Landscape and Urban Planning 227 (2022) 104539

Contents lists available at ScienceDirect

Landscape and Urban Planning

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

The relationship between natural environments and subjective well-being as measured by sentiment expressed on Twitter

Yian Lin[a,*], Spencer A. Wood[b,c,d], Joshua J. Lawler[d,c]

[a] Quantitative Ecology and Resource Management, University of Washington, Seattle, WA, USA
[b] Natural Capital Project, Stanford University, Stanford, CA, USA
[c] Nature and Health, University of Washington, Seattle, WA, USA
[d] School of Environmental and Forest Sciences, University of Washington, Seattle, WA, USA

HIGHLIGHTS

- Sentiment expressed in tweets was associated with surrounding environments.
- Some types of nature improve sentiment of people and some do not.
- The strength and the direction of the relationships varied by land-use type.
- Nuanced analyses reveal more nuanced sentiment–environment relationships.
- Geolocated social media data can help quantify sentiment–environment relationships.

ARTICLE INFO

Keywords:
Sentiment
Well-being
Land cover
Tree canopy
Urban parks
Twitter

ABSTRACT

There is growing evidence that time spent in nature can affect well-being. Nonetheless, assessing this relationship can be difficult. We used 1,971,045 geolocated tweets sent by 81,140 users from Seattle, Washington, USA to advance our understanding of the relationship between subjective well-being and natural environments. Specifically, we quantified the relationships between sentiment (negative/neutral/positive) expressed in geolocated tweets and their surrounding environments, focusing on three environmental indicators: land-cover type, tree-canopy density, and urban parks. We allowed the relationships to vary according to the broader type of environment (i.e., land-use zoning). We estimated three random-intercept partial proportional odds models corresponding to the three environmental indicators while controlling for multiple covariates. Our results suggested that for a given land-use type, tweets sent from some natural land-cover types were less likely to be negative compared to tweets sent from the urban built land-cover type. For tweets sent in industrial zones, an increase in tree-canopy cover was associated with a lower probability of having negative sentiments and with a higher probability of having positive sentiments; but for tweets sent in commercial/mixed zones, the association was reversed. Also, urban parks were generally associated with a lower probability of having negative sentiments, but tweets sent from large natural parks in residential zones were less likely to be positive. Our results suggest that the relationship between subjective well-being and natural environments depends on where people are situated in the built environment and may be more complex than previously thought. The more nuanced understanding provided by analyzing geolocated social media has potential to inform urban planning and land management.

1. Introduction

By 2050, 68 % of the world's total population is expected to live in cities, and the urban population is projected to increase by 2.5 B (United Nations, 2019). Although urban living may provide more job opportunities, higher income, and more convenience, an urban lifestyle that leads to more time spent indoors and less contact with nature may negatively affect mental health and well-being (Soga & Gaston, 2016; White et al., 2013; Cox et al., 2018; Hartig et al., 2011). There are several possible mechanisms through which natural environments can

* Corresponding author.
E-mail addresses: yianlin@uw.edu (Y. Lin), spwood@uw.edu (S.A. Wood), jlawler@uw.edu (J.J. Lawler).

https://doi.org/10.1016/j.landurbplan.2022.104539
Received 21 February 2022; Received in revised form 1 July 2022; Accepted 9 August 2022
Available online 22 August 2022

reverse adverse health effects by providing psychological benefits to urban populations (Frumkin et al., 2017; Hartig et al., 2014). Some of the most direct mechanisms that have been proposed include the Biophilia Hypothesis (Kellert & Wilson, 1995; Wilson, 1984), Stress Reduction Theory (SRT) (Ulrich et al., 1991) and Attention Restoration Theory (ART) (Kaplan & Kaplan, 1989; Kaplan, 1995). According to these theories, exposure to natural environments satisfies innate emotional needs, reduces stress levels, and helps people recover from attention fatigue. In addition, psychological benefits may be mediated by physical activity (Richardson et al., 2013), social interactions (Maas et al., 2009), and the mitigation of environmental hazards (Dadvand et al., 2012).

A growing body of research suggests that various types of exposure to the natural environment are associated with a wide range of mental health benefits, including higher levels of psychological well-being (McMahan & Estes, 2015) and lower risks of certain types of mental illness (Bratman et al., 2019). Laboratory experiments have demonstrated that exposure to images, videos, and sounds of the natural environment can reduce stress levels and help people recover from attention fatigue (Berto, 2005; Wang et al., 2016). Field experiments that assigned participants actual exposure to natural environments have demonstrated that walking in natural versus urban environments can improve positive affect, memory span, mood, and directed-attention abilities (Berman et al., 2008, 2012; Bratman et al., 2015a), and reduce anger, stress, negative affect, and rumination (Bratman et al., 2015b; Hartig et al., 2003; Marselle et al., 2013). Observational studies using traditional survey data also provide a range of evidence for the positive association between nature exposure and mental well-being. For example, it has been shown that people living in areas with a large amount of green space are less affected by stressful life events (van den Berg et al., 2010). Higher greenness in residential areas has been negatively associated with the risk of depression for people with diabetes (Gariepy et al., 2015). The probability of reporting positive well-being significantly increases for people who spend at least two hours a week in nature (White et al., 2019).

Different approaches used in previous studies to advance our understanding of the psychological benefits of exposure to nature have their strengths and weaknesses. Laboratory experiments have the greatest potential to make causal inference and test underlying mechanisms, but have the lowest ecological validity as they can only examine the exposure to simulated or virtual natural environments rather than real nature experiences. Field experiments have more ecological validity, but have less control over potential confounders such as weather conditions (Browning et al., 2020). Also, field experiments require a relatively large amount of time and effort from both researchers and participants, as well as research funding. Observational studies that focus on the proximity of homes to natural environments and the local amount of greenness can only approximate real exposure (Frumkin et al., 2017). Furthermore, observational studies relying on retrospective interviews and surveys are subject to recall bias and the lack of objective data on the participants' actual location and of data on potential confounders. In addition, these studies can be expensive and time-consuming if they have large sample sizes, broad geographic extents, or aim to collect longitudinal data.

Studies have tried to overcome challenges to empirical research by crowd-sourcing data from people who volunteer information through social media platforms (Wood et al. 2013). The large volume of social media shared through Twitter has been shown to inform the understanding of people's diurnal and seasonal mood patterns (Golder & Macy, 2011), predict stock market trends (Bollen et al., 2011), estimate park visitation (Donahue et al. 2018, Hamstead et al. 2018), and track influenza (Lampos & Cristianini, 2010). Also, some Twitter-based studies provide evidence supporting the benefits of visiting urban green space (Lim et al., 2018; Plunz et al., 2019; Roberts et al., 2019; Schwartz et al., 2019). However, most of the Twitter-based studies have focused on urban parks only and have not accounted for important factors, such as the dependence among data points collected from the same user and the impact of potential confounders such as weather conditions. The failure to account for these factors could bias the estimates of the effects of nature.

Here, we rely on geolocated tweets as a data source in order to address some of the challenges faced by previous research using more traditional methods, including low ecological validity, difficulties assessing exposure, study costs, and lack of objective location data and data on potential confounders. We also address many of the limitations of previous passively crowd-sourced projects (e.g., relatively narrow scope and the failure to account for confounding factors and dependence among data) by examining the relationship between subjective momentary sentiment and people's surrounding environments with a focus on land cover, tree canopy, and urban parks. We allow the relationships to vary according to the broader type of environment (i.e., land-use zoning) because the larger environmental context of a location and the activities in which people are engaged may influence how people interact with, and feel about, their surrounding environments (MacKerron & Mourato, 2013). We quantify the relationships using statistical models that allow us to adjust for potential confounders as well as heterogeneity among data contributors. Specifically, we ask three questions. How does subjective momentary sentiment expressed in tweets vary across urban-built and more natural land-cover types? Within a given land-cover type, is the amount of tree canopy associated with the sentiment expressed in tweets? Are people more likely to express more positive sentiment in urban parks than they are elsewhere in the urban landscape?

