Ecosystem Services 76 (2025) 101777

Contents lists available at ScienceDirect

Ecosystem Services

journal homepage: www.elsevier.com/locate/ecoser

Full Length Article

Quantifying effects of demographic biases on estimation of cultural ecosystem services using social media in Japan

Michio Oguro [a,1,2,*], Rei Shibata [b,1,3,*]

[a] Forestry and Forest Products Research Institute, Matsunosato 1, Tsukuba, Ibaraki 305-8687, Japan
[b] Faculty of Agriculture, Niigata University, 8050 Ikarashi 2-no-cho, Nishi-Ku, Niigata, Niigata 950-2181, Japan

ARTICLE INFO

Keywords:
Cultural ecosystem services
Social media
Travel cost
Sentiment analysis
National park

ABSTRACT

Quantifying cultural ecosystem services (CESs) has been difficult using traditional methods such as questionnaires, so increasing numbers of studies are utilizing big data obtained from social media for evaluation of CESs. Although data obtained from social media are often considered to be biased and non-representative, the actual effects of biases on evaluations of CESs are rarely quantified. In this study, we sampled posts from a microblogging service, Twitter, and investigated the effects of demographic biases on three indicators capturing aspects of social media which can be used for evaluations of CESs (spatial distribution, sentiment of post, and travel cost), focusing on the Japanese national parks as a case study. We found that Twitter users obtained from the randomly sampled 1% posts had a different distribution of demographic attributes (age, sex, and residential prefecture) from the Japanese population. More importantly, we found that activities on the platform such as estimated frequency of georeferencing and posting are different among demographic groups. This indicates that spatial distribution of CESs commonly evaluated by the number of georeferenced posts can be affected by additional demographic biases which cannot be captured by statistics based on all users on social media platforms. On the other hand, among the three indicators, only spatial distribution was strongly affected by the demographic attributes of users, indicating that the potential effects of biases are different for each indicator tested. These results clearly show that consideration of demographic attributes and their potential effects would be necessary to obtain representative and inclusive evaluations of CESs.

1. Introduction

Cultural ecosystem services (CESs) are the ‘nonmaterial benefits people obtain from ecosystems through spiritual enrichment, cognitive development, reflection, recreation, and aesthetic experience, including, e.g., knowledge systems, social relations, and aesthetic values’ (Millennium Ecosystem Assessment, 2005, p. 894). CESs are important in a wide range of settings (Milcu et al., 2013) and play several essential roles in human-ecosystem interactions, contributing to human well-being (Nowak-Olejnik et al., 2022; Plieninger et al., 2015) and willingness to conserve nature (Soga et al., 2016). In developed countries, CESs are sometimes valued more than other kinds of ecosystem services (Plieninger et al., 2015; Quétier et al., 2010) and demand for CESs is increasing (Guo et al., 2010). However, because of their intangible nature, CESs are less likely to be incorporated in ecosystem management or decision making than other ecosystem services (Daniel et al., 2012; Milcu et al., 2013; Plieninger et al., 2015; Raymond et al., 2013).

On the other hand, the recent development of smartphones with global positioning system (GPS) capability, mobile networks, and user generated content on social media platforms enable researchers to quantify cultural use of ecosystems at larger scales than traditional survey methods such as interviews and questionnaires (Ghermandi et al., 2023; Havinga et al., 2020; Langemeyer et al., 2023; Wilkins et al., 2020; Zhang et al., 2020). Using social media data, researchers have evaluated and/or mapped many kinds of CESs including recreation and tourism (Barros et al., 2019; Hausmann et al., 2018; Orsi and Geneletti, 2013; Ruiz-Frau et al., 2020), aesthetic values (van Zanten et al., 2016; Yoshimura and Hiura, 2017), and nature experiences (Wartmann et al.,

* Corresponding authors.
E-mail addresses: oguro_michio630@ffpri.go.jp, oguro.michio@gmail.com (M. Oguro), rshibata@agr.niigata-u.ac.jp, ray.shibata@gmail.com (R. Shibata).
[1] Contributed equally.
[2] Michio Oguro ORCID: https://orcid.org/0000-0003-0974-7971
[3] Rei Shibata ORCID: https://orcid.org/0000-0002-6251-3237

https://doi.org/10.1016/j.ecoser.2025.101777
Received 27 July 2024; Received in revised form 22 August 2025; Accepted 11 September 2025
Available online 24 September 2025

2019). In other cases, studies evaluated sentiment of national park visitors (Hausmann et al., 2020) or conducted travel cost analyses (Ghermandi, 2018; Keeler et al., 2015; Sinclair et al., 2020, 2018) using social media. By combining social media and other sources of information, studies also investigated factors explaining spatial distribution of CESs (Aiba et al., 2023, 2019; van Zanten et al., 2016; reviewed in Zhang et al., 2020).

While social media data have many potential applications for understanding cultural relationships between humans and nature, the representativeness of evaluations of CESs using social media is affected by multiple biases caused by the nature of social media platforms (Ghermandi et al., 2023; Ruths and Pfeffer, 2014; Wilkins et al., 2020). For example, activities of social media users are not equal among users. Therefore, the results of analyses could be affected by a small number of very active users (Hollenstein and Purves, 2010). It is also known that social media users are not representative of target populations and often have skewed distributions of demographic attributes such as sex (Wang et al., 2022), age (Sloan et al., 2015; Wang et al., 2022), social occupation (Sloan et al., 2015), or race/ethnicity (Chen et al., 2022). Although it is rarely evaluated, users' behavior on social media platforms can be different among demographic groups. For example, the frequency of posting can be different among demographic groups (Cui and He, 2021). Also, spatial distributions of CESs are commonly evaluated using geolocations (e.g., Zhang et al., 2020), which are sometimes manually added by users. If the frequency of posting georeferenced posts differs among demographic groups, it produces additional bias which cannot be captured by the demographic information of users on the platform (but see Sloan and Morgan, 2015). To achieve inclusive environmental management, quantifying biases and their effects and acquiring representativeness for the focal population is crucial (Ghermandi et al., 2023). However, although the problem of highly active users is commonly remedied by counting only one photo/user/day (Wood et al., 2013), direct evaluation of the effects of other biases is rarely done because socio-demographic information is rarely described by users (Ruiz-Frau et al., 2020; Sinclair et al., 2020; but see Venter et al., 2023). Therefore, to achieve inclusive management of CESs, it is crucial to know the biases in the dataset and estimate their effects.

In this work, we estimated demographic attributes of social media users and quantified biases and their effects on indicators which capture aspects of social media data with potential importance for estimations of CESs using georeferenced Twitter posts in the 34 national parks in Japan (Fig. S1) as a case study. We focused on three indicators used in previous studies for evaluation of CESs: 1) spatial distribution of numbers of posts, 2) travel cost, and 3) sentiment. The spatial distribution of numbers of posts is the most used indicator for evaluations of CESs. Many researchers mapped the number of posts over their research areas and considered them as a proxy for landscape values (e.g., aesthetic and recreational values) or used them for usage monitoring of protected areas and national parks (reviewed in Havinga et al., 2020). Another indicator used for evaluations of CESs is the travel cost. Using georeferenced posts on social media and estimated locations of users' residences, studies calculated traveling costs from the residences to the ecosystems and used them to derive estimates of the value of nature-based recreational experiences (Keeler et al., 2015) and estimated economic values of ecosystems (Ghermandi, 2018; Sinclair et al., 2020, 2018). Although it is not as common as the other indicators, sentiment of posts has also been used for evaluation of CESs. For example, Fox et al. (2021a) used estimated sentiment of texts attached to images on Flickr to filter relevant images for the evaluation of CESs. Also, Hausmann et al. (2020) used it to understand sentiment of national park visitors from social media data. Using these indicators, we asked the following specific questions:

1. What demographic biases do Twitter users have?
2. Do the demographic biases differ between all posts and georeferenced posts?
3. What effects do demographic biases have on the estimation of the three indicators?

We found unrepresentativeness of the spatial distribution of number of posts (see results), so we explored a way to remedy this problem.

2. Materials and methods

2.1. Obtaining Twitter data

Studies evaluating CESs commonly use the photo sharing service Flickr as their data source (e.g., Ghermandi, 2018; Hausmann et al., 2018; Sinclair et al., 2020), but in Japan people tend to share their images on Instagram. However, because of multiple changes in the application programming interface (API), the availability of georeferenced posts from Instagram is now restricted (Ghermandi et al., 2023). On the other hand, 46.2 % of Japanese people were using Twitter in 2021 (Institute for Information and Communications Policy, Ministry of Internal Affairs and Communications, Japan, 2022). Also, georeferenced posts could be relatively easily obtained from Twitter using API (Ghermandi et al., 2023) at the time when the study was being conducted. Therefore, we focused on Twitter in this study.

Twitter (formerly https://twitter.com/; rebranded as X (https://x.com/) in 2023) is a micro-blogging service on which users can share short texts (called 'tweets'), images, and videos. Users can add various kinds of geolocation to their tweets including coordinates of a point (longitude and latitude) stored in the *coordinates* attribute and/or name and bounding box of a geographic entity such as a country, city, point of interest (POI), etc. stored in the *place* attribute. On average, about 0.06 % of tweets were georeferenced in 2019 (Kruspe et al., 2021). In this study we only used geolocations stored in the *coordinates* attribute and POIs stored in the *place* attribute as the sources of geolocation because other kinds of geolocations are too coarse to detect posts within the national parks and are not suitable for the evaluation of CESs. We only used tweets posted from the official Twitter and Foursquare (a social network service based on geolocation) clients because geolocations attached to a tweet posted from these clients were obtained from the device that the user posted the tweet from (Kruspe et al., 2021), which is suitable for estimation of travel costs. The filtering of clients was conducted using the *source* attribute of tweets. The list of values in the *source* attribute which were regarded as indicating official Twitter and Foursquare clients are shown in Table S1.

When estimating CESs provided by a given ecosystem using georeferenced posts on social media, data can have biases introduced at multiple steps. Therefore, quantification and adjustment of biases require consideration of these steps (Fig. 1, bottom part). First, because social media users are a sample of a population, they can have different distributions of demographic attributes compared to the focal population the data should represent (Fig. 1a). However, considering this bias is not sufficient because different demographic groups can have different proportions of users making georeferenced posts, which can produce additional biases (Fig. 1b). In addition, because users in different demographic groups can make different numbers of posts, this should also be considered when analyzing numbers of posts.

To consider these multiple steps, we constructed three datasets, namely 1) randomly sampled tweets, 2) georeferenced tweets, and 3) national park tweets to estimate demographic biases and their effects on the three indicators. An overview of the data retrieval and processing is shown in Figs. 1d and a detailed explanation of data retrieval is provided in Appendix 1. To find demographic biases in Japanese Twitter users, we collected randomly sampled tweets by Japanese users using the Streaming API of Twitter (statuses/sample endpoint). We obtained roughly 1 % of tweets in real time. The data download was conducted from January 1st, 2017 to December 31st, 2022 and the resulting tweets were used for subsequent analyses (Dataset 1, randomly sampled tweets).

