On the relationship between crowdsourced sentiments and mobility trends during COVID-19: A case study of Kyoto

Wenzhe Sun[a,*], Hironori Kobayashi[a], Satoshi Nakao[a], and Jan-Dirk Schmöcker[a]

[a]Department of Urban Management, Graduate School of Engineering, Kyoto University, Kyoto, Japan

*Corresponding author, Wenzhe Sun
Email address: wz.sun@trans.kuciv.kyoto-u.ac.jp

ABSTRACT

COVID-19 has significantly changed people's attitudes and behavior in Japan. In this study, we firstly conducted a sentiment analysis on the contents of tweets posted on Twitter to understand the changes in people's emotions due to the COVID-19 pandemic and related restriction policies. ML-ask, an open-source analysis system for textual input in Japanese, is used for this sentiment analysis. Its emotion identification is lexicon-based and can identify ten different emotions including joy, sadness, dislike, and anger. Secondly, we investigated the impacts of such "crowdsourced" sentiments on mobility and activity participation by regression models. Publicly available, also crowdsourced, Google's COVID-19 Community Mobility Reports are used as dependent variables, and the identified tweet sentiments are used as independent variables together with other conventional variables such as dummies for policies, weekends, and holidays. As a result, it was confirmed that the changes in mobility and activity participation could be explained more accurately by the counts of tweet sentiments. The model fit was significantly improved if sentiments are added as variables, compared to the cases in which only the number of tweets or the number of daily new cases was used. In addition, we test the effect of time lag. More specifically, the model fit was higher when using the averaged emotion counts of the past seven days than using the count of emotions on the same day. We conclude that using tweet sentiments offers the feasibility to improve the prediction levels of urban mobility.

Keywords: Tweet sentiment, Emotion identification, Mobility and activity, COVID-19, Crowdsourced data

1. INTRODUCTION

COVID-19 has had a significant long-term impact worldwide. Governments were struggling to contain the spread of the virus with a variety of restriction policies including lockdowns, closure of public services, canceling of public gatherings, and so forth. The resulting human behavioral changes in mobility and activity participation are desired by policymakers. Some studies use surveys to understand the behavioral changes at a disaggregated level, such as the attitudes and stated or revealed behavior for a specific activity and travel mode (Abdullah et al., 2020; Parady et al. 2020; Shakibaei et al., 2021).

With the advancement in Information and Communication Technology (ICT) and the availability of crowdsourced big data, the aggregated behavioral changes have become more visible and traceable in time series. A large number of studies capture the aggregate mobility changes using urban big data analytics. Benjamins et al (2022) title their editorial "mobile big data in the fight against COVID-19". Exact location data of individuals from mobile phones have been at the forefront to capture mobility changes. An exemplary study we refer to is Hu et al. (2021). They use mobile device location data to track the mobility changes and measure the impacts of COVID-19 restriction policies.

For activity participation, an important source for a number of researchers has been Google Location History (GLS). Using the data from the users that opt-in to GLS, Google provides aggregated indicators of human mobility and activity participation via two services: COVID-19 Community Mobility Reports (CMRs) for many countries and "Popular times" for selected places of a city (Google, 2020a; Google, 2020b). CMRs provide the percentage changes of visits compared to a baseline before COVID-19 for activity categories such as retail and recreation, workplace, and parks. Google Popular Times information furthermore indicates the real-time business of a venue. Sun et al. (2022) quantified the impacts of the policies on people's mobility and activity participation trends in Kyoto, Japan using Google's CMRs and a regression model with time series errors. Mahajan et al. (2021) demonstrate the potential of Google Popular Times in indicating the changes in visits to a specific venue and therefore the changes in activity participation before and after COVID-19 lockdowns in Munich, Germany. These emerging datasets of globally available services offer worldwide policymakers the opportunity of observing the impacts of the pandemic and the policies they impose over time. However, for the measurement of policy impacts and the prediction of mobility and activity trends, these data sources do not provide a direct explanation as to what motivates behavioral changes. In particular, whether fear of COVID or the policy itself drives the change cannot be distinguished.

Examining people's sentiments and social concerns across the different stages of the pandemic will provide useful information for predicting the social impact of future quarantine measures. Some studies have been analyzing the content of social networking service (SNS) postings. Among them, Quercia et al. (2012) suggested that tweets have real-time self-reporting functions that can partially reveal the subjective well-being or happiness of people. SNS data are expected to analyze and understand social concerns and "city mood" about such events with profound impacts such as COVID-19. Compared to questionnaires, tweet sentiments can capture the temporal dynamics of a wide range of population groups. Although some studies notice the possibility of tweet data in analyzing people's sentiments during COVID-19, this paper puts the focus on the relationship between tweet sentiments and mobility/activity indicators. More specifically, it incorporates tweet sentiments as key independent variables to explain dependent indicators of mobility and activity participation in Google's CMRs.

The contribution of this paper is two-fold. Firstly, it explores the feasibility of identifying emotions from tweet data in Japanese. It investigates the reliability of the results by checking their consistency with COVID-19 infections. Secondly, it shows the usefulness of crowdsourced tweet sentiments in explaining the future trends of aggregated crowdsourced mobility and activity participation. It finds that better performance can be achieved if massively collected tweets are refined by such keywords as "mobility" and lagged tweet sentiments are used. We also note that this study proposes a transferable methodology by mapping the data of two globally available services. That is, the Google and Twitter data used in this study can be downloaded and collected by researchers in any city.

The remainder of this paper is as follows. Section 2 provides a literature review of studies using tweet data to understand the impact of a disaster or this COVID-19 pandemic on human behavior. Section 3 explains the methodology and the details to obtain tweet sentiments. We include data collection in this section to imply that creating a tweet database is a key step. Section 4 proposes the regression models using tweet sentiments and external policy variables to forecast mobility trends. Section 5 reports the results of the sentiment analysis and regression analysis. We compare the model fit and model interpretability of the regression models with different combinations of independent variables to present the usefulness of tweet sentiments in explaining the trends of human mobility and activity participation during the COVID-19 pandemic. Finally, we discuss key findings and future research directions in Section 6.

2. LITERATURE REVIEW

Among the social network services (SNSs), Twitter, in particular, has been attracting attention as a research target and various findings have been obtained because the number of users has been rapidly increasing and user behavior can be observed in real-time. Furthermore, the APIs provided by Twitter allow researchers to collect a sample of the Tweets as we will discuss in more detail later.

The posts on Twitter, also known as Tweets, in general have two functions: Information sharing and self-reporting. Several studies focused on the reliability of their function of information gathering and sharing. Castillo et al. (2011) proposed a method to calculate the reliability of tweets based on the length and subject of the posted tweets and the information of the user who posted the tweet, and discussed the reliability of information flowing on Twitter. Suda et al. (2013) attempted to estimate whether the Tweet information was a hoax or not by analyzing approximately 300 million tweets posted at the time of the Great East Japan Earthquake, with the analysis results as features. They extracted tweets that were retweeted frequently and narrowed them down to estimate hoaxes by clustering them using retweet depth information and emotional polarity. In Japan, Twitter usage during disasters, especially in the Great East Japan Earthquake, has been studied. Miura (2012) analyzed Twitter communication during the Great East Japan Earthquake. Miura showed the characteristics of human information behavior during a major disaster through the analysis of Twitter logs. It was shown that users disclosed a lot of their own strong feelings of anxiety and tried to share a lot of the information they obtained with others. Kazama et al. (2012) proposed a method to analyze the similarity of temporal changes in the frequency of words appearing in tweets by using the Earth Mover's Distance (EMD) method. In addition, text mining was conducted on tweets during the earthquake disaster to analyze the time-series changes in the two words "earthquake" and "nuclear power plant", their related nouns, and the relationships between them, to clarify the impact of actual incidents and other events on Twitter. The usefulness of Twitter data in times of disaster has thus been discussed.