2. Methods

We investigated the association between subjective momentary sentiment and the surrounding environments using 1,971,045 geolocated tweets sent between September 2010 and February 2020 by 81,140 Twitter users at locations throughout Seattle, WA. Using geolocated tweets, we were able to objectively locate users and infer their real-time momentary sentiment at that location.

2.1. Sample

From Twitter's streaming application programming interface (API), we obtained 2.6 M random English geolocated tweets sent between November 2012 and January 2019 in Seattle (Fig. S1). Each tweet retrieved from the API was a JavaScript object notation (JSON) object that included the text content of the tweet, the geolocation, the timestamp, the source type, and other attributes. Only tweets containing a "coordinate" value, representing the exact latitude and longitude, instead of just a generic "place" value, were considered as geolocated tweets. Retweets did not have a "coordinate" value according to Twitter's API documentation, and thus were not included.

The 2.6 M geolocated tweets were sent by 172 K Twitter user accounts. In February 2020, we retrieved from the Twitter API up to 3200 of the most recent tweets posted by each user to their timeline for the accounts that still existed and were public. The API returned 2.1 M tweets that were sent within Seattle from 93 K of the 172 K user-accounts. Therefore, in total, we had 4.7 M geolocated tweets from 172 K user-accounts before any further filtering.

In the filtering process, we first removed duplicate tweets, non-English tweets, and tweets with imprecise geolocations whose latitude or longitude coordinates had less than four decimal places. We then excluded tweets sent via bots, such as advertisements and weather condition reports. We identified bots by checking the "tweet source labels" attached to each tweet and removed all tweets sent from suspicious sources (such as "TweetMyJOBS") that were not listed in Box S1. In addition, tweets cross-posted from other social media platforms such as Instagram were excluded because the geo-coordinates associated with cross-posted tweets may not have been the exact location of the tweets.

Cross-posted tweets were also identified through the "tweet source labels". After filtering, there were 1.97 M tweets sent between September 2010, and February 2020, by 81 K user-accounts throughout Seattle. The spatial distribution of tweets is shown in Fig. 1A. The median number of tweets per user was three (first quartile = 1, third quartile = 11). Most of the tweets (96.48 %) were sent between 2012 and 2015.

2.2. Measures

2.2.1. Momentary sentiment

To extract the real-time subjective momentary sentiment of each tweet based on its content, we employed the widely used sentiment analysis tool—Valence Aware Dictionary for Sentiment Reasoning (VADER; Hutto & Gilbert, 2014). VADER is a lexicon and rule-based sentiment analysis approach that uses natural language processing to combine lexical features with certain grammatical and syntactical rules.

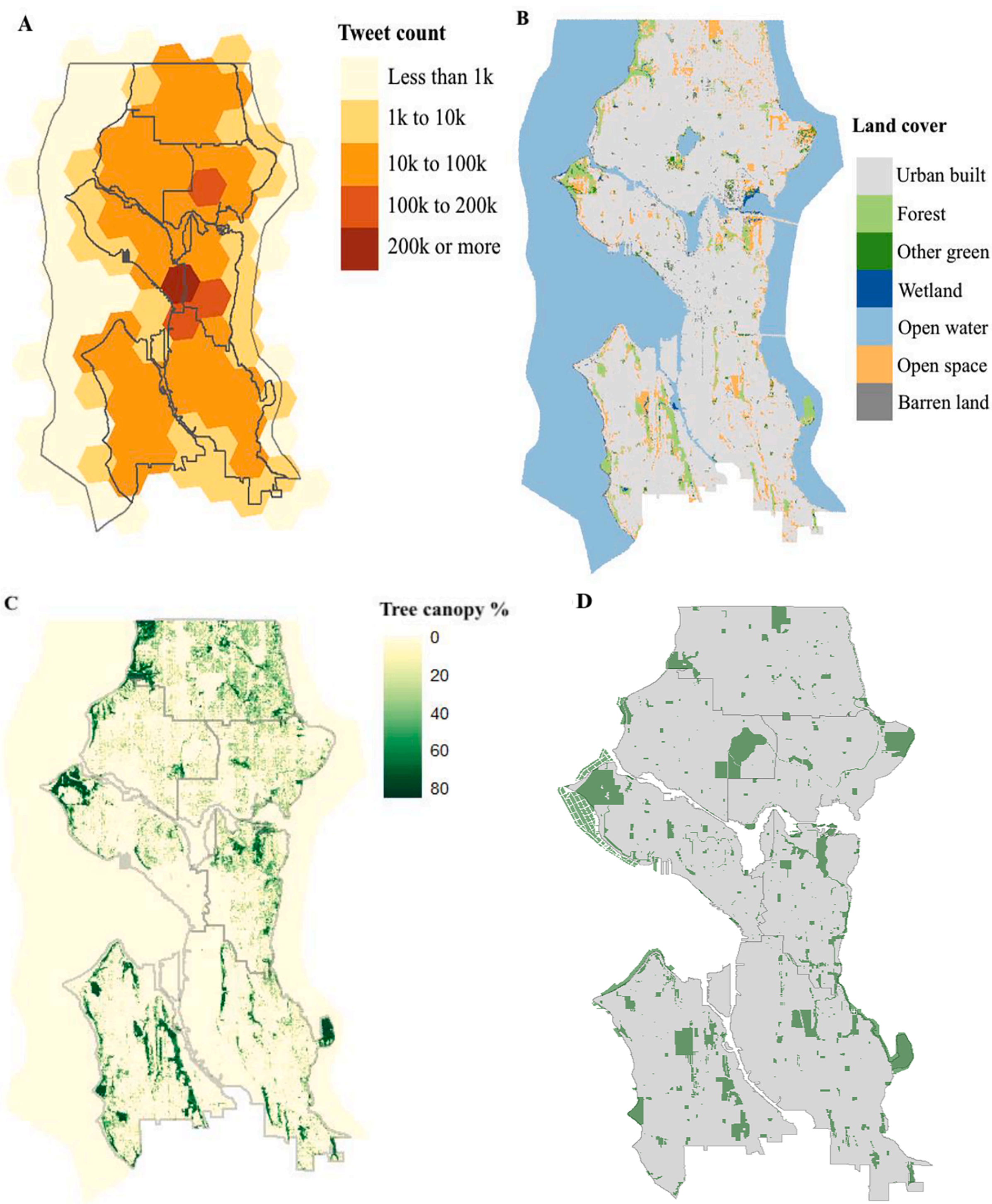

Fig. 1. (A) The spatial distribution of tweets in the sample. (B) Land cover in Seattle. (C) Tree canopy coverage in Seattle. (D) Parks in Seattle. Land-cover and tree-canopy data were taken from the 2016 National Land Cover Database (NLCD) (Dewitz, 2019). Polygons representing urban parks within Seattle were downloaded from Seattle GeoData portal (Seattle Parks, 2012).

VADER comes tailored to extract sentiments expressed in tweets as well as texts in other domains. VADER's success (with classification thresholds set at −0.05 and 0.05) detecting three-class sentiment polarity (positive vs neutral vs negative) is among the best for social media posts such as tweets (Hutto & Gilbert, 2014; Ribeiro et al., 2016). In addition, VADER is fully open-source, self-contained, and computationally efficient. Studies in various disciplines such as environmental management, computer engineering, and computational Linguistics have employed VADER (Althoff et al., 2016; Becken et al., 2017; Kim et al., 2016).