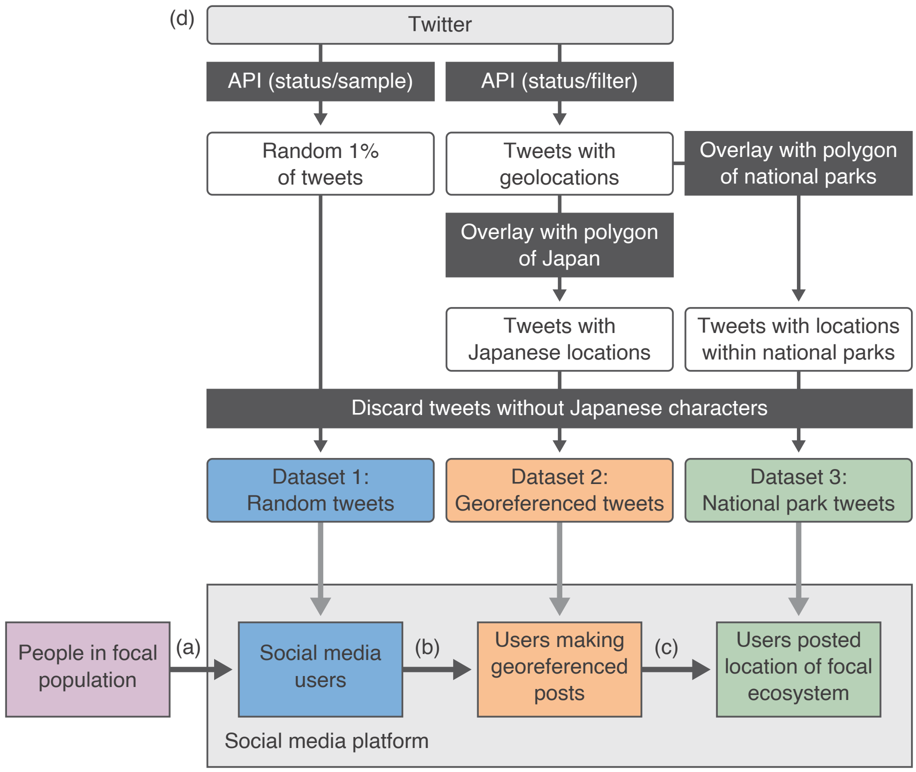

Fig. 1. Schematic flow of biases introduced during evaluation of CESs using georeferenced posts on social media and overview of the data retrieval and processing to capture them. Biases can be introduced by (a) differences between the focal population and social media users and (b) unequal frequency of georeferencing. Also, we can estimate the relative frequency of visiting the focal ecosystem by comparing users making georeferenced posts and users posting the location of the focal ecosystem (c). To investigate this flow, we conducted a series of data acquisitions and processing (d).

To find demographic biases in users who posted georeferenced tweets, tweets with location information in Japan were obtained by specifying the bounding box of Japan (122.9336, 20.4253, 153.9864, 45.5572) for the Streaming API (statuses/filter endpoint). The data collection was conducted from March 19th, 2017 to December 31st, 2022. After obtaining georeferenced tweets, tweets with location information outside of the terrestrial area of Japan defined by the standard 10 km grid (the National Standard Area 2nd Mesh, https://www.geospatial.jp/ckan/dataset/biodic-mesh, accessed on July 13th, 2023) were discarded because the specified bounding box contains other countries and the API occasionally returned tweets with locations outside of the bounding box. From the resultant tweets, we selected tweets having Japanese characters using the same method as for Dataset 1 and used them for further analyses (Dataset 2, georeferenced tweets).

To find demographic biases in users who posted tweets with locations within the 34 national parks and to evaluate their effects on the three indicators, we used a subset of the tweets in the collection of georeferenced tweets (Dataset 2). The national parks are located throughout Japan (Fig. S1) and offer a wide variety of CESs, attracting many tourists for driving, hiking, winter activities, marine recreation, and historical and cultural experiences. From the georeferenced tweets within the bounding box, we selected tweets having geolocations within the areas of the 34 national parks defined by a polygon provided by the Ministry of Environment, Japan (http://gis.biodic.go.jp/webgis/sc-026.html?kind=nps, accessed on July 13th, 2023). Note that because some areas of Japanese national parks contain parts of the ocean, which are not covered by the 10 km grid used for Dataset 2, this dataset contains a small number of tweets which are not contained in Dataset 2. After selecting the tweets within the national parks, we also selected tweets with Japanese characters and used them for further analyses (Dataset 3, national park tweets).

2.2. Estimation of demographic attributes of Twitter users

Although existing estimation methods for demographic attributes of Twitter users are available (e.g., Sloan et al., 2013; Sloan and Morgan, 2015; Yang et al., 2023), these methods could only be applied to English speaking users, who are uncommon in Japan. Therefore, we used a commercially available demographic attributes estimation service (NTT DATA Twitter User Profile DB; https://nttdata-nazuki.jp/data/value.html, accessed on July 13th, 2023). This company has been acquiring all public tweets from Twitter and the service estimates demographic attributes of the users by analyzing historical posts of users using Rich Indexer, a text mining engine extracting information from natural Japanese text, developed by NTT Service Innovation Laboratory Group (https://www.nttdata.com/jp/ja/trends/data-insight/2023/032790/, https://www.nttcoms.com/news/files/15/file.pdf, https://web.archive.org/web/20190201124417/https://www.ntt.co.jp/svlab/activity/category_2/product2_07.html, accessed on April 14th, 2025; in Japanese). Because we can obtain publicly available validation datasets (see below), we used age (<20, 20–29, 30–39, 40–49, 50–59, >=60), sex (male, female), and prefecture (one of the 47 prefectures of Japan) of residence in this study. Estimation of the demographic attributes was conducted for the three datasets noted above. For Dataset 1, we estimated 50,000 randomly selected users for each year because the total number of users in Dataset 1 was too many to analyze (See results). For Datasets 2 and 3, we analyzed all users in the datasets.

The quality of the estimation was verified by comparing the proportion of users in each demographic group with estimations from other existing statistics. For the statistics, we used the Survey on Information and Communication Media Usage Time and Information Behavior conducted by the Institute for Information and Communications Policy, Ministry of Internal Affairs and Communications, Japan (hereafter, MIC survey) and the Twitter Usage survey conducted by MyVoice

Communications, Inc. (hereafter, MyVoice survey). Using these statistics, we estimated the proportion of users in each demographic group on Twitter by estimating the number of Twitter users for each group by multiplying the proportion of users who answered to post tweets in each group by the Japanese population in each group obtained from the 2020 national census. The results were compared with the proportion of each demographic group calculated from the estimation results for Dataset 1 because the demographic attributes of the users in randomly sampled tweets would be comparable to the existing statistics. Congruency of the estimations for sex and age was checked using Cramér's V (Cramér, 1946), which measures associations between categorical variables in a contingency table and is commonly used as a measure of effect size for contingency tables. It ranges from 0 to 1 where 0 means no association (i.e., proportions of groups on rows are not different between estimation methods on columns) and 1 means strong association (i.e., proportions of groups on rows are different between estimation methods on columns). Values of V at 0.1, 0.3, and 0.5 can be considered to be small, medium, and large effects, respectively, for 2×2 tables (in our study, sex) and those of 0.05, 0.13, and 0.22 can be considered to be small, medium, and large effects, respectively, for 2×6 tables (in our study, age), according to Cohen (1988).

For the estimation of prefecture, congruency of the estimations was evaluated by Pearson's correlation coefficients between estimated proportions of users in each prefecture based on Twitter and other statistics. The detailed method for this verification is described in Appendix 2.

2.3. Estimation of travel costs

Travel costs to the locations indicated by the georeferenced tweets from users' residential areas were estimated using the Level of Transportation Services dataset provided by the Ministry of Land, Infrastructure, Transport and Tourism of Japan (https://www.mlit.go.jp/sogoseisaku/soukou/sogoseisaku_soukou_fr_000018.html; accessed on July 22nd, 2023) which contains distances, travel times, and travel costs by airplane, train, bus, and private car between all combinations of the 207 regions in Japan. To estimate travel costs, we obtained the dataset for 2015, which was the most recent at the time of the analysis, and calculated the minimum travel costs from prefectural government offices to the 207 regions by all available transportation modes. Using this dataset, we assigned the minimum travel cost from the prefectural government office of the user's residential prefecture, which was estimated by the method described above, to the location in a national park indicated by the geolocation associated with the tweet for each georeferenced tweet in Dataset 3 (national park tweets). Note that the travel cost was calculated as a one-way trip.

2.4. Estimation of sentiment

Sentiments of tweet texts in Dataset 3 (national park tweets) were estimated using a commercially available machine learning service, IBM Watson Natural Language Understanding (https://www.ibm.com/cloud/watson-natural-language-understanding; accessed on July 22nd, 2023). The service estimates the sentiment of a text and returns a sentiment score ranging from −1 (negative) to 1 (positive) as well as a sentiment label (negative/neutral/positive). It has relatively good performance for estimating the sentiment of a short Japanese text (Appendix 3). Before the analysis, we applied a text preprocessing suite and removed expressions which can disrupt the sentiment estimation (Appendix 4). Then using the service, a sentiment label and score were assigned to each tweet in Dataset 3.

2.5. Differences in travel costs and sentiment among demographic groups

To see the differences in estimated travel costs and sentiment among demographic groups, we constructed linear mixed models which investigate differences in these two indicators among demographic groups using the *lmer* function in the *lmerTest* package (Kuznetsova et al., 2017) with R 4.3.1 (R Core Team, 2023). For the response variables of the models, minimum travel cost and sentiment score were used. For the explanatory variables, sex and age were used for both models, but prefecture was included only for the sentiment model because travel cost was calculated from estimated prefecture, therefore it should have strong effects on the estimated travel cost. Because Twitter usage could be different between types of clients, type of client (Twitter official clients or Foursquare clients) was also included as an explanatory variable. Finally, to consider inter-annual differences in the travel costs and sentiment, year, which was also intended to capture effects of the COVID-19 pandemic, was also included. To reduce the dominance of massively tweeting users, only one tweet per day per one 1 km grid for each user (one tweet/user/day/grid) was randomly selected. User ID was used for the random effect of the models to control differences among the users. The significance of explanatory variables was tested by type II analysis of deviance using the *Anova* function in the *car* package (Fox and Weisberg, 2019). For these analyses, the unit of replication was a tweet and tweets for which we could not estimate all the demographic attributes used for each analysis were removed. After constructing the models, the estimated marginal mean of each response variable in each group was calculated using the *emmeans* function in the *emmeans* package (Lenth, 2023). Also, R^2 values of the models were calculated using the method of Nakagawa et al. (2017) implemented in the *r.squaredGLMM* function in the *MuMIn* package (Bartoń, 2023).

