Computers, Environment and Urban Systems 130 (2026) 102489

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

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

How multi-sensory experience shapes place-based human sentiments: A hyperlocal analysis using environmental sensing and social media data in the Bronx, New York City

Xiaotong Zhang[a], Yuhao Kang[b,c], Simone Mora[c], Fábio Duarte[c], Carlo Ratti[c,d], An Wang[a,*]

[a] Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong Special Administrative Region, China
[b] GISense Lab, Department of Geography and the Environment, The University of Texas at Austin, TX 78712, United States
[c] Senseable City Lab, Department of Urban Studies and Planning, Massachusetts Institute of Technology, MA 02139, United States
[d] Politecnico di Milano, Milan 20133, Italy

ARTICLE INFO

Keywords:
Mobile sensing
Multi-sensory sentiments
Social media
Urban physical environments

ABSTRACT

People experience cities through their senses of sight, hearing, smell, touch, and taste. However, most research focuses only on visual stimuli. In this study, we develop a straightforward multi-sensory, data-based linear regression model to explain human sentiment in cities through visual, acoustic, olfactory, and tactile explanatory features. We use social media to reflect people's sentiments, and concurrent urban environmental data ranging from street view imagery to air quality, collected in the Bronx, New York. We observe correlations between non-visual features and urban experience, where noise and thermal comfort are identified as significant impact factors. The adjusted coefficients of determination for the simple linear models can reach 0.80 and 0.60 for positive and negative sentiments, respectively, indicating moderate to strong interactions between multi-sensory urban features and human sentiments.

1. Introduction

Among the myriad factors that impact human sentiment, the physical environment plays a crucial role, encompassing visual, acoustic, olfactory, and tactile features. A substantial body of literature has examined the impact of a single sense at a time on urban experience. Zhang et al. (2018) developed a deep learning framework to analyze street view images and demonstrated that specific urban visual elements can significantly affect human sentiment in different neighborhoods. Noise and unpleasant smells have long been identified as negative contributors to human sentiment in urban areas (Herrera & Cabrera-Barona, 2022; Simion et al., 2024). Meanwhile, the adverse effects of extreme temperature, perceived through the sense of touch, on public sentiments have been verified in various contexts (Wang et al., 2024). Recent studies highlight the combined effects of senses in shaping human sentiment. The cross-modal effect of tactile and visual perceptions is uncovered, indicating thermal comfort is jointly influenced by heat and sun brightness (Lam et al., 2020). Noise survey and street view data show that the interaction between the visual and acoustic environment influences the perceived noise level (Huang et al., 2024; Zhang et al., 2024). Yet, there has been very limited literature that seeks to understand the holistic human sentiment in urban spaces through multi-sensory channels, including the senses of sight, hearing, smell, touch, and taste, as well as their interactions.

In recent years, the emergence of social media and online reviews has introduced novel data sources for understanding human sentiments (Liu et al., 2023). These platforms allow for the extraction of two types of knowledge, including semantics and sentiments, particularly from neighborhood reviews (Hu et al., 2019). Restaurant reviews on platforms such as Google Maps and Dianping, the Chinese equivalent of Yelp, have been analyzed to gain further insights into public perceptions (Niu & Xing, 2024; Shin et al., 2022). In addition to online reviews, social media also plays a vital role in understanding human sentiment. For example, by examining tweets from four different cities, researchers can identify the positive and negative sentiments associated with specific geographic locations (Lanotte & Siracusa, 2022). Similarly, the human sense of place can be extracted from Twitter using Latent Dirichlet Allocation (LDA) (Siragusa & Leone, 2018).

When it comes to the multisensory sentiment in cities, it is difficult to collect concurrent, collocated environmental indicator data to represent

* Corresponding author.
E-mail address: an.wang@polyu.edu.hk (A. Wang).

https://doi.org/10.1016/j.compenvurbsys.2026.102489
Received 21 November 2025; Received in revised form 5 June 2026; Accepted 29 June 2026
Available online 17 July 2026

the physical environment. Existing environmental sensing efforts mostly focus on a single aspect, such as urban air quality, micro-climate, noise, and green space (Guo et al., 2022; Hu et al., 2024; Zhang et al., 2021; Zou et al., 2025). The emerging miniature and integrated sensing technologies have provided new opportunities, which enable mobile environmental sensing to cover a large spatio-temporal scope while measuring multiple environmental indicators concurrently (O'Keeffe et al., 2019).

Our study proposes a framework that incorporates drive-by sensing and social media, thereby enhancing our understanding of the physical environment via multi-sensory human sentiment interactions (Anjomshoaa et al., 2018). We employ drive-by sensing to collect data on urban environments, including street views, soundscape, micro-climate, and air quality, that are associated with visual, auditory, tactile, and olfactory senses, respectively. The sense of taste is omitted in our study as it is less commonly experienced in a general visit not involving eating or drinking. Meanwhile, social media is utilized as a surrogate for multi-sensory human sentiment. Both urban environmental data and social media are collected during a concurrent period, aiming to provide relatively stable conditions of human multi-sensory experience at a place. This study seeks to provide a deeper understanding of how the urban physical environment interacts with human sentiments through multi-sensory pathways, driven by two major research questions:

1. How are human sentiments affected by various senses concurrently present at places?
2. What are their relative importances in the human sentiment in cities?

2. Methodology

As illustrated in Fig. 1, the framework of this research includes three main parts. (a) Multi-sensory integration using mobile environmental sensing and street view images; (b) Human sentiment measurement based on social media; (c) Analysis for physical environment—human sentiments interaction.

2.1. Study scope and urban environmental factors

The study area covers the Bronx borough in New York City, as shown in Fig. 2(b). It is located to the north of Manhattan Island in the south-eastern part of New York State. The Bronx is well known for its dense built environment with a population of around 1.4 million, where people from various races and income levels live. Hispanic or Latino and Black Americans comprise the majority of the local population, accounting for 55.0% and 45.1%, respectively (U.S. Census Bureau, 2024). The timeframe for this research spans four months, from September to December 2021.

We consider four senses in the study: sight, hearing, smell, and touch. Street view images obtained from Google Maps are utilized to measure an average state of visual sense, as they can reflect people's vision from