We assessed and calibrated classification thresholds for VADER and then evaluated its performance using a random sample of 1,000 tweets from our full dataset (before filtering). We (the three authors) manually labeled the polarity of every tweet. Of the 1000 tweets, 171 were detected as spam and removed by the human raters. Within the remaining 829 tweets, 98.2 % were labeled with the same polarity by at least two raters, and 67 % with the same polarity by all three raters. The values of Fleiss's Kappa and Kendall's coefficient of concordance (after correction for ties) were 0.62 and 0.80 respectively, which suggests good inter-annotator agreement (Landis and Koch, 1977). We aggregated the annotation of the three human raters by assigning each tweet the sentiment polarity with the majority of the votes. The 15 tweets for which there was no agreement between raters were considered neutral. As a result of the aggregation, 52 % of the 829 tweets were classified as neutral, 35 % as positive, and 13 % as negative. Some examples of tweets are shown in Box S2 with sentiment labels. Based on the aggregated human annotation, we found that VADER performed best with classification thresholds set at −0.15 and 0.35 for our dataset. At these thresholds, the overall accuracy and weighted F1-score were 0.69 (Table S1, formulas used to compute these comparison metrics are provided below Table S1). We also compared VADER's performance with that of other commonly used sentiment analysis tools such as SentiStrength (Thelwall et al., 2012), Google's Natural Language API, Sentiment140 (Go et al., 2009), and TextBlob using the aggregated human annotation. VADER outperformed these other tools in our tests based on the overall accuracy, macro F1-score, and weighted F1-score (Table S1).

2.2.2. *Predictors of main interest: Land-cover, Tree-canopy cover, parks*

We used three spatial data layers representing the environments from which tweets were sent including data on land-cover type, tree-canopy cover, and urban park boundaries. Land-cover and tree-canopy data were taken from the 2016 National Land Cover Database (NLCD) (Dewitz, 2019) downloaded from the Multi-Resolution Land Characteristics Consortium. NLCD provides nationwide data on land-cover classes and tree-canopy cover at 30-m resolution. The land-cover data has 15 subclasses based on a modified Anderson Level II classification system. We grouped the 15 subclasses into 7 broad land-cover categories (Table S2; Fig. 1B). For tree-canopy cover, the value assigned to every 30-m by 30-m grid cell in the NLCD represents the percentage of tree-canopy coverage (Fig. 1C). Polygons representing urban parks within Seattle (Fig. 1D) were downloaded from Seattle GeoData portal (Seattle Parks, 2012). For each tweet, we assigned values to the three predictors—*Land-cover*, *Tree-canopy*, *Park*—according to the extracted values from each of the three spatial layers based on the tweet's coordinates. To facilitate analysis, we scaled *Tree-canopy* by dividing it by 10 and created a categorical variable *Park* with three categories—within a large natural park (parks with less than 30 % impervious surface and covering at least 3716 m^2 (40 K ft^2)), within one of the other parks (parks with >30 % impervious surface area and/or smaller than 3716 m^2), and not in any park (Table S6).

2.2.3. *Covariates*

In addition to the predictors of main interest, we included a number of time-varying covariates that may obscure the relationship between momentary sentiment expressed in tweets and the predictors of main interest. These were divided into four groups: location type, weather conditions, time, and tweet type. All of these variables covary with momentary sentiment, and many of them may also be correlated with the predictors of main interest, thus potentially confounding their effects. We adjusted for these covariates in the regression analyses presented in Section 2.3.

For location-type covariates, we used *Zoning* and *Outdoor*. *Zoning* indicates the type of land development that is allowed (e.g., residential, industrial, etc.) and is likely indicative of the types of activities people are performing in a given location. The distributions of land cover, tree canopy, and park type within each zoning category (i.e., commercial/mixed, residential, industrial) are shown in Table S3 to S5. *Outdoor* is a binary variable that indicates whether a tweet was sent from an outdoor (*Outdoor* = 0) or indoor (*Outdoor* = 1) location. The data for the zoning (Current Land Use Zoning Detail, 2020) and for determining indoor/outdoor location (Building Outlines 2015, 2018) were downloaded from the Seattle GeoData portal. For weather conditions, we included five categorical covariates: *Temperature*, *Dew point*, *Visibility*, *Rain*, and *Sky cover*. Weather data were obtained from the NOAA Integrated Surface Dataset (Integrated Global Surface Hourly Data, n.d.). We associated each geolocated tweet with the weather conditions from the station closest to the tweet both spatially and temporally. Furthermore, based on the timestamp associated with each tweet, three time-related covariates were created: *Time*, *Day of week*, and *Day*. *Time* was continuous with month as its unit. For each user, Time was set to 0 for his/her earliest tweet included in the sample. *Day* was a binary covariate that indicated whether a tweet was sent during the day (6 a.m. to 5 p.m.) or at night. Lastly, according to the tweet-type indicators associated with each tweet, we created one binary tweet-type covariate: *Tweet type*. *Tweet type* was set to 1 when a tweet was a reply or a quote tweet and to 0 when a tweet was an original tweet. Descriptive statistics on all categorical covariates are shown in Table S6. Distribution of tweets by land-cover type and zoning, as well as by park category and zoning are shown in Table S7 and S8.

2.3. *Statistical analysis*

To investigate the association between sentiment polarity expressed in tweets and the surrounding environments, we used random-intercept partial proportional odds models with *User* as the random effect. Ordinal logistic regression was chosen because the outcome variable—sentiment polarity—has ordinal categories (1 = negative, 2 = neutral, and 3 = positive). We allowed users to have random intercepts because about 70 % of the 81 K users have more than one tweet included in the sample and tweets sent by the same user may not be independent. Using mixed models allowed us to take the unobserved heterogeneity between users into account when estimating the fixed effects coefficients (Rabe-Hesketh & Skrondal, 2012). We also explored random-slope models that allow users to have subject-specific trends over time. Because the random-slope models failed to converge, we chose to use the more parsimonious random-intercept models. Finally, we chose partial proportional odds models instead of proportional odds models because for most covariates, except for *Outdoor*, *Time*, *Tweet type*, and *Tree-canopy*, the proportional odds assumption is not valid based on likelihood ratio tests and Akaike Information Criterion (AIC; Table S9). We tested the proportional odds assumption for all fixed-effect predictors in preliminary univariate analyses in which we fitted models with only one fixed-effect predictor each time. And for each predictor, we compared the two models: one with the proportional odds assumption and the other with the proportional odds assumption relaxed. To avoid overcomplicating the models, we only relaxed the proportional odds assumption when both likelihood ratio test and AIC values suggest that the assumption should be relaxed.

We fit three main models corresponding to each of our three research questions. The first main model used *Land-cover* as the predictor of main interest to investigate the relationship between different land-cover types and the sentiment polarity of tweets while controlling for other covariates. We also tested the interaction between *Land-cover* and *Zoning* and the interaction between *Land-cover* and *Outdoor*. We kept only the interaction between *Land-cover* and *Zoning* in the final model because the interaction between *Land-cover* and *Outdoor* was not significant based on the likelihood ratio test and AIC. Model 1 was specified as

Sentiment ~ *Land-cover* * *Zoning* + *Outdoor* + *Time* + *Day of week* + *Day* + *Temperature* + *Dew point* + *Visibility* + *Sky cover* + *Rain* + *Tweet type* + (1| *User*),

where Sentiment (1 = Negative, 2 = Neutral, 3 = Positive) was an ordinal categorical variable and *Land-cover*, its interaction with *Zoning*, and other covariates were fixed effects. All predictors except *Time* were included as categorical variables, and their categories are shown in Table S6. Covariates that are underlined in the model are those for which the proportional odds assumption holds. The last term is the random effect *User* indicating who sent the tweet.

A second model examined the association between tree-canopy coverage and sentiment polarity while adjusting for land-cover type and all other covariates included in Model 1. We adjusted for land-cover type because it was a potential confounder, and we wanted to isolate the tree-canopy effects from the land-cover effects. In Model 2, we allowed the interaction between *Tree-canopy* and *Zoning*. Other potential interactions, between *Tree-canopy* and *Land-cover*, *Tree-canopy* and *Outdoor*, and *Land-cover* and *Zoning* were tested, but none were significant based on likelihood ratio tests and AIC. Therefore, we specified Model 2 as

Sentiment ~ *Tree-canopy* * *Zoning* + *Land-cover* + *Outdoor* + *Time* + *Day of week* + *Day* + *Temperature* + *Dew point* + *Visibility* + *Sky cover* + *Rain* + *Tweet type* + (1| *User*).