2.6. Differences in spatial distribution of CES among demographic groups

To detect differences in the location of posts among demographic groups, we checked representativeness of the spatial distribution of all Twitter posts, i.e., we tested whether the spatial distribution of all Twitter posts can represent that of each demographic group. For this analysis, we considered only age and sex as demographic groups because the residential prefectures of users should affect the spatial distribution of tweets. First, we calculated the number of tweets for each 1 km grid for all tweets, including tweets for which demographic attribute(s) could not be estimated and tweets posted by users in each demographic group using Dataset 3 (i.e., all tweets in the national parks). To reduce the dominance of massively tweeting users, only one tweet/user/day/grid was counted. Next, we calculated Pearson's correlation coefficients between the total number of tweets and the number of tweets in each demographic group. Because the spatial distributions of the number of tweets showed weak but significant spatial autocorrelations, we removed them by a spatial filtering approach using spatial lag variables before calculating the correlation coefficients (Appendix 5). We considered that the spatial distribution of all tweets could represent that of a demographic group if the correlation coefficient between all tweets and the tweets posted by the group was high.

The weak correlations between the total number of tweets and the number of tweets for each demographic group could be caused by differences among the groups in 1) locations where tweets are posted as well as 2) the number of tweets by each user. To separate the effects of these factors, we conducted two analyses. To test for differences in the location of posts, we created a pairwise correlation matrix for the number of tweets for each 1 km grid for each demographic group, while considering spatial autocorrelation using the method described above. To test for differences in the frequency of posts, a generalized linear mixed model (GLMM) with a Poisson error distribution was developed using the *glmmTMB* package (Brooks et al., 2017). For the response variable of the model, the number of tweets for a user was used. For the explanatory variables, age and sex as well as their interaction were used. For the random effect, user ID was used. The significance of explanatory variables and R^2 were calculated using the method described above. For this analysis, users whose gender or age could not be estimated were removed.

On the other hand, lower values of correlation coefficients could be due to a lower number of tweets for a demographic group. To test whether lower numbers of tweets can explain the observed values of correlation coefficients, we calculated the correlation coefficients between the number of tweets calculated using all tweets and subsamples of all tweets irrespective of demographic groups. For this purpose, we randomly selected specific numbers of tweets from Dataset 3 and calculated the correlation coefficients using the method described above. For the number of tweets in subsamples, we used 500, 1,000, 2,500, 5,000, 10,000, 25,000, 50,000, 100,000, 250,000, 500,000, and 1,000,000. After subsampling, we selected one tweet per user per 1 km grid in one day. Calculation of correlation coefficients was repeated 1,000 times for each number of subsamples. Note that when calculating correlation coefficients, spatial autocorrelation was considered as described above.

2.7. Exploring methods to correct biased spatial distribution

Because we found that the spatial distribution of the total number of tweets could not represent all demographic groups (see results), we explored methods to remedy the situation. To adjust the biases and obtain more representative estimations, users in underrepresented and overrepresented demographic groups should be assigned higher and lower weights, respectively (i.e., calibration; Lavallée and Beaumont (2015)). The weight of a user in a demographic group d can be calculated as follows:

$$User\ weight_d = \frac{P_{P,d}}{P_{S,d}} \times \frac{P_{S,d}}{P_{G,d}} = \frac{P_{P,d}}{P_{G,d}}$$

where $P_{P,d}$ denotes the proportion of people in group d in the focal population, $P_{S,d}$ denotes the proportion of users in group d on the platform, and $P_{G,d}$ denotes the proportion of users in group d who make georeferenced posts. For a group with $P_{G,d} < P_{P,d}$ (i.e., an underrepresented group), $\frac{P_{P,d}}{P_{G,d}}$ becomes > 1 and a higher weight is put on users in that group, and vice versa. Also, the weight of a post from a demographic group can be calculated as:

$$Post\ weight_d = \frac{1}{Mean\ number\ of\ posts_d}$$

By multiplying samples by these weights, we can obtain more representative estimations. For example, when a distribution of numbers of posts is analyzed, we can obtain a more representative distribution by multiplying the user weight and post weight by the numbers of posts.

First, we tried adjusting the user weight and post weight for each demographic group and checked their effects on the spatial distribution. For the adjustment, it is desirable to use the proportion of people visiting the national parks in each demographic group as the reference. However, because such statistics were not available in Japan, we used the Japanese population as the reference population in this study. Since people in different demographic groups can have different frequencies of visiting the national parks, the Japanese population was used as the reference for users in Dataset 2 rather than users in Dataset 3. This is because the ratio of the proportion of users in Dataset 3 to Dataset 2 (Fig. 1c) can be considered as the relative frequency of visitation to national parks and this frequency is incorporated into the estimation by adjusting the difference between Dataset 2 and the Japanese population. We first calculated the proportion of the Japanese population for each age group and sex (2 × 6 matrix) using population data obtained from the national census conducted in 2020 by the Statistics Bureau, Ministry of Internal Affairs and Communications (https://www.stat.go.jp/data/kokusei/2020/index.html, accessed on July 14th, 2023). Also, we calculated the proportions of the number of users who posted georeferenced tweets for each sex and age group using georeferenced tweets (Dataset 2). Using these proportions, the user weights were calculated for each demographic group (2 × 6 matrix). We also calculated the post weight by calculating 1/mean number of georeferenced tweets per user for each demographic group (2 × 6 matrix). By combining these two weights, we calculated three weights (adjustment of the user weight, post weight, and both). Then we calculated the number of tweets for each sex and age group for each 1 km grid while counting only one tweet/user/day/grid and multiplied them by the three weights noted above to obtain weighted spatial distributions. The differences between the spatial distributions were evaluated by calculating the Pearson's correlation coefficients considering spatial autocorrelation between the unweighted and weighted numbers of tweets for each grid. Second, because a small number of users posted a vast number of tweets even after adjusting for effects of massively tweeting users by using only one tweet/user/day/grid, we removed the top 1 % of users with the highest number of tweets from the dataset and repeated all the analyses used for the spatial distribution.

3. Results

3.1. Numbers of tweets and users

In total, we obtained 553,331,000 tweets posted by 70,215,513 users containing Japanese characters for randomly sampled tweets (Dataset 1). We also obtained 112,886,453 georeferenced tweets posted by 3,828,912 users. Of these, 34,206,689 tweets by 2,269,023 users were posted from the official Twitter and Foursquare clients, were located on the Japanese land area defined by the polygon of the 10 km grid covering the land area of Japan, and contained Japanese characters (Dataset 2). Finally, we obtained 2,013,166 tweets by 418,720 users which were located within the national parks. Of these, 793,721 tweets by 226,965 users were posted by the official Twitter and Foursquare clients and contained Japanese characters (Dataset 3). The numbers of tweets and users for each year and dataset are shown in Table S7.

3.2. Demographic attributes of Twitter users

The estimated proportions of each sex and prefecture based on Twitter data were congruent with the estimations based on existing statistics (Fig. 2; for the sex, Cramér's V were 0.002, 0.006, and 0.014 for MyVoice, MIC (Max), and MIC (Min), respectively, and Pearson's R was 0.952 for the prefecture) yet the proportion of users in Tokyo was slightly overestimated. On the other hand, estimation of age yielded proportions of older people that are lower than the proportions in the Japanese population (Fig. 2, Appendix 2) though small values of Cramér's V (0.053 for MIC (Max) and 0.063 for MIC (Min)) indicate the differences were not large taken as a whole. Results for each year are shown in Appendix 2.

Users' sex could be estimated for ca. 60 % of users and the proportion was higher for users who posted georeferenced and national park tweets (Fig. S7a). The estimated proportions of each sex for the randomly sampled tweets showed a skewed tendency toward males compared to the actual Japanese population (Fig. 3a). Also, the estimated proportions were different among the three datasets (Fig. 3a): users who posted georeferenced and national park tweets were more dominated by males than the randomly sampled users. Users' age could be estimated for ca. 40 % of randomly sampled users and the proportion was slightly higher for users who posted georeferenced tweets or national park tweets (Fig. S7b). The estimated proportions of each age group in randomly sampled tweets were much lower for elder users (50′s and more) compared to the actual Japanese population (Fig. 3b). Also, the estimated proportions were slightly different between the datasets (Fig. 3b). The estimated proportion of users younger than 30 were ca. 40 % for randomly sampled users whereas it was ca. 20 % for users who posted georeferenced tweets or national park tweets. User's residential prefecture could be estimated for only ca. 10 % of users for randomly sampled tweets whereas these proportions were ca. 30 % to 40 % for georeferenced tweets and national park tweets, respectively (Fig. S7c).

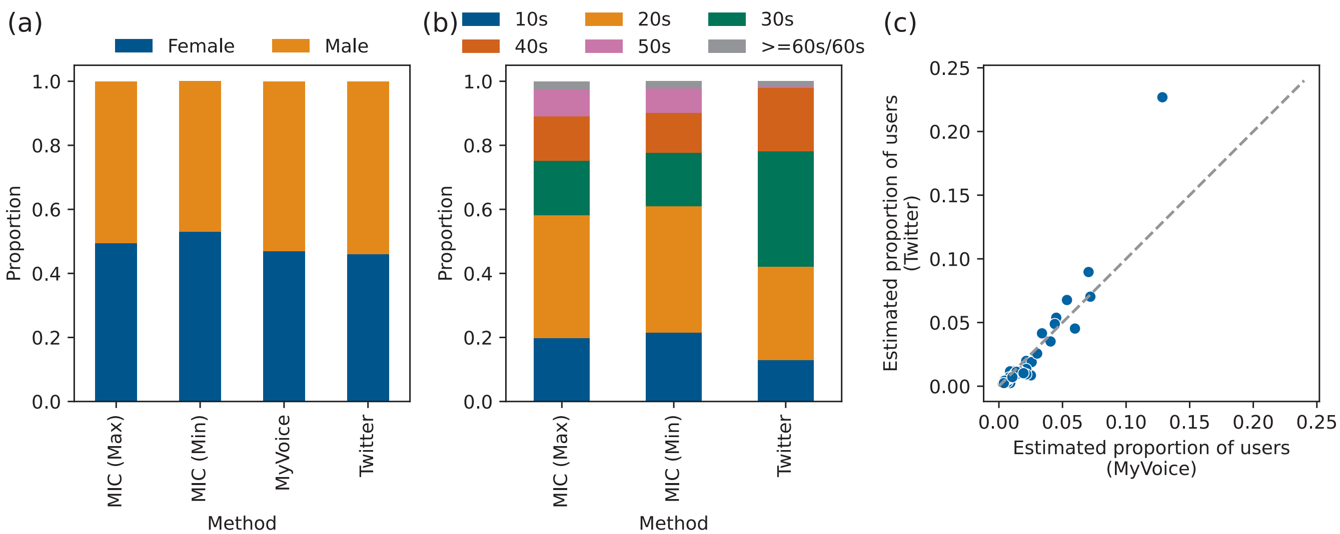

Fig. 2. Proportions of Twitter users in each (a) sex and (b) age group estimated from Twitter and the existing statistics and relationships between proportions of Twitter users in each prefecture estimated using Twitter and that using the existing statistics (c). For the statistics, we used the Survey on Information and Communication Media Usage Time and Information Behavior conducted by the Institute for Information and Communications Policy, Ministry of Internal Affairs and Communications, Japan (MIC) and the Twitter Usage survey conducted by MyVoice Communications, Inc. (MyVoice). Because the MIC survey separately asked respondents whether they post tweets from smartphones and/or computers, we calculated two estimates, assuming maximum or minimum frequency of posts for each demographic group. Note that the diagonal line shown in panel (c) represents y = x and the outlier near the center top represents Tokyo.