More relevant to this paper is the function of self-reporting as indicated by Quercia et al. (2012). They found a positive correlation between the sentiment scores obtained from tweets and socio-demographic well-being levels at the community level. The sentiment score of an individual is based on comparing the number of positive words over negative words in the tweets of the user's profile. Furthermore, Quercia et al. (2011) captured the relationship between personality and the publicly available profile information of a user such as following, followers, and listed counts. Cui and He (2021) inferred users' socio-demographic attributes such as age, gender, ethnicity, and education level using texts, hashtags, and emojis of tweets and Twitter's social network. Some studies evaluated whether tweets are a good indicator of one's attitude to an issue and the situation faced at the time of posting. Kabbani et al. (2022) used tweets to detect the incidents of public transit services in real-time. They successfully captured transit incidents of several categories such as delay and denied boarding and identified passengers' negative or positive attitudes to the incidents. Also using the tweets related to public transport, Haghighi et al. (2018) proposed topic models and sentiment analysis to understand the topics and attitudes of passengers regarding the service quality. Similarly, Collins et al (2013) captured transit users' sentiments with Twitter data and discuss that total tweets about the transit system and sentiments can be used as alternative performance measures.

The techniques of sentiment identification from textual inputs have matured in the past decade. For readers having an interest in the development of these techniques, the comprehensive review and comparison carried out by Ribeiro et al. (2016) will be beneficial. There are two mainstream methods: Lexicon Dictionary and Supervised Machine Learning. A combination of both also exists. Their comparison also includes paid software. They compared 24 methods over 18 datasets in terms of 2-class (positive, negative) and 3-class (positive, negative, neutral) identification. According to their comparison, VADER (Hutto and Gilbert, 2014) got the highest mean rank for 3-class tasks for all datasets and Umigon (Levallois, 2013) won most of the 3-class tasks for the datasets of social networks including tweets. SentiStrength (Thelwall et al. 2010) achieved the highest mean rank for 2-class tasks over all datasets and the dataset of social networks. Both Umigon and VADER are lexicon-based methods, and SentiStrength uses a combination of lexicon and machine learning. They also found that the performance can vary considerably from one dataset to another. For example, the 2-class accuracy of VADER drops from 99% of one tweet dataset to 84% of another. Besides, the 2-class accuracy is usually higher than the 3-class accuracy due to the difficulty of identifying neutral sentiments. In general, VADER, Umigon, and SentiStrength are consistently accurate for social media datasets. They are considered reliable techniques for short and informal texts such as social media posts. Bonta and Janardhan (2019) considered VADER as a favorable option for sentiment analysis using texts from social media as its lexicon is enhanced with a set of linguistic and grammatical rules, marks, and emojis. However, these tools are established and improved based on lexicons and rules of the English language. Similar comprehensive comparisons are not available for sentiment analysis tools using Japanese. ML-ask, a lexicon-based open-source software developed by Ptaszynski et al. (2017), is used in this paper and will be explained in Section 3.

So far, the utilization of tweet data and derived sentiments in transportation studies mainly focuses on the measurement of user satisfaction on service quality. The usefulness of them for city-wide mobility forecasting and modeling appears to have not been explored much yet. In the past three years, two mainstream methodological approaches of using Tweet data to analyze human behavior responsive to the COVID-19 pandemic and related restriction policies can be found in the literature. One is topic analysis (Chandrasekaran et al., 2020; Mutanga et al., 2022), and the other is sentiment analysis (Chandrasekaran et al., 2020; Long et al., 2020; Yao et al., 2021). Chandrasekaran et al. (2020) and Long et al. (2020) both processed tweet data into Bag of Words (BoW) and extracted features from tweets. These features were used as input to a classification model to identify if a tweet is positive or negative. Naive Bayes, logistic regression, and neural networks are the tested classifiers. The ground truth was based on a judgment using a lexicon. For a tweet, if it contains more words defined as positive in the lexicon than those defined as negative, it is labeled as positive. Yao et al. (2021) further investigated the correlation between tweet sentiment and COVID-19 policies. They found that the tweet sentiments were positively correlated with stricter quarantine measures, this correlation is not significant with new cases and hospitalization. We conclude this section by noting that the utilization of these uncovered tweet sentiments for real-life prediction and policy-making is less noticed and explored. Therefore, in this paper, we conduct a sentiment analysis and then employ the derived sentiments to explain and forecast mobility trends.

3. DATA COLLECTION AND SENTIMENT ANALYSIS

To begin with, the overview of the methodology of this research is provided in Figure 1. The methodology can be separated into three steps. The first step is to collect data. Considering that the Google CMRs in Japan are available only at the prefecture level, the COVID infection cases and other external factor data are also aggregated or obtained for the same geographical unit. As Google CMRs and external factors are publicly available, the main obstacle is to collect tweet data. The details to collect tweet data will be elaborated in this section, followed by the second step of sentiment analysis. With the results obtained from the first and second steps, we can establish the regression models to analyze the impacts of policies and crowdsourced sentiments, as will be discussed in Section 4.

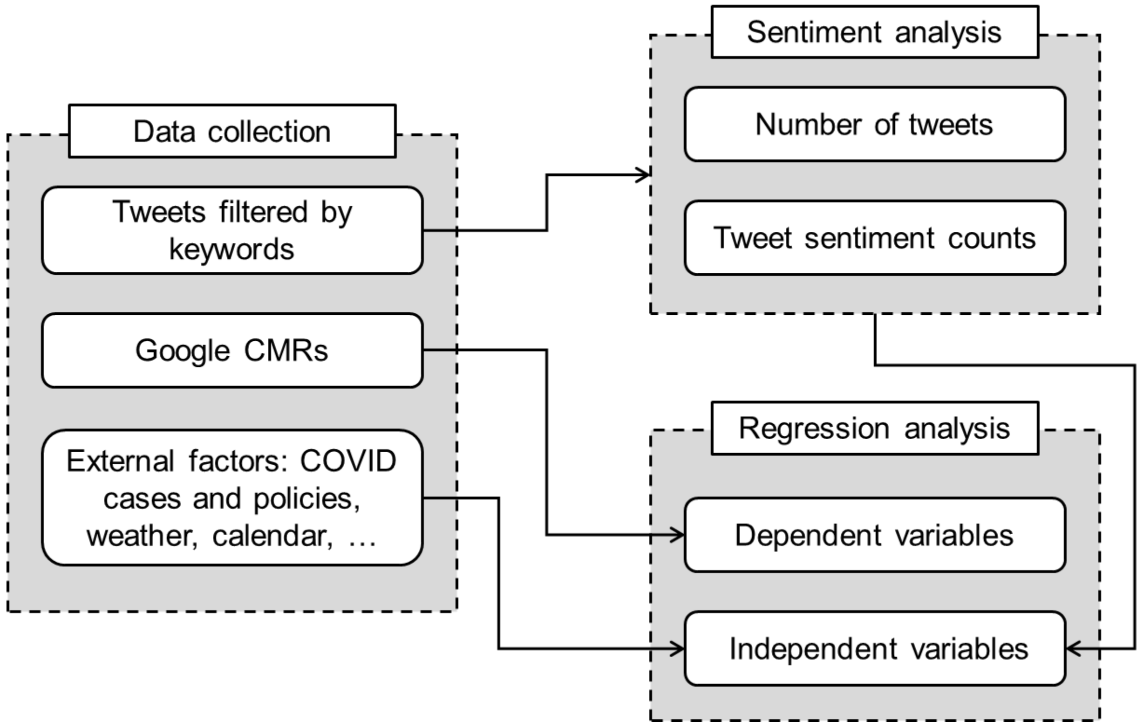

Figure 1 Flowchart of the methodology

3.1 Tweet data collection

Due to the potential data scale as well as the limitations imposed by Twitter API to collect data, it is essential to create a pipeline that only includes relevant tweets in the dataset. In order to capture the "city mood" and match it with the mobility indicator of this city, the tweets posted concerning the target city or which are posted by persons in this city are desired. Some studies collect tweets with geographical tags or "check-ins" at the postings. García-Palomares et al. (2018) used this type of tweet data to capture the connections between urban dynamics and land use. Hasan and Ukkusuri (2014, 2015) used Foursquare check-in data to describe the lifestyle of users and more generally the urban dynamics. Using geo-tags has the advantage that the derived tweet sentiments should be more in line with the dynamics of residents within the city. However, Bi et al. (2023) also found that the geo-tagged tweets in Kyoto City lack text and do not allow in-depth sentiment analysis. Another commonly-used effective way is to specify keywords. Keywords are highly dependent on the research purpose. In this study, we collected tweets that include the word "Kyoto (京都)", as we aim to create a collection of comprehensive tweets mentioning "Kyoto". Since Kyoto is a tourism city, we did not want to exclude the sentiments of people planning or considering to visit this city. This is another reason why we used keywords instead of geographical references to create the tweet datasets.