A third model with *Park* as the predictor of main interest investigated the differences in sentiment polarity of tweets between being sent from within large natural parks, within other parks in Seattle, and sent from non-park areas, while controlling for other covariates. We included the interaction between *Park* and *Zoning*. We explored the interaction between *Park* and *Outdoor*, but found it not significant based on the likelihood ratio test and AIC. Therefore, Model 3 was specified as

Sentiment ~ *Park* * *Zoning* + *Outdoor* + *Time* + *Day of week* + *Day* + *Temperature* + *Dew point* + *Visibility* + *Sky cover* + *Rain* + *Tweet type* + (1| *User*).

All models were fitted using the “mixor” package in R, and model parameters were estimated via the maximum likelihood method (Archer et al., n.d.). To make interpretation more straightforward, we present model results as odds ratios (OR = exp(coefficient)) and 95 % confidence intervals of odds ratios in Section 3. In addition, we adjusted p-values for multiple testing for all two-way interactions following the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995). All explanatory variables except *User* were time-varying variables and their effects were subject-specific (Rabe-Hesketh & Skrondal, 2012). As mentioned above, to allow the fixed effects to vary according to which cumulative logits of the model we were considering (logit(Pr(Y > 1)) or logit(Pr(Y > 2)), we relaxed the proportional odds assumption for some fixed-effect variables (*Land-cover, Parks, Zoning, Day of week, Day, Temperature, Dew point, Visibility, Sky cover, Rain*) based on likelihood ratio tests and AIC, which resulted in partial proportional odds models. As a result, we produced two sets of estimated odds ratios for these variables when they were included in a model: the odds of being neutral or positive (Y = 2 or 3) as opposed to being negative (Y = 1), and the odds of being positive (Y = 3) as opposed to being neutral or negative (Y = 2 or 1). Thus, for each model, there were two odds ratios estimated for each of these explanatory variables, corresponding to each of the two odds contrasts. Hereafter, we refer to the first odds contrast as the “non-negative odds contrast” and the other as the “positive odds contrast.”

The large size of our sample provides us more power to detect small and complex effects. However, large samples also drive p-values to zero quickly (Sullivan & Feinn, 2012). Therefore, gauging the practical significance of statistically significant effects based on both effect size and p-values is more appropriate (Lin et al., 2013). Because odds ratios can be hard to interpret, we also computed predicted probabilities to understand the effect size of our statistically significant primary predictors (i.e., the land-cover types that were estimated to be significantly different from the urban-built class, the tree canopy effects that were estimated to be significantly different from 0, and the types of parks that were estimated to be significantly different from non-park areas), holding all other variables at their sample means. We calculated these predicted probabilities as average marginal probabilities by averaging over predicted probabilities that corresponded to simulated values of random intercepts. As a result, the predicted probabilities have population-level interpretation (Steele, 2009).

3. Results

3.1. Sentiment and land-cover

Based on VADER’s classification, 35.18 % of the tweets in our sample were positive, 45.15 % were neutral, and 19.67 % were negative. We found that tweets sent from some natural land-cover types were less likely to be negative than those sent from urban built environments across commercial/mixed, industrial, and residential zones of the city, according to the OR estimated for the non-negative odds contrast (Fig. 2; Table S10; see page 14 of the supplementary material for a brief explanation of the relationship between the estimated OR and the probability of being negative or the probability of being positive). However, we found no evidence that tweets sent from places with more natural land cover were more likely to be positive than tweets sent from urban built environments, based on the OR estimated for the positive odds contrast. More specifically, within commercial/mixed zones, tweets sent from other green, barren land, and open water were less likely to be negative; the odds of being non-negative rather than negative was 7.32 % (95 % CI = [1.46 %, 13.52 %]), 3.55 % ([1.13 %, 6.02 %]), and 4.40 % ([0.55 %, 8.40 %]) higher, respectively, than that of tweets sent from urban built environments by the same user. In industrial zones, a given user was less likely to send negative tweets in open space, barren land, and other green land-cover types than in the urban built land-cover type. Tweets sent from the three land-cover types had 22.29 % ([13.84 %, 31.38 %]), 9.82 % ([1.58 %, 18.72 %]), and 24.33 % ([6.93 %, 44.57 %]) higher odds, respectively, of being non-negative compared to tweets sent from urban built environments. We also found that tweets sent in wetlands were more likely to be negative compared to tweets sent from urban built environments. However, it is worth noting that only 53 tweets in the sample were sent from wetlands in industrial zones. Lastly, in residential zones, tweets sent in open space and barren land were less likely to be negative: the odds of being non-negative were 5.60 % ([1.37 %, 10.01 %]) and 13.22 % ([6.64 %, 20.20 %]) higher than those of tweets sent from urban built environments respectively.

We found associations between the sentiment of tweets and several of the covariates included in the model (Table S10). For instance, tweets sent on Thursdays, Fridays, and weekends were all less likely to be negative than tweets sent between Monday and Wednesday, with Saturdays being the least negative days of the week. However, tweets sent on weekends were also less likely to be positive than tweets sent from Monday to Wednesday. Tweets sent at night were more likely to be negative and less likely to be positive than tweets sent during the day. The absence of rain was associated with higher odds of being positive. Higher temperatures (>=18 °C) were associated with lower probability of being negative relative to low temperatures (<0°C). Tweets sent when temperatures were between 0 and 7 °C were more likely to be positive than tweets sent while temperatures were under 0 °C. Higher dew point (2–16 or >16 °C) were associated with higher probability of being negative when compared to low dew point (<= 2 °C). Having fog was associated with higher probability of being positive compared to having clear visibility. Tweets sent under overcast skies were more likely to be negative and less likely to be positive relative to tweets sent on days with

