Development of Beach Quality Assessment Parameters Based on Tourist Reviews

Myunghwa Kang
Hospitality, Restaurant, and Tourism Management, University of Nebraska - Lincoln

Hanyoung Go
Hospitality, Restaurant, and Tourism Management, University of Nebraska - Lincoln

Development of Beach Quality Assessment Parameters Based on Tourist Reviews

Introduction

Beaches are popular natural attractions for holiday vacations and recreational activities among tourists (Choudri, Baawain, Al-Sidairi, Al-Nadabi, & Al-Zeidi, 2016; Holden, 2000). The coastal environment influences human wellbeing in multiple ways, such as recreational, natural, and cultural services (Costanza et al.,1997; Lozoya, Sardá, & Jiménez, 2014; Roca, Villares, & Ortego, 2009). Tourism communities have been considered beaches as important tourism resources that generate valuable economic value (Botero, Pereira, Tosic, & Manjarrez, 2015; Houston, 2002). The ongoing increase in beach tourism demand negatively impacts the coastal beach environment and tourist quality experiences (Roca et al., 2009), especially at and around well-known beaches. Beach management agencies strive to maintain the quality of recreational resources, facilities, and services to increase the attractiveness of beaches, achieving sustainable management. This effort has brought about beach assessment tools and systems that consist of multiple parameters of recreational and environmental quality, based on beach visitors' perceptions and expectations, and the description of beach's natural characteristics, including physical, chemical and biological (Rocca, Villares, Fragell, & Junyent, 2008; Botero et al., 2015).

Massive amounts of tourist data, such as online reviews and consumer behavioral data on social media, have vastly grown and become a valuable source of information to understand tourists and support the environmental and marketing decisions of tourism destinations (Hajli & Laroche, 2019; Tussyadiah, & Zach, 2017; Smyth, Wu, & Greene, 2010). To date, measurement parameters of beach quality perception, as expressed by actual or potential beach visitors, have been developed mainly by relying on the survey questionnaires on physical, social, environmental, and service-related aspects (Rocca et al., 2008). Therefore, there is a critical need to define and assess the quality of tourist beaches based on visitors' experiences presented in online reviews on social media. The purpose of this study was to compare commonalities and variations of beach quality between assessment parameters from questionnaire surveys and those from tourist online review analysis on social media and, ultimately, to propose measurement parameters of beach quality and prioritizes based on beach visitors' perceptions and attitudes presented in reviews on social media.

Literature Review

Beach quality is determined by both the quality of the environment and tourist experience (Duvat, 2011). The assessment of the quality of beaches uses biological and environmental measurements with visitors' perceptions and expectations and includes beach visitors' perceptions and attitudes. Peña-Alonso, Ariza, Hernández-Calvento, and Pérez-Chacón (2018), for example, analyzed the recreational quality of urban, semi-urban, and natural beaches as a tourist experience measurement, as well as natural beaches, and developed a system of seven indicators of accessibility, environmental quality, water quality, comfort, scenic quality, human activity, and infrastructure. Roca et al. (2009) assessed public perception and attitude towards beach quality. Lucrezi, Saayman, and Van der Merwe (2016) developed an assessment tool for sandy beaches, integrating beach description, human dimensions, and economic factors to identify priority management issues. Phillips and House (2009) constructed a beach rating checklist comprising 50 physical, biological and human use factors, making weightings in response to priorities of three tourism markets: surfing, eco-tourism, and family travel. Chen and Bau (2016) developed a multi-criteria evaluation tool for tourist beaches in Taiwan for sustainable beach tourism management. Botero et al. (2014)

proposed recreational parameters as an assessment tool for beach quality, finding that the environmental quality of tourism beaches is a function of the beach performance as an ecosystem and a satisfier of human needs. Other scholars have established a scenic evaluation checklist system to rate coastal scenic quality (Botero, Anfuso, Williams, & Palacios 2013; Phillips, Edwards, & Williams, 2010).

These authors proposed a technical instrument, summarizing multiple criteria for analyzing environment parameters with a management tool and concluded that beach quality could be assessed by recreational parameters, which were scientifically designed based on beach visitors' perception. To date, the scholarly literature shows that measurement tools and systems of beach quality perception expressed by actual or potential beach visitors have developed mainly through the survey questionnaires on physical, social, environmental, and service-related aspects.

