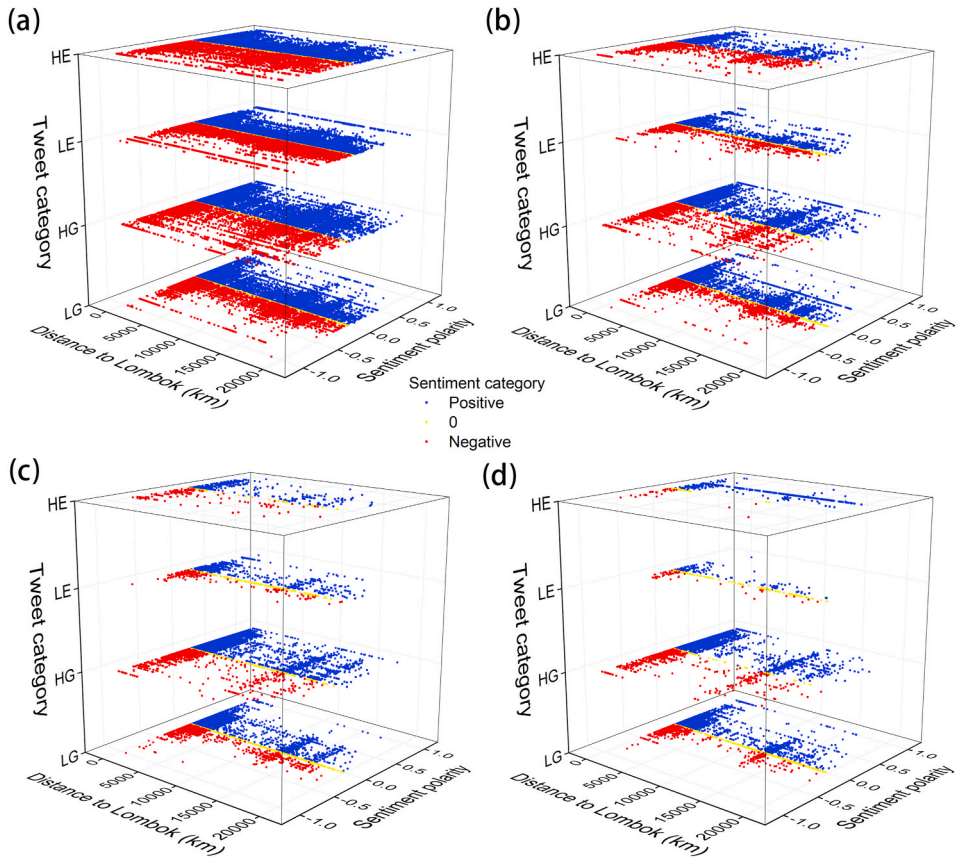

Fig. 5. Sentiment cubes encapsulating the scatter plots of the geotagged Tweets mentioning Lombok posted in (a) August 2018, (b) October 2018, (c) January 2019, and (d) April 2019.

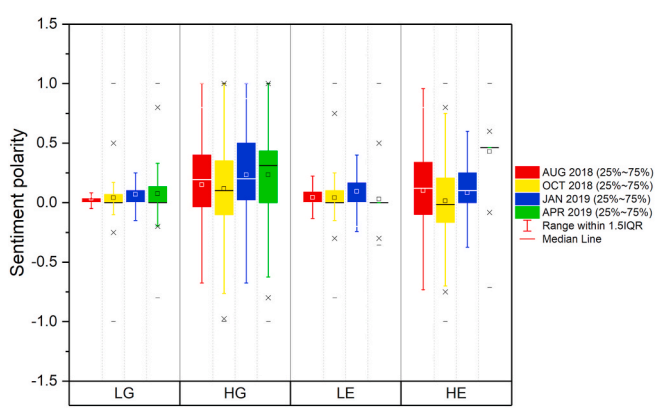

Fig. 6. Boxplot showing the sentiment polarities' central tendencies of the four clusters of Tweets mentioning Lombok in the four observation months.

et al., 2020), a descriptor is provided for each topic based on our human interpretation of the most relevant Tweet and the topic keywords. A scatter plot of the four clusters of Tweets (i.e., LG, HG, LE, HE) associated with the top five topics is also created for each of the four observation months. Lastly, the sentiment polarities' central tendencies of the four clusters of Tweets associated with the top five topics are visualized for the four observation months using box plot.

5. Results and interpretations

5.1. Overall sentiments of tweets

For all Tweets of the four observation months, the distribution of sentiment subjectivity and sentiment polarity and the correlation between them is shown in Fig. 4. It is observed that a big portion of the Tweets have a low subjectivity between 0 and 0.05. It is also found that low subjectivity tweets also have low variance in polarity, suggesting that low subjectivity Tweets are more emotionally stable.

The scatter plots of the Tweets mentioning Lombok are shown in Fig. 5. It is observed that August 2018 (the month that the earthquakes occurred) is associated with the highest density of Tweets, and a descending trend of Tweets density can be seen in the following three observation months. This trend is because, after the earthquakes, Twitter users' attention to Lombok has decreased as time goes on. In general, the density of Tweets with negative sentiments decreases more rapidly than those with positive sentiments. Especially for the HE and LE Tweets in April 2019 (nine months after the earthquakes), the density of Tweets with negative sentiments is much lower than those with positive sentiments. These observations imply that people have become less negative about Lombok over time.

The sentiment polarities' central tendencies of the four clusters of Tweets mentioning Lombok are presented in Fig. 6. It is discovered that the HG Tweets generally have the greatest sentiment polarity means across the four observation months; this indicates that the HG Tweets are, albeit less objective, the most positive about Lombok. The sentiment polarity means of the HE Tweets peak in April 2019; this implies that the HE Tweets are the most positive about the earthquake-related situations in Lombok around nine months later. The HG and HE Tweets are both emotionally less stable over the four observation months compared to the LG and LE Tweets, due to their high subjectivity. An ascending trend of the sentiment polarity means of the LG Tweets over the four observation months can be seen, although the variation is relatively small; this suggests that the LG Tweets have become more positive about Lombok over time. As for the LE Tweets, their sentiment polarity means