In this study, tweets containing specific keywords were extracted using the Twitter Search API. 100 tweets can be retrieved with a single request using the Twitter Search API. The request can be repeated, but there is a limit of 180 requests in 15 minutes. Therefore, 18,000 tweets can be retrieved at one time. We created a Python code to repeatedly retrieve tweets for the keyword "Kyoto" every hour considering that the number of tweets containing "Kyoto" posted every hour is, on average, 1000 tweets per hour and we never found more than 10,000 tweets in an hour during a pre-experiment. Therefore, it is sufficient that tweets for our keyword are retrieved only once every hour. The studied timeline for which data are collected and analyzed in this paper is from 28 June 2021 to 21 January 2022. For a few days during this period the Tweet collection was not complete due to network issues or auto-scheduled computer reboots.

Given the textual tweet database, one has the flexibility to refine the tweet dataset that can better serve the specific purpose. For example, the sentiments based on comprehensive tweets may convey too much irrelevant and unnecessary information to the regression model. For our purposes we only use a subset of the "Kyoto" tweets that are related to mobility. More specifically, tweets containing one of the following words: "going (行く)", "coming (来る)" or "sightseeing (観光)" within the "Kyoto" tweets are filtered in. We refer to this tweet dataset as "Kyoto & Mobility" tweets.

In order to map tweets and Google CMRs, which are daily data, we treated a day as missing data if tweets could not be retrieved for the whole day. We had 15 such days with missing data in the studied timeline. The missing data were estimated by taking the average of the tweet numbers and estimated sentiments on the previous and following day. Figure 2 shows the daily number of tweets containing the keyword. Strong weekly patterns can be observed and some weak seasonal ones. Autumn is a peak tourist time in Kyoto and the number of tweets is correspondingly increasing. As we aim to show in the following the relation between the derived tweet sentiments and the mobility during the COVID-19 pandemic, we further add daily new cases in Kyoto Prefecture to Figure 2 as a reference. It can be noticed that the number of comprehensive tweets barely presents a correlation to the number of daily new confirmed cases. The number of "Kyoto & Mobility" tweets shows a stronger negative correlation to the COVID cases. The correlation coefficient changes from -0.14 to -0.45 with only mobility-related tweets filtered in.

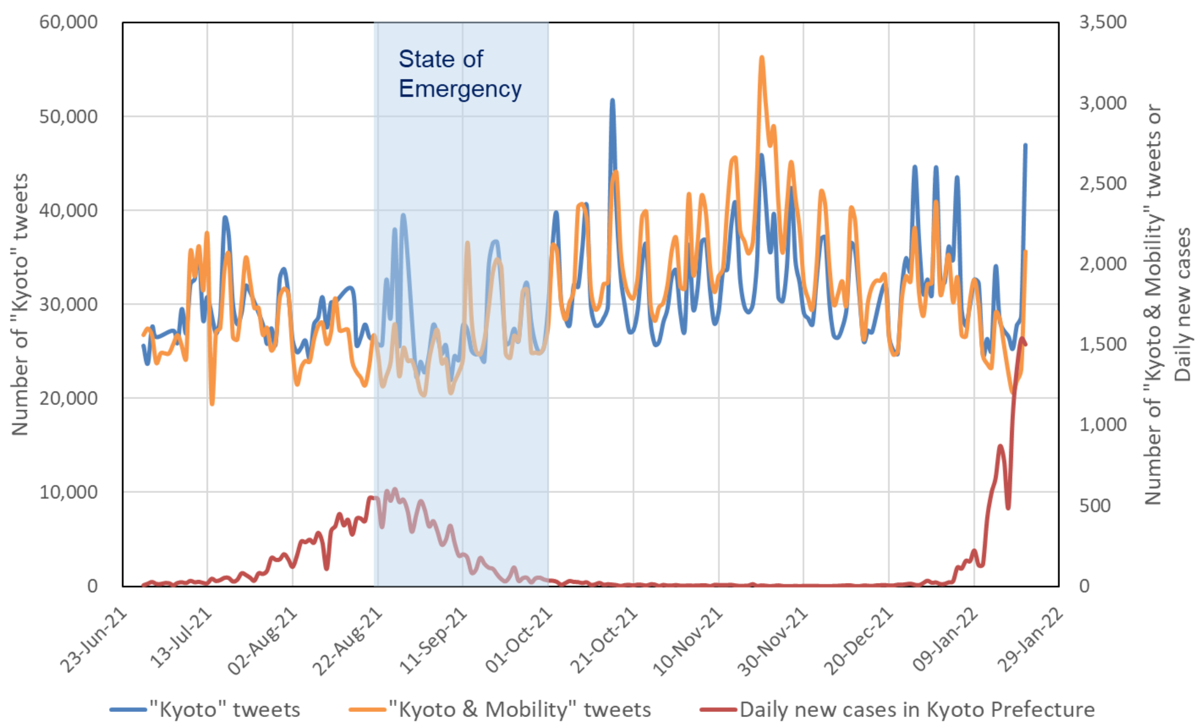

Figure 2 Daily number of “Kyoto” or “Kyoto & Mobility” tweets and new COVID-19 cases in Kyoto Prefecture

3.2 Sentiment analysis and emotion identification

Sentiment analysis was conducted on these extracted tweets. Ptaszynski et al. (2009, 2017) proposed ML-ask (eMotive eLement and Expression Analysis System), the first open-source affect analysis system for textual input in Japanese. The sentiment analysis conducted by ML-ask is based on a pre-determined lexicon. For one Japanese sentence, ML-ask attempts to capture the emotive expressions at first. If the sentence contains any of the keywords that are associated with a specific emotion type, it is classified accordingly. Ten types of emotion can be identified by ML-ask, which are Joy, Fondness, Relief, Sadness, Dislike, Anger, Fear, Shame, Excitement, and Surprise. The lexicon of emotive expressions used in ML-ask consists of 2,100 expressions, based on the Emotive Expression Dictionary edited by Nakamura (1993). The number of expressions per emotion type and examples of the expressions can be found in Ptaszynski et al. (2017). According to Ptaszynski et al. (2017), ML-ask incorporates the negation type of Contextual Valence Shifters that were proposed by Polanyi and Zaenen (2006). Negations such as “not” and “never” may reverse the semantic and sentiment orientation of the sentence so that a total of 108 negation structures are included in ML-ask to avoid wrongly identifying the emotion type.

Regarding the accuracy of ML-ask, the F1-score of binary emotive/non-emotive identification is 83% and that of emotion type identification is 47% in Ptaszynski et al. (2010). The results were obtained from textual inputs of separate sentences and conversations. Highly accurate identification performance was reported in Ptaszynski et al. (2014) by a sentiment analysis of Japanese blogs which contain similar short and informal characteristics to Japanese tweets. The emotiveness identification achieved 99% accuracy and emotion type identification was 73% correct for emotive sentences and 89% for all sentences. The F1-score of emotion type identification is 85% for all sentences. This level of accuracy is considered acceptable in this study.

In this paper, we use ML-ask to obtain the counts of emotive sentences per emotion in time series format by feeding the tweets collected every day. The time series for each emotion is then considered as an explanatory variable for the level of mobility.

4. REGRESSION ANALYSIS

In this section, using the results of the sentiment analysis conducted in the previous section and the "Transit Station" data from CMR, a multiple regression analysis was conducted to understand the degree of correlation between mobility and emotions of tweet contents.