In the past decade, user-generated content (UGC) on social media (e.g., reviews on TripAdvisor) has provided rich and trustworthy information of visitors' own actual experiences with tourism attractions, and products and services (Alexander, Blank, & Hale, 2018; Cenni, & Goethals, 2017; Levy, Duan, & Boo, 2013; Valdivia, Luzón, & Herrera, 2017). UGC has increasingly been utilized as new sources of information for tourism destination research (Brunner, Palmer, Togher, Dann, & Hemsley, 2019; Chang, Ku, & Chen, 2019; Jiang & Mondschein, 2019; Leung, Law, Van Hoof, & Buhalis, 2013). A growing number of travelers share actual travel experiences during and after their trip on social media. As a result, massive amounts of subjective travelers' opinions, recommendations, and ratings, and consumer behavioral data on social media, called UGC, has enormously increased. This has become a valuable source of information to support the marketing activities of tourism destinations, but only if they are analyzed in meaningful ways (Hajli & Laroche, 2019; Tussyadiah, & Zach, 2017). Given the prevalent use of online reviews, travelers' actual experiences presented in online reviews on social media have been studied extensively (Smyth et al., 2010).

TripAdvisor has been considered as a leading website among a large number of travel-related social media (Fong, Lei, & Law, 2017; Yu, Li, & Jai, 2017). Emerging in 2004 for the tourism domain, TripAdvisor has been become the most popular site for trip planning among tourists (TripAdvisor, 2020; Valdivia et al., 2017). TripAdvisor is the world's largest travel platform, drawing nearly 460 million travelers each month in 2019 (TripAdvisor, 2020). In 2019, tourists across the globe used the TripAdvisor website and app to browse more than 830 million reviews and opinions of 8.6 million attractions, accommodations, restaurants, experiences, airlines, and cruises (TripAdvisor, 2020). TripAdvisor reviewers provide a short review with a title, and a ranking from one star (lowest) to five stars (highest) based on the reviewer's perception of the quality of the hotels, restaurants, and attractions (Xie, Chen, & Wu, 2016), as well as metadata such as their origin, trip purpose and type, and even like votes. TripAdvisor reviews, as short stories about visitors' experiences (Vasquez, 2012), constitute narrative appraisals of tourism attraction sites, that is, visitor-authored stories about places (Alexander, Blank, & Hale, 2018). Thus, TripAdvisor has been used in a wide variety of studies on tourism attractions and tourist perception (Cormier-MacBurnie, Mombourquette, Sneddon, & Young, 2018; Fang, Ye, Kucukusta, & Law, 2016).

Although the significance of TripAdvisor users' review data has been recognized as a valuable information resource, no study has used reviews on TripAdvisor to investigate and extract beach users' perceptions and experiences on beach quality and expectations on beach characteristics. Therefore, there is a need to develop a tool or system of beach quality assessment and measurement

based on actual visitors' experiences presented in reviews on TripAdvisor. This study investigates beach quality based on beach visitors' perceptions and attitudes presented in reviews of tourist beaches on TripAdvisor and then develops assessment parameters for beach quality and prioritizes based on those perceptions and attitudes.

Methodology

This study analyzes visitors' reviews of the most popular beaches worldwide to extract the criteria of beach quality assessment. The study involves three procedures. First, data are collected from the review texts of the top 25 beaches on TripAdvisor (Table 1). Second, the collected textual review data are pre-processed, applying tokenization and stop-words removal. Third, topic modeling is applied to extract topic words of beach quality, which represent themes of beach quality that will be used for developing a beach assessment tool.

Data Collection

TripAdvisor was chosen as the source of the review data. The reviewed textual data were collected from the reviews of the top 25 beaches posted by reviewers to TripAdvisor. The top 25 beaches were ranked by TripAdvisor in 2019 as the most popular beaches in the world. A total of 141,468 reviews of the top 25 beaches were extracted (Table 1). 5-star and 4-star reviews, considered positive reviews, constituted 74% and 20% of the reviews respectively, whereas 1-star and 2-star (negative) reviews, represented only 0.7% and 1.2%, respectively. 3-star ratings accounted for 5% of reviews. With the textual review data, metadata of reviews and its reviewers are collected – reviewers' star ratings, number of like votes, and other review postings; demographic/profile data of reviewers (e.g., origin), trip behavior (e.g., trip purpose, trip type, trip date). The metadata is publicly available on TripAdvisor. The reviewers' origins and beach location coordination were classified into spatial metadata and the posting date of reviews and trip dates of reviewers into temporal metadata.