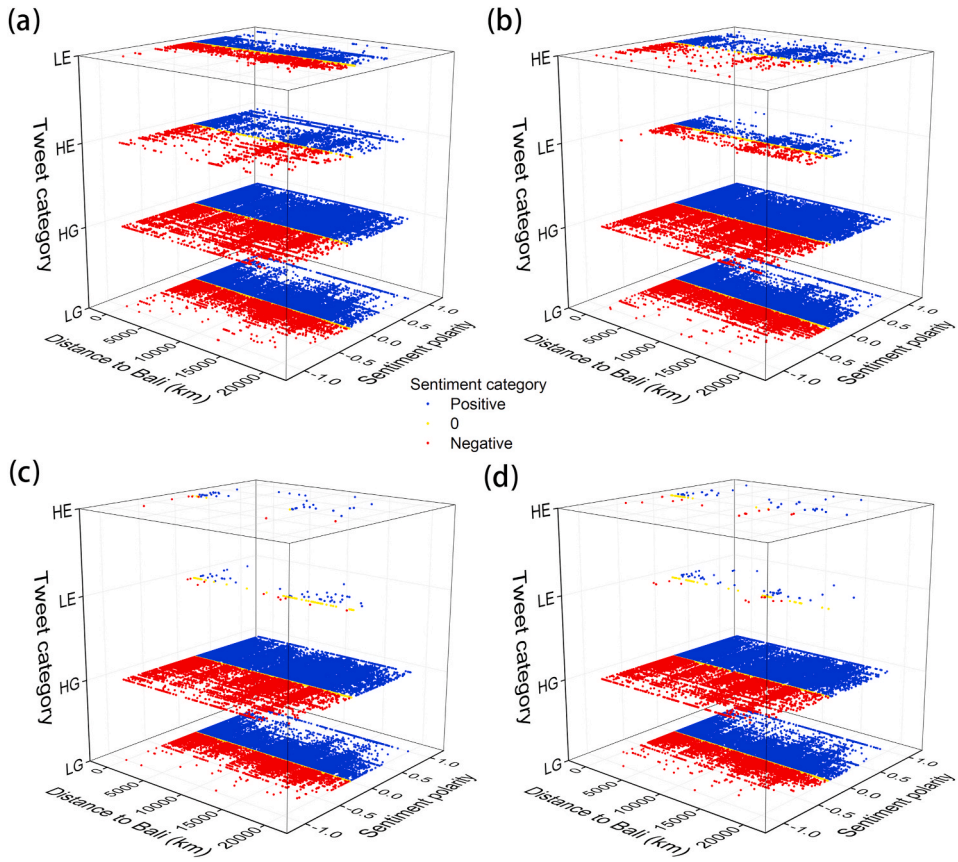

Fig. 7. Sentiment cubes encapsulating the scatter plots of the geotagged Tweets mentioning Bali posted in (a) August 2018, (b) October 2018, (c) January 2019, and (d) April 2019.

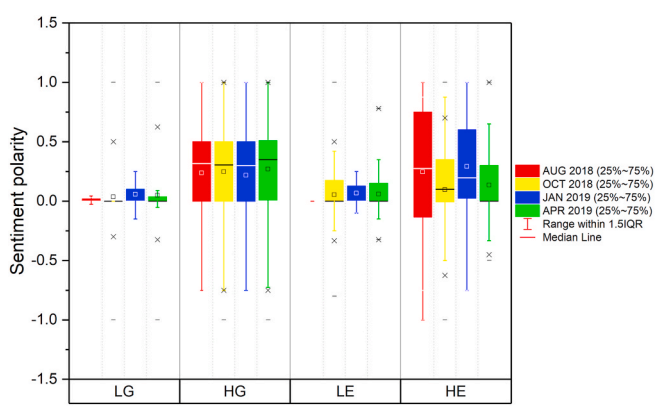

Fig. 8. Boxplot showing the sentiment polarities' central tendencies of the four clusters of Tweets mentioning Bali in the four observation months.

fluctuate over time, which could be an objective sentiment reflection on the post-disaster recovery status.

The scatter plots of the Tweets mentioning Bali are shown in Fig. 7. The patterns and trends exhibited in Fig. 7 resemble those observed in Fig. 5, with two exceptions: (1) no obvious decay of the densities of the LG and HG Tweets over the four observation months is observed; (2) a considerably lower number of earthquake-related Tweets (i.e., the LE and HE Tweets) can be seen in January 2019 and April 2019. The first exception can be explained by Bali's higher popularity than Lombok as a tourism destination, as a result of which the number of LG and HG Tweets mentioning Bali is generally stable during the four observation months. The second exception suggests that the impact of earthquakes on Bali has become less an issue and has been paid far less attention in 2019. Compared with Lombok, Bali is more likely to receive greater attention and support from the international community. People may be distracted by other events and issues as Bali's reconstruction proceeds.

The sentiment polarities' central tendencies of the four clusters of Tweets mentioning Bali are presented in Fig. 8. Similar to Fig. 6, the HG Tweets are generally with the greatest sentiment polarity means across the four observation months. In addition, the sentiment polarity means of the LE Tweets in August 2018 are the lowest across the four observation months, which have increased in the subsequent months; this suggests that the Tweets are the least optimistic about the earthquake-related situations in Bali in August 2018. The sentiment polarity means of the LG and HE Tweets fluctuate over time.

5.2. Topic-based sentiments of tweets

Table 2 and Appendix 2 show the top five LDA-derived topics of the Tweets mentioning Lombok (with "earthquake" in the text of each Tweet) posted in the four observation months together with the corresponding sets of representative topic keywords and the most relevant Tweet in each topic. The keywords and the most relevant Tweets suggest that the topics in the first month are mainly about earthquake response (e.g., impact evaluations, rescue actions, volunteer activities, donations, shelters, and response speed). The topics in following three observation months are mainly about post-earthquake aids and recovery (e.g., measures and progress). Regarding people's perspectives, the help of Red Cross has been appreciated. Dissatisfactions about the recovery progress have also been reported. For example, slow housing reconstruction pace and unfavorable living conditions in Lombok after more

Table 2
Keywords describing the top five LDA-derived topics of the Tweets mentioning Lombok (with "earthquake" in the text of each Tweet) posted in the four observation months and a brief description of the most relevant Tweet in each topic.