When aggregating users' residential prefectures into seven regions, the estimated proportions were not different among the three datasets but the estimated proportion of users in Kanto was higher than the actual Japanese population (Fig. 3c), maybe because estimations based on Twitter tend to yield higher proportions of users in Tokyo (Fig. 2c). The proportions of users in each prefecture are shown in Fig. S8.

3.3. Differences in the three indicators among the demographic groups

3.3.1. Travel cost

The results of the GLMM investigating differences in travel cost among age and sex groups showed significant effects of age and year terms (Table S8, Fig. 4ab, distributions of travel cost for each sex and age group are shown in Fig. S9). However, marginal R^2 (i.e., performance of explanatory variables) of the model was 0.001, indicating the explanatory variables have limited effects on the estimated travel cost.

3.3.2. Sentiment

The results of the GLMM testing differences in sentiment score among age, sex, and residential prefecture showed that all explanatory variables had significant effects on the sentiment score (Table S9, Fig. 4c-f, Fig. S10). However, similar to travel cost, the marginal R^2 of the model was 0.025, indicating the explanatory variables have limited effects on the sentiment score. A model estimating categorical sentiment classes (negative/neutral/positive) also could not predict sentiment using the same explanatory variables (Appendix 6).

3.3.3. Spatial distribution

The total number of tweets in each 1 km grid was strongly correlated with the numbers of tweets posted by male and female users (Fig. 5a). On the other hand, although the total number of tweets in each grid was strongly correlated with the numbers of tweets posted by users younger than 50 (Fig. 5c), it had weak correlation with the numbers of tweets posted by users aged 50 and over.

Further analyses showed that the different distributions among ages were caused both by different spatial distributions and different numbers of posts among ages (Fig. 6a, Fig. 7). The number of tweets posted by users younger than 50 showed weak correlations with those by users aged 50 and over (Fig. 6a), indicating different distributions between younger and elder users. In addition, the number of tweets posted by those in their 50′s showed weak correlation with those in their 60′s (Fig. 6a). Also, the GLMM found significant differences in number of georeferenced tweets per user among sex and age groups, yet the R^2 of the model was low (marginal $R^2 = 0.03$; Table S10). The estimated marginal means of the number of tweets showed that male users had significantly higher numbers of tweets than female users and elder users tend to post more tweets (Fig. 7).

The simulation changing the number of tweets in each subsample showed that the observed low values of correlation coefficients could not be explained solely by the reduced sample size for each demographic group (Figs. 8a, c). The actual correlation coefficients were lower than those calculated using subsamples with reduced numbers of tweets.

3.4. Exploring methods to correct biased spatial distributions

The adjustment by the user weights resulted in a different spatial distribution than the unweighted distribution of the number of tweets (Fig. 9a). However, the resultant spatial distribution seemed to be strongly affected by some users posting massive numbers of tweets, i.e., the adjustment brought many outliers. On the other hand, adjustment of the post weights did not have a large effect (Fig. 9b). When both weights were adjusted, the result was similar to the one with adjustment of the user weights only (Fig. 9c).

After removing the 1 % of users with the highest number of tweets, the spatial distribution of the total number of tweets per 1 km grid showed stronger correlations with those of each demographic group (Fig. 5). Also, the removal increased the correlations among the number of tweets per grid for each demographic group (Fig. 6), yet it had a limited effect on the number of tweets per user (Fig. 7, Table S11). Finally, the adjustment of the user weights did not bring outliers, and their effects became smaller (Fig. 9d-f), suggesting that the total number of tweets could represent all demographic groups well.

4. Discussion

We found that Japanese Twitter users obtained from the randomly sampled tweets were dominated by young (Fig. 3b) and slightly dominated by male (Fig. 3a) users. Moreover, we found that the estimated

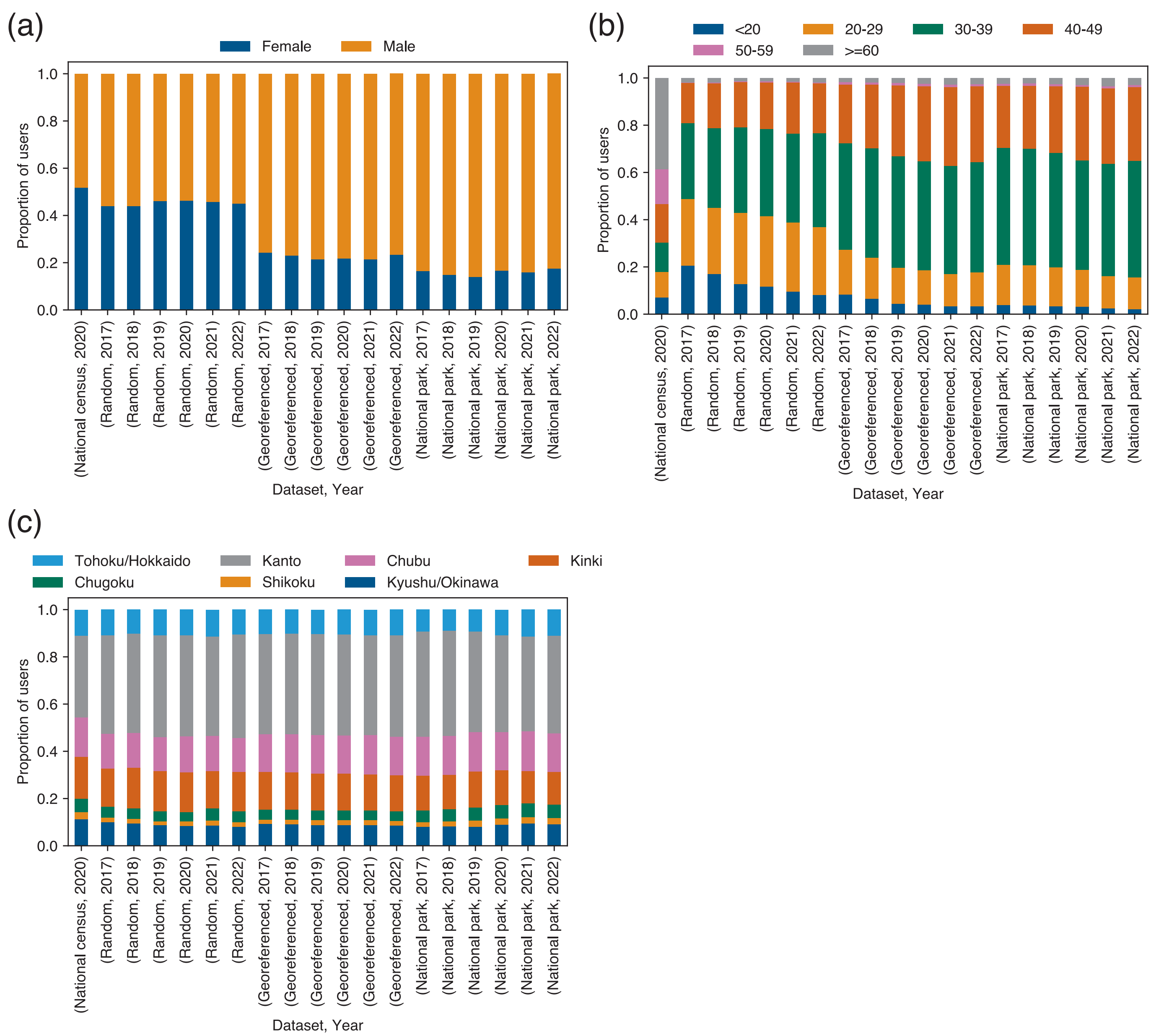

Fig. 3. Estimated proportion of (a) sex, (b) age, and (c) residential region of Twitter users for each dataset in each year. Actual proportions of the Japanese population obtained from the 2020 national census are also shown. Prefectures located in each region and the estimated proportions of users in each prefecture are shown in Fig. S8.

proportion of users posting georeferenced tweets were more dominated by male users (Fig. 3a) and the estimated numbers of posts per user were different among demographic groups (Fig. 7a), indicating that users' behavior on the social media platform can produce additional biases which cannot be represented by demographic biases of the users on the platform itself. On the other hand, although the estimated travel costs and sentiment of the tweets did not vary strongly among the demographic groups (Fig. 4, Appendix 6), users with different estimated age groups showed different spatial distributions of the number of tweets (Fig. 5) and this could not be explained by the small number of tweets for each group (Fig. 8). Therefore, the effects of demographic biases on evaluations of CESs using social media can vary among indicators used and it would be better to estimate them for each indicator. However, considering that spatial distributions of social media posts are frequently used for evaluation of spatial distribution of CESs (e.g., Barros et al., 2019; Hausmann et al., 2018; Orsi and Geneletti, 2013; van Zanten et al., 2016; Wartmann et al., 2019; Yoshimura and Hiura, 2017), incorporation and control of demographic attributes and their effects in analyses of CESs would be important for most cases.