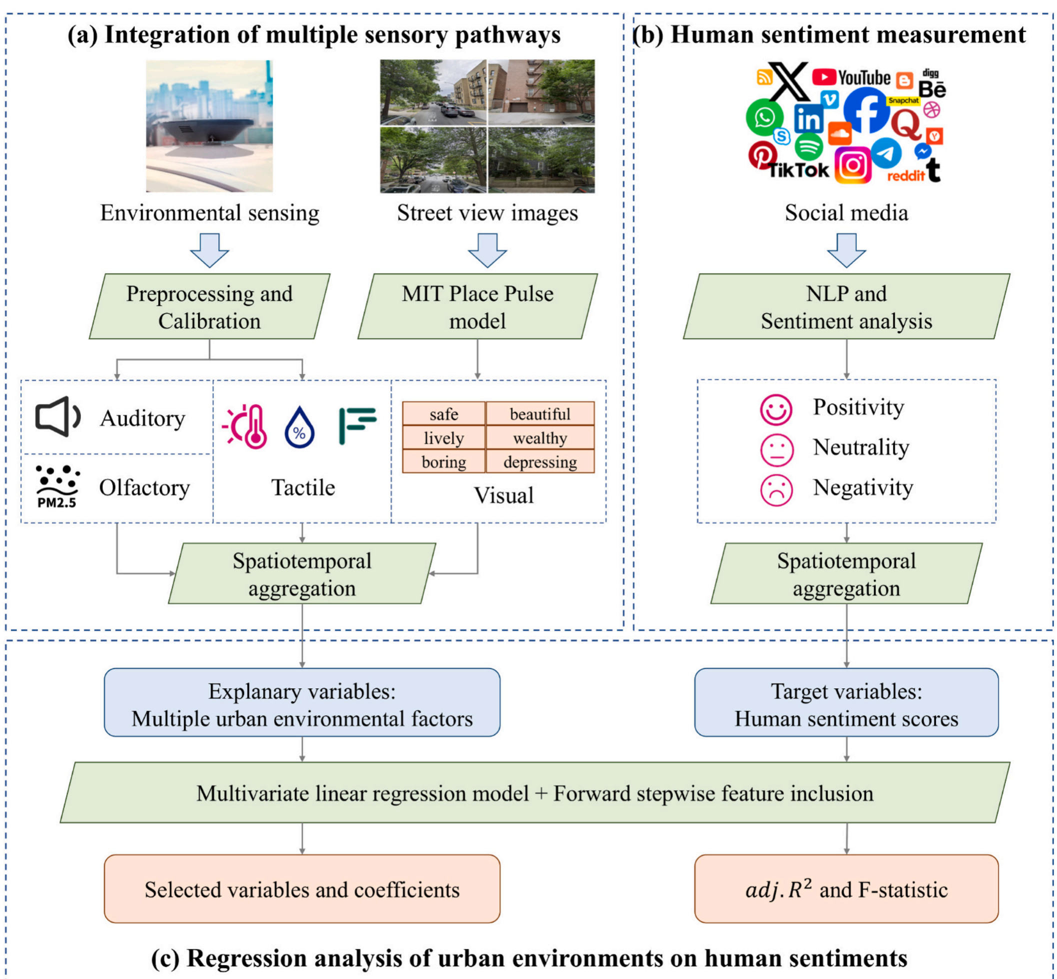

Fig. 1. Research framework. (a) demonstrates the integration of multiple sensory pathways in urban environments, (b) introduces the measurement of human sentiments. (c) shows the regression analysis for the interaction between the physical environments and human sentiments.

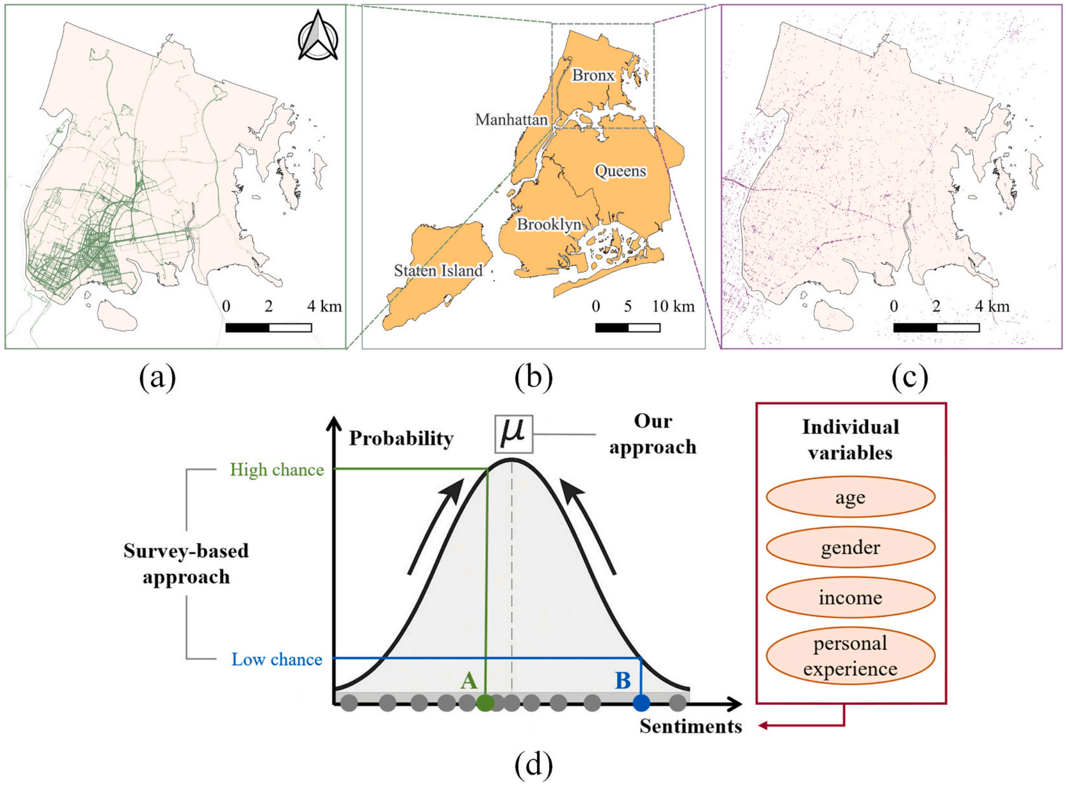

Fig. 2. Study scope and data source. (a) is the spatial distribution of environmental sensing data. (b) is the map of New York City. (c) is the distribution of tweets. (d) Illustrations of population-level and survey-based individual-level sentiment sampling. The bell-shaped curve illustrates the probability distribution of sentiments, where μ denotes the average state of sentiments, the primary focus of our approach. In contrast, survey-based approaches examine individual-level variance, represented by individuals A and B.

eye level, including buildings, greenery, and landscapes (Dai et al., 2021; Ito et al., 2024; Yan et al., 2023). To overcome the challenge of measuring sensation from smell, we measure a typical air pollutant, PM2.5, as an indirect surrogate, since it is an important carrier of natural or anthropogenic smells, in the form of combustion exhaust, cooking fumes, pollens, catkins, dust, and sea salt particles, in coastal regions like New York City. In addition, the formation of secondary PM aerosol is known as part of the photochemical reaction, which is an important contributor to odors (Chen et al., 2022). Volatile organic compounds (VOCs) are not measured as the existing low-cost sensors cannot distinguish them. Auditory sense is shaped by the urban soundscape, which includes both noise and natural sounds, measured in decibels in this study. Amplified by climate change and the urban heat island effect, urban tactile sense is more relevant than ever to microclimate conditions. Therefore, ambient temperature, relative humidity, and wind speed are measured, as they are direct indicators of microclimate.

We combine multiple data collection channels to compile our urban environmental dataset. Six visual environmental indicators reflecting the average state of visual perception are aligned with the areas covered by the environmental sensing deployments during the concurrent period (Zhang et al., 2018). The street view images of each location point in the Bronx include views from four different directions. Based on the methodology proposed in previous studies (Ito et al., 2024; Zhang et al., 2018), we process the street view images to assess six perception indicators: safe, lively, boring, wealthy, depressing, and beautiful. In the initial stage, a deep convolutional neural network (DCNN) is used to extract high-dimensional features from these street view images. Following that, each perceptual score is labeled with binary values as the target variable, and the support vector machine (SVM) classifier is trained for all the perceptual scores. Finally, by deploying the pre-trained model to street view images in the Bronx, six visual perceptual scores are derived from images and output by the model. Each score ranges from 0 to 10, with 0 representing the least safe, lively, boring, wealthy, depressing, and beautiful, and 10 representing the most. Nonetheless, these perceptual scores are not time-dependent, reflecting how people perceive the street view in an average state without changing from time to time.