4.1 Dependent variables

We aim to explain the time series of Kyoto CMR data. They are available for six activity/mobility categories: Retail & recreation, Grocery & pharmacy, Parks, Transit station, Workplace, and Residential. For many countries, the spatial resolution is at the level of regions and sub-regions. For Japan, the spatial resolution is to the prefecture level. The temporal resolution is the day. For each day since mid-February 2020, a percentage change in the visits to a specific category of places is provided by this data, and the percentage change is compared with a baseline shortly before the outbreak of the pandemic. More specifically, the published percentage change for a specific date is the difference in visits between this day and the average visits on the corresponding day of the week during the five weeks between 3 January and 6 February 2020. We will show the results for "Transit station" only as it is most closely related to travel that one might tweet and/or that one might not conduct due to negative sentiments such as COVID fears.

4.2 Independent variables

For comparison purposes, five different combinations of independent variables are considered. The base model only uses dummy variables to explain the changes in the visits to transit stations. Dummy variables include a holiday dummy, weekend dummy, rainy day dummy, and two policy dummies which consider two important non-pharmaceutical restriction policies implemented in Japan: Quasi-Emergency measures and State of Emergency. The second model uses the total number of tweets and dummies. The third model uses the number of daily new cases and dummies. COVID-19 cases, policy dummies, and weather variables are commonly used in studies modeling the travel demand during COVID-19, e.g. Lei and Ukkusuri (2022), Sun et al. (2022). The recommended models are the fourth and fifth models, in which tweet sentiments are used. These two models differ from each other in the usage of sentiment results. The fourth one uses the counts of sentiments and dummies while the fifth uses the counts of sentiments, daily new cases, and dummies. By examining the model fit improvements among the models, we can obtain insights into the added value of using tweet sentiments for explaining mobility trends.

Furthermore, for each model, we distinguish two patterns to use the non-dummy variables. “Real-time” models relate the unsmoothed independent variables to the dependent variable of the same day, and “Time lag” models always use the smoothed lagged data which is the moving average of the seven days before the day of the dependent variable. For example, the values for the number of tweets, the number of infected persons, and the number of emotions were averaged from 1 July to 7 July and fitted to the mobility change on 8 July. The implicit assumption of the time-lag formulation is that human behavior is the consequence of past short-term “city mood” instead of real-time sentiments.

As illustrated in Figure 1, two types of tweet variables are used. The first is the number of tweets. The second type includes different sentiment counts obtained from the sentiment analysis. Considering that the dependent variable is a relative change to a baseline in January 2020, the tweet variables are supposed to be also relative changes compared to the situation before COVID-19. However, our tweet collection only started after the outbreak of the pandemic. To overcome the data unavailability, we average the number of tweets and sentiment counts for the days having CMR percentage changes between -5% and 5%, to simulate the pre-COVID baseline. The tweet-related model inputs then are processed into relative values compared to the baseline. By doing so, the changes in the total number of tweets or sentiment counts are related to percentage changes in mobility. As the baselines of CMR and tweets do not perfectly match, an intercept should be included. The intercept also includes the long-term constant mobility change due to COVID-19, everything else be equal to the situation before the pandemic.

The logarithm of daily new cases is used for the independent variable of infection cases. It considers the non-linear effect of the number of cases on the perceived risk by people and human mobility behavior (Sun et al., 2022). The number of cases in the later wave is much larger than that in the first wave though, the negative impacts on mobility might not increase in proportion. Our results also show that the logarithm of daily new cases improves overall model fit and variable significance, compared to using raw daily new cases.

The “Holiday” dummy was set to 1 for national holidays and “three-day weekends” which are two normal weekend days followed by a national holiday on Monday. As an example, 9 August 2021 was a holiday in Japan so the dummy for the three days 7 to 9 August is set to 1. The “Holiday” dummy is also applied to two long holidays. For 13 August to 16 August, the summer holiday dummy is set to one, though it is not a public holiday. These days are associated with a religious event and the Japanese often take these days for a summer holiday. Secondly, 29 to 31 December are added to public holidays. Though only 1 to 3 January are public holidays, workplaces, schools, and universities are commonly closed on the last three days of the year. The “Weekend” dummy applies to the weekends not followed by a holiday.

Two types of COVID-19 policy dummies are employed in this paper. One is the “Quasi Emergency Measures” dummy. The periods from 28 June to 11 July and from 2 August to 19 August 2021, when Kyoto

Prefecture and Kyoto City requested their citizens to take measures to prevent the spread of the disease, are set to 1. This “Quasi Emergency” state was followed by the declaration of “State of Emergency” from 20 August to 30 September 2021, which imposed more strict restrictions in Kyoto Prefecture. “State of Emergency” is the strongest restriction policy in Japan, and this declaration was the fourth time over the Japanese COVID-19 timeline. This period then is marked by the “State of Emergency” dummy. Generally, during these periods, business hours of restaurants were shortened, some public facilities were closed and large events were canceled or only held with limited public participation. More details about the differences between “Quasi Emergency Measure” and “State of Emergency” can be found in Sun et al. (2022).

The “Rainy day” dummy has been introduced to consider that the number of visitors may be affected by the weather. In order to consider rainfall that is greater than the Japan Meteorological Agency's (JMA) standard for light rainfall, which is “rain that lasts for several hours but does not reach 1 mm”, this study uses data from JMA as a reference and assumes the dummy takes one on days when the daily rainfall in Kyoto City is 1.5 mm or more.

5. RESULTS

5.1 Sentiments obtained from tweets

Figure 3 shows the weekly moving average of sentences containing these sentiments. For comparison, the weekly moving averages for transit stations from CMRs are also illustrated in the figure. CMRs are used as dependent variables of our regression models. At a glance, we can observe that changes in the “Joy” emotion appear to coincide with the dynamics of the visits to a transit station. Besides, considering both Figures 2 and 3, the “Joy” sentiment appears to have some negative correlation with the number of daily new cases in Kyoto Prefecture. By giving a closer look at the text of “Kyoto” tweets identified as “Joy”, we can find many of them talk about the sightseeing experience, e.g. “It is fun just to stroll around the streets of Kyoto! (京都の街並みブラブラするだけで楽しい！)”. For the negative sentiments such as “fear” and “dislike”, we found that many of them mention crowdedness on public transport and in public places, especially during some holidays. Crowdedness will be always disliked but especially during the COVID pandemic.

The one-week moving average curves of the tweet sentiments about “Kyoto” and “Mobility” are shown in Figure 4. By comparing Figure 4 to Figure 3, the peaks of “Joy” and “Dislike” sentiments appear to be more obvious and their dynamics are more in line with the dynamics of the visits to transit stations. The reason why the “Dislike” sentiment becomes more correlated to the mobility trend is that including “Mobility” filters out the (in some cases ironic) tweets related to the stereotypes of Kyoto. For example, a tweet that is filtered out is “I'm afraid of the sarcasm of Kyoto people (京都の人の嫌味は怖いけど).” Therefore, in particular the negative sentiments are refined by adding the “Mobility” keywords.