Time period	Topic number	Keywords describing the topic	Brief description of the most relevant Tweet
August 2018	1st	island, magnitude, Indonesian, people, hit, shake, strike, strong, Sunday, quake	Response (satellite observations)
	2nd	donation, donate, raise, fund, distribute, care, august, collect, link, channel	Response (donations)
	3rd	time, people, Bali, day, lose, affect, shelter, market, building, open	Response (volunteer activities and shelter preparation)
	4th	death, toll, island, Indonesian, news, rise, people, hit, magnitude, official	Response (death toll)
	5th	disaster, national, status, government, set, president, request, Lombok, parliament, official	Response (slow response)
October 2018	1st	home, Lombok, rebuild, time, disaster, central, people, hit, follow, god	Recovery (home rebuilding)
	2nd	agriculture, sell, shoulder, touch, share, show, country, victim, burden, feel	Recovery (agricultural activities)
	3rd	aid, rehabilitation, disbursement, review, reconstruction, back, process, ensure, government, follow	Recovery (government aids)
	4th	north, promise, today, yesterday, president, break, realization, reconstruction, building, heavily	Recovery (no reconstruction of ground-level building)
	5th	demonstration, fund, demand, aid, government, relief, back, history, disaster, severe	Recovery (government aids)
January 2019	1st	home, president, tarpaulin, set, sun, protect, helpful, durian, fruit, run	Aids for recovery (Red Cross)
	2nd	finish, house, construction, month, remember, reach, character, disappoint, buzzer, difference	Recovery (slow housing reconstruction)
	3rd	home, side, read, cry, mental, Allah, remember, agar, Tweet, masjid	Recovery (housing reconstruction plan)
	4th	recover, water, tourism, food, industry, top, support, today, establishment, visit	Recovery (the tourism is improving slowly)
	5th	aid, hand, coverage, silence, praise, fly, hold, charity, victim, anniversary	Aids for recovery (charity aids)
April 2019	1st	family, life, beautiful, natural, disaster, drastically, moment, house, magic, unpaid	Recovery (life changed)
	2nd	north, follow, bar, west, regulation, school, fund, hygiene, people, affect	Recovery (unfavorable living conditions)
	3rd	region, impact, polling, station, contact, help, flood, north, twitter, high	Response (flood)
	4th	magnitude, central, west, rattle, feel, good, Bali, box, shake, august	Aids for recovery (sustenance sharing)
	5th	year, hit, rise, massive, learn, world, develop, recover, tourism, earth	Recovery (the tourism is still a concern)

than eight months following the earthquake have been reported (Appendix 2). It has also been reported in January 2019 that tourism began to recover at a slow pace, which was still a concern in April 2019 as echoed by media releases (e.g., Aisyah, 2019).

Table 3 and Appendix 3 show the top five LDA-derived topics of the Tweets mentioning Lombok (without "earthquake" in the text of each Tweet) posted in the four observation months together with the corresponding sets of representative topic keywords and the most relevant Tweet in each topic. Since these Tweets are not directly related to earthquakes, the topics are more diverse. The topics in the first month are also mainly about earthquake response and impact (e.g., government responses and damage reports). The focus of the topics in October 2018 have become mainly about aids (e.g., public donations) and tourism (e.g., the Halal Tourism of Lombok). The topics in January 2019 pertain to aids (e.g., government's financial aid), festival (e.g., the Bau Nyale Charm Festival of Lombok), business (e.g., international business investment), and tourism (e.g., the Mount Rinjani tourism), while the topics in April 2019 center around politics (e.g., elections), food (e.g., cooking), and tourism (e.g., the culture village tourism of Lombok). The patterns and trends of the Twitter topics imply that people's attention to the disasters have gradually been distracted by other matters as reconstruction proceeds.

The scatter plots of the Tweets mentioning Lombok involved in the top five LDA topics are shown in Fig. 9. The overall density of the Tweets

Table 3
Keywords describing the top five LDA-derived topics of the Tweets mentioning Lombok (without "earthquake" in the text of each Tweet) posted in the four observation months and a brief description of the most relevant Tweet in each topic.

Time period	Topic number	Keywords describing the topic	Abstract of the most relevant Tweet
August 2018	1st	north, epicenter, mag, northwest, feel, sea, west, central, land, east	Earthquake impact (magnitude and epicenter)
	2nd	Lombok, people, good, government, pack, work, disaster, hope, make, order	Response (government aids)
	3rd	quake, island, rubble, god, cry, instant, pray, magnitude, toll, death	Response (death toll)
	4th	Bali, travel, island, Lombok, beach, day, week, time, trip, beauty	Mourning
	5th	stay, safe, pray, prayer, Bali, president, god, surrounding, strong, grateful	Prayer
October 2018	1st	feel, sea, epicenter, mag, northeast, west, southeast, unrivaled, collectively, spectator	Earthquake impact (magnitude and epicenter)
	2nd	Lombok, strike, disaster, care, nation, big, people, make, write, manage	Comment
	3rd	Lombok, donation, god, land, servant, money, people, ill, auxiliary, continue	Compassion
	4th	travel, beach, chili, island, Bali, marathon, call, day, trip, good	Tourism
	5th	loading, disaster, prayer, owe, central, food, set, kitchen, bit, area	Public aids for victims
January 2019	1st	festival, beautiful, charm, island, west, February, unique, daughter, incarnation, public	Festival
	2nd	Indonesia, top, east, north, proof, aid, governor, provide, fly, handle	Government aids
	3rd	business, cool, time, product, hotel, wait, local, good, sell, like	Business
	4th	work, good, market, morning, pack, people, real, leader, hard, love	Business
	5th	island, travel, make, west, nice, roam, prepare, Bali, hub, mount	Tourism
April 2019	1st	support, circuit, continue, host, country, family, day, national, east, west	Politics
	2nd	April, vote, optimistic, win, victory, heart, ready, society, steady, community	Politics
	3rd	chili, eat, rice, fry, green, sauce, chicken, tasty, breakfast, coffee	Food
	4th	Bali, beach, travel, island, location, beautiful, trip, love, day, visit	Tourism
	5th	village, Bali, tourism, trip, attack, bung, build, price, java, good	Tourism