For air quality, temperature, noise, and humidity, we use more than 80,000 records collected in the Bronx from September to December 2021, Fig. 2(a). The selection of this period is attributed to the common data collection arrangement in large-scale travel survey, as well as relatively mild weather of New York City during this period, allowing us to minimize the impacts of extreme thermal stress (Bricka et al., 2024). Before deployment, the PM and meteorological sensors were first collocated and calibrated at a local reference air quality station to ensure their measuring performance (Wang et al., 2023). In data quality assurance, we observed that noise levels were significantly affected by the speed of the vehicle. To mitigate this bias, a two-step data filtering process was implemented. First, we removed outliers by eliminating measurements that exceeded three standard deviations from the mean in either vehicle speed or noise level. Second, records were excluded when the combined influence of vehicle speed and wind speed exceeded 5 m/s. This threshold was determined based on empirical observations of noise measurement stability under different motion conditions, accounting for aerodynamic effects. Wind speed data is acquired from a local weather station.

Given that different data sources have varying spatial and temporal resolutions, we aggregate environmental indicators data spatially and temporally to generate an average condition in a concurrent period. We first aggregate all the urban environmental variables in the study period into grid cells with resolutions ranging from 100 m to 400 m. In addition, these data are organized into temporal bins by day of the week. Specifically, according to a rule of thumb, at least ten records from four unique days are sufficient to represent the overall environmental conditions within a grid cell (Wang et al., 2023). The grid cells containing fewer records are eliminated to ensure the robustness of environmental condition measurements. For each grid cell and time slot, we calculate the mean values of all data records. The aggregation compiles the explanatory variable dataset, containing the average environmental conditions by day of the week, at various spatial resolutions. Subsequently, it enables systematic analysis of how the visual, olfactory, auditory, and tactile environments are perceived and then influence human sentiments in various places.

2.2. Measurement of human sentiments using social media

The complexity and diversity of human sentiments make it challenging to quantify them comprehensively. Social media offers a valuable opportunity to measure human sentiments in urban environments from the perspective of big data. We collect geotagged Twitter data within the Bronx borough, covering the concurrent period from September to December 2021. Tweets' spatial distribution is demonstrated in Fig. 2(c). We use text sentiment as a proxy for urban experience in this study, which measures the emotional valence (i.e., positive or negative) expressed in social media posts.

To quantify place-based subjective sentiments, sentiment analysis is applied to extract emotional valence from geotagged tweets. The process begins with text preprocessing: (1) special characters and hashtags are removed, (2) stop words, including determiners, conjunctions, and prepositions, are eliminated, and (3) all the texts are converted to lowercase. Finally, the processed messages serve as input to the sentiment analysis model. Specifically, the sentiment extraction from tweets of the Bronx is powered by a robustly optimized BERT approach (RoBERTa) (Barbieri et al., 2020; Camacho-Collados et al., 2022; Loureiro et al., 2022). This approach is based on the bidirectional encoder structure from Transformers, which enhances its capabilities in natural language processing tasks. Since it is pre-trained on large-scale Twitter data of 584 million tokens, it can be transferred directly for analyzing our Twitter dataset. The model generates probabilistic sentiment scores across three dimensions for each tweet, including positivity, neutrality, and negativity. Specifically, the three scores all range from 0% to 100%, and sum up to 100%, with higher values of positivity and negativity indicating a stronger tendency toward positive and negative expressions, respectively. In contrast, a higher value of neutrality means that tweet messages have a relatively weak emotional valence. Since this research aims to move beyond individual-level variance toward population-level human sentiment trends using crowd-sourced social media data, we aggregate sentiment scores from all geotagged tweets within each grid cell and day of the week. These mean values represent average states of sentiments in certain grid cells on a given day of the week. This approach may mitigate the biases introduced by individual-level variables, such as age, gender, and income, by achieving spatial representativeness at the population level, as illustrated in Fig. 2 (d). This is conceptually aligned with the Neighborhood Effect Averaging Problem (NEAP) in environmental exposure, which suggests that individual-level exposures tend to be around the population mean (Kwan, 2018). Nonetheless, these individual-level variables remain valuable for sentiment studies from an individual-level perspective.

2.3. Analysis for physical environment-human sentiments interaction

We evaluate the impacts of all environmental indicators at the grid cell level using forward stepwise feature inclusion in linear models, considering the unique role of non-visual indicators and cross-sensory interactions. In the proposed framework, all urban environmental indicators are treated as explanatory variables, while the three sentiment scores—positivity, neutrality, and negativity—serve as target variables, as shown in Table 1. Besides, cross-sensory composite variables, such as 'Beautiful × Sound level', are yielded by multiplying two representative indicators for each sense, and subsequently included as explanatory variables. Specifically, we limit the number of individual and cross-sensory variables containing the same indicator to two, given their high collinearity. We select the models with robust performance and sufficient sample size for joint and individual variable interpretation. In addition to the abovementioned environmental indicators as explanatory variables, a categorical indicator for holidays in New York City, 'holiday or not', is included to explore the impact of temporal context on human sentiments, as people may feel happier and more relaxed on holidays (Stone et al., 2012; Suk et al., 2021). In view of the potential impacts of socio-economic characteristics on Twitter sentiments, confounding variables, including population density, the percentage of ethnic minorities, mean household income, unemployment rate, the proportion of people with a bachelor's degree or higher, and the sex ratio, are controlled. To address the high variance in tweet numbers across grid cells, we incorporated weights into the regression framework, where the number of tweets in each grid cell serves as a weight, allowing the model to emphasize the grid cells containing more sentiment observations. Specifically, a log-scale transformation is applied to the tweet counts, as it can mitigate grid cells with super-active tweeting activity from dominating the model. This analysis enables us to examine the effectiveness of non-visual indicators and cross-sensory interactions, which in turn tests how other senses may contribute to human sentiments.

Table 1
Explanatory, control, and target variables for regression.

Variable types	Corresponding senses	Indicators	Unit	Data source
Explanatory variables	Auditory	Sound Level	*dB*	Mobile Environmental Sensing
	Olfactory	$PM_{2.5}$ Concentration	$\mu g/m^3$	
	Tactile	Temperature	°C	
		Relative Humidity	%	
		Wind speed	*m/s*	Local Weather Station
	Visual	Six indicators (safe, lively, boring, wealthy, beautiful, and depressing)	None	Street View Images
	Multi-sensory interaction	Multiplying two indicators for each sense, such as 'Beautiful × Sound level'.	None	All above
	None	Holiday or not	None	N/A
Control variables	None	Population density	Persons/ km^2	American Community Survey (ACS) 2022
		Mean household income	Dollars	
		Unemployment rate	%	
		Proportion of people with a bachelor's degree or higher	%	
		Percentage of ethnic minorities	%	
		Sex ratio	None	
Target variables	Sentiments	Positivity	%	Twitter
		Neutrality		
		Negativity		

2.4. Uncertainty and sensitivity analysis of confounding variables

We consider potential uncertainties that might affect the validity of our findings. The impacts of grid cell resolution on model effectiveness are first evaluated, with particular attention to the common modifiable areal unit problem (MAUP) in geo-spatial analysis. We generate grid cells at three different spatial resolutions: 100 m, 200 m, and 400 m, respectively. Subsequently, the explanatory and target variables are averaged at these resolutions. A similar stepwise regression approach is then applied to grid cells of different resolutions.