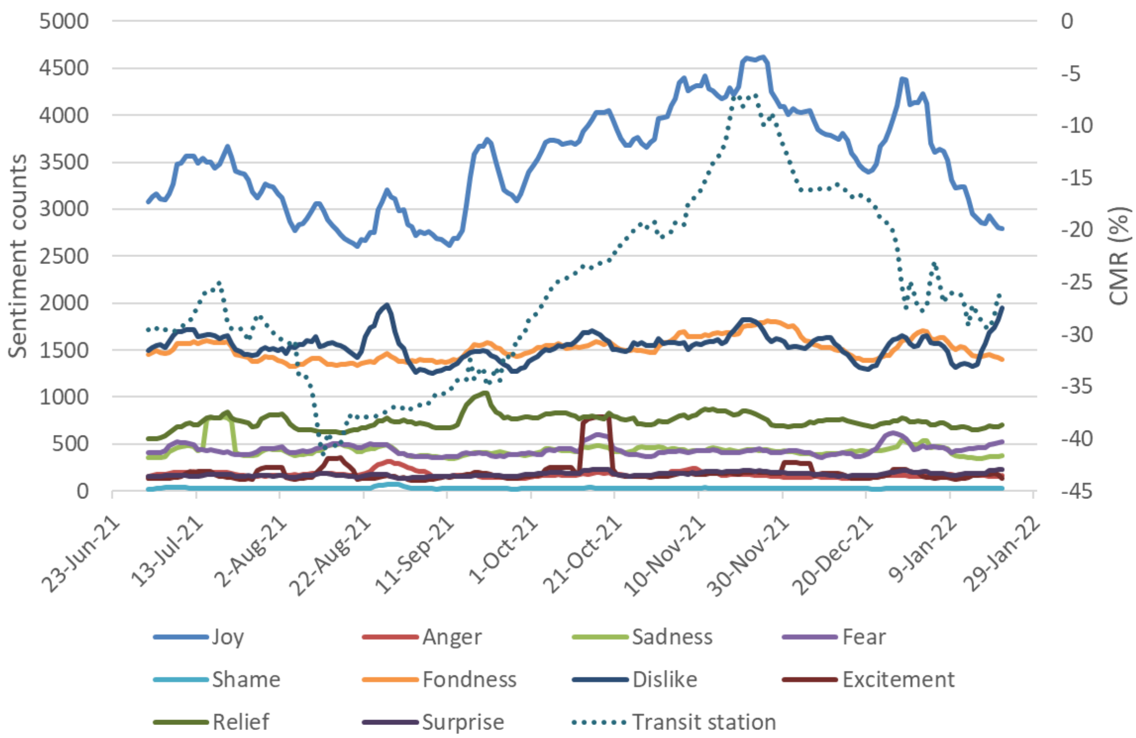

Figure 3 Weekly moving average of "Kyoto"-mentioned tweets containing sentiments

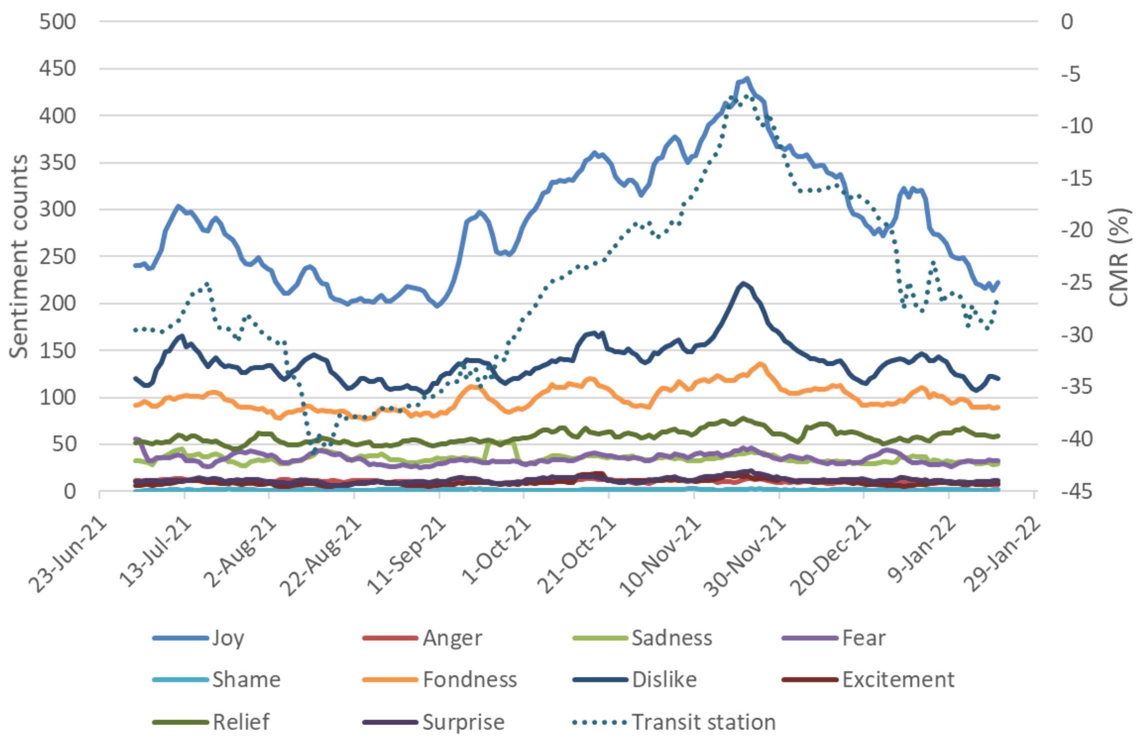

Figure 4 Weekly moving average of "Kyoto & Mobility"-mentioned tweets containing sentiments

5.2 Correlation analysis of the independent variables

In order to avoid multicollinearity, a correlation analysis was carried out for the daily new cases, the total number of tweets, and the ten types of emotions. The results using “Kyoto” tweets and “Kyoto & Mobility” tweets are shown in Table 1 and Table 2 respectively. The correlation coefficients higher than or equal to 0.6 are highlighted in bold.

Table 1 shows the results using all “Kyoto” tweets. The high correlation between sentiments and the total number of tweets indicates stable proportions of emotional tweets. Therefore, we avoid using the number of tweets and sentiment counts in one model. Moreover, we attempt to compare the usefulness of total tweet numbers and tweet sentiment counts, so we have “Tweet numbers & Dummies” and “Tweet sentiments & Dummies” models. A high correlation is also found between some pairs of sentiments. According to the correlation coefficients, the three emotions “Fondness”, “Excitement” and “Surprise” are excluded, leaving the seven emotions of “Joy”, “Anger”, “Sadness”, “Fear”, “Shame”, “Dislike” and “Relief”. Besides, we exclude “Shame” due to its extremely low counts. These seven sentiment variables are employed in the models using “Kyoto” tweets.

Table 2 shows the results using “Kyoto & Mobility” tweets. The correlation between the number of tweets and sentiments remains high. Besides, the stronger negative correlation between mobility-related “Joy” and daily new cases is noteworthy. Another noteworthy finding is that the opposite emotions of “Joy” and “Dislike” show a fairly high positive correlation of -0.72 for the tweets with mobility-related content. In Figure 4, we can observe a peak of both sentiments at the end of November, which was a tourism peak for viewing fall foliage in Kyoto. The concurrence of joyful sightseeing experiences and complaints about crowdedness can account for this correlation. We also confirmed the concurrence by analyzing the tweet texts. The four emotions “Shame”, “Fondness”, “Excitement”, and “Surprise” are also excluded from the models using “Kyoto & Mobility” tweets, to eliminate the multicollinearity to “Joy” and “Dislike” and allow for comparison to the models using “Kyoto” tweets.

Table 1 Correlation between independent variables adapted from COVID cases and “Kyoto” tweets. |Coefficient| ≥ 0.6 in bold.

	Number of tweets	Daily new cases	Joy	Anger	Sadness	Fear	Shame	Fondness	Dislike	Excitement	Relief	Surprise
Number of tweets	1											
Daily new cases	-0.28	1										
Joy	**0.85**	-0.56	1									
Anger	0.23	0.22	0.1	1								
Sadness	0.43	-0.02	0.35	0.24	1							
Fear	**0.65**	0.07	0.33	0.25	0.28	1						
Shame	0.24	0.17	0.13	0.54	0.16	0.22	1					
Fondness	**0.79**	-0.42	**0.88**	0.14	0.4	0.31	0.14	1				
Dislike	**0.71**	-0.07	0.49	0.41	0.39	0.55	0.35	0.5	1			
Excitement	0.4	-0.06	0.22	0.11	0.17	**0.63**	0.09	0.2	0.22	1		
Relief	**0.62**	-0.1	0.57	0.16	0.29	0.3	0.16	0.53	0.39	0.24	1	
Surprise	**0.8**	-0.15	0.58	0.18	0.37	**0.67**	0.12	0.58	0.56	0.5	0.4	1

Table 2 Correlation between independent variables adapted from COVID cases and “Kyoto & Mobility” tweets. |Coefficient| ≥ 0.6 in bold.