Data Processing

The collected textual review data are pre-processed through the following (Tussyadiah & Zach; 2017; Alexander et al., 2018; Vinithra, Selvan, Kumar, & Soman, 2015): (1) tokenization, breaking a stream of texts into words, phrases, symbols, and other meaningful elements called tokens; (2) stop word removing, eliminating stop words (frequently occurring non-context-bearing, common words), such as definite or indefinite articles and auxiliary verbs (e.g., 'a', 'an', 'and', 'the', and 'or'); (3) stemming words, collapsing synonyms into a single word, such as 'entertaining', 'entertains', and other each variation into 'entertain'; (4) expanding contractions, changing 'can't' and 'didn't' into 'cannot' and 'did not'; and (5) transforming some phrases, all variations on N-years-old (e.g., 10 years old) to 'year_old', a constructed word.

Table 1. Top 25 Beaches and the Number of Star Ratings

Rank	Beach	Number of Reviews by Star Ranking					Total
		1	2	3	4	5	
1	Baia do Sancho	9	15	63	514	7,824	8,425
2	Varadero Beach	156	204	921	4,011	1,5239	20,531
3	Eagle Beach	29	65	365	1,685	7,619	9,763
4	La Concha Beach	28	38	314	2,245	6,290	8,915
5	Grace Bay Beach	15	23	122	612	6,288	7,060
6	Clearwater Beach	127	166	618	2,464	8,759	12,134
7	Spiaggia Dei Conigli	40	42	108	508	5,283	5,981
8	Seven Mile Beach	41	68	272	965	4,419	5,765
9	Playa Norte	56	92	402	1,432	5,820	7,802
10	Seven Mile Beach	64	134	501	1,617	5,505	7,821
11	Falesia Beach	12	26	273	1,458	3,678	5,447
12	Prainhas do Pontal do Atalaia	12	32	126	805	4,867	5,842
13	Playa de Ses Illetes	40	61	217	868	4,339	5,525
14	Ka'anapali Beach	30	86	462	1,965	6,543	9,086
15	Balos Lagoon	112	216	589	1,733	7,194	9,844
16	Radhanager Beach	7	9	143	798	3,026	3,983
17	Playa Manuel Antonio	208	372	1,216	2,999	6,937	11,732
18	Manly Beach	36	64	509	2,352	4,582	7,543
19	Kelingking Beach	12	15	58	226	1,066	1,377
20	Bournemouth Beach	69	112	397	2,206	5,695	8,479
21	Elafonissi Beach	179	261	989	2,848	9,628	13,905
22	Fig Tree Bay	31	77	382	1,016	3,304	4,810
23	Surfers Paradise Beach	50	67	361	1,647	3,679	5,804
24	Anse Lazio	6	30	183	832	2,997	4,048
25	Elafonissi Beach	5	4	27	116	887	1,039
Total		1,374	2,279	9,618	37,922	141,468	192,661
Percent		0.7%	1.2%	5.0%	19.7%	73.4%	100%

Data Analysis

Since the collected textual data is massive, it is not feasible to use qualitative content analysis that requires manual-hand coding of beach quality criteria in each text of the collected reviews. Therefore, the present study uses a topic modeling approach, using a computer-assisted qualitative data analysis software (QDAS), NVivo 12 Plus. Topic modeling is a statistical technique designed to discover the main topics in a large text document, where the volume of text is too massive to be analyzed by manual categorization (Blei, Ng, & Jordan, 2003). The topic modeling technique presents groups of words that commonly appear together. QDAS automatically groups meaning-rich words into topics. In this study, two tourism experts validate the number of beach quality themes by assessing how appropriately the topics are clustered based on a group of words (Alexander et al., 2018). The two experts decide on the number of topics using an interactive approach in which they seek topics that are internally consistent, sufficiently specific, and distinct from other topics. Topic modeling is, therefore, an exploratory technique useful for generating ideas. This combination of computational methods and human interpretation allows us to analyze a massive data by taking advantage of two complementary strengths. By combining these strengths,

this study can extract meaningful information from large bodies of review texts that comprise our data.