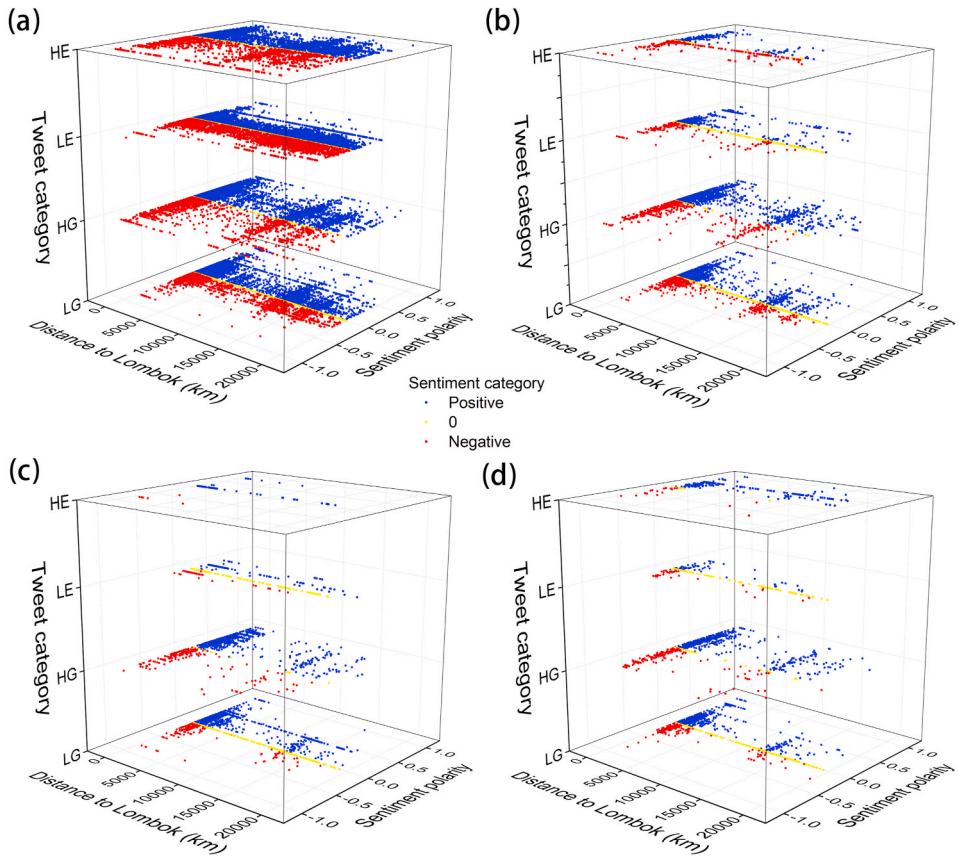

Fig. 9. Sentiment cubes encapsulating the scatter plots of the geotagged Tweets involved in the top five LDA topics mentioning Lombok posted in (a) August 2018, (b) October 2018, (c) January 2019, and (d) April 2019.

is lower, but similar patterns to those in Fig. 5 are observed. The sentiment polarities' central tendencies of the four clusters of Tweets involved in the top five LDA topics mentioning Lombok in the four observation months are illustrated in Fig. 10. Similar to Fig. 6, the HG Tweets, in general, are with the greatest sentiment polarity means across the four observation months. The sentiment polarity means of the LG Tweets in August 2018 is the lowest, suggesting that the LG Tweets are the least positive about Lombok in August 2018. Similar to Fig. 6, the sentiment polarity means of the HE Tweets peak in April 2019; the sentiment polarity means of the LE Tweets fluctuate over time.

Table 4 and Appendix 4 show the top five LDA-derived topics of the

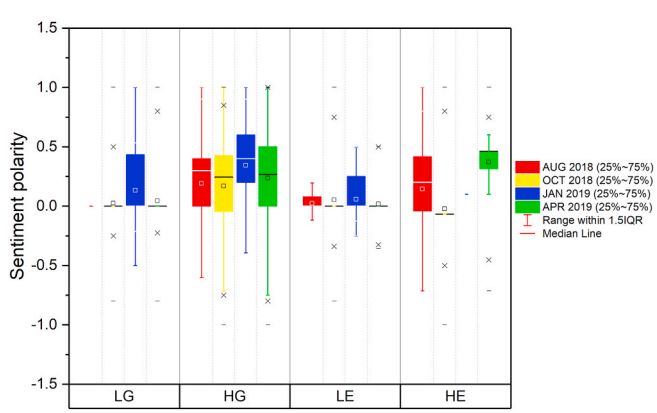

Fig. 10. Boxplot showing the sentiment polarities' central tendencies of the four clusters of Tweets involved in the top five LDA topics mentioning Lombok in the four observation months.

Tweets mentioning Bali (with "earthquake" in the text of each Tweet) posted in the four observation months together with the corresponding sets of representative topic keywords and the most relevant Tweet in each topic. According to the keywords and the most relevant Tweets, the topics in the first month are mainly about earthquake impact and response; the topics in October 2018 are mainly about earthquake impact, response, and aids; the topics in January 2019 have changed to earthquake impact and recovery; the topics in April 2019 mainly pertain to earthquake impact. Regarding people's perspectives, it is interesting to note that both supporting and opposing opinions (i.e., contestations) have been posted with regard to the huge funds for hosting the 2018 Bali International Monetary Fund (IMF)-World Bank meeting after the earthquakes. It is further noticed that multiple disasters, including not only earthquakes but also tsunamis and volcano eruptions, were reported after the major earthquakes in August 2018. People questioned the distribution of food aid in October 2019. Dramatic loss of tourism revenues and tourists due to the earthquakes were reported in April.

Table 5 and Appendix 5 show the top five LDA-derived topics of the Tweets mentioning Bali (without "earthquake" in the text of each Tweet) posted in the four observation months together with the corresponding sets of representative topic keywords and the most relevant Tweet in each topic. Similar to the case of Lombok, the topics are diverse compared with those Tweets including the keyword "earthquake". In addition, the topics of Tweets regarding Bali are more diverse than those about Lombok (Table 3 and Appendix 3). This is partly explained by Bali's higher popularity as a tourism destination than Lombok. In August 2018, the topics are about tourism, religion (e.g., the Onam celebrations in Kerala), environment (e.g., the coal waste), and business, while in October 2018, the topics are mainly related to tourism, crime (e.g., frauds), and international meeting (e.g., the IMF World Bank meeting).

Table 4
Keywords describing the top five LDA-derived topics of the Tweets mentioning Bali (with "earthquake" in the text of each Tweet) posted in the four observation months and a brief description of the most relevant Tweet in each topic.