	Number of tweets	Daily new cases	Joy	Anger	Sadness	Fear	Shame	Fondness	Dislike	Excitement	Relief	Surprise
Number of tweets	1											
Daily new cases	**-0.64**	1										
Joy	**0.93**	**-0.72**	1									
Anger	0.27	-0.07	0.22	1								
Sadness	0.27	-0.01	0.22	0.19	1							
Fear	0.36	-0.15	0.26	0.26	0.22	1						
Shame	0.15	-0.02	0.11	-0.07	-0.02	-0.03	1					
Fondness	**0.73**	-0.48	**0.75**	0.21	0.23	0.23	0.14	1				
Dislike	**0.85**	-0.44	**0.77**	0.32	0.33	0.39	0.14	**0.62**	1			
Excitement	0.55	-0.32	0.55	0.22	0.17	0.16	0.09	0.39	**0.6**	1		
Relief	**0.6**	-0.27	0.56	0.21	0.23	0.22	0.03	0.46	0.51	0.38	1	
Surprise	**0.67**	-0.32	**0.62**	0.3	0.2	0.24	0.07	0.5	**0.61**	0.47	0.45	1

5.3 Regression analysis results using "Kyoto" tweets

Table 3 presents the results of the multiple linear regression analysis. First, focusing on the key indicator of model fit "Adjusted R^2", it can be observed that it is lowest when the explanatory variable is "Only dummies", and that its value increases as the number of tweets, number of daily new cases, and emotions are included. For all models, except for the "Daily new cases & Dummies" case, the Adjusted R^2 is significantly higher in "Time lag" models than in "Real-time" models. This indicates that the lagged tweet numbers and sentiments are more accurate and effective in explaining mobility trends. It further is good news as it allows one to predict mobility based on past sentiment trends.

Considering the overall model fit, the "Time lag" models using "Tweet sentiments & Dummies" and "Daily new cases & Tweet sentiments & Dummies" are the recommended models. Both of them use lagged tweet sentiments to fit the mobility change and achieve an adjusted R^2 higher than 0.8. In the model using "Tweet sentiments & Dummies", the improvement of introducing a time lag is fairly obvious, increasing the model fit from 0.71 to 0.80. In these two cases, most of the used sentiment variables are found statistically significant with explainable negative/positive signs. "Joy", "Sadness", and "Fear" sentiments are consistently found significant and of substantial magnitudes with positive/negative signs in line with our intuition. Namely, positive "Joy" emotion has a positive correlation to mobility, and negative emotions "Sadness" and "Fear" have a negative correlation. "Dislike" and "Relief" have no significant correlation to mobility changes and are estimated with a small magnitude of impact.

More specifically, we can interpret the impact of "Joy" sentiment in the "Time lag" version of the "Daily new cases & Tweet sentiments & Dummies" model as follows: Our model suggests that an increase of 0.06% in visits to the transit station in Kyoto on a specific day can be related to an increase of 10 joyful tweets mentioning something about "Kyoto" in the past week. Similarly, a 0.13% decrease in visits to the transit stations in Kyoto can be expected if fearful "Kyoto" tweets increased by 10 in the past seven days. We note again that the baseline of tweet changes and mobility changes does not match due to the unavailability of tweet data before the outbreak of COVID-19. The intercept indicates that the visits to the train stations will be 9.18% lower than the baseline, if no COVID case is found, no policy is imposed, and the tweet sentiment numbers are at a similar level to that before the pandemic. This can be considered as a long-term or permanent reduction in mobility due to the pandemic, which is not captured by the model variables.

5.4 Regression analysis results using "Kyoto & Mobility" tweets

We repeat the multiple regression analysis with the alternative, reduced tweet dataset that only includes tweets related to mobility. Recall that tweets about mobility are defined as those containing one of the words "going", "coming" or "sightseeing". The hypothesis is that using these tweets only might be more strongly related to the number of persons observed at transit stations. For emotions, the same seven types of emotions as in Section 5.3 were used: Joy, Anger, Sadness, Fear, Shame, Dislike, and Relief. The same data was used for the dummies and the number of infected persons, so the results of “Only Dummies” and “Daily new cases & Dummies” are identical to that in Table 1.

Comparing Tables 3 and 4, it can be observed that the model fit of our preferred models, the “Time lag” models using “Tweet sentiments & Dummies” and “Daily new cases & Tweet sentiments & Dummies”, are improved with this refined tweet data set. The adjusted R^2 increases from 0.80 to 0.81 for the former model and from 0.81 to 0.84 for the latter one. In the model using time lag and “Tweet sentiments & Dummies”, sentiment variables such as “Joy”, “Anger”, and “Sadness” remain statistically significant when the tweets are refined by mobility keywords. Besides, the significance of “Joy” sentiment variables vanishes, and the coefficient changes to a counterintuitive, negative value, if daily new cases are introduced as independent variables. The strong negative correlation between them as shown in Table 2 can account for this. The significance, sign, and magnitude of “Anger” and “Sadness” are stable with the introduction of infection cases. With angry tweets increasing by 10 in the past week, the mobility will decrease by 7.94% the next day.

Notably, “Dislike” and “Relief” tweets become significant variables in the time lag model of “Daily new cases & Tweet sentiments & Dummies”. We leave “Joy”, “Dislike”, and “Daily new cases”, which have strong correlations with each other, all as independent variables though, a selection among them will help to eliminate the multicollinearity. We recall that the correlation between mobility-related “Joy” and “Dislike” tweets results from that joyful sightseeing activities coincide with complaints about crowdedness. Contrary to the estimation results of strong negative emotions such as “Anger” and “Sadness”, mild “Dislike” tweets show a correlation to mobility increase. An implication could be, that the unpreferred crowdedness did not prevent people from traveling, although they kept complaining. This can be observed in real life when the pandemic is of less concern.

Moreover, we found that an acceptable model fit can be obtained by using the total number of mobility-related tweets (“Tweet numbers & Dummies”). Compared to the results using derived sentiments, the “real-time” model has the same model fit. For the “time lag” model, the model fit is 0.81, much higher than 0.67 for the corresponding case reported in Table 3. The value is also merely 0.01 lower than using “Tweet sentiments & Dummies” and 0.03 lower than the best model. An implication could be, that the tweets containing words related to mobility show an active tendency to travel so that the visits to transit stations can be explained by these refined tweets to a larger degree and a sentiment analysis is not as necessary as in the case using a more comprehensive tweets data set. Besides, the decreased intercept in “Kyoto & Mobility” models compared to “Kyoto” models indicate that, the uncaptured loss in mobility is reduced if the tweet and “city mood” baselines are confined to mobility-related postings. That is to say, we can expect a level of mobility closer to the “old normal” if the crowdsourced sentiments regarding mobility recover to the level before COVID. As shown in the results of our best model, the long-term loss will be -3.44%, everything else equal to the situation before the pandemic.

Both Tables 3 and 4 show that a time lag of tweet sentiments improves the model fit and therefore strengthens the model interpretability and predictability on future mobility trends. In an additional analysis not presented in this paper, we investigated also different time lags of up to one month. The model results are in general fairly stable with a time lag of two to three weeks. In some cases, an even slightly higher model fit is obtained. The model fit drops with time lags longer than one month. This finding indicates that the observed real-time mobility of a city is likely to be highly dependent on the "city mood" of the past two or three weeks.

Table 3 Multiple regression analysis using “Kyoto” tweets. Significance codes: p-Value ≤ 0.001***, 0.01**, 0.05*.