Expected Results and Implications

We expect that our findings have implications for both beach visitors and beach management. Past studies using a questionnaire-based survey within the context of beach quality have been limited to specific geographical locations, visitor groups, and time periods. The present study extends the research on beach quality based on the reviews posted on TripAdvisor and contributes to the literature on beach quality by describing new methods that present massive numbers of actual tourists' experiences and stories. The present study will provide a new opportunity to compare the commonalities and variations of different geographic regions and time periods with regard to beach quality and its assessment based on social media reviews. The future competitiveness of destinations will be based on the extent to which they are concerned about the sustainability of their natural, economic, and cultural resources (Laws, 1995). This study will help to increase the sustainable management of beaches, thereby positively influencing human wellbeing.

References

Alexander, V. D., Blank, G., & Hale, S. A. (2018). "Digital traces of distinction? Popular orientation and user-engagement with status hierarchies in TripAdvisor reviews of cultural organizations." New Media & Society, 20(11): 4218-4236.

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

Botero, C., Anfuso, G., Williams, A. T., & Palacios, A. (2013). "Perception of coastal scenery along the Caribbean littoral of Colombia. Journal of Coastal Research", 65(sp2): 1733-1739.

Botero, C. M., Pereira, C., Anfuso, G., Cervantes, O., Williams, A. T., Pranzini, E., & Silva, C. P. (2014). "Recreational parameters as an assessment tool for beach quality." Journal of Coastal Research, 70(sp1): 556-563.

Botero, C., Pereira, C., Tosic, M., & Manjarrez, G. (2015). "Design of an index for monitoring the environmental quality of tourist beaches from a holistic approach." Ocean & Coastal Management, 108: 65-73.

Brunner, M., Palmer, S., Togher, L., Dann, S., & Hemsley, B. (2019, January). Content analysis of tweets by people with Traumatic Brain Injury (TBI): implications for rehabilitation and social media goals. In Proceedings of the 52nd Hawaii International Conference on System Sciences.

Cenni, I., & Goethals, P. (2017). "Negative hotel reviews on TripAdvisor: A cross-linguistic analysis." Discourse, Context & Media, 16: 22-30.

Chang, Y. C., Ku, C. H., & Chen, C. H. (2019). "Social media analytics: Extracting and visualizing Hilton hotel ratings and reviews from TripAdvisor." International Journal of Information Management, 48: 263-279.

Chen, C. L., & Bau, Y. P. (2016). "Establishing a multi-criteria evaluation structure for tourist beaches in Taiwan: A foundation for sustainable beach tourism." Ocean & Coastal Management, 121: 88-96.

Choudri, B. S., Baawain, M., Al-Sidairi, A., Al-Nadabi, H., & Al-Zeidi, K. (2016). "A study of beach use and perceptions of people towards better management in Oman." Indian Journal of Geo-Marine Sciences, 45(10): 1327-1333.

Cormier-MacBurnie, P., Mombourquette, P., Sneddon, G., & Young, J. (2018). "The B&B sector in Nova Scotia: Some preliminary evidence from Tripadvisor reviews." Small Business Institute Journal, 14(2): 41-60.

Costanza, R., d'Arge, R., De Groot, R., Farber, S., Grasso, M., Hannon, B., & Raskin, R. G. (1997). "The value of the world's ecosystem services and natural capital." Nature, 387(6630): 253.

Duvat, V. (2011). "Interest of quality-based policies for Integrated Coastal Zone Management implementation: Lessons learnt from a French case study." Ocean & Coastal Management, 54(11): 831-843.

Fang, B., Ye, Q., Kucukusta, D., & Law, R. (2016). "Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics." Tourism Management, 52: 498-506.

Fong, L. H. N., Lei, S. S. I., & Law, R. (2017). "Asymmetry of hotel ratings on TripAdvisor: Evidence from single-versus dual-valence reviews." Journal of Hospitality Marketing & Management, 26(1): 67-82.