Time period	Topic number	Keywords describing the topic	Brief description of the most relevant Tweet
August 2018	1st	agree, world, budget, bank, meeting, Bali, assessment, Bali, survivor, victim	Response (fiscal expenditures)
	2nd	hit, imam, video, viral, pray, mosque, show, remarkable, Bali, gather	Prayer
	3rd	magnitude, video, center, Allah, Bali, mosque, speechless, sustained, pray, sad	Response (situational awareness)
	4th	Bali, magnitude, tsunami, rattle, quake, hit, warn, island, potential, alert	Earthquake impact (magnitude and epicenter)
	5th	Bali, people, island, resort, hit, magnitude, kill, Indonesian, tourist, dead	Response (death toll)
October 2018	1st	conference, tsunami, happiness, affect, fair, supporter, deal, marriage, Bali, clubbing	Perspectives about IMF Bali conference
	2nd	Bali, java, east, magnitude, rattle, people, panic, update, death, run	Earthquake impact (magnitude)
	3rd	Bali, java, east, magnitude, feel, pray, surrounding, safety, give, protection	Earthquake impact (magnitude and feelings)
	4th	Bali, magnitude, vibration, rattle, java, video, deadly, madly, news, quake	Earthquake comments
	5th	Bali, mag, food, understand, refugee, aid, central, opinion, market, strengthen	Response (delayed food aids)
January 2019	1st	volcanic, ash, throw, volcan, column, today, island, Bali, date, advice	Volcanic eruption
	2nd	expertise, mag, Bali, southwest, tsunami, central, northeast, post, skills, recovery	Recovery (volunteer activities)
	3rd	Bali, magnitude, damage, January, numerous, trigger, quake, sequence, province, session	Earthquake impact (magnitude and damage assessment)
	4th	side, feel, day, Bali, place, connection, lead, note, legend, pray	Earthquake impact and prayer
	5th	Bali, mount, volcano, yoga, barn, experience, night, class, waking, late	User experience during an earthquake
April 2019	1st	tsunami, potential, lose, hammer, drastically, strait, loss, Bali, chili, seismic	Earthquake impact on tourism industry
	2nd	magnitude, Bali, tectonic, depth, rattle, time, region, volcanic, location, automatic	Earthquake impact (magnitude and epicenter)
	3rd	people, volcano, eruption, manila, affect, late, India, snapshot, Philippine, Bali	Impacts of disasters
	4th	sea, Bali, report, ago, area, center, tue, regional, shallow, metropolis	Earthquake comments
	5th	Bali, southeast, April, mag, epicenter, hit, august, depth, strike, local	Earthquake impact (magnitude and epicenter)

The major topics of Tweets are about tourism and crime in January 2019, and about tourism, animal protection, politics, entertainment news, and religions instead of earthquake in April 2019. The topics imply that various daily routines (e.g., the tourism, business, religions, politics, crimes controls, and international meetings) functioned well in Bali after the earthquakes, despite the loss of tourism revenues and tourists as mentioned above.

The scatter plots of the Tweets mentioning Bali involved in the top five LDA topics are shown in Fig. 11. The density of the Tweets is lower but similar patterns to those in Fig. 7 are observed. The sentiment polarities' central tendencies of the four clusters of Tweets involved in the top five LDA topics mentioning Bali in the four observation months are shown in Fig. 12. Similar to Fig. 8, the HG Tweets, in general, are with the greatest sentiment polarity means across the four observation months. The sentiment polarity means of the HG and HE Tweets exhibit a descending trend over time, implying that some people had become less optimistic about Bali. One of the reasons may be the multiple disasters following the major earthquakes in August 2018 (Table 4 and Appendix 4). Measures should be taken by Bali to improve its destination image and reputation.

6. Concluding remarks

6.1. Research implications

In this study, a mixed methods approach involving sentiment analysis and LDA topic modeling is designed for mining sheer volume of Tweets about Lombok and Bali. The Tweets were generated by nonlocal Twitter users after a series of earthquakes occurred in the two places in August 2018. The multiplicity and variation of people's sentiments and perspectives after the earthquakes have been uncovered, which has implications for stakeholders' appraisal of the recovery status of the two islands. Specifically, the findings of this study mainly suggest that people have generally become less negative about Lombok and Bali over time, despite fluctuations in their sentiment polarities' central tendencies.

Table 5
Keywords describing the top five LDA-derived topics of the Tweets mentioning Bali (without "earthquake" in the text of each Tweet) posted in the four observation months and a brief description of the most relevant Tweet in each topic.

Time period	Topic number	Keywords describing the topic	Brief description of the most relevant Tweet
August 2018	1st	Bali, travel, beach, beautiful, trip, good, island, holiday, love, vacation	Tourism
	2nd	Bali, birthday, happy, hope, enjoy, story, love, good, great, celebrate	Religion
	3rd	Bali, travel, style, bora, japan, wife, hope, canary, swing, Paris	Tourism
	4th	Bali, beach, business, lie, money, island, foreign, people, month, bank	Environment
	5th	Bali, good, day, time, back, morning, holiday, today, flight, home	Business
October 2018	1st	Bali, goal, event, beach, vacation, spend, clean, relax, sand, cover	Tourism
	2nd	Bali, bank, dear, meeting, video, India, give, feel, testimony, impact	Crime
	3rd	Bali, Halloween, luxury, twitter, holiday, event, post, party, big, lot	Tourism
	4th	Bali, travel, beach, sunset, beautiful, island, temple, resort, beauty, holiday	Tourism
	5th	Bali, world, bank, annual, meeting, event, support, international, success, economy	International meeting
January 2019	1st	Bali, year, time, prison, court, spend, sentence, British, visa, prosecutor	Crime
	2nd	Bali, night, hotel, travel, book, day, good, bo, ready, price	Tourism
	3rd	Bali, travel, beach, good, sunset, beautiful, island, temple, nature, trip	Tourism
	4th	Bali, villa, resort, dog, coffee, hotel, vacation, pool, travel, design	Tourism
	5th	Bali, cheap, cool, travel, money, good, time, stop, person, fun	Tourism
April 2019	1st	Bali, travel, beach, island, good, hotel, resort, trip, beautiful, post	Tourism
	2nd	Bali, back, miss, fracture, amazing, life, elephant, sanctuary, feel, return	Animal protection
	3rd	Bali, make, work, party, parliament, government, festival, speech, people, April	Politics
	4th	Bali, time, send, hope, honey, video, photo, live, trip, part	Entertainment news
	5th	day, Bali, temple, father, child, pray, focus, interfere, disrupt, religion	Religion