Although adjustment of demographic biases would be a straightforward approach to control the biases, this approach produced many outliers (Fig. 9a-c) depending on the existence of a small number of massively tweeting users even after applying the commonly-used one post/user/day/grid limit. This result suggests that the commonly-used one photo/user/day method (Wood et al., 2013) did not have enough ability to control the effects of massively posting users (e.g. users who tweet every day) in our dataset. On the other hand, after removing the top 1 % of users with the highest number of tweets, the spatial distribution of the total number of tweets per 1 km grid had a strong correlation with the weighted spatial distribution after adjustment of demographic biases in the number of users and the number of tweets per user, which would be the best representative distribution (Fig. 9f). This would mean that controlling the effects of massively posting users is also important for the sake of reducing demographic biases in results,

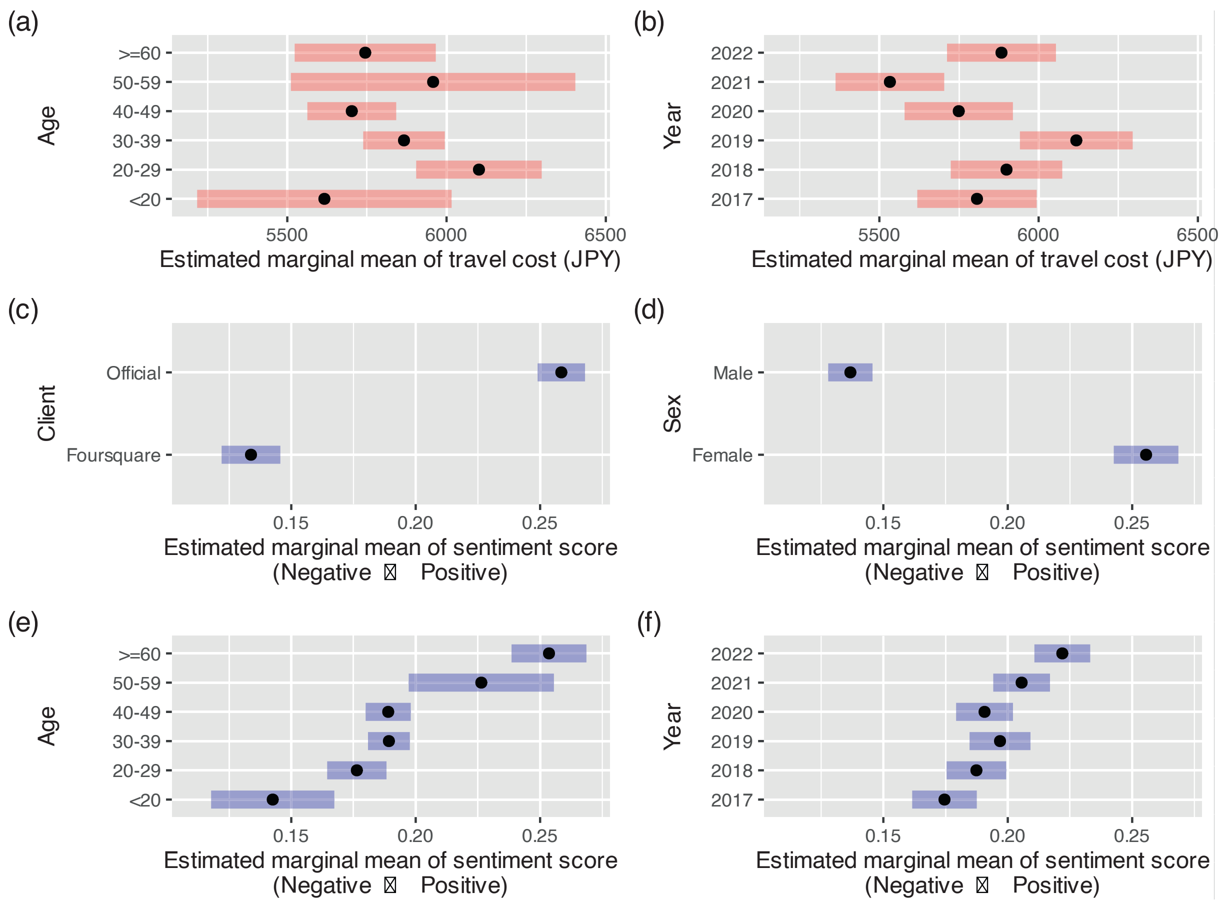

Fig. 4. Estimated marginal means of (a-b) travel cost and (c-f) sentiment score for (a, e) each age group, (b, f) year, (c) client, and (d) sex. Dots represent estimated marginal means and bars represent their 95% confidence intervals. Note that the travel cost was calculated as a one-way trip.

potentially caused by differences in posting behavior among demographic groups (Fig. 7).

4.1. Demographic biases in Twitter data

Quantifying demographic biases and their effects to acquire representativeness for the focal population is crucial for achieving inclusive environmental management (Ghermandi et al., 2023). Our analysis showed that the Twitter users obtained from the randomly sampled tweets were dominated by younger male users and the estimated bias from the actual Japanese population was stronger for age than for sex (Figs. 3a and b). These results were confirmed by the estimations using the existing statistics (Appendix 2, Fig. S6a and b). Previous studies also showed equivalent results for Twitter and other social media services. For user age, dominance of younger users was also observed in other studies on Twitter (Cui and He, 2021; Sloan et al., 2015; Yang et al., 2023) as well as other social media services (Venter et al., 2023). Also, other studies showed smaller sex biases than age biases for Twitter though male or female prevalence varied by study (Sloan et al., 2013; Yang et al., 2023).

More importantly, we found a stronger bias toward male users who tend to post georeferenced tweets in our dataset (Fig. 3a). This implies that evaluations of CESs using geolocation of Twitter posts are affected more by the behavior of male users and this overrepresentation cannot be captured by the demographic information of Twitter users. On the other hand, Sloan & Morgan (2015) demonstrated no clear difference in the frequency of posting georeferenced tweets between sexes for worldwide Twitter users. Although the difference between our study and Sloan & Morgan (2015) could be caused by the different methods for estimation of gender or the rapidly evolving internet environment, it is also possible that differences in the frequency of geolocation sharing between genders may vary across regions, countries, or cultures. Also, it is possible that demographic groups on other social media platforms add geolocation to their posts differently. Therefore, estimating demographic attributes of users and directly testing their effects on evaluations of CESs would improve results and contribute to more inclusive environmental management.

4.2. Effects of demographic attributes on estimations of CESs

Although we found demographic biases in the Twitter users and the biases were more prominent for users posting their geolocations (Fig. 3), we did not find strong effects of the biases on the estimated travel costs (Fig. 4 and S9, Table S8). On the other hand, the spatial distributions of the number of tweets within the national parks were different between younger and elder users (Fig. 5c). Even after removing the top 1 % of users who post frequently, elder people tend to post different geolocations than younger people (Fig. 6b). These results indicate that the patterns of movements from residential areas to the national parks, which are represented by the travel costs, were not affected by demographic attributes, but the patterns of movements within the national parks were different among the age groups. Although we do not have data to support understanding the reason behind this pattern, there are several possible explanations. One possible explanation would be a difference in physical fitness among age groups. It is possible that while travel to the national park from the residence by some form of transportation is not related to the user's physical fitness, travel within the parks might be dependent on it. Also, purposes and interests could be different between younger and elder visitors, and this could cause age-specific distributions of the number of tweets for each 1 km grid.

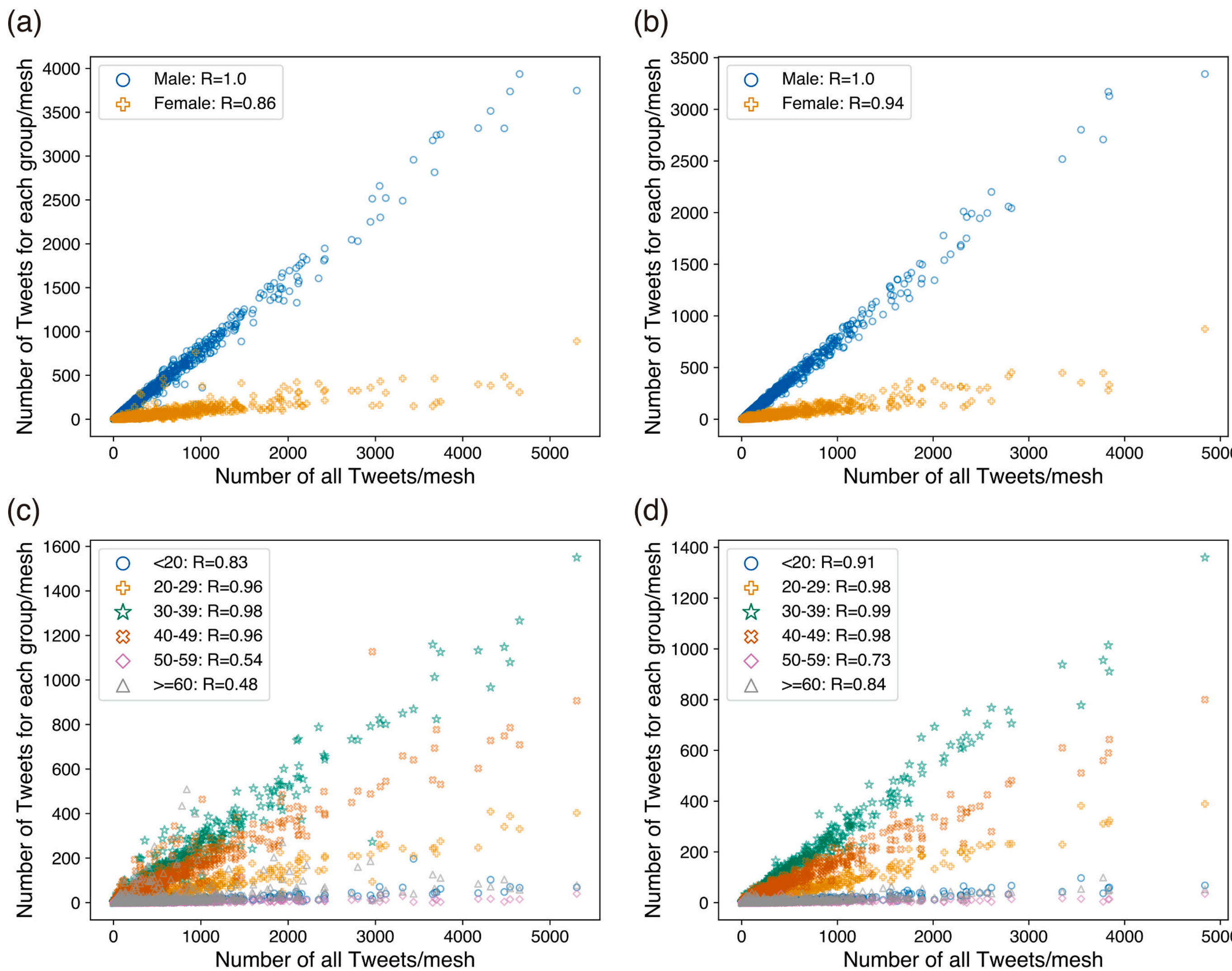

Fig. 5. Relationships between number of tweets in each 1 km grid covering all the national parks for all georeferenced tweets and number of georeferenced tweets posted by (a, b) each sex or (c, d) age group (a, c) with or (b, d) without the top 1 % of users with the highest number of tweets. The spatial autocorrelation-adjusted correlation coefficients between the number of total tweets and the number of tweets posted by each demographic group are shown in the legends.