Additionally, we also investigate the impact of the number of tweets aggregated in a single grid cell on the model outcome, which reflects the emotional representativeness of these tweet observations. Unlike environmental factors, for which multiple studies have summarized rules of thumb for representative observations (Apte et al., 2017; Hatzopoulou et al., 2017; Wang et al., 2023). There are few explorations on this matter in the literature for measuring human sentiment in locations. Thus, we test a series of tweet thresholds within each grid cell to gain insights into the appropriate number of tweet observations for sentiment analysis. We train stepwise regression models using 100 m by

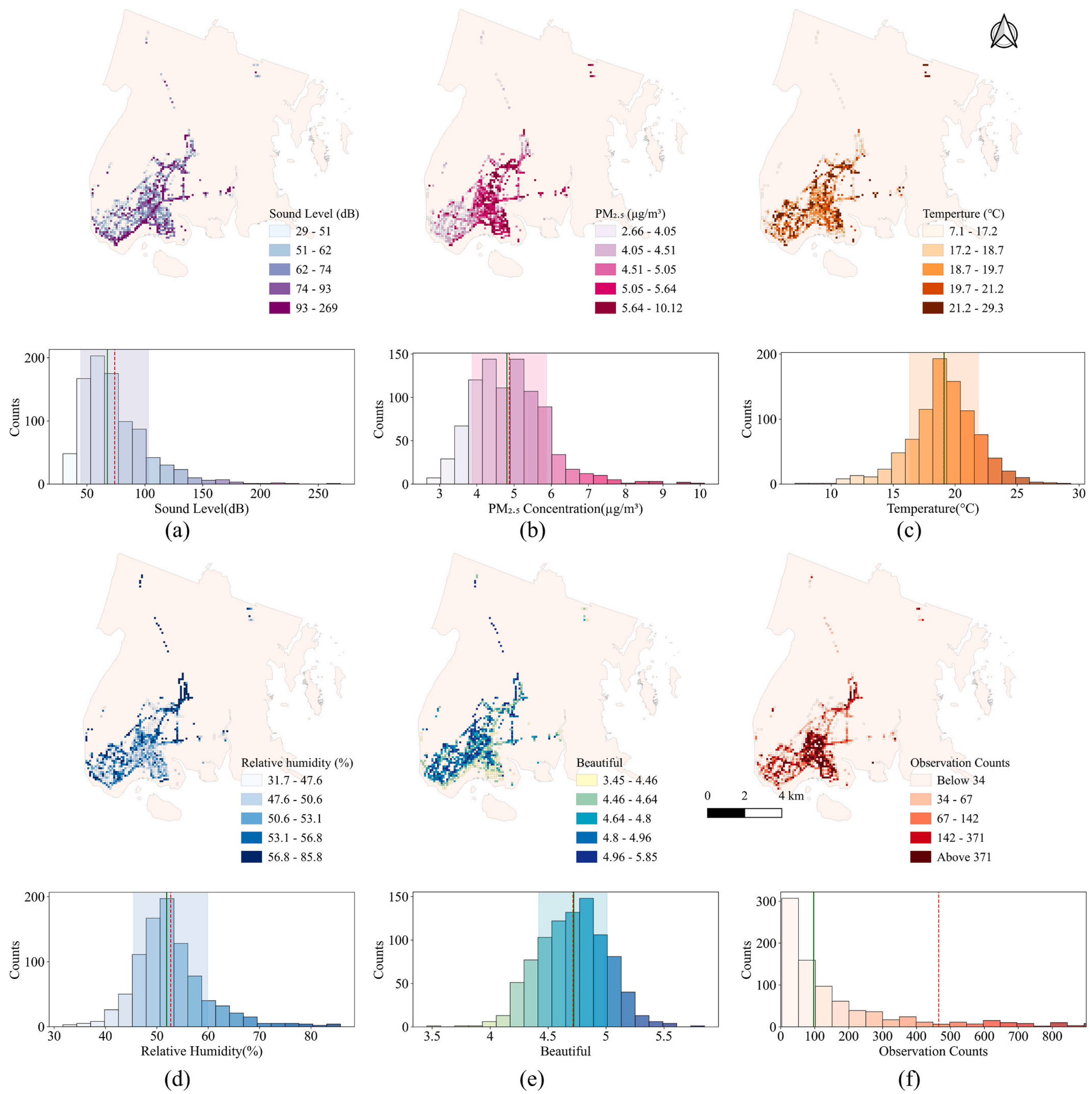

Fig. 3. Distribution of aggregated urban environmental factors. The upper parts of (a), (b), (c), (d), (e), and (f) illustrate the average values of noise, $PM_{2.5}$ concentration, ambient temperature, relative humidity, and a representative visual environmental indicator, 'beautiful', as well as observation counts across grid cells. In these illustrations, darker colors represent higher values of these variables. The bottom sections of these six subfigures present histograms for the six variables, aggregated by grid cells. The height of each bar represents the number of statistical units, specifically a grid cell whose values fall within the specified intervals. As the values of these variables increase, the shade of the bars becomes darker, providing a visual cue for higher values. In addition, the red dashed line and green solid line represent the mean and median values of each variable, respectively. To highlight the extent of variation, the shaded area indicates the range between the mean minus one standard deviation (Std.) and the mean plus one Std. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

100 m cells that contain at least 1 tweet, up to a maximum of 15 tweets. The number of available cells for regression decreases as the threshold improves, resulting in sample size differences up to tenfold. We investigate whether the performance difference between the 1-tweet-threshold model and the 15-tweet-threshold model is due to the change in sample size or the different number of tweets. We adopt a Monte Carlo sampling approach that repeatedly samples a constant number of 53 grid cells 100 times with replacement from the available cells of different tweet thresholds (from 1 to 15 tweets). In each round, all 53 grid cells are treated as a chunk and drawn at once, so that models developed with smaller sample sizes will have smaller variation in their performance. By controlling sample size as a confounding variable, we focus specifically on the impacts of the tweet count threshold on model performance and estimation robustness.

3. Results

3.1. Spatiotemporal analysis of single sensory data and sentiments

As illustrated in Fig. 3, multiple urban environmental factors exhibit distinct spatial distribution patterns across the Bronx. Notably, $PM_{2.5}$ concentrations are exceptionally high in the southeastern Bronx, particularly near major highways, transportation facilities, and industrial areas, which may lead to adverse visual and olfactory effects. Additionally, noise levels along major highways often exceed those in adjacent areas. These adverse acoustic conditions would impair local sensations from sound. It is worth noting that the raw noise data and ego speed of the mobile sensing vehicle exhibited a strong positive correlation, with Pearson's r up to 0.342, which could introduce bias into the model. This is due to the wind noise induced by higher vehicle speed. After excluding noise data collected at speeds greater than 5 m/s, we observe that the correlation between noise levels and the vehicle's speed relative to the air is significantly decreased to 0.013, enabling the direct use of noise data in the model to indicate the urban soundscape. In contrast, high-temperature and high-humidity zones are sparsely distributed across the Bronx, without clear spatial clustering. Meanwhile, the visual environment has a higher 'beautiful' score in the northwest Bronx and a lower score in the southeast Bronx. Other visual scores are documented in Fig. S1, Supplementary Information (SI). In addition, most of the urban environment factors are normally distributed, fulfilling a key assumption of linear regression. The detailed statistics of all explanatory and target variables, including mean and median values, as well as standard deviation (Std.), are shown in the upper section of Table S1 in the SI.