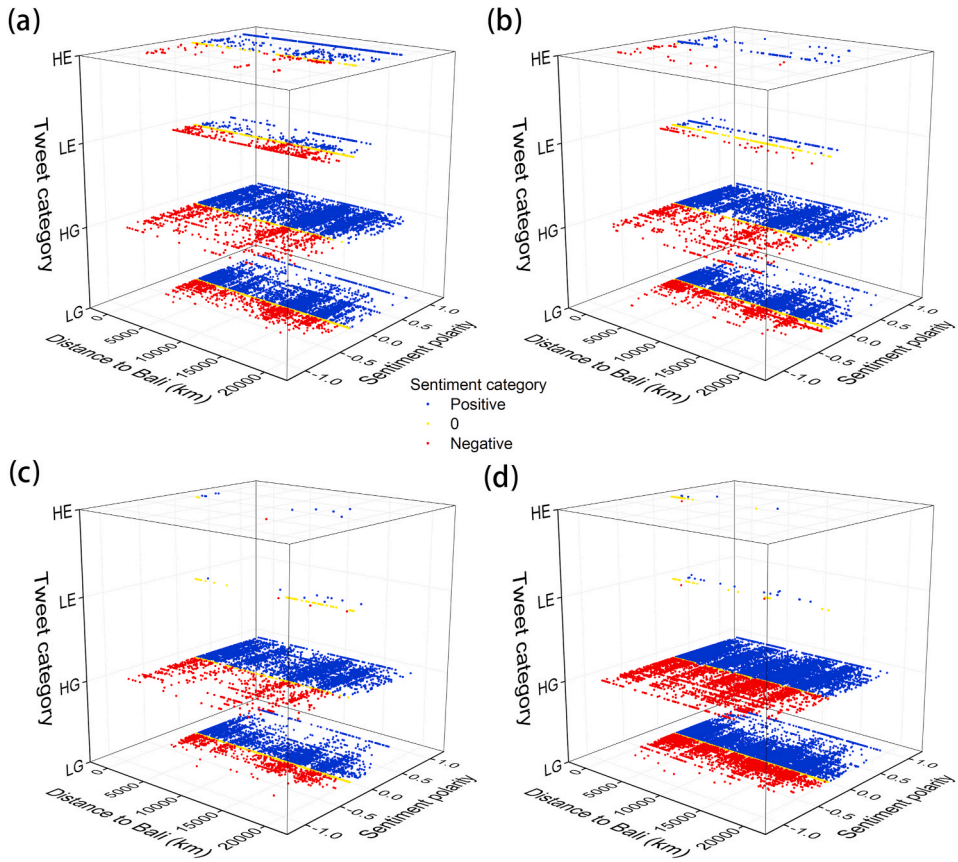

Fig. 11. Sentiment cubes encapsulating the scatter plots of the geotagged Tweets involved in the top five LDA topics mentioning Bali posted in (a) August 2018, (b) October 2018, (c) January 2019, and (d) April 2019.

Some concerns regarding the post-disaster recovery have also been discovered. For instance, dissatisfactions about the housing reconstruction progress, living conditions, and tourism recovery status in the affected areas of Lombok still existed in 2019 (Appendix 2); contestations about the huge funds for hosting the 2018 Bali IMF-World Bank meeting after the earthquakes have also been found (Appendix 4). Therefore, unlike many existing studies which have focused purely on the reconstruction status of the physical components (i.e., infrastructures) and re-influx of tourists in tourism destinations (Yan et al., 2017a), the text mining of this study can reveal people's feelings during the recovery process of Lombok and Bali. The negative views about the Lombok recovery can urge policy-makers to take targeted measures (e.g., measures for enhancing housing reconstructions) to fix the image and reputation of the tourism destination. As a matter of fact, fixing physical damage to tourism destinations after catastrophic events can be planned effectively. However, it is a complicated and long-term project to repair the image and reputation of an affected destination, which may benefit from the approach and workflow designed in this study. On the bright side, this study has also uncovered Tweets with positive sentiments and perspectives, such as people's positive views about having holidays in Bali (Table 5 and Appendix 5). These Tweets are in fact beneficial for the post-earthquake repairment of the place image and reputation of Bali as well as for the promotion of Bali tourism.

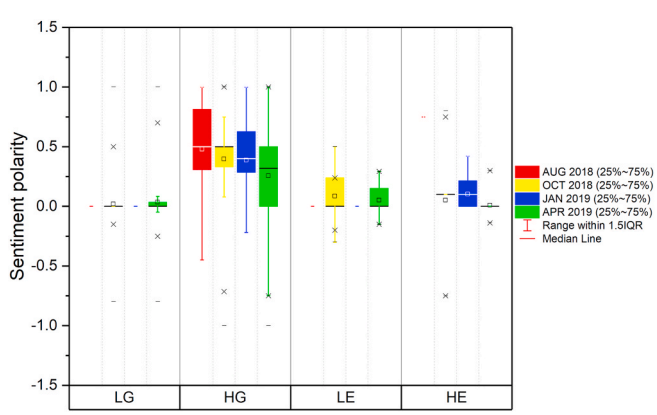

Fig. 12. Boxplot showing the sentiment polarities' central tendencies of the four clusters of Tweets involved in the top five LDA topics mentioning Bali in the four observation months.

Overall, the value of social media for monitoring and assessing the post-disaster recovery of tourism destinations has been demonstrated. That is, the text mining in this study can effectively reveal the variations of people's sentiments and perspectives of general and specific issues regarding the post-disaster recovery process of Lombok and Bali.

6.2. *Limitations and future works*

Three limitations of this study must be acknowledged. First, although Tweets posted by news channels are excluded in this study through a simple keyword-based approach (Section 4.1), better methods should be adopted to remove diverse non-human-generated (bot-generated) Tweets from the database, so as to analyze only Tweets from the general public. Second, as a data source, the sampling issue and bias of social media related to culture, demographics, spatial distribution, user behavior (e.g., not all user geotag their Tweets), and even its own API (at most 1% sample of all Tweets are retrievable) are widely acknowledged

and left unaddressed in this research (Yan et al. 2020). Third, note that the text mining of the Tweets from the nonlocal Twitter users may only reflect their sentiments and perspectives rather than the ground truth. Many of them do have visited the islands after the earthquakes, but many of them may not be aware of the actual local situation and rely on international media who frequently paint inaccurate pictures of the situation on the ground after a disaster. As mentioned in Section 2.2, "given enough eyes all bugs are shallow" according the Linus' Law. Therefore, it is justifiable that the overall sentiment and perspective patterns (the general consensuses) from the Tweets are still helpful for understanding how people feel about the recovery and for delineating the image and reputation of the tourism destinations. Understanding people's psychology as opposed to authoritative voices is an important aspect worth studying, but robust approaches are needed to validate the trustworthiness and credibility of Tweets as well as separate Tweets with situational awareness from those without situational awareness.