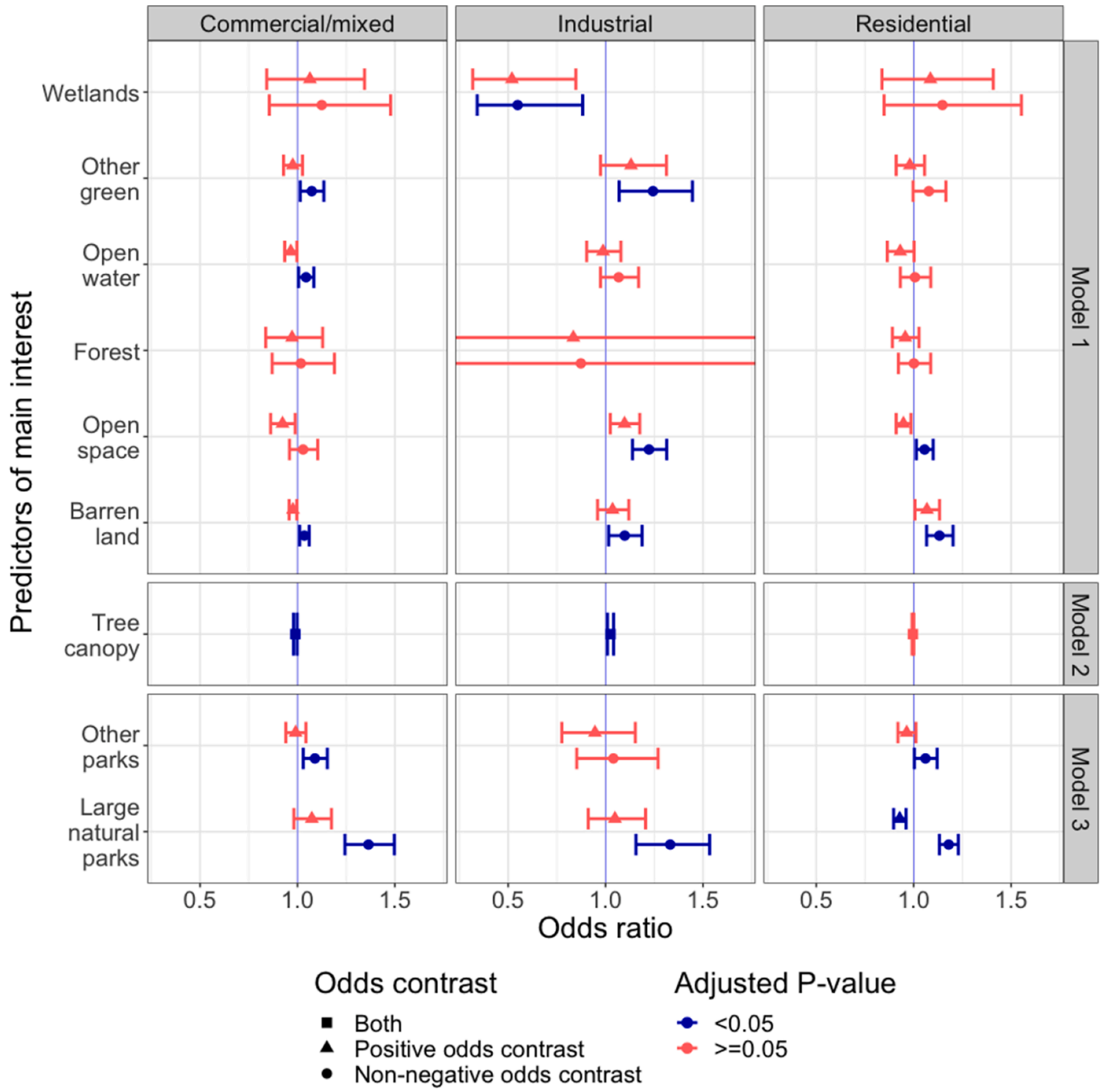

Fig. 2. Estimated odds ratios and 95% CI for the predictors of main interest in Model 1, 2, and 3, which explore the potential relationships of sentiment with land cover (Model 1), tree canopy (Model 2), and parks (Model 3), respectively, in a given zone of Seattle (i.e., commercial/mixed, industrial, and residential). Each of the three models is a random-intercept partial proportional odds model with *User* as the random effect. The reference class for *Land-cover* in Model 1is urban built. The reference class for *Park* in Model 3 is non-park areas. The p-values were adjusted for multiple testing following the Benjamini-Hochberg procedure.

clear skies or scattered clouds. Tweets sent indoor were less likely to be negative and more likely to be positive than those sent outdoor. Finally reply or quote tweets were less likely to be negative and more likely to be positive than original tweets.

3.2. Sentiment and tree canopy

We found a significant relationship between sentiment and tree canopy in commercial/mixed and industrial zones (Fig. 2; Table S11). In commercial/mixed zones, tweets were more likely to be negative and less likely to be positive when there was more tree canopy. By contrast, in industrial zones, tweets were more likely to be positive and less likely to be negative when there was more tree canopy. It should be noted that the proportional odds assumption was kept for *Tree-canopy* in Model 2 based on the likelihood ratio test and AIC. Therefore, the effect of tree-canopy coverage was the same for both odds contrasts. For instance, for tweets sent in commercial/mixed zones, each 10-percentage-point increase in tree canopy was associated with 1.10 % ([0.16 %, 2.03 %]) lower odds of being non-negative and also 1.10 % ([0.16 %, 2.03 %]) lower odds of being positive. The estimated effects for other covariates under Model 2 were similar to those estimated under Model 1.

3.3. Sentiment and urban parks

We found a significant relationship between parks and sentiment across all zoning types. For example, in commercial/mixed zones and residential zones, both tweets sent within large natural parks and those within other parks were less likely to be negative compared to tweets sent from non-park areas (Fig. 2; Table S12). In industrial zones, only tweets sent within large natural parks were less likely to be negative. We also found that in residential zones tweets sent within large natural parks were less likely to be positive compared to tweets sent from non-park areas.

For tweets sent within large natural parks or other parks in commercial/mixed zones, the odds of being non-negative was 36.45 % ([24.35 %, 49.73 %]) and 8.95 % ([2.95 %, 15.30 %]) higher, respectively, compared to tweets sent from non-park areas. For tweets sent in industrial zones, being within large natural parks was associated with

33.18 % ([15.62 %, 53.40 %]) higher odds of being non-negative, but being within other parks was not significantly different from being within non-park areas in terms of the probability of being negative. It is worth noting that the number of tweets sent from other parks in industrial zones (369) was relatively small in our study. For tweets sent from large natural parks and other parks in residential zones, the odds of being non-negative was 17.97 % ([13.22 %, 22.93 %]) and 6.06 % ([0.37 %, 12.08 %]) higher respectively compared to tweets sent from non-park areas. However, being in large natural parks in residential zones was also associated with 7.18 % ([3.95 %, 10.30 %]) lower odds of being positive. In other words, tweets sent from large natural parks in residential zones were less likely to be negative, but also less likely to be positive. The estimated effects for other covariates under Model 3 were similar to those estimated under Model 1 and Model 2.

3.4. Discrete changes in predicted probabilities

For tweets sent from locations with natural land-cover types that were estimated to be significantly different from urban built cover, the predicted probabilities of being negative were between 0.38 and 3.17 percentage-points lower when compared to tweets sent from urban built environments in the same zones, except in wetlands within industrial zones (Fig. S2). For tree canopy in commercial/mixed zones, a 20-percentage-point increase (from 0 % to 20 % for all land cover except forest; from 20 % to 40 % for forest) was associated with 0.06 to 0.73 percentage-point increases in the probability of being negative (Fig. S3A) and 0.06 to 1.36 percentage-point decreases in the probability of being positive (Fig. S3B). By contrast, the same amount of increase in tree canopy in industrial zones was associated with 0.15 to 1.01 percentage-point decreases in the probability of being negative (Fig. S3A) and 0.37 to 1.81 percentage-points increases in the probability of being positive (Fig. S3B). Lastly, being in parks, except for the other parks in industrial zones, was associated with 1.14 to 3.92 percentage-point decreases in the probability of being negative (Fig. S4A). Being in large natural parks in residential zones was associated with a 1.56-percentage-point decrease in the probability of being positive (Fig. S4B).

4. Discussion

4.1. Main findings

Our analysis of 2 M tweets sent by people in Seattle suggests that the relationship between people's surroundings and their subjective momentary sentiment is nuanced, with some types of nature positively affecting sentiment, but only under certain conditions. We observe that the psychological benefit of nature, according to Twitter, is dependent not only on the type of nature that people experience, but also on the broader type of environment (i.e., land-use zoning) where they are situated in the moment. For instance, barren land covers improve sentiment (i.e., tweets were less likely to be negative) in all three zoning types, but open space land cover is beneficial (i.e., tweets were less likely to be negative) only around residential and industrial areas of Seattle. Also, with the exception of other parks in industrial zones, urban parks were generally associated with a lower probability of having negative sentiments when compared to non-park areas. However, large natural parks in residential zones were also associated with a lower probability of being positive. These findings highlight the importance of considering the larger environmental context of a location as well as the experiences and activities that people are doing in different types of built environments as represented by zoning in our study.

In general, our findings provide some support for the growing body of evidence that time spent in more natural settings can have positive effects on psychological well-being and mental health (Bratman et al., 2021; Zhang et al., 2020). As urban populations continue to grow, such findings highlight one potentially important role of natural spaces in cities. Although the effects sizes of most of our findings were relatively small, the discrete changes in predicted probabilities presented in Section 3.4 demonstrate that even these small effects could have a positive impact. For city planners, the small differences that we detected for individuals can translate into larger differences for a large population. For every 100,000 tweets, for instance, a one percentage-point decrease in the probability of being negative results in 1,000 fewer negative tweets sent by people in the region.