The three sentiment scores, positivity, neutrality, and negativity, collectively characterize the human sentiments in urban environments. As shown in Fig. 4, three distinct color maps, red, orange, and blue, are used to display the three sentiments, respectively. Analysis reveals a strong predominance of neutrality in emotional expressions on social media, with an average value of 74%. By contrast, average values for positive and negative sentiments are comparatively lower, at about 14% and 11%, respectively. This pattern indicates that most users post matter-of-fact, emotionally neutral content rather than tweets with strong emotional valence. Grid cells with high positivity are distributed in the southwestern Bronx, particularly in areas proximal to Manhattan. In contrast, grid cells along major railway corridors demonstrate suppressed positivity and elevated negativity simultaneously. Additionally, sentiments expressed on social media exhibit distinct day-of-week rhythms. It is observed that posts on weekdays tend to convey frustrations and stress more frequently. Conversely, leisure time during

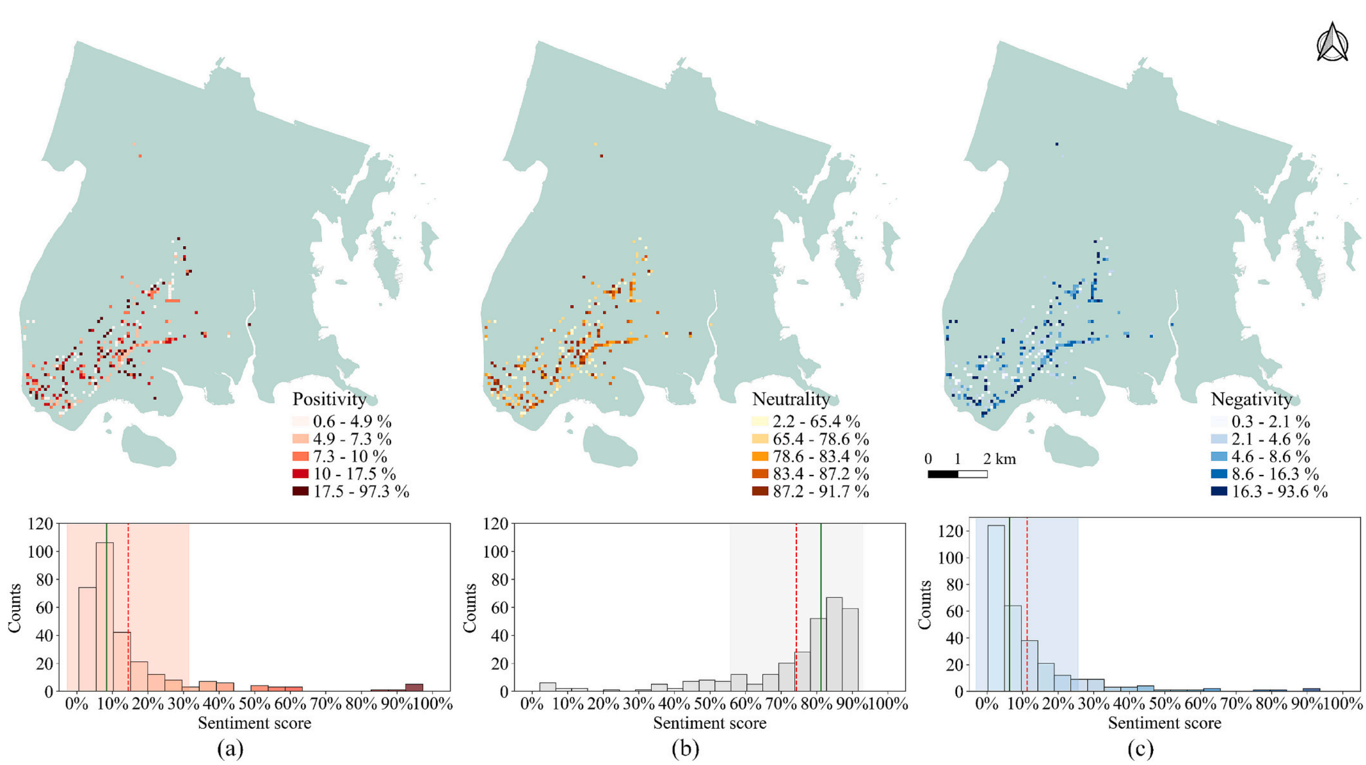

Fig. 4. Distribution of human sentiments across grid cells. The top sections of (a), (b), and (c) illustrate the spatial distribution of average positive, neutral, and negative sentiments across grid cells. Specifically, in the grid cells on the map, lighter shades of red represent lower positivity, while darker shades indicate higher positivity. Similarly, the darker orange and blue shades indicate higher neutrality and negativity, respectively. The bottom sections of the three subfigures present histograms for the three sentiments, aggregated by both grid cells and days of the week. In addition, the red dashed line and the green solid line represent the mean and median values of each sentiment, respectively. The shaded area indicates the range between the mean minus one standard deviation (Std.) and the mean plus one Std. Specifically, darker font colors in the table indicate higher values, corresponding to the color map associated with each sentiment. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

holidays may facilitate people's emotional recovery, leading to more positive expressions on social media.

3.2. Relationship between urban environments and human sentiment

The correlations between indicators for three of the four senses, visual, auditory, and olfactory, are relatively low, as shown in Fig. S2 in Section 1 of the SI. Among these indicators, the highest correlation is between the average $PM_{2.5}$ concentration and temperature, with a positive value of 0.59. Following that, there is a positive correlation of 0.15 between sound levels and $PM_{2.5}$. However, the six visual environmental indicators are highly correlated with each other, which may result in multicollinearity problems. The two unfavorable visual environmental indicators, boring and depressing, exhibit a high level of correlation of 0.83. Both are negatively correlated with indicators of a favorable visual environment. Similarly, the correlation between safe and boring is −0.93, as shown in dark green. Therefore, it is imperative to address the multicollinearity of visual environmental indicators using stepwise feature inclusion.