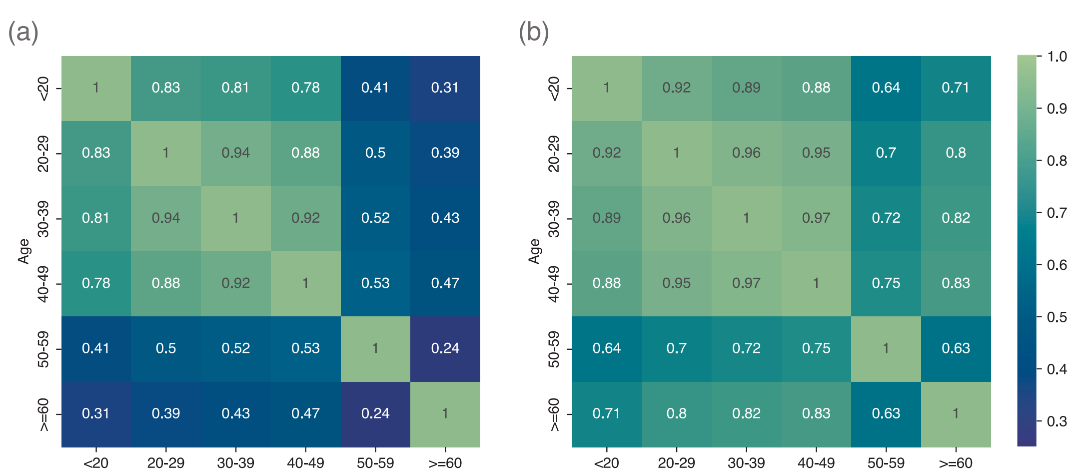

Fig. 6. Matrix of spatial autocorrelation-adjusted Pearson's correlation coefficients among the number of georeferenced tweets in each 1 km grid covering all the national parks for each age group, calculated for (a) with and (b) without the top 1 % of users with the highest number of tweets.

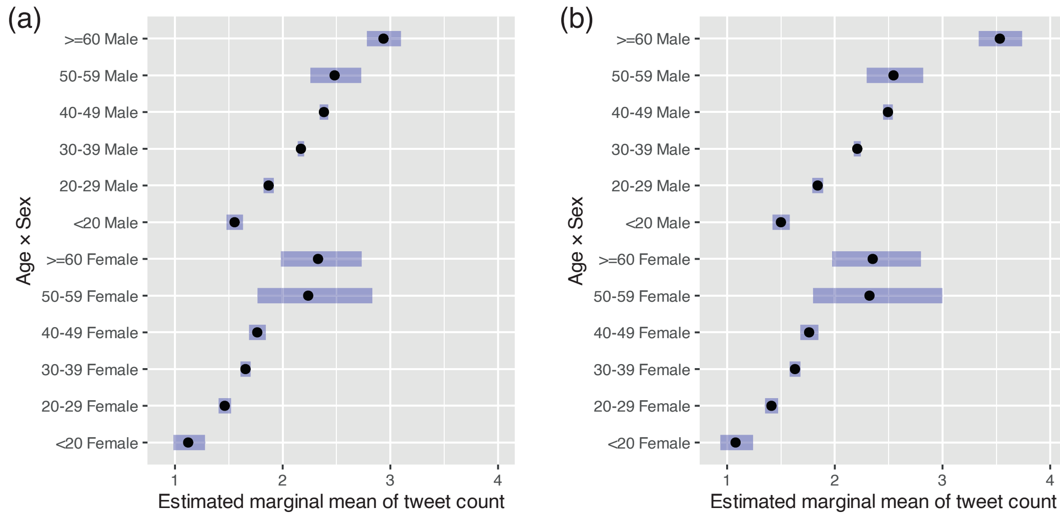

Fig. 7. Estimated marginal means of the number of georeferenced tweets per user for each age group and sex (a) with and (b) without the top 1% of users with the highest number of tweets. Dots represent estimated marginal means for each group and bars represent their 95% confidence intervals.

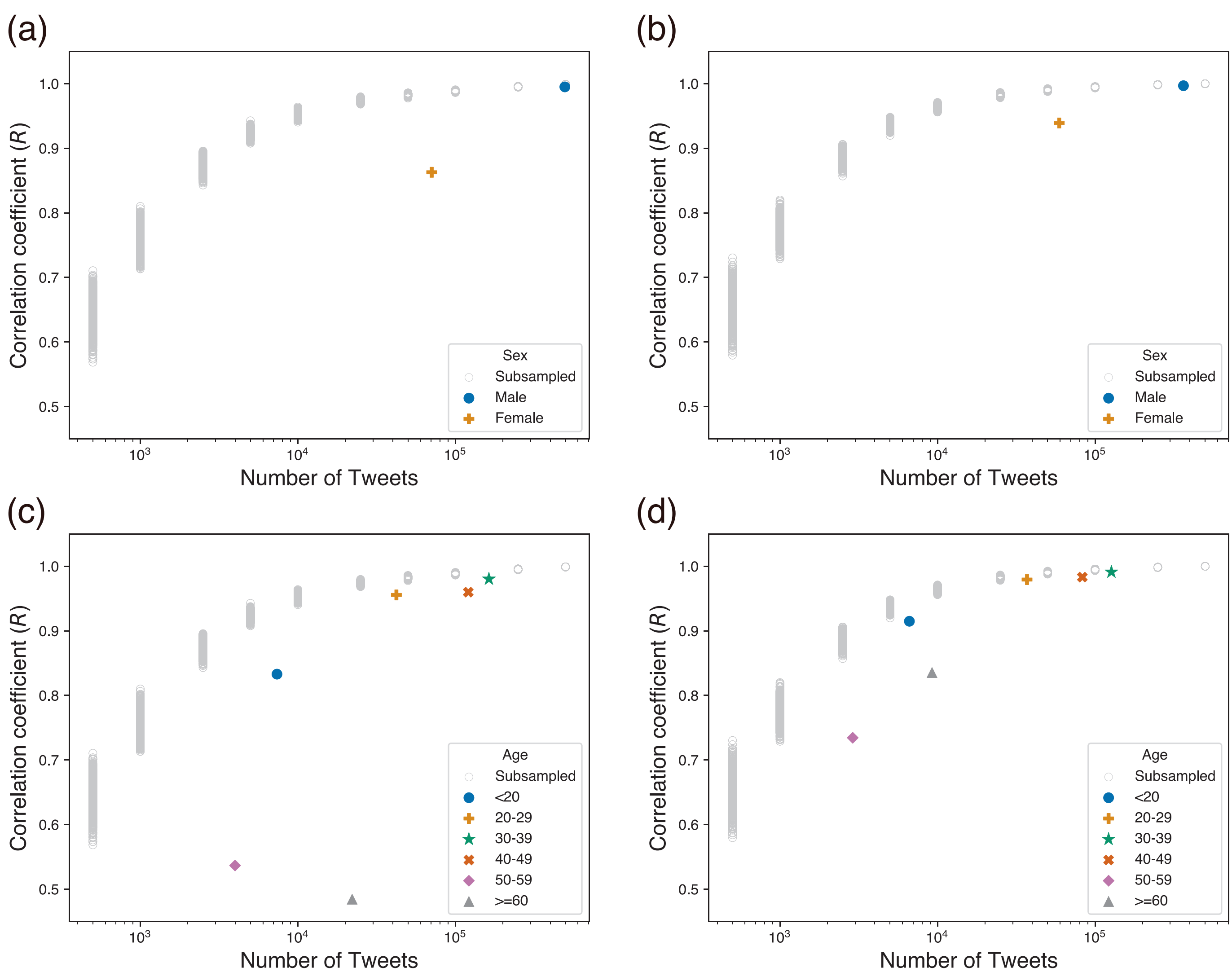

Fig. 8. Spatial autocorrelation-adjusted correlation coefficients between the numbers of georeferenced tweets in each 1 km grid covering all national parks for all tweets and subsampled tweets having reduced total numbers of tweets. The closed symbols represent the actual correlation coefficients for (a, b) each sex and (c, d) each age group for datasets (a, c) with and (b, d) without the top 1 % of users with the highest number of tweets.

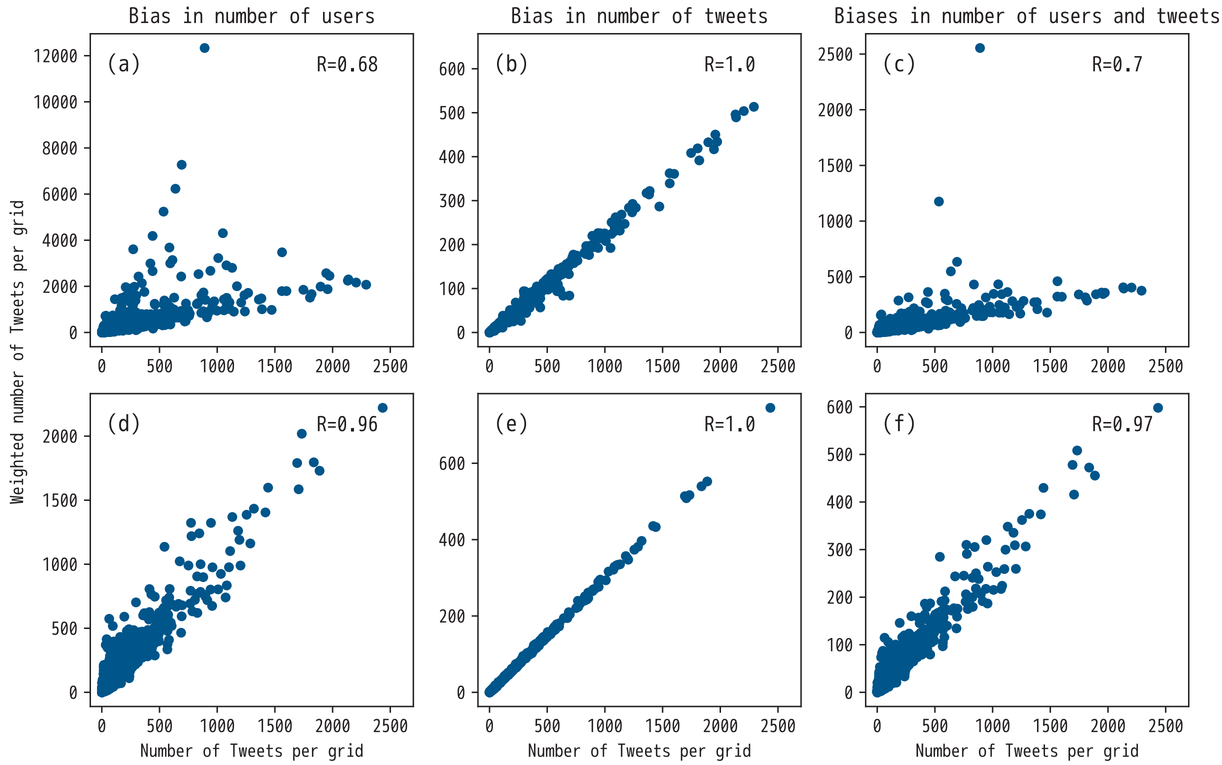

Fig. 9. Relationships between the total number of georeferenced tweets in each 1 km grid covering all the national parks and the weighted number of georeferenced tweets adjusted by (a, d) the user weights, (b, e) the post weights, and (c, f) both. The upper and lower rows represent results with and without the top 1 % of users with the highest number of tweets, respectively. Spatial autocorrelation-adjusted Pearson's correlation coefficients between the variables are also shown.