For future works, the limitations mentioned above should be addressed primarily. Additionally, the number of nonlocal Bali-related Tweets with "earthquake" is small in January and April 2020 (Table 1). Although the LDA modeling generated satisfactory results in this study, in many cases LDA may suffer from so-called "small sample size problem" (Lu et al. 2016) which should be taken into consideration in future research. Moreover, photo-based social media data (e.g., Instagram and Flickr) can be involved in the data mining for providing more visual contexts for the Twitter text-mining. Furthermore, in-depth analysis and classification of sentiments and perspectives (e.g., anger, sadness, and joy) should also be performed. Lastly, the political, social, economic, religious, and cultural context of tourism destinations have a role in people's sentiments and perspectives during post-disaster recovery, future research should investigate the social media data in close connection with each of these aspects and their interdependence.

CRediT authorship contribution statement

Yingwei Yan: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing - original draft, Writing - review & editing, Visualization, Supervision, Project administration, Funding acquisition. **Jingfu Chen:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Writing - original draft, Writing - review & editing, Data curation, Project administration, Funding acquisition. **Zhiyong Wang:** Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - original draft, Writing - review & editing.

Declaration competing of interest

We declare that there are no real or perceived conflicts of interest involved in the submission and/or publication of this manuscript.

Acknowledgements

This work has been supported by the GDAS′ Project of Science and Technology Development (grant number: 2020GDASYL-20200103005 and 2019GDASYL-0501001), the National Natural Science Foundation of China (grant number: 41901330 and 41901165), the Key Special Project for Introduced Talents Team of Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou) (grant number: GML2019ZD0301), and Guangdong Innovative and Entrepreneurial Research Team Program (grant number: 2016ZT06D336).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.apgeog.2020.102306.

References

spaCY. (2018). Industrial-strength natural language processing in Python [online]. Available from: https://spacy.io/ Accessed 23rd August 2018.

Aisyah, R. (2019). Lombok tourism taking baby steps toward recovery after earthquake [online]. Available from: https://www.thejakartapost.com/news/2019/02/28/lombok-tourism-taking-baby-steps-toward-recovery-after-earthquake.html Accessed 11th October 2019.

Al Walid, M., Anisuzzaman, D., & Saifuddin Saif, A. (2019). Data analysis and visualization of continental cancer situation by Twitter scraping. *International Journal of Modern Education and Computer Science, 11*(7).

BBC. (2018). Lombok earthquake: Strong tremors shake Indonesian island [online]. Available from: https://www.bbc.com/news/world-asia-45238018 Accessed 19th August 2018.

Beech, S., & LeBien, M. (2018). At least 91 dead after powerful earthquake strikes Indonesian resort islands of Bali. *Lombok* [online]. Available from: https://edition.cnn.com/2018/08/05/asia/indonesia-earthquake/index.html Accessed 6th August 2018.

Bird, S., & Loper, E. (2004). NLTK: The natural language toolkit. In *Proceedings of the association for computational linguistics (ACL) 2004 on interactive poster and demonstration sessions* (p. 31). Barcelona, Spain: ACL.

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

Boulton, C. A., Shotton, H., & Williams, H. T. (2016). Using social media to detect and locate wildfires. In *Proceedings of the workshops of the 10th international AAAI conference on Web and social media*. Cologne, Germany: The Association for the Advancement of Artificial Intelligence.

Brown, D., Saito, K., Liu, M., Spence, R., So, E., & Ramage, M. (2012). The use of remotely sensed data and ground survey tools to assess damage and monitor early recovery following the 12.5.2008 Wenchuan earthquake in China. *Bulletin of Earthquake Engineering, 10*(3), 741–764.

Budiharto, W., & Meiliana, M. (2018). Prediction and analysis of Indonesia Presidential election from Twitter using sentiment analysis. *Journal of Big Data, 5*(1), 51.

Carriere, D. (2019). Geocoder: Simple, consistent [online]. Available from: https://geocoder.readthedocs.io/ Accessed 1st August 2019.

Clips. (2018). Pattern [online]. Available from: https://www.clips.uantwerpen.be/pattern Accessed 23rd July 2019.

de Albuquerque, J. P., Herfort, B., Brenning, A., & Zipf, A. (2015). A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management. *International Journal of Geographical Information Science, 29*(4), 667–689.

Du, E., Cai, X., Sun, Z., & Minsker, B. (2017). Exploring the role of social media and individual behaviors in flood evacuation processes: An agent-based modeling approach. *Water Resources Research, 53*(11), 9164–9180.

Fohringer, J., Dransch, D., Kreibich, H., & Schröter, K. (2015). Social media as an information source for rapid flood inundation mapping. *Natural Hazards and Earth System Sciences, 15*(12), 2725–2738.

Frisby, E. (2002). Communicating in a crisis: The British Tourist Authority's responses to the foot-and-mouth outbreak and 11th September, 2001. *Journal of Vacation Marketing, 9*(1), 89–100.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management, 35*(2), 137–144.

Gauba, H., Kumar, P., Roy, P. P., Singh, P., Dogra, D. P., & Raman, B. (2017). Prediction of advertisement preference by fusing EEG response and sentiment analysis. *Neural Networks, 92*, 77–88.

Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *Geojournal, 69*(4), 211–221.

Goodchild, M. F., & Glennon, J. A. (2010). Crowdsourcing geographic information for disaster response: A research frontier. *International Journal of Digital Earth, 3*(3), 231–241.

Goodchild, M. F., & Li, L. (2012). Assuring the quality of volunteered geographic information. *Spatial Statistics, 1*, 110–120.

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences, 101*(suppl 1), 5228–5235.

Haworth, B. T. (2018). Implications of volunteered geographic information for disaster management and GIScience: A more complex world of volunteered geography. *Annals of the Association of American Geographers, 108*(1), 226–240.

Hoshi, T., Murao, O., Yoshino, K., Yamazaki, F., & Estrada, M. (2014). Post-disaster urban recovery monitoring in Pisco after the 2007 Peru earthquake using satellite image. *Journal of Disaster Research, 9*(6), 1059–1068.

Huang, X., Li, Z., Wang, C., & Ning, H. (2019). Identifying disaster related social media for rapid response: A visual-textual fused CNN architecture. *International Journal of Digital Earth*, 1–23.

Huang, X., Wang, C., & Li, Z. (2018). Reconstructing flood inundation probability by enhancing near real-time imagery with real-time gauges and tweets. *IEEE Transactions on Geoscience and Remote Sensing, 56*(8), 4691–4701.

Indonesia. (2020). Statistic of international visitor arrivals in Indonesia [online]. *Ministry of Tourism and Economy Creative, Republic of Indonesia*. Available from https://www.kemenpar.go.id/categories/statistic-of-international-visitor Accessed 2 February 2020.