There is an uneven spatial distribution of tweeting activity within the study area. As illustrated in Fig. 5, the number of selected grid cells decreases sharply as the threshold for tweet counts within each cell increases. When the threshold is set to a minimal level, i.e., at least one tweet per cell, 558 aggregation units meet this criterion. As the threshold increases, the number of qualifying grid cells drops sharply. For instance, when the threshold of tweets rises to at least two tweets per

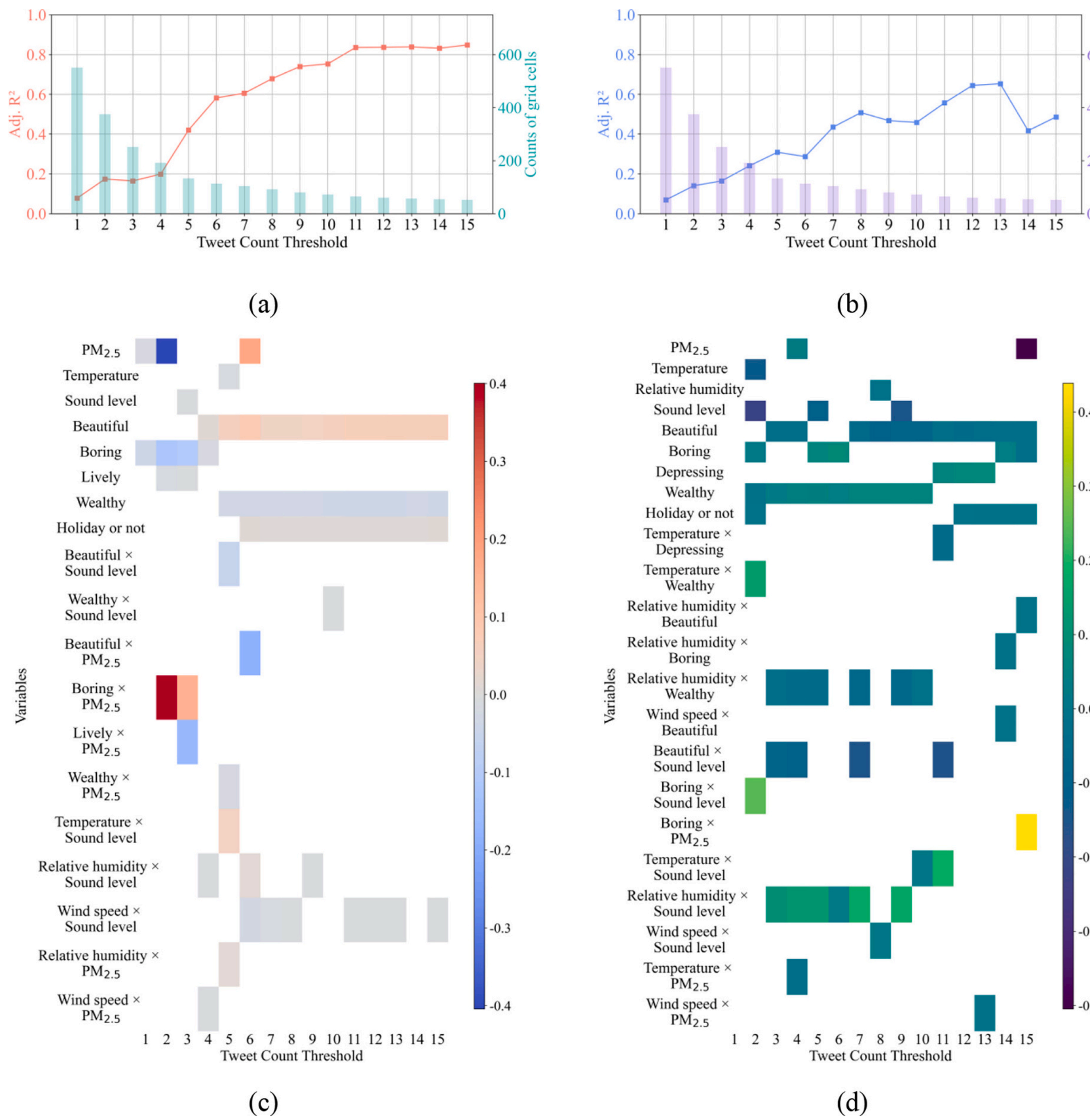

Fig. 5. The $adj.R^2$ for positivity and negativity with varying tweet count thresholds. (a) and (b) illustrate how the $adj.R^2$ changes for positive and negative sentiments, respectively, as the thresholds increase. (c) and (d) display the urban environmental variables selected using stepwise techniques. In these figures, each column represents a model based on grid cells containing at least the threshold indicated by the corresponding x-tick, while each row represents a variable. Within each column, blank spaces indicate variables not selected, whereas shaded cells denote variables included in a certain model. Red shading in (c) indicates a positive contribution to explaining human sentiments, while blue shading indicates variables that have an inverse contribution. Similarly, yellow to blue shading in (d) corresponds to coefficients ranging from low to high. In each row, as the tweet count threshold increases, changes in the shade colors within each row reflect variations in the coefficients of the corresponding variable in the model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

cell, only 380 units are retained. It is observed that about one-third of the aggregation units consist of only a single tweet. Besides, in the at-least-15-tweets-per-cell model, only 53 samples are included, which constrains the usage of complex regression models, such as gradient boosting trees or neural networks.

The $adj.R^2$ for the weighted multivariate linear regression model with cross-sensory interactions are shown in Fig. 5. We emphasize positivity and negativity results but have not included results for neutrality because there is no clear trend for its variation. Moreover, the adj.R^2 of the model increases with rising thresholds, eventually stabilizing at a plateau when the threshold reaches approximately eleven. At this elbow point, $adj.R^2$ for positivity remains above 0.80, while that for negativity fluctuates around 0.55. Beyond this point, further increases in the threshold do not significantly enhance the adj.R^2. Moreover, as the threshold increases, the coefficients for 'beautiful' and 'depressing', the two most prominent indicators of positive and negative human sentiments, stabilize at around 0.07 and 0.08, respectively. This observation suggests that more than ten unique sentiment observations from social media are a relatively reliable benchmark for capturing representative sentiments at a given location. Moreover, it suggests that the proposed model can effectively capture the interactions between urban physical environments and both the positive and negative aspects of grouped human sentiment. In contrast, the interactions are relatively indistinct when there are fewer tweets in grid cells.

Finally, the contribution of each environmental factor is detailed by incorporating them stepwise, as shown in Fig. 6. Specifically, we select the model using grid cells that contain at least 11 tweets, as this threshold marks the common elbow points of our models. By including indicators using stepwise feature inclusion, we discover that one of the visual environmental indicators, 'beautiful', contributes the most to explaining positive human sentiments. This indicator is included at the first step, with a coefficient of 0.04 and a significance level of 0.001, and its impact increases to a coefficient of 0.07 in the final step. Another visual indicator, 'wealthy', while highly correlated with 'beautiful', exhibits an adverse effect on positivity, with negative coefficients remaining −0.04, throughout all the steps. In contrast, holidays have a positive relationship with favorable human sentiments, with a small coefficient of 0.01. For negative sentiments, the visual environmental indicator, 'depressing', is the first variable included in the model, and maintains a coefficient of approximately 0.05 at the final step. Conversely, 'beautiful' significantly reduces negative sentiments even when combined with sound level, yielding a coefficient around −0.15 and a significance level of 0.01 in the final model. Meanwhile, the tactile environment indicator, 'temperature', exhibits a minor correlation with negativity, with relatively small coefficients of −0.04, while its interaction with sound level strongly amplifies negative sentiments, producing a substantial coefficient of 0.17 at the final step. This finding underscores the combined impacts of poor soundscapes and thermal stress on deteriorating human sentiment. Nevertheless, all the selected visual environmental indicators are static in a grid cell across the study period, making it infeasible to evaluate how visual sense influences human sentiments temporally. Moreover, as shown in (a) and (b), we can observe that the value of adj.R^2 is notably lower when only two visual environmental indicators, 'beautiful' and 'wealthy' are considered, compared to including environmental indicators of other sensory channels. It means that in addition to visual environments, other acoustic and tactile variables contribute significantly to explaining the variation of human sentiments, highlighting the significance of understanding environmental factors beyond just visual aspects.