Irrespective of the underlying mechanisms, these results indicate that results of large-scale travel cost analyses using social media where the movements are largely done by transportation are not affected by demographic biases of users, but demographic attributes, in particular age of users, should be considered when analyzing small-scale travel costs or within-park distributions of CESs using social media services.

The conclusions above should be considered with some caution owing to the coarse estimation of travel cost in this study. Because the method we used only estimates residential prefecture of the users, within-prefecture movements around users' residences and/or movements to locations in national parks within the users' residential prefectures could not be estimated correctly. Therefore, if there were demographic biases in spatial distribution of residential locations within prefectures, our analysis could not capture their potential effects. Also, because we used a pre-calculated database of travel costs between the 207 regions for performance reasons, our analysis could not capture differences in small scale movement within the national parks among the demographic groups. On the other hand, previous studies established methods for estimating the location of a user's residence for another photography sharing service (Flickr) (Ghermandi, 2018; Keeler et al., 2015; Sinclair et al., 2020, 2018). Using a similar method would improve estimation of short-range movement and resolution of the problem.

Although elder females tended to post positive tweets (Fig. 4c-f), the explanatory power of the model was very low (marginal $R^2 = 0.022$). This was improved when considering the differences among users using the random effect of user ID (conditional $R^2 = 0.135$). This result may indicate that personal attributes such as personality have stronger effects on sentiment than demographic attributes but even the personal attributes cannot explain temporal fluctuation of users' sentiment. However, because the performance of the model for estimating sentiment was not very high (Appendix 3, Table S5), there is a possibility that the explanatory power of the demographic attributes could be improved if the estimation of sentiment is improved. Future work incorporating emerging large-scale language models or other means to improve the estimation of sentiment would also improve the analyses of the effects of demographic biases on the sentiment.

4.3. Toward inclusive evaluations of CESs

We found age-biased spatial distribution of the number of tweets within the national parks (Fig. 5c). Therefore, simple aggregation of spatial distribution of georeferenced tweets cannot provide a representative evaluation of CESs. As far as we know, no studies have shown such different spatial distributions of the number of posts, which was frequently used for evaluations of CESs using social media, for different demographic groups. This is problematic because decisions dependent on such results may not be inclusive since the results are biased toward specific demographic groups (younger generations, in our case). Resolution of the problem may depend on the purpose of the evaluation of CESs. If the purpose of an evaluation is simply to know the spatial distribution of CESs or to find culturally important regions, weighted aggregation of the number of posts by sampling weight (Lavallée and Beaumont, 2015) can improve the situation. However, removing the top 1 % of users had stronger effects than adjustment of demographic biases for our study while a weighting-only method caused another problem (Fig. 9). Therefore, suitable methods should be selected for each dataset used for evaluations of CESs. On the other hand, managers can implement measures specific for some age groups that frequently visit a location using estimated spatial distributions of CESs for each age group. For example, measures improving safety of visitors in parks would be more effective if such measures were implemented in locations where elderly people tend to visit. Whatever the purposes, knowing demographic attributes and incorporating them into analyses is the first step for the inclusive evaluation of CESs.

4.4. Limitations and future directions

Although social media data has many potential applications for studies of human-nature interactions, misuse of it can violate users' privacy and cause ethical issues (reviewed by Ghermandi et al., 2023). Therefore, research must be conducted in an appropreate manner minimizing such issues (e.g., presenting results at an abstract and aggregate level; Sloan et al., 2015). Also, social media data has specific limitations in continuous data availability. Social media services owned by private companies frequently change their service contents, specifications of API, and policy for data access which can affect research activities (Ghermandi et al., 2023). For example, Panoramio, a photo sharing service which had been used for evaluation of CESs (Figueroa-Alfaro and Tang, 2017; van Zanten et al., 2016), was terminated in 2016 and the data is no longer available for research (Ghermandi et al., 2023; Zhang et al., 2020). As for Twitter, georeferencing functionality was gradually restricted similarly to the change in the Instagram API noted above (see 2.1): on June 18th 2019, the official client stopped the functionality of adding exact coordinates to tweets (Kruspe et al., 2021) and around February 22nd 2023, it also stopped the functionality of adding place of POI to tweets, at least in Japan (https://x.com/jumbo_744/status/1638858023874756613, https://x.com/tkatochin/status/1638554084205789186, in Japanese, accessed at April 15th, 2025). Moreover, Twitter had been known to have relatively open policy for the use of API for academic data access (Ghermandi et al., 2023). However, during its rebranding to X, free academic access to the API was stopped in 2023 (Kupferschmidt, 2023). Although the historical archive of tweets is still accessible, these situations hinder continuous data availability, particularly for georeferenced tweets, and limit the reproducibility and applicability of our study due to the high cost of the API license. On the other hand, continuous efforts explored new data sources such as Reddit (Fox et al., 2021b), TripAdvisor, and Google Maps (Owuor et al., 2023) for evaluations of CESs. Estimation and adjustment of demographic attributes as shown by this study would be useful when analyzing such data to achieve inclusive evaluations of CESs. In addition, integrating other methods such as citizen science and development of custom applications like Mappiness (MacKerron and Mourato, 2013) would be more important in future studies.

5. Conclusions

We found that biases in users' demographic attributes and spatial distributions of posts, which are commonly used for estimation of spatial distributions of CESs, are different among age groups. These results illustrate the importance of estimating and incorporating demographic attributes of users into evaluation of CESs. Of course, different results would be obtained using different social media platforms. Therefore, similar studies using different social media platforms or using multiple social media platforms should be encouraged. To achieve inclusive evaluation of CESs using social media, considering demographic attributes and obtaining representativeness would be an important step.

6. Data access

The data and programs are available on Zenodo (https://doi.org/10.5281/zenodo.11366621) and Mendeley Data (https://doi.org/10.17632/2hvh5yzcbd.1).

CRediT authorship contribution statement

Michio Oguro: Writing – review & editing, Writing – original draft, Project administration, Software, Methodology, Formal analysis, Conceptualization. **Rei Shibata:** Writing – review & editing, Writing – original draft, Software, Project administration, Methodology, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Michio Oguro reports financial support was provided by Environmental Restoration and Conservation Agency. Rei Shibata reports financial support was provided by Environmental Restoration and Conservation Agency. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank Takehiro Sasaki, Masashi Soga, Thoru Nakashizuka, and two anonymous reviewers for their valuable comments, Jun Oba for his help with the travel cost analysis and Jane Glazebrook for English editing. This study was supported by the Environment Research and Technology Development Fund (JPMEERF16S11520; JPMEERF20204R02; PMEERF23S12100) of the Environmental Restoration and Conservation Agency provided by the Ministry of the Environment of Japan.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecoser.2025.101777.

Data availability

The authors do not have permission to share original data.

References

Aiba, M., Shibata, R., Oguro, M., Nakashizuka, T., 2019. The seasonal and scale-dependent associations between vegetation quality and hiking activities as a recreation service. Sustain. Sci. 14, 119–129. https://doi.org/10.1007/s11625-018-0609-7.

Aiba, M., Shibata, R., Oguro, M., Nakashizuka, T., 2023. Variable effects of vegetation characteristics on a recreation service depending on natural and social environment. Sci. Rep. 13, 684. https://doi.org/10.1038/s41598-023-27799-7.

Barros, C., Moya-Gómez, B., Gutiérrez, J., 2019. Using geotagged photographs and GPS tracks from social networks to analyse visitor behaviour in national parks. Curr. Issues Tour. 23 (10), 1291–1310. https://doi.org/10.1080/13683500.2019.1619674.

Bartoń, K., 2023. MuMIn: Multi-model inference. R package version 1.48.4. https://doi.org/10.32614/CRAN.package.MuMIn.

Brooks, M.E., Kristensen, K., van Benthem, K.J., Magnusson, A., Berg, C.W., Nielsen, A., Skaug, H.J., Mächler, M., Bolker, B.M., 2017. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. The R J. 9 (2), 378–400. https://doi.org/10.32614/RJ-2017-066.

Chen, Y., He, S., Zhou, Z., 2022. Investigation of social media representation bias in disasters: towards a systematic framework. Int. J. Disaster Risk Reduct. 81, 103312. https://doi.org/10.1016/j.ijdrr.2022.103312.

Cohen, J., 1988. Chi-square tests for goodness of fit and contingency tables. In: Statistical Power Analysis for the Behavioral Sciences, 2nd ed. Routledge, p. 215. 10.4324/9780203771587.

Cramér, H., 1946. Mathematical Methods of Statistics. Princeton University Press, Princeton, 10.1515/9781400883868.

Cui, Y., He, Q., 2021. Inferring Twitters' socio-demographics to correct sampling bias of social media data for augmenting travel behavior analysis. J. Big Data Anal. Transp. 3, 159–174. https://doi.org/10.1007/s42421-021-00037-0.

Daniel, T.C., Muhar, A., Arnberger, A., Aznar, O., Boyd, J.W., Chan, K.M.A., Costanza, R., Elmqvist, T., Flint, C.G., Gobster, P.H., Grêt-Regamey, A., Lave, R., Muhar, S., Penker, M., Ribe, R.G., Schauppenlehner, T., Sikor, T., Soloviy, I., Spierenburg, M., Taczanowska, K., Tam, J., von der Dunk, A., 2012. Contributions of cultural services to the ecosystem services agenda. PNAS 109 (23), 8812–8819. https://doi.org/10.1073/pnas.1114773109.

Figueroa-Alfaro, R.W., Tang, Z., 2017. Evaluating the aesthetic value of cultural ecosystem services by mapping geo-tagged photographs from social media data on Panoramio and Flickr. J. Environ. Plan. Manag. 60 (2), 266–281. https://doi.org/10.1080/09640568.2016.1151772.

Fox, J., Weisberg, S., 2019. An R Companion to Applied Regression, 3rd ed. Sage, Thousand Oaks, CA. https://www.john-fox.ca/Companion/index.html.

Fox, N., Graham, L.J., Eigenbrod, F., Bullock, J.M., Parks, K.E., 2021a. Enriching social media data allows a more robust representation of cultural ecosystem services. Ecosyst. Serv. 50, 101328. https://doi.org/10.1016/j.ecoser.2021.101328.

Fox, N., Graham, L.J., Eigenbrod, F., Bullock, J.M., Parks, K.E., 2021b. Reddit: a novel data source for cultural ecosystem service studies. Ecosyst. Serv. 50, 101331. https://doi.org/10.1016/j.ecoser.2021.101331.

Ghermandi, A., 2018. Integrating social media analysis and revealed preference methods to value the recreation services of ecologically engineered wetlands. Ecosyst. Serv. 31, 351–357. https://doi.org/10.1016/j.ecoser.2017.12.012.