Kumar, A., & Singh, J. P. (2019). Location reference identification from tweets during emergencies: A deep learning approach. *International Journal of Disaster Risk Reduction, 33*, 365–375.

Longueville, B. D., Luraschi, G., Smits, P., Peedell, S., & Groeve, T. D (2010). Citizens as sensors for natural hazards: A VGI integration workflow. *Geomatica, 64*(1), 41–59.

Loria, S. (2014). Textblob: Simplified text processing [online]. Available from: https://textblob.readthedocs.io/en/dev/ Accessed 23rd July 2019.

Lu, G.-F., Zou, J., & Wang, Y. (2016). A new and fast implementation of orthogonal LDA algorithm and its incremental extension. *Neural Processing Letters, 43*(3), 687–707.

Mair, J., Ritchie, B. W., & Walters, G. (2016). Towards a research agenda for post-disaster and post-crisis recovery strategies for tourist destinations: A narrative review. *Current Issues in Tourism, 19*(1), 1–26.

Martín, Y., Li, Z., & Cutter, S. L. (2017). Leveraging twitter to gauge evacuation compliance: Spatiotemporal analysis of hurricane matthew. *PloS One, 12*(7), Article e0181701.

Maskrey, A. (2011). Revisiting community-based disaster risk management. *Environmental Hazards, 10*(1), 42–52.

McCallum, A. K. (2002). Mallet: A machine learning for language toolkit [online]. Available from: http://mallet.cs.umass.edu/ Accessed 15th August 2018.

Metke-Jimenez, A., Raymond, K., & MacColl, I. (2011). Information extraction from web services: A comparison of tokenisation algorithms. In *Proceedings of the SKY2011 workshop: Discovery and representation of runnable knowledge* (pp. 12–23). Paris, France: Queensland University of Technology.

Platt, S., Brown, D., & Hughes, M. (2016). Measuring resilience and recovery. *International Journal of Disaster Risk Reduction, 19*, 447–460.

Prabhakaran, S. (2020). Topic modeling with gensim (Python) [online]. Available from: https://www.machinelearningplus.com/nlp/topic-modeling-gensim-python/ Accessed 14 July 2020.

Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the space of topic coherence measures. In *Proceedings of the 8th association for computing machinery (ACM) international conference on Web search and data mining* (pp. 399–408). Shanghai, China: ACM.

Sahni, T., Chandak, C., Chedeti, N. R., & Singh, M. (2017). Efficient Twitter sentiment classification using subjective distant supervision. In *2017 9th international conference on communication systems and networks (COMSNETS), 4–8* pp. 548–553). IEEE. Jan. 2017.

Schroeder, R. (2014). Big Data and the brave new world of social media research. *Big Data & Society, 1*(2), 2053951714563194.

Schultz, F., Utz, S., & Göritz, A. (2011). Is the medium the message? Perceptions of and reactions to crisis communication via twitter, blogs and traditional media. *Public Relations Review, 37*(1), 20–27.

Schumann, R. L. (2018). Ground truthing spatial disaster recovery metrics with participatory mapping in post-Katrina Mississippi. *Applied Geography, 99*, 63–76.

Smith, G. P., & Wenger, D. (2007). Sustainable disaster recovery: Operationalizing an existing agenda. In *Handbook of disaster research*. New York, NY: Springer. New York.

Steiger, E., Resch, B., & Zipf, A. (2016). Exploration of spatiotemporal and semantic clusters of Twitter data using unsupervised neural networks. *International Journal of Geographical Information Science, 30*(9), 1694–1716.

Steyvers, M., & Griffiths, T. (2007). Probabilistic topic models. *Handbook of latent semantic analysis, 427*(7), 424–440.

Takahashi, B., Tandoc, E. C., & Carmichael, C. (2015). Communicating on twitter during a disaster: An analysis of tweets during typhoon haiyan in the Philippines. *Computers in Human Behavior, 50*, 392–398.

Walden, M. (2019). Lombok earthquake rattles Indonesia tourism [online]. Available from: https://www.voanews.com/east-asia-pacific/lombok-earthquake-rattles-indonesia-tourism Accessed 2019 11th October.

Wang, Z., & Ye, X. (2018). Social media analytics for natural disaster management. *International Journal of Geographical Information Science, 32*(1), 49–72.

Yan, Y., Kuo, C.-L., Feng, C.-C., Huang, W., Fan, H., & Zipf, A. (2018). Coupling maximum entropy modeling with geotagged social media data to determine the geographic distribution of tourists. *International Journal of Geographical Information Science, 32*(9), 1699–1736.

Yan, Y., Feng, C.-C., Huang, W., Fan, H., Wang, Y.-C., & Zipf, A. (2020). Volunteered geographic information research in the first decade: A narrative review of selected journal articles in GIScience. *International Journal of Geographical Information Science*, 1–27.

Yan, Y., Eckle, M., Kuo, C.-L., Herfort, B., Fan, H., & Zipf, A. (2017a). Monitoring and assessing post-disaster tourism recovery using geotagged social media data. *ISPRS International Journal of Geo-Information, 6*(5), 144.

Yan, Y., Feng, C.-C., & Chang, K. T.-T. (2017b). Towards enhancing integrated pest management based on volunteered geographic information. *ISPRS International Journal of Geo-Information, 6*(7), 224.

Yan, Y., Feng, C.-C., & Wang, Y.-C. (2017c). Utilizing fuzzy set theory to assure the quality of volunteered geographic information. *Geojournal, 82*, 517–532.

Yaqub, U., Sharma, N., Pabreja, R., Chun, S. A., Atluri, V., & Vaidya, J. (2020). Location-based sentiment analyses and visualization of Twitter election data. *Digital Government: Research and Practice, 1*(2). Article 14.

Yates, D., & Paquette, S. (2010). Emergency knowledge management and social media technologies: A case study of the 2010 Haitian earthquake. In *Proceedings of the 73rd ASIS&T annual meeting on navigating streams in an information ecosystem - volume 47*. Pittsburgh, Pennsylvania: American Society for Information Science.

Rathfon, D., Davidson, R., Bevington, J., Vicini, A., & Hill, A. (2013). Quantitative assessment of post-disaster housing recovery: A case study of punta gorda, Florida, after hurricane charley. *Disasters, 37*(2), 333–355.

Řehůřek, R. (2019). Gensim topic modelling for humans: Models.*ldamallet – latent dirichlet Allocation via mallet* [online]. Available from: https://radimrehurek.com/gensim/models/wrappers/ldamallet.html Accessed 24th July 2019.