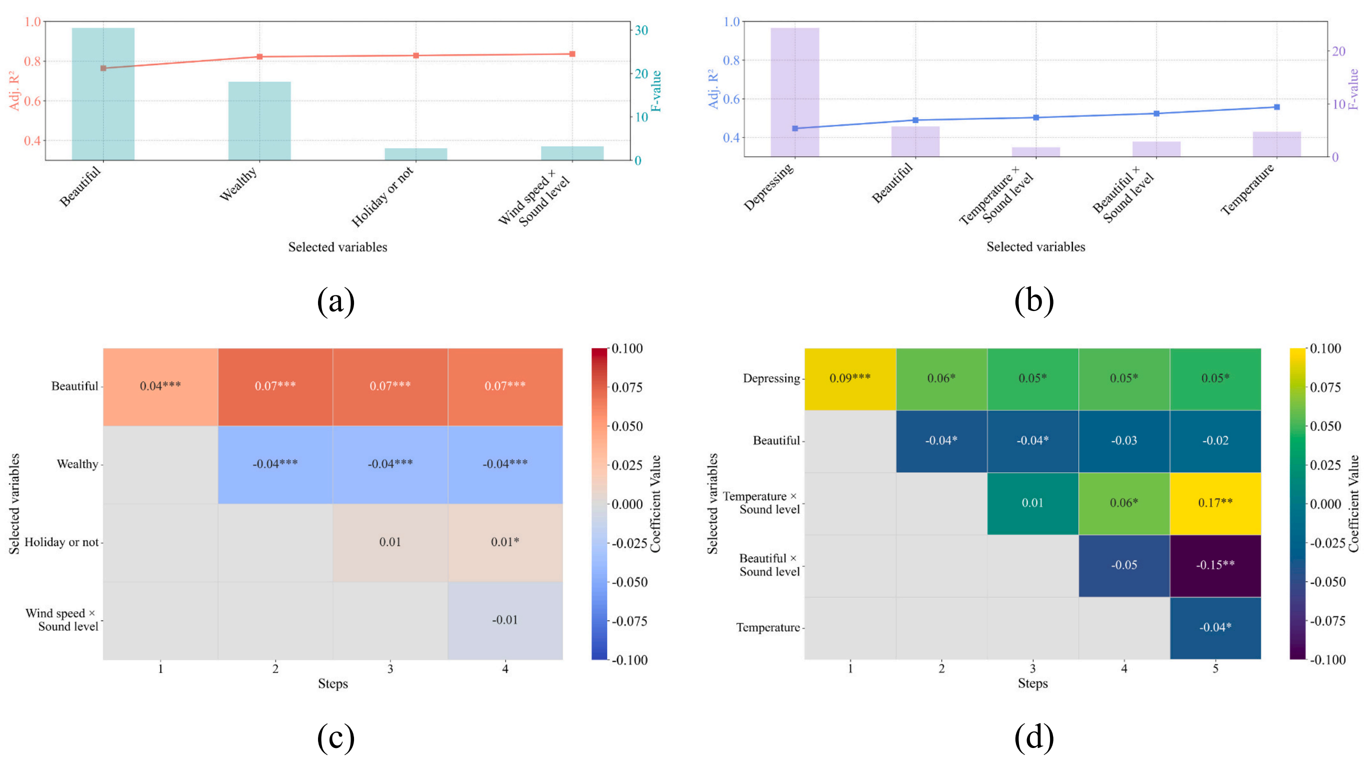

Fig. 6. Effectiveness of the model with stepwise variable inclusion for grid cells containing at least 11 tweets. (a) and (b) illustrate the increase in $adj.R^2$ for both positive and negative sentiments. The y-axis on the left represents $adj.R^2$, and the right y-axis shows F-values of the stepwise model. The x-axis sequentially lists the variables included in each step. This sequence emphasizes the order of importance and contribution of each variable in explaining human sentiments. Green and purple bars indicate the F-values of models for the two scenarios, positivity and negativity. (c) and (d) detail the variable inclusion process corresponding to (a) and (b), respectively. The color shading reflects the coefficients, with warmer colors indicating higher coefficients above zero and cooler colors indicating coefficients below zero with greater absolute value. Specifically, the number of stars within each grid cell denotes the significance level for the variables. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.3. Sensitivity analysis results

3.3.1. Impact of MAUP

The impact of the MAUP is first evaluated at three different grid cell sizes, 100 m, 200 m, and 400 m, respectively. Fig. 7 illustrates the variation in $adj.R^2$ and the number of grid cells at sizes of 100 m, 200 m, and 400 m. As the grid cell size becomes larger, the number of grid cells containing more than a certain threshold of tweets also increases. When the grid cell size reaches 400 m, the total number of samples is fewer than that of 100 m, due to fewer grid cells being drawn out. Thus, increasing the size of grid cells does not necessarily increase or decrease the sample size. In addition, the trend of $adj.R^2$ remains consistent across grid cell sizes. Compared to 100-m grid cells, the performance elbow points come with larger tweet thresholds when larger grid cells are used, demonstrating that more tweets are needed for spatial representativeness of a larger geographic unit.

3.3.2. Model performance impact of tweet count threshold

In our previous analysis, we observed that when the tweet count threshold exceeds 10, $adj.R^2$ for all models reach a plateau for 100 m grid cells, with only moderate increases thereafter. Whether this effect is an artifact of the decreasing sample sizes or really represents a more robust reflection of human sentiment with more tweet observations still needs to be determined. Using Monte Carlo sampling, we find that the increases in $adj.R^2$ are not because of the decreased sample size. As illustrated in Fig. 8, similar trends for both positivity and negativity emerge even when the sample size remains constant. Consequently, these results suggest that the determinative factor for the effectiveness of the model is the tweet count thresholds rather than how many grid cells are sampled.

4. Discussion

4.1. Enhancing urban environments through targeted sustainable planning strategies

We observe that acoustic environments play a significant role in shaping human sentiments across various places. This finding is in line with Ma and Thompson (2015) as well as Zhuang et al. (2024), but provides quantitative estimates of its impact on place-based human sentiment. Specifically, the soundscape has a comparable impact on negative sentiments to that of visual environments. This is probably