4.2. Comparisons with previous research

The few studies examining the relationship between sentiment expressed in tweets and natural environments have primarily focused on urban parks and green spaces. The negative correlation between parks and the probability of having negative sentiment detected in our study is generally in line with previous studies suggesting a positive association between subjective sentiment expressed in tweets and visits to urban parks (Lim et al. 2018, Plunz et al. 2019, Roberts et al. 2019, Schwartz et al. 2019). Lim et al. (2018), for instance, found that tweets sent in green spaces expressed less negative emotion and more positive emotion compared to tweets sent outside green spaces. In contrast to Lim et al. (2018) and other previous studies, however, we do not find any evidence of tweets sent from parks being more likely to be positive across commercial/mixed, industrial, and residential areas of the city, after controlling for factors such as day of week and weather. Rather, our results even suggest a negative effect of large natural parks in residential areas on the probability of being positive. Among the previous studies mentioned above, Plunz et al. (2019) is the only one that also found a negative association between sentiment expressed in tweets and being within parks, in Manhattan, NY, where daily average of sentiment expressed in parks was less positive than those expressed in non-park areas.

Although we have attempted to compare our work to previous studies, it is often difficult to make direct comparisons among passive crowd-sourcing studies due to the heterogeneous methods applied across the relatively small number of existing studies. This is particularly true about the different methods used to measure and classify the sentiment of tweets. Plunz et al. (2019), for instance, used naive Bayes logistic regression with embedding features to assign each tweet a sentiment score (1 = positive; 0 = neutral; -1 = negative). Lim et al. (2018) computed eight scores, each corresponding to one sentiment category introduced by Plutchik's theory of emotions, and also the positive score, negative score, and polarity for each tweet. Schwartz et al. (2019) employed a bag-of-words approach to calculate sentiment for hourly bins of tweets, instead of for individual tweets. In Roberts et al. (2019), tweets were manually assigned to the three sentiment categories (positive/neutral/negative), then positive and negative tweets were further categorized into eight classes based on Ekman's six basic emotions and other previous research. In this study we classified tweets into the three sentiment categories using VADER.

The second type of heterogeneity among studies is related to how sentiment scores are analyzed. Previous studies rely mainly on descriptive statistics and t-tests. To our knowledge, our study is one of the first to construct statistical models describing how variability in sentiment expressed in tweets is related to nature while adjusting for multiple covariates. Our results suggest that previous studies making simple comparisons of sentiment inside versus outside of urban parks are likely missing important context and the role of covariates such as weather and day of week that are necessary to interpret the effects of parks. Furthermore, given that we find that the effect of parks on visitors' sentiment varies depending on how their location is zoned (e.g., industrial, residential), it may be inappropriate and ineffective to draw too many conclusions from comparisons with earlier studies.

The positive connections that we find between sentiment and some of the natural land-cover types are generally consistent with previous studies using active crowd-sourced data (e.g., MacKerron & Mourato,

2013) and survey data (Alcock et al., 2015). Nonetheless, contrary to expectations, we do not detect higher probabilities of being positive in natural land-cover types. These results most likely differ from previous studies because of our use of categorical compound sentiment scores. Instead of finding that tweets are likely to be higher on a continuous index of sentiment, our analytical approach allowed us to disentangle instances in which tweets were less likely to be negative from those in which tweets were more likely to be positive. Additionally, tweets from natural wetlands are more likely to be negative within industrial areas of the city (based on a small sample of 53 tweets). This could be a result of the small sample size or the condition of industrial wetlands.

Methodological differences may also contribute to the seeming inconsistency between our results and existing research that has found broad psychological benefits across all types of natural environments. Our passive sampling was completely non-intrusive: the data were derived from subjects who did not choose to participate with knowledge of the research. By contrast, most studies based on survey data or actively crowd-sourced data provide participants with information about the goals and research questions concerning the psychological effects of green spaces. The possibility that a large proportion of data contributors chose to participate because they enjoy spending time in nature can lead to selection bias. It is also possible that knowing the purpose of the study primes the participants to respond in a particular way (Orne, 1962). Therefore, it is not surprising that this study (and other passive crowd-sourcing studies) might reveal weaker positive relationships between sentiment and the natural environment.

4.3. Implications, limitations and future studies

This study demonstrates the potential to use passive crowd-sourced data such as tweets to investigate the association between people's momentary subjective sentiment and various surrounding environments with wide temporal and spatial coverage. The development of technologies for GPS positioning, remote sensing, and natural language processing creates new ways for researchers to explore links between nature and psychological well-being, and may help inform city planning and land management. Compared to traditional survey data, tweets sent from various environments are easier to obtain and more cost-effective. Large sample sizes (millions of observations) provide more power to detect and quantify nuanced, complex, and small effects. For city planners, the small effects that we detected—such as the small changes in predicted probabilities shown in section 3.4—may translate into important differences for a large population. For every 100,000 tweets, for instance, a one percentage-point increase in the probability of being negative results in 1000 more negative tweets by people in the region. Furthermore, with timestamps and geolocations, accurate information on potential confounders such as weather conditions and specific data on the physical environment, such as land-cover types and tree-canopy coverage, can be linked to each tweet and accounted for in statistical analyses. Compared to actively crowd-sourced data that shares many of the strengths mentioned above, tweets are not subject to the same demand effects (Orne, 1962), recall bias, and the selection bias as subjects who are recruited according to their interests (Bubalo et al., 2019).

Despite the strengths of passively crowd-sourced data, there are some limitations. Twitter users who send geolocated tweets are likely not representative of the whole population, particularly older individuals (Wojcik & Hughes, 2019). The lack of demographic information on the users hinders our ability to understand the representativeness of the data and to account for some other potential tweet-level confounders such as users' income levels and health conditions that can change over time. The lack of demographic factors in our study may explain why some of our results were different from some previous studies, some of which were able to adjust for demographic factors in their analyses. As a second limitation, tweets are neither anonymous nor direct reports of users' sentiment. Sentiment expressed publicly on social media platforms may not reflect people's true emotional states. Also, because the tweets are rarely a direct report of sentiment, the language must be interpreted. Regardless of whether the interpretations are done manually or by natural language processing techniques such as VADER, high accuracy can be difficult to achieve due to the limited information provided by the short texts and lack of context (Hutto & Gilbert, 2014). A standard way to assess the performance of a sentiment analysis classifier such as VADER is to compare it with human raters, assuming that human raters' interpretation of the crowd-sourced data is correct. Nonetheless, how well human raters interpret the passively crowd-sourced data such as tweets is still an unanswered question. One possible way to investigate this question would be to compare human raters' interpretations with Twitter users' self-reported answers, which could be obtained by developing a mobile application that asks participants to report their sentiment at the same time that they send a tweet. Despite the potential limitations, the effects of covariates such as weather generally match our expectations, increasing our confidence in the estimated relationships between sentiment and nature.

We assume in this study that the likelihood of sending a tweet is not significantly impacted by people's sentiment and surrounding environments. This assumption is a necessary condition for us to approximate the relationship between environments and people's sentiment using the estimated relationship between environments and sentiment expressed in tweets. Although to our knowledge there is no evidence suggesting that this assumption is wrong, it is still important to acknowledge that what we detect based on Twitter data is the relationship between natural environments and the probability of tweets being positive or negative, instead of the probability of people feeling positive or negative.

5. Conclusions

Our results suggest that the relationship between natural environments and peoples' sentiment can be complex. How humans respond to their surrounding environments depends on various factors including but not limited to the degree of naturalness and greenness. Additionally, the general type of naturalness, the amount of tree canopy, how well the green spaces are maintained, and the ways in which people use the space are also likely to be important. It is, therefore, not enough to know that more green space in a city will generally benefit people. We must know under what circumstances and in what contexts those benefits are manifest. Our nuanced understanding based on geolocated social media data has the potential to inform urban planning and land management. More in-depth analyses that consider multiple factors and span urban environments are needed.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Natural Capital Project and The Nature Conservancy. The authors would like to thank Sunny Jardine for her contribution.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.landurbplan.2022.104539.