Hajli, N., & Laroche, M. (2019). "Applications of business intelligence and analytics in social media marketing." International Journal of Information Management, 48: 226-227.

Holden, A. (2000). Environment and Tourism. Poutledge. London, UK.

Houston, J. R. (2008). "The economic value of beaches: A 2008 update." Shore and beach, 76(3): 22-26.

Jiang, Z., & Mondschein, A. (2019). "Examining the effects of proximity to rail transit on travel to non-work destinations: Evidence from Yelp data for cities in North America and Europe." Journal of Transport and Land Use, 12(1).

Leung, D., Law, R., Van Hoof, H., & Buhalis, D. (2013). "Social media in tourism and hospitality: A literature review." Journal of Travel & Tourism Marketing, 30(1-2): 3-22.

Levy, S. E., Duan, W., & Boo, S. (2013). "An analysis of one-star online reviews and responses in the Washington, DC, lodging market." Cornell Hospitality Quarterly, 54(1): 49-63.

Lozoya, J. P., Sardá, R., & Jiménez, J. A. (2014). "Users expectations and the need for differential beach management frameworks along the Costa Brava: Urban vs. natural protected beaches" Land Use Policy, 38: 397-414.

Lucrezi, S., Saayman, M., & Van der Merwe, P. (2016). "An assessment tool for sandy beaches: A case study for integrating beach description, human dimension, and economic factors to identify priority management issues." Ocean & Coastal Management, 121: 1-22.

Phillips, M. R., & House, C. (2009). "An evaluation of priorities for beach tourism: Case studies from South Wales, UK." Tourism Management, 30(2): 176-183.

Phillips, M. R., Edwards, A. M., & Williams, A. T. (2010). “An incremental scenic assessment of the Glamorgan Heritage Coast, UK.” Geographical Journal, 176(4): 291-303.

Peña-Alonso, C., Ariza, E., Hernández-Calvento, L., & Pérez-Chacón, E. (2018). “Exploring multi-dimensional recreational quality of beach socio-ecological systems in the Canary Islands (Spain).” Tourism Management, 64: 303-313.

Roca, E., Villares, M., Ortego, M.I. (2009). “Assessing public perceptions on beach quality according to beach users’ profile: A case study in the Costa Brava (Spain).” Tourism Management, 30: 598-607.

Roca, E., Riera, C., Villares, M., Fragell, R., & Junyent, R. (2008). “A combined assessment of beach occupancy and public perceptions of beach quality: a case study in the Costa Brava, Spain.” Ocean & Coastal Management, 51(12), 839-846.

Smyth, P. C. B., Wu, G., & Greene, D. (2010). Does TripAdvisor make hotels better. Derek Greene School of Computer Science & Informatics, University College Dublin Belfield.

TripAdvisor (2020, January 6). About TripAdvisor. Retrieved from https://tripadvisor.mediaroom.com/us-about-us

Tussyadiah, I. P., & Zach, F. (2017). “Identifying salient attributes of peer-to-peer accommodation experience.” Journal of Travel & Tourism Marketing, 34(5): 636-652.

Valdivia, A., Luzón, M. V., & Herrera, F. (2017). “Sentiment analysis in Tripadvisor.” IEEE Intelligent Systems, 32(4): 72-77.

Vásquez, C. (2012). “Narrativity and involvement in online consumer reviews: The case of TripAdvisor.” Narrative Inquiry, 22(1): 105-121.

Vinithra, S. N., Selvan, S. A., Kumar, M. A., & Soman, K. P. (2015). “Simulated and self-sustained classification of Twitter data based on its sentiment.” Indian Journal of Science and Technology, 8(24), 1.

Xie, K. L., Chen, C., & Wu, S. (2016). “Online consumer review factors affecting offline hotel popularity: evidence from TripAdvisor.” Journal of Travel & Tourism Marketing, 33(2): 211-223.

Yu, Y., Li, X., & Jai, T. M. (2017). “The impact of green experience on customer satisfaction: Evidence from TripAdvisor.” International Journal of Contemporary Hospitality Management, 29(5): 1340-1361.