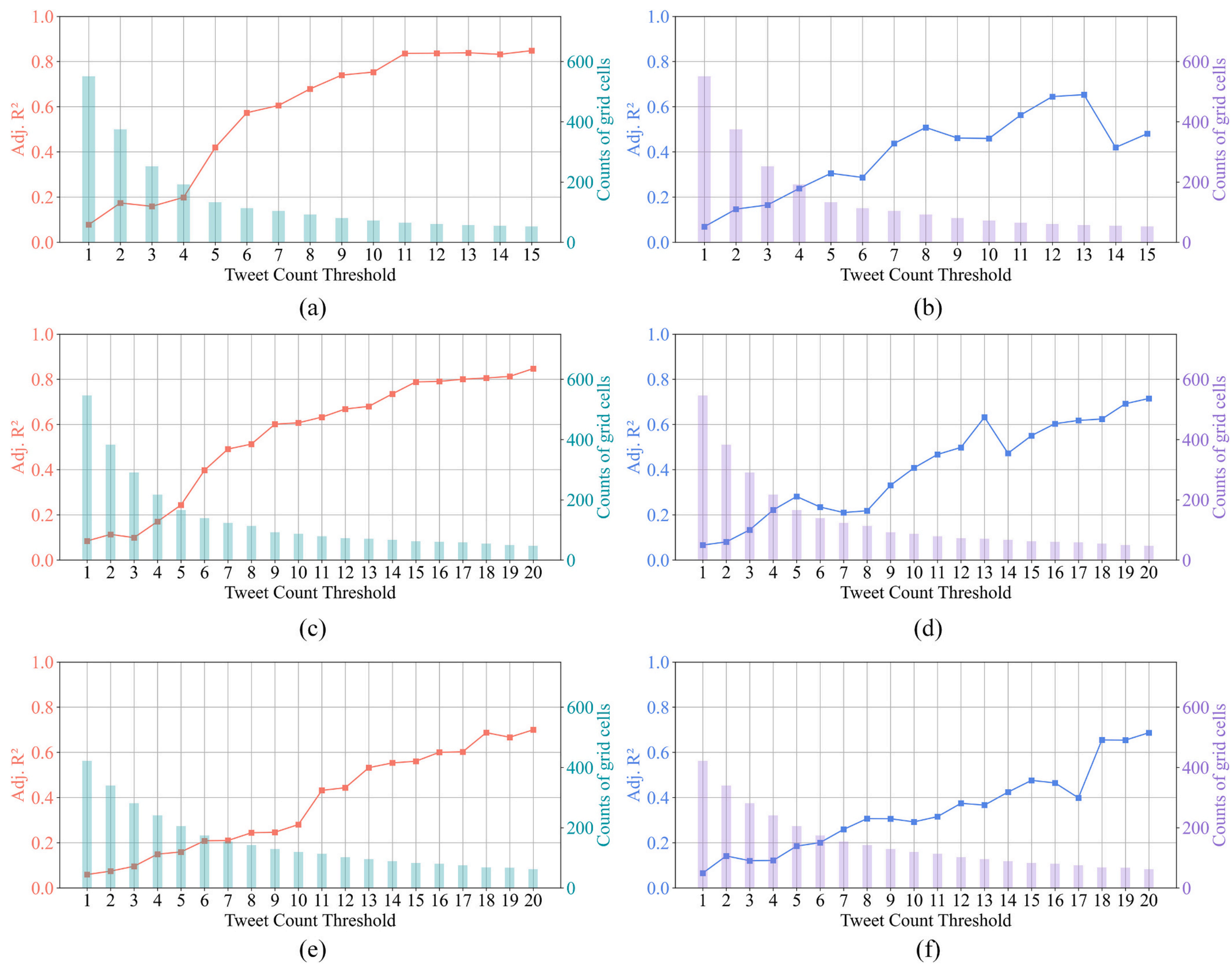

Fig. 7. Variations of $adj.R^2$ and grid cell counts across multiple sizes. Sub-figures (a), (c), and (e) on the left part of this graph demonstrate the results for positivity in human sentiments. In these sub-figures, red lines represent the variations of the values of $adj.R^2$, while green bars denote the numbers of eligible grid cells corresponding to a series of thresholds. The remaining sub-figures on the right part demonstrate the results for negativity. The blue lines and purple bars represent the values of $adj.R^2$ and counts of grid cells, respectively. Each row in this figure represents a distinct grid cell size. (a) and (b) show the results for 100 m, (c) and (d) for 200 m, (e) and (f) for 400 m. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

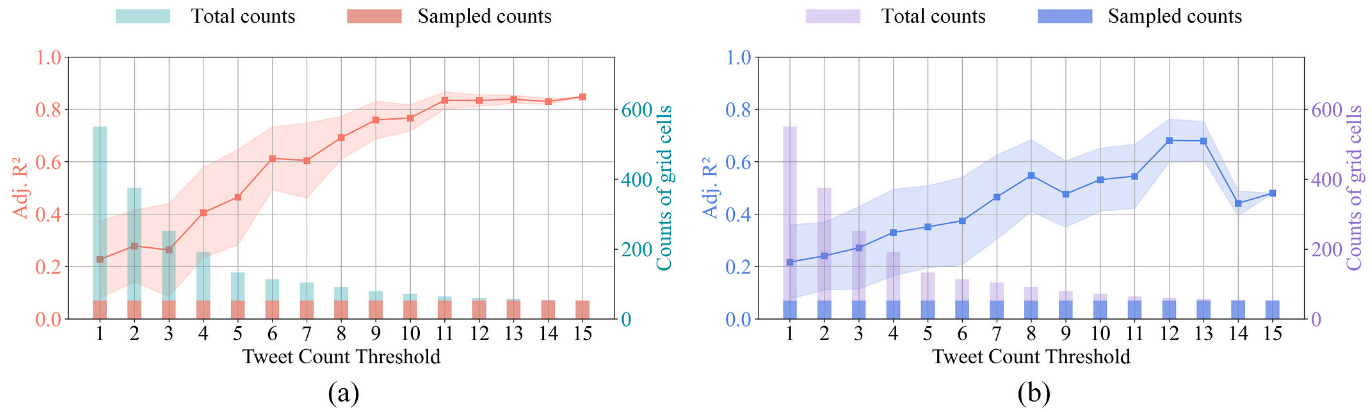

Fig. 8. Sensitivity analysis of sample sizes. The two subfigures illustrate how the $adj.R^2$ varies with different tweet count thresholds, while maintaining an identical number of grid cells, as sampled by Monte Carlo. The red and blue curves represent the variation in $adj.R^2$ for positive and negative sentiments, respectively. The shaded areas around these lines indicate the range between the mean minus one standard deviation (Std.) and the mean plus one Std., based on repeated Monte Carlo sampling. (a) and (b) display trends for both positive and negative sentiments when 53 grid cells, i.e., those containing at least 15 tweets with a size of 100 m, are sampled from all eligible cells, as shown by the red and blue bars, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

because soundscapes become more prominent in noisy situations compared to visual stimuli, drawing more attention and potentially leading to annoyance (Buxton et al., 2021; Li et al., 2022). Accordingly, roadside greenery can not only improve the aesthetic appeal of streets visually but also serve as a buffer to prevent roadside neighborhoods from traffic-related noise exposure, thereby improving the experiences of both pedestrians and community residents (Zhao et al., 2021). The impact of tactile stimuli is multifaceted. Higher ambient temperature significantly reduces negative sentiments, consistent with findings in the cooler climate zones (Wang et al., 2025), whereas relative humidity does not have a prominent effect on either positivity or negativity. This demonstrates that thermal comfort is crucial for mitigating unpleasant feelings, particularly in mild seasons and temperate regions. Nevertheless, the findings of this research may be partially applicable to cities with similar climates and mild seasons, but not generalizable to tactile sense in hot summers or cold winters, during which providing shelter from heat waves could further enhance positive sentiments (Gu et al., 2026; Park et al., 2023). While the influence of the olfactory sense is rather limited, this is likely because the study measures only $PM_{2.5}$, which people cannot directly sense through smelling, leading to an underestimation of its actual effects. Other odorous air includes volatile organic compounds (VOCs), sulfur compounds like hydrogen sulfide (H_2S), nitrogen compounds like ammonia (NH_3). Specifically, VOCs are the primary source of odors in urban environments, such as fragrant VOCs released by flowers and leaves that evoke pleasant feelings (Cai et al., 2025). Proper management and a diversified selection of urban flora enhance visual and olfactory senses, while