References

Alcock, I., White, M. P., Lovell, R., Higgins, S. L., Osborne, N. J., Husk, K., & Wheeler, B. W. (2015). What accounts for 'England's green and pleasant land'? A panel data analysis of mental health and land cover types in rural England. *Landscape*

and Urban Planning, 142, 38–46. https://doi.org/10.1016/j.landurbplan.2015.05.008
Althoff, T., Clark, K., & Leskovec, J. (2016). Large-scale analysis of counseling conversations: An application of natural language processing to mental health. *Transactions of the Association for Computational Linguistics, 4*, 463–476. https://doi.org/10.1162/tacl_a_00111
Archer, K. J., Hedeker, D., Nordgren, R., & Gibbons, R. (n.d.). *mixor: An R Package for Longitudinal and Clustered Ordinal Response Modeling*. https://cran.r-project.org/web/packages/mixor/vignettes/mixor.pdf.
Becken, S., Stantic, B., Chen, J., Alaei, A. R., & Connolly, R. M. (2017). Monitoring the environment and human sentiment on the Great Barrier Reef: Assessing the potential of collective sensing. *Journal of Environmental Management, 203*, 87–97. https://doi.org/10.1016/j.jenvman.2017.07.007
Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological), 57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x
Berman, M. G., Jonides, J., & Kaplan, S. (2008). *The cognitive benefits of interacting with nature*. Psychological Science.
Berman, M. G., Kross, E., Krpan, K. M., Askren, M. K., Burson, A., Deldin, P. J., Kaplan, S., Sherdell, L., Gotlib, I. H., & Jonides, J. (2012). Interacting with nature improves cognition and affect for individuals with depression. *Journal of Affective Disorders, 140*(3), 300–305. https://doi.org/10.1016/j.jad.2012.03.012
Berto, R. (2005). Exposure to restorative environments helps restore attentional capacity. *Journal of Environmental Psychology, 25*(3), 249–259. https://doi.org/10.1016/j.jenvp.2005.07.001
Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science, 2*(1), 1–8. https://doi.org/10.1016/j.jocs.2010.12.007
Bratman, G. N., Anderson, C. B., Berman, M. G., Cochran, B., de Vries, S., Flanders, J., … Daily, G. C. (2019). Nature and mental health: An ecosystem service perspective. *Science. Advances, 5*(7), eaax0903. https://doi.org/10.1126/sciadv.aax0903
Bratman, G. N., Daily, G. C., Levy, B. J., & Gross, J. J. (2015a). The benefits of nature experience: Improved affect and cognition. *Landscape and Urban Planning, 138*, 41–50. https://doi.org/10.1016/j.landurbplan.2015.02.005
Bratman, G. N., Hamilton, J. P., Hahn, K. S., Daily, G. C., & Gross, J. J. (2015b). Nature experience reduces rumination and subgenual prefrontal cortex activation. *Proceedings of the National Academy of Sciences, 112*(28), 8567–8572. https://doi.org/10.1073/pnas.1510459112
Bratman, G. N., Olvera-Alvarez, H. A., & Gross, J. J. (2021). The affective benefits of nature exposure. *Social and Personality Psychology Compass*, *15*(8), e12630.
Browning, M. H. E. M., Saeidi-Rizi, F., McAnirlin, O., Yoon, H., & Pei, Y. (2020). The role of methodological choices in the effects of experimental exposure to simulated natural landscapes on human health and cognitive performance: A systematic review. *Environment and Behavior*, 0013916520906481.
Bubalo, M., van Zanten, B. T., & Verburg, P. H. (2019). Crowdsourcing geo-information on landscape perceptions and preferences: A review. *Landscape and Urban Planning, 184*, 101–111. https://doi.org/10.1016/j.landurbplan.2019.01.001
Building Outlines 2015. (2018). [Map]. City of Seattle GIS Program. Retrieved August 1, 2020 from https://data-seattlecitygis.opendata.arcgis.com/datasets/building-outlines-2015?geo metry=-122.335%2C47.616%2C-122.313%2C47.621.
Cox, D. T. C., Shanahan, D. F., Hudson, H. L., Fuller, R. A., & Gaston, K. J. (2018). The impact of urbanisation on nature dose and the implications for human health. *Landscape and Urban Planning, 179*, 72–80. https://doi.org/10.1016/j.landurbplan.2018.07.013
Current Land Use Zoning Detail. (2020). [Map]. City of Seattle GIS Program. Retrieved March 1, 2020 from https://data-seattlecitygis.opendata.arcgis.com/datasets/current-land-use-zoning-detail?geometry=-123.092%2C47.454%2C-121.584%2C47.778&orderBy=CATEGORY_DESC&selectedAttribute=CLASS_DESC.
Dadvand, P., de Nazelle, A., Triguero-Mas, M., Schembari, A., Cirach, M., Amoly, E., Figueras, F., Basagaña, X., Ostro, B., & Nieuwenhuijsen, M. (2012). Surrounding greenness and exposure to air pollution during pregnancy: An analysis of personal monitoring data. *Environmental Health Perspectives, 120*(9), 1286–1290. https://doi.org/10.1289/ehp.1104609
Dewitz, J. (2019). *National Land Cover Database (NLCD) 2016 Products: U.S. Geological Survey data release*. 10.5066/P96HHBIE.
Donahue, M. L., Keeler, B. L., Wood, S. A., Fisher, D. M., Hamstead, Z. A., & McPhearson, T. (2018). Using social media to understand drivers of urban park visitation in the Twin Cities, MN. *Landscape and Urban Planning, 175*, 1–10. https://doi.org/10.1016/j.landurbplan.2018.02.006
Frumkin, H., Bratman, G. N., Breslow, S. J., Cochran, B., Kahn, P. H., Jr, Lawler, J. J., Levin, P. S., Tandon, P. S., Varanasi, U., Wolf, K. L., & Wood, S. A. (2017). Nature contact and human health: A research agenda. *Environmental Health Perspectives, 125*(7), Article 075001. https://doi.org/10.1289/EHP1663
Gariepy, G., Kaufman, J. S., Blair, A., Kestens, Y., & Schmitz, N. (2015). Place and health in diabetes: The neighbourhood environment and risk of depression in adults with Type 2 diabetes. *Diabetic Medicine, 32*(7), 944–950. https://doi.org/10.1111/dme.12650
Go, A., Bhayani, R., & Huang, L. (2009). *Twitter Sentiment Classification using Distant Supervision*.
Golder, S. A., & Macy, M. W. (2011). Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science, 333*(6051), 1878–1881. https://doi.org/10.1126/science.1202775
Hamstead, Z. A., Fisher, D., Ilieva, R. T., Wood, S. A., McPhearson, T., & Kremer, P. (2018). Geolocated social media as a rapid indicator of park visitation and equitable park access. *Computers, Environment and Urban Systems, 72*, 38–50. https://doi.org/10.1016/j.compenvurbsys.2018.01.007
Hartig, T., Evans, G. W., Jamner, L. D., Davis, D. S., & Gärling, T. (2003). Tracking restoration in natural and urban field settings. *Journal of Environmental Psychology, 23*(2), 109–123. https://doi.org/10.1016/S0272-4944(02)00109-3
Hartig, T., Mitchell, R., de Vries, S., & Frumkin, H. (2014). Nature and health. *Annual Review of Public Health, 35*(1), 207–228. https://doi.org/10.1146/annurev-publhealth-032013-182443
Hartig, T., van den Berg, A. E., Hagerhall, C. M., Tomalak, M., Bauer, N., Hansmann, R., Ojala, A., Syngollitou, E., Carrus, G., van Herzele, A., Bell, S., Podesta, M. T. C., & Waaseth, G. (2011). Health benefits of nature experience: Psychological, social and cultural processes. In K. Nilsson, M. Sangster, C. Gallis, T. Hartig, S. de Vries, K. Seeland, & J. Schipperijn (Eds.), *Forests, trees and human health* (pp. 127–168). Netherlands: Springer. https://doi.org/10.1007/978-90-481-9806-1_5.
Hutto, C. J., & Gilbert, E. (2014). *VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text*. 10.
Integrated Global Surface Hourly Data. (n.d.). NOAA National Centers for Environmental Information. Retrieved August 13 2019 from https://www7.ncdc.noaa.gov/CDO/cdopoemain.cmd?datasetabbv=DS3505&countryabbv=&georegionabbv=&resolution=40.
Kaplan, R., & Kaplan, S. (1989). *The experience of nature: A psychological perspective* (pp. xii, 340). Cambridge University Press.
Kaplan, S. (1995). The restorative benefits of nature: Toward an integrative framework. *Journal of Environmental Psychology, 15*(3), 169–182. https://doi.org/10.1016/0272-4944(95)90001-2
Kellert, S., & Wilson, E. (1995). *The biophilia hypothesis*. Island Press.
Kim, Y. B., Kim, J. G., Kim, W., Im, J. H., Kim, T. H., Kang, S. J., & Kim, C. H. (2016). Predicting fluctuations in cryptocurrency transactions based on user comments and replies. *PLoS ONE, 11*(8), e0161197.
Lampos, V., & Cristianini, N. (2010). Tracking the flu pandemic by monitoring the social web. In *2010 2Nd International Workshop on Cognitive Information Processing* (pp. 411–416). https://doi.org/10.1109/CIP.2010.5604088
Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. JSTOR. 10.2307/2529310.
Lim, K. H., Lee, K. E., Kendal, D., Rashidi, L., Naghizade, E., Winter, S., & Vasardani, M. (2018). The grass is greener on the other side: Understanding the effects of green spaces on twitter user sentiments. *Companion of the The Web Conference 2018 on The Web Conference 2018 - WWW '18*, 275–282. 10.1145/3184558.3186337.
Lin, M., Lucas, H. C., & Shmueli, G. (2013). Research commentary - too big to fail: Large samples and the p-value problem. *Information Systems Research, 24*, 906–917. https://doi.org/10.1287/isre.2013.0480
Maas, J., van Dillen, S. M. E., Verheij, R. A., & Groenewegen, P. P. (2009). Social contacts as a possible mechanism behind the relation between green space and health. *Health & Place, 15*(2), 586–595. https://doi.org/10.1016/j.healthplace.2008.09.006
MacKerron, G., & Mourato, S. (2013). Happiness is greater in natural environments. *Global Environmental Change, 23*(5), 992–1000. https://doi.org/10.1016/j.gloenvcha.2013.03.010
Marselle, M., Irvine, K., & Warber, S. (2013). Walking for well-being: Are group walks in certain types of natural environments better for well-being than group walks in urban environments? *International Journal of Environmental Research and Public Health, 10*(11), 5603–5628. https://doi.org/10.3390/ijerph10115603
McMahan, E. A., & Estes, D. (2015). The effect of contact with natural environments on positive and negative affect: A meta-analysis. *The Journal of Positive Psychology, 10*(6), 507–519. https://doi.org/10.1080/17439760.2014.994224
Orne, M. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist, 17*, 776–783. https://doi.org/10.1037/h0043424
Plunz, R. A., Zhou, Y., Carrasco Vintimilla, M. I., Mckeown, K., Yu, T., Uguccioni, L., & Sutto, M. P. (2019). Twitter sentiment in New York City parks as measure of well-being. *Landscape and Urban Planning, 189*, 235–246. https://doi.org/10.1016/j.landurbplan.2019.04.024
Rabe-Hesketh, S., & Skrondal, A. (2012). *Multilevel and longitudinal modeling using Stata* (3rd ed, Vol. 2nd). Stata Press Publication.
Ribeiro, F. N., Araújo, M., Gonçalves, P., Gonçalves, M. A., & Benevenuto, F. (2016). SentiBench—A benchmark comparison of state-of-the-practice sentiment analysis methods. *EPJ Data Science, 5*(1), 1–29. https://doi.org/10.1140/epjds/s13688-016-0085-1
Richardson, E. A., Pearce, J., Mitchell, R., & Kingham, S. (2013). Role of physical activity in the relationship between urban green space and health. *Public Health, 127*(4), 318–324. https://doi.org/10.1016/j.puhe.2013.01.004
Roberts, H., Sadler, J., & Chapman, L. (2019). The value of Twitter data for determining the emotional responses of people to urban green spaces: A case study and critical evaluation. *Urban Studies, 56*(4), 818–835. https://doi.org/10.1177/0042098017748544
Schwartz, A. J., Dodds, P. S., O'Neil-Dunne, J. P. M., Danforth, C. M., & Ricketts, T. H. (2019). Visitors to urban greenspace have higher sentiment and lower negativity on Twitter. *People and Nature, pan3.10045*. https://doi.org/10.1002/pan3.10045
Seattle Parks. (2012). [Map]. City of Seattle GIS Program. Retrieved August 26, 2019 from https://data-seattlecitygis.opendata.arcgis.com/datasets/seattle-parks.
Soga, M., & Gaston, K. J. (2016). Extinction of experience: The loss of human–nature interactions. *Frontiers in Ecology and the Environment, 14*(2), 94–101. https://doi.org/10.1002/fee.1225
Steele, F. (2009). *Module 7: Multilevel Models for Binary Responses*. University of Bristol, Centre for Multilevel Modelling. https://www.cmm.bris.ac.uk/lemma/pluginfile.php/2281/mod_resource/content/2/C7%20Multilevel%20models%20for%20binary%20responses.pdf.