Variable name	Unit of the variable	Only Dummies	Tweet numbers & Dummies		Daily new cases & Dummies		Tweet sentiments & Dummies		Daily new cases & Tweet sentiments & Dummies	
			Real-time	Time lag	Real-time	Time lag	Real-time	Time lag	Real-time	Time lag
Intercept	/	-17.52***	-12.85***	-14.04***	-10.64***	-9.25***	-6.97***	-11.1***	-7.36***	-9.18***
Number of tweets	1000	/	0.34**	1.1***	/	/	/	/	/	/
Daily new cases	ln	/	/	/	-2.61***	-3.11***	/	/	-2.23***	-1.46***
Joy	10	/	/	/	/	/	0.06***	0.11***	0.02	0.06***
Anger	10	/	/	/	/	/	-0.18*	-0.32*	-0.06	-0.15
Sadness	10	/	/	/	/	/	-0.06**	-0.11*	-0.05*	-0.13*
Fear	10	/	/	/	/	/	-0.13**	-0.27***	-0.06	-0.19**
Dislike	10	/	/	/	/	/	0.03	0.07	0.04*	0.07
Relief	10	/	/	/	/	/	-0.05	-0.1	-0.02	-0.02
Holiday	Dummy	-11.87***	-13.32***	-13.73***	-10.09***	-10.06***	-15.63***	-12***	-11.25***	-10.88***
Weekend	Dummy	-6.79***	-8.9***	-6.79***	-7.19***	-7.45***	-11.18***	-7.1***	-8.16***	-7.29***
Quasi Emergency Measures	Dummy	-10.97***	-9.78***	-6.9***	-5.88***	-5.21***	-6.2***	-3.21*	-5.53***	-3.74**
State of Emergency	Dummy	-13.81***	-12.37***	-2.67	-7.32***	-0.3	-8.99***	-1.29	-6.89***	-1.15
Rainy Day	Dummy	-3.13**	-3.45**	-3.07**	-2.57**	-2.44**	-2.55**	-2.38**	-2.54**	-2.38**
Ordinary R^2		0.61	0.63	0.67	0.78	0.78	0.73	0.81	0.79	**0.82**
Adjusted R^2		0.60	0.62	0.66	0.77	0.78	0.71	0.80	0.78	**0.81**

Table 4 Multiple regression analysis using "Kyoto & Mobility" tweets. Significance codes: p-Value ≤ 0.001***, 0.01**, 0.05*.

Variable name	Unit of the variable	Only Dummies	Tweet numbers & Dummies		Daily new cases & Dummies		Tweet sentiments & Dummies		Daily new cases & Tweet sentiments & Dummies	
			Real-time	Time lag	Real-time	Time lag	Real-time	Time lag	Real-time	Time lag
Intercept		-17.52***	-0.21	-7.16***	-10.64***	-9.25***	-3.56	-6.82***	-4*	-3.44**
Number of tweets	1000	/	13.12***	19.66***	/	/	/	/	/	/
Daily new cases	ln	/	/	/	-2.61***	-3.11***	/	/	-2.2***	-2.44***
Joy	10	/	/	/	/	/	0.56***	0.93***	0.01	-0.12
Anger	10	/	/	/	/	/	-1.17	-8.71**	-0.89	-7.94**
Sadness	10	/	/	/	/	/	-0.34	-1.93*	-0.41	-1.85*
Fear	10	/	/	/	/	/	-0.64	0.37	-0.39	-0.16
Dislike	10	/	/	/	/	/	0.4	0.44	0.64***	1.35***
Relief	10	/	/	/	/	/	0.13	0.78	0.52	2.59**
Holiday	Dummy	-11.87***	-14.68***	-12.36***	-10.09***	-10.06***	-14.54***	-11.06***	-11.47***	-9.9***
Weekend	Dummy	-6.79***	-11.93***	-7.16***	-7.19***	-7.45***	-11.58***	-7.27***	-8.94***	-7.51***
Quasi Emergency Measures	Dummy	-10.97***	-6.01***	-3.17**	-5.88***	-5.21***	-5.39***	-1.32	-4.99***	-2.93*
State of Emergency	Dummy	-13.81***	-8.18***	-5.03***	-7.32***	-0.3	-7.94***	-2.66*	-6.14***	-2.17*
Rainy Day	Dummy	-3.13**	-3***	-2.79***	-2.57**	-2.44**	-3.1***	-2.43**	-3.13***	-2.64***
Ordinary R^2		0.61	0.75	0.81	0.78	0.78	0.75	0.83	0.80	**0.85**
Adjusted R^2		0.60	0.74	0.80	0.77	0.78	0.74	0.81	0.79	**0.84**

6. DISCUSSION AND CONCLUSIONS

In this study, the correlation between tweet sentiment and mobility observed at an urban level is explored. Tweets containing the keyword “Kyoto” were collected over half a year and urban mobility was measured through Google’s Community Mobility Reports. Sentiment analysis of the content of the tweets was conducted to understand the changes in people's attitudes and the “city mood” during COVID-19. This analysis was conducted by using the “Affect Analysis System” ML-Ask for textual inputs in Japanese. Even though for around 80% of the tweets collected we cannot assign an emotion, it was found that the proportion of tweets that do convey an emotion is meaningful. This large proportion of unemotional tweets also indicates the need for an advanced emotion-identification system for Japanese inputs. The improvement of Japanese lexicons, the consideration of Japanese grammatical and linguistic rules, and the combination of lexicon-based and machine-learning-based methods can be future research directions.

Our results showed that the emotion of “Joy” shows the most similar pattern to mobility trends. We interpret this result as follows: A general trend for positive feelings in tweets about Kyoto expresses that there is a positive mood associated with coming to Kyoto, or, for residents, to conduct activities in the city. In additional analysis we confirmed that the model fit for the “Retails & Recreation” category in Google’s Community Mobility Reports is also improved with tweet sentiments introduced and the “Joy” variable is significant and has a positive impact. This result is likely to be highly associated with the COVID situation. We speculate that in non-COVID times, “Joy” might remain influential whereas “Fear” will be less related to mobility patterns. Confirming this is one direction of further work. We emphasize, that in our analysis, these results were found controlling for COVID cases and mobility restraining policies during the analysis period.

We conclude that the lagged sentiments obtained from tweets mentioning “Kyoto” city are found useful to fit the mobility indicators. The “time lag” model shows that one can improve future mobility demand predictions by monitoring tweets. The model fit is further improved if the tweets used for sentiment analysis are confined to those mentioning mobility-related activities. This also implies an assumption that people’s emotions and attitudes regarding mobility might determine their travel decision in the short future. Specifically during the COVID period, reverse causalities are also possible. Further work therefore could take sentiments as the dependent variable, investigating how mobility can affect people’s state of mind.

The regression analysis showed further that general Twitter activeness, measured in terms of the number of tweets is a suitable explanatory variable. In particular for a tourist city such as Kyoto this is not a surprising result in that visitors are likely to tweet about upcoming and experienced events in the city or the seasonal changes such as autumn foliage, which attracts many visitors. Further comparative work with the here conducted analysis being repeated in other cities could confirm whether indeed tweets are more suitable to explain mobility in touristic cities.

Related to this, we also note that we find that there is less need to conduct a sentiment analysis for the purposes of predicting mobility levels, if one can define a good keyword set according to which the Twitter data are collected. We showed that the model fit increases if only Kyoto tweets that also mention "going", "coming" or "sightseeing" are used. These expressions, or rather the "Kanji" (Chinese characters) expressing the action of a movement appeared to be obvious choices for us. Another further work direction could be, however, to aim to find an "optimal set" of keywords. This would require the integration of the sentiment analysis and the regression analysis into an optimization framework. The set might also be adjusted for defining specific travel purposes or activities, such as recreational activities.

As a final further work direction, we mention the collection and analysis of geo-tagged tweets to tackle the same problem considered in this paper, including the effect of the number of geo-tagged tweets and the added value of sentiment analysis. In conjunction with refined mobility data, this could also allow one to understand the relation between tweets and activity levels for smaller geographical analysis zones. In ongoing work, where we have been starting to collect geo-tagged tweets, we find, however, that these tweets contain fewer sentiments as the majority of those tweeting only want to convey their current location to others.

DECLARATIONS

A. Author contribution

Wenzhe Sun: Content design, data analysis, manuscript writing and editing. Hironori Kobayashi: Content design, data analysis, manuscript writing. Satoshi Nakao: Content design, data analysis, manuscript writing. Jan-Dirk Schmöcker: Content design, manuscript editing.

B. Availability of data and material

The authors do not have the permission to share the data.

C. Funding

This work was supported by the EIG CONCERT-Japan DARUMA project, Grant Number JPMJSC20C4 funded by JST SICORP, Japan.

D. Conflict of interests

The authors declared that there is no conflict of interest.