Ghermandi, A., Langemeyer, J., Van Berkel, D., Calcagni, F., Depietri, Y., Egarter Vigl, L., Fox, N., Havinga, I., Jäger, H., Kaiser, N., Karasov, O., McPhearson, T., Podschun, S., Ruiz-Frau, A., Sinclair, M., Venohr, M., Wood, S.A., 2023. Social media data for environmental sustainability: a critical review of opportunities, threats, and ethical use. One Earth 6 (3), 236–250. https://doi.org/10.1016/j.oneear.2023.02.008.

Guo, Z., Zhang, L., Li, Y., 2010. Increased dependence of humans on ecosystem services and biodiversity. PLoS One 5 (10), e13113. https://doi.org/10.1371/journal.pone.0013113.

Hausmann, A., Toivonen, T., Slotow, R., Tenkanen, H., Moilanen, A., Heikinheimo, V., Di Minin, E., 2018. Social media data can be used to understand tourists' preferences for nature-based experiences in protected areas. Conserv. Lett. 11 (1), e12343. https://doi.org/10.1111/conl.12343.

Hausmann, A., Toivonen, T., Fink, C., Heikinheimo, V., Kulkarni, R., Tenkanen, H., Di Minin, E., 2020. Understanding sentiment of national park visitors from social media data. People Nat. 2 (3), 750–760. https://doi.org/10.1002/pan3.10130.

Havinga, I., Bogaart, P.W., Hein, L., Tuia, D., 2020. Defining and spatially modelling cultural ecosystem services using crowdsourced data. Ecosyst. Serv. 43, 101091. https://doi.org/10.1016/j.ecoser.2020.101091.

Hollenstein, L., Purves, R., 2010. Exploring place through user-generated content: using Flickr tags to describe city cores. J. Spat. Inf. Sci. 1, 21–48. https://doi.org/10.5311/JOSIS.2010.1.3.

Institute for Information and Communications Policy, Ministry of Internal Affairs and Communications, Japan, 2022. Survey on Information and Communication Media Usage Time and Information Behavior in FY2021. Institute for Information and Communications Policy, Ministry of Internal Affairs and Communications, Japan, Tokyo, Japan.

Keeler, B.L., Wood, S.A., Polasky, S., Kling, C., Filstrup, C.T., Downing, J.A., 2015. Recreational demand for clean water: evidence from geotagged photographs by visitors to lakes. Front. Ecol. Environ. 13 (2), 76–81. https://doi.org/10.1890/140124.

Kruspe, A., Häberle, M., Hoffmann, E.J., Rode-Hasinger, S., Abdulahhad, K., Zhu, X.X., 2021. Changes in Twitter geolocations: Insights and suggestions for future usage. In Proceedings of the Seventh Workshop on Noisy User-generated Text (W-NUT 2021), 212–221, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.wnut-1.24.

Kupferschmidt, K., 2023. Twitter's threat to curtail free data access angers scientists. Science 379, 624–625. https://doi.org/10.1126/science.adh1696.

Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B., 2017. lmerTest package: Tests in linear mixed effects models. J. Stat. Softw. 82 (13), 1–26. https://doi.org/10.18637/jss.v082.i13.

Langemeyer, J., Ghermandi, A., Keeler, B., van Berkel, D., 2023. The future of crowd-sourced cultural ecosystem services assessments. Ecosyst. Serv. 60, 101518. https://doi.org/10.1016/j.ecoser.2023.101518.

Lavallée, P., Beaumont, J.-F., 2015. Why we should put some weight on weights. Survey methods: Insights from the field. https://doi.org/10.13094/SMIF-2015-00001.

Lenth, R.V., 2023. emmeans: Estimated marginal means, aka least-squares means. R package version 1.10.5. https://rvlenth.github.io/emmeans/.

MacKerron, G., Mourato, S., 2013. Happiness is greater in natural environments. Glob. Environ. Change 23 (5), 992–1000. https://doi.org/10.1016/j.gloenvcha.2013.03.010.

Milcu, A.I., Hanspach, J., Abson, D., Fischer, J., 2013. Cultural ecosystem services: a literature review and prospects for future research. Ecol. Soc. 18 (3), 44. https://doi.org/10.5751/ES-05790-180344.

Millennium Ecosystem Assessment, 2005. Ecosystems and Human Well-being: Current State and Trends. Island Press, NW, Washington, DC.

Nakagawa, S., Johnson, P.C.D., Schielzeth, H., 2017. The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. J. R. Soc. Interface 14, 20170213. https://doi.org/10.1098/rsif.2017.0213.

Nowak-Olejnik, A., Schirpke, U., Tappeiner, U., 2022. A systematic review on subjective well-being benefits associated with cultural ecosystem services. Ecosyst. Serv. 57, 101467. https://doi.org/10.1016/j.ecoser.2022.101467.

Orsi, F., Geneletti, D., 2013. Using geotagged photographs and GIS analysis to estimate visitor flows in natural areas. J. Nat. Conserv. 21 (5), 359–368. https://doi.org/10.1016/j.jnc.2013.03.001.

Owuor, I., Hochmair, H.H., Paulus, G., 2023. Use of social media data, online reviews and wikipedia page views to measure visitation patterns of outdoor attractions. J. Outdoor Recreat. Tour. 44, 100681. https://doi.org/10.1016/j.jort.2023.100681.

Plieninger, T., Bieling, C., Fagerholm, N., Byg, A., Hartel, T., Hurley, P., López-Santiago, C.A., Nagabhatla, N., Oteros-Rozas, E., Raymond, C.M., van der Horst, D., Huntsinger, L., 2015. The role of cultural ecosystem services in landscape management and planning. Curr. Opin. Environ. Sustain. 14, 28–33. https://doi.org/10.1016/j.cosust.2015.02.006.

Quétier, F., Rivoal, F., Marty, P., De Chazal, J., Thuiller, W., Lavorel, S., 2010. Social representations of an alpine grassland landscape and socio-political discourses on rural development. Reg. Environ. Chang. 10, 119–130. https://doi.org/10.1007/s10113-009-0099-3.

Raymond, C.M., Singh, G.G., Benessaiah, K., Bernhardt, J.R., Levine, J., Nelson, H., Turner, N.J., Norton, B., Tam, J., Chan, K.M.A., 2013. Ecosystem services and beyond: using multiple metaphors to understand human–environment relationships. Bioscience 63 (7), 536–546. https://doi.org/10.1525/bio.2013.63.7.7.

R Core Team, 2023. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Ruiz-Frau, A., Ospina-Alvarez, A., Villasante, S., Pita, P., Maya-Jariego, I., de Juan, S., 2020. Using graph theory and social media data to assess cultural ecosystem services in coastal areas: Method development and application. Ecosyst. Serv. 45, 101176. https://doi.org/10.1016/j.ecoser.2020.101176.

Ruths, D., Pfeffer, J., 2014. Social media for large studies of behavior. Science 346 (6213), 1063–1064. https://doi.org/10.1126/science.346.6213.1063.

Sinclair, M., Ghermandi, A., Sheela, A.M., 2018. A crowdsourced valuation of recreational ecosystem services using social media data: an application to a tropical wetland in India. Sci. Total Environ. 642, 356–365. https://doi.org/10.1016/j.scitotenv.2018.06.056.

Sinclair, M., Mayer, M., Woltering, M., Ghermandi, A., 2020. Valuing nature-based recreation using a crowdsourced travel cost method: a comparison to onsite survey data and value transfer. Ecosyst. Serv. 45, 101165. https://doi.org/10.1016/j.ecoser.2020.101165.

Sloan, L., Morgan, J., 2015. Who tweets with their location? Understanding the relationship between demographic characteristics and the use of geoservices and geotagging on Twitter. PLoS One 10 (11), e0142209. https://doi.org/10.1371/journal.pone.0142209.

Sloan, L., Morgan, J., Housley, W., Williams, M., Edwards, A., Burnap, P., Rana, O., 2013. Knowing the tweeters: Deriving sociologically relevant demographics from Twitter. Sociol. Res. Online 18 (3), 74–84. https://doi.org/10.5153/sro.3001.

Sloan, L., Morgan, J., Burnap, P., Williams, M., 2015. Who tweets? Deriving the demographic characteristics of age, occupation and social class from Twitter user meta-data. PLoS One 10 (3), e0115545. https://doi.org/10.1371/journal.pone.0115545.

Soga, M., Gaston, K.J., Yamaura, Y., Kurisu, K., Hanaki, K., 2016. Both direct and vicarious experiences of nature affect children's willingness to conserve biodiversity. Int. J. Environ. Res. Public Health 13 (6), 529. https://doi.org/10.3390/ijerph13060529.

van Zanten, B.T., Van Berkel, D.B., Meentemeyer, R.K., Smith, J.W., Tieskens, K.F., Verburg, P.H., 2016. Continental-scale quantification of landscape values using social media data. PNAS 113 (46), 12974–12979. https://doi.org/10.1073/pnas.1614158113.

Venter, Z.S., Gundersen, V., Scott, S.L., Barton, D.N., 2023. Bias and precision of crowdsourced recreational activity data from Strava. Landsc. Urban Plan. 232, 104686. https://doi.org/10.1016/j.landurbplan.2023.104686.

Wang, Z., Fu, H., Jian, Y., Qureshi, S., Jie, H., Wang, L., 2022. On the comparative use of social media data and survey data in prioritizing ecosystem services for cost-effective governance. Ecosyst. Serv. 56, 101446. https://doi.org/10.1016/j.ecoser.2022.101446.

Wartmann, F.M., Tieskens, K.F., van Zanten, B.T., Verburg, P.H., 2019. Exploring tranquillity experienced in landscapes based on social media. Appl. Geogr. 113, 102112. https://doi.org/10.1016/j.apgeog.2019.102112.

Wilkins, E.J., Wood, S.A., Smith, J.W., 2020. Uses and limitations of social media to inform visitor use management in parks and protected areas: a systematic review. Environ. Manag. 67, 120–132. https://doi.org/10.1007/s00267-020-01373-7.

Wood, S.A., Guerry, A.D., Silver, J.M., Lacayo, M., 2013. Using social media to quantify nature-based tourism and recreation. Sci. Rep. 3, 2976. https://doi.org/10.1038/srep02976.

Yang, Y.-C., Al-Garadi, M.A., Love, J.S., Cooper, H.L.F., Perrone, J., Sarker, A., 2023. Can accurate demographic information about people who use prescription medications nonmedically be derived from Twitter? PNAS 120 (8), e2207391120. https://doi.org/10.1073/pnas.2207391120.

Yoshimura, N., Hiura, T., 2017. Demand and supply of cultural ecosystem services: use of geotagged photos to map the aesthetic value of landscapes in Hokkaido. Ecosyst. Serv. 24, 68–78. https://doi.org/10.1016/j.ecoser.2017.02.009.

Zhang, H., Huang, R., Zhang, Y., Buhalis, D., 2020. Cultural ecosystem services evaluation using geolocated social media data: a review. Tour. Geogr. 24 (4–5), 646–668. https://doi.org/10.1080/14616688.2020.1801828.