Sullivan, G. M., & Feinn, R. (2012). Using effect size—Or why the P value is not enough. *Journal of Graduate Medical Education, 4*(3), 279–282. https://doi.org/10.4300/JGME-D-12-00156.1

Thelwall, M., Buckley, K., & Paltoglou, G. (2012). Sentiment strength detection for the social web. *Journal of the American Society for Information Science and Technology, 63*(1), 163–173. https://doi.org/10.1002/asi.21662

Ulrich, R. S., Simons, R. F., Losito, B. D., Fiorito, E., Miles, M. A., & Zelson, M. (1991). Stress recovery during exposure to natural and urban environments. *Journal of Environmental Psychology, 11*(3), 201–230. https://doi.org/10.1016/S0272-4944(05)80184-7

United Nations, Department of Economic and Social Affairs, & Population Division. (2019). *World urbanization prospects: The 2018 revision*. New York: United Nations.

van den Berg, A. E., Maas, J., Verheij, R. A., & Groenewegen, P. P. (2010). Green space as a buffer between stressful life events and health. *Social Science & Medicine, 70*(8), 1203–1210. https://doi.org/10.1016/j.socscimed.2010.01.002

Wang, X., Rodiek, S., Wu, C., Chen, Y., & Li, Y. (2016). Stress recovery and restorative effects of viewing different urban park scenes in Shanghai, China. *Urban Forestry & Urban Greening, 15*, 112–122. https://doi.org/10.1016/j.ufug.2015.12.003

White, M. P., Alcock, I., Grellier, J., Wheeler, B. W., Hartig, T., Warber, S. L., Bone, A., Depledge, M. H., & Fleming, L. E. (2019). Spending at least 120 minutes a week in nature is associated with good health and wellbeing. *Scientific Reports, 9*(1), 7730. https://doi.org/10.1038/s41598-019-44097-3

White, M. P., Alcock, I., Wheeler, B. W., & Depledge, M. H. (2013). Would you be happier living in a greener urban area? A fixed-effects analysis of panel data. *Psychological Science, 24*(6), 920–928. https://doi.org/10.1177/0956797612464659

Wilson, E. (1984). *Biophilia*. Harvard University Press.

Wojcik, S., & Hughes, A. (2019, April 24). How Twitter Users Compare to the General Public. *Pew Research Center: Internet, Science & Tech*. https://www.pewresearch.org/internet/2019/04/24/sizing-up-twitter-users/.

Wood, S. A., Guerry, A. D., Silver, J. M., Lacayo, M., Guerry, A. D., Silver, J. M., & Lacayo, M. (2013). Using social media to quantify nature-based tourism and recreation. *Scientific Reports, 3*(1), 2976. https://doi.org/10.1038/srep02976

Zhang, Y., Mavoa, S., Zhao, J., Raphael, D., & Smith, M. (2020). The Association between green space and adolescents' mental well-being: A systematic review. *International Journal of Environmental Research and Public Health*, *17*(18), 6640.