REFERENCES

Abdullah, M., Dias, C., Muley, D., & Shahin, M. (2020). Exploring the impacts of COVID-19 on travel behavior and mode preferences. Transportation research interdisciplinary perspectives, 8, 100255.

Benjamins, R., Vos, J., & Verhulst, S. (2022). Mobile Big Data in the fight against COVID-19. Data & Policy, 4, E9. doi:10.1017/dap.2021.39.

Bi, M., Sun, W., Schmöcker J.D., Ma Y., Moya-Gómez B., Nakao S., & Yamada, T. (2023) Using geo-tagged tweets for understanding temporal and spatial activity distributions in Kyoto. The 15th International Conference of Eastern Asia Society for Transportation (EASTS 2023), Kuala Lumpur, Malaysia.

Bonta, V., & Janardhan, N. K. N. (2019). A comprehensive study on lexicon based approaches for sentiment analysis. Asian Journal of Computer Science and Technology, 8(S2), 1-6.

Castillo, C., Mendoza, M., & Poblete, B. Information Credibility on Twitter, Proceedings of the 20th international conference on World wide web, pp.675-684, 2011.

Chandrasekaran, R., Mehta, V., Valkunde, T., & Moustakas, E. (2020). Topics, trends, and sentiments of tweets about the COVID-19 pandemic: Temporal infoveillance study. Journal of medical Internet research, 22(10), e22624.

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

Cui, Y., & He, Q. (2021). Inferring Twitters' socio-demographics to correct sampling bias of social media data for augmenting travel behavior analysis. Journal of big data analytics in transportation, 3, 159-174.

García-Palomares, J. C., Salas-Olmedo, M. H., Moya-Gomez, B., Condeco-Melhorado, A., & Gutierrez, J. (2018). City dynamics through Twitter: Relationships between land use and spatiotemporal demographics. Cities, 72, 310-319.

Google (2020a). COVID-19 Community Mobility Reports. Retrieved from: https://www.google.com/covid19/mobility/.

Google (2020b). Popular times, wait times, and visit duration. Retrieved from: https://support.google.com/business/answer/6263531?hl=en.

Haghighi, N. N., Liu, X. C., Wei, R., Li, W., & Shao, H. (2018). Using Twitter data for transit performance assessment: a framework for evaluating transit riders' opinions about quality of service. Public Transport, 10(2), 363-377.

Hasan, S., & Ukkusuri, S.V. (2015) Location contexts of user check-ins to model urban geo life-style patterns. PLOS One, 10(5), e0124819. doi: 10.1371/journal.pone.0124819.

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

Hu, S., Xiong, C., Yang, M., Younes, H., Luo, W., & Zhang, L. (2021). A big-data driven approach to analyzing and modeling human mobility trend under non-pharmaceutical interventions during COVID-19 pandemic. Transportation Research Part C: Emerging Technologies, 124, 102955.

Hutto, C., & Gilbert, E. (2014, May). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In Proceedings of the international AAAI conference on web and social media (Vol. 8, No. 1, pp. 216-225).

Kabbani, O., Klumpenhouwer, W., El-Diraby, T., & Shalaby, A. (2022). What do riders say and where? The detection and analysis of eyewitness transit tweets. Journal of Intelligent Transportation Systems, 1-17.

Kazama, K., Toriumi, F., Sakaki, T., Shinoda, K., Kurihara S., & Noda, I. (2012) Time-Series Variation Analysis of Relationships between Terms Using Tweets in East Japan Earthquake (in Japanese), The 26th Annual Conference of the Japanese Society for Artificial Intelligence.

Lei, Z., & Ukkusuri, S. V. (2022). Understanding the Recovery of On-Demand Mobility Services in the COVID-19 Era. Journal of Big Data Analytics in Transportation, 4(1), 1-21.

Levallois, C. (2013). Umigon: sentiment analysis for tweets based on lexicons and heuristics. https://clementlevallois.net/download/umigon.pdf

Long, Z., Alharthi, R., & El Saddik, A. (2020). Needfull–A tweet analysis platform to study human needs during the COVID-19 pandemic in New York State. IEEE Access, 8, 136046-136055.

Mahajan, V., Cantelmo, G., & Antoniou, C. (2021). Explaining demand patterns during COVID-19 using opportunistic data: a case study of the city of Munich. European Transport Research Review, 13(1), 1-14.

Miura, A. (2012) Social Psychology of Online communication on 3.11 disasters in Japan (in Japanese). The Institute of Electronics, Information and Communication Engineers (IEICE), 95(3).

Mutanga, M. B., & Abayomi, A. (2022). Tweeting on COVID-19 pandemic in South Africa: LDA-based topic modelling approach. African Journal of Science, Technology, Innovation and Development, 14(1), 163-172.

Nakamura, A (1993). Kanjo hyogen jiten (in Japanese). Tokyodo Publishing. [Dictionary of Emotive Expressions]. https://www.worldcat.org/title/kanjo-hyogen-jiten/oclc/215306489

Parady, G., Taniguchi, A., & Takami, K. (2020). Travel behavior changes during the COVID-19 pandemic in Japan: Analyzing the effects of risk perception and social influence on going-out self-restriction. Transportation Research Interdisciplinary Perspectives, 7, 100181.

Polanyi, L., & Zaenen, A. (2006). Contextual valence shifters. In Computing attitude and affect in text: Theory and applications (pp. 1-10). Springer, Dordrecht.

Ptaszynski, M., Dybala, P., Shi, W., Rzepka, R., & Araki, K. (2009). A system for affect analysis of utterances in Japanese supported with web mining. Journal of Japan Society for Fuzzy Theory and Intelligent Informatics, 21(2), 194-213.

Ptaszynski, M., Dybala, P., Shi, W., Rzepka, R., & Araki, K. (2010). Contextual affect analysis: a system for verification of emotion appropriateness supported with Contextual Valence Shifters. International Journal of Biometrics, 2(2), 134-154.

Ptaszynski, M., Rzepka, R., Araki, K., & Momouchi, Y. (2014). Automatically annotating a five-billion-word corpus of Japanese blogs for sentiment and affect analysis. Computer Speech & Language, 28(1), 38-55.

Ptaszynski, M., Dybala, P., Rzepka, R., Araki, K., & Masui, F. (2017). ML-Ask: Open source affect analysis software for textual input in Japanese. Journal of Open Research Software, 5(1).

Quercia, D., Kosinski, M., Stillwell, D., & Crowcroft, J. (2011, October). Our twitter profiles, our selves: Predicting personality with twitter. In 2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing (pp. 180-185). IEEE.

Quercia, D., Ellis, J., Capra, L., & Crowcroft, J. (2012, February). Tracking" gross community happiness" from tweets. In Proceedings of the ACM 2012 conference on computer supported cooperative work (pp. 965-968).

Ribeiro, F. N., Araújo, M., Gonçalves, P., André Gonçalves, M., & Benevenuto, F. (2016). Sentibench-a benchmark comparison of state-of-the-practice sentiment analysis methods. EPJ Data Science, 5, 1-29.

Shakibaei, S., De Jong, G. C., Alpkökin, P., & Rashidi, T. H. (2021). Impact of the COVID-19 pandemic on travel behavior in Istanbul: A panel data analysis. Sustainable cities and society, 65, 102619.

Suda, T., Kojima, K., Itoh, Y., Ishigame, M., & Toriumi, F. A cosideration of detecting false rumors using burst and spread information on Twitter during the East Japan Earthquake (in Japanese). 75th Ingormation processing society of Japan, 2013.

Sun, W., Schmöcker, J. D., & Nakao, S. (2022). Restrictive and stimulative impacts of COVID-19 policies on activity trends: A case study of Kyoto. Transportation research interdisciplinary perspectives, 13, 100551.

Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., & Kappas, A. (2010). Sentiment strength detection in short informal text. Journal of the American society for information science and technology, 61(12), 2544-2558.

Yao, Z., Yang, J., Liu, J., Keith, M., & Guan, C. (2021). Comparing tweet sentiments in megacities using machine learning techniques: In the midst of COVID-19. Cities, 116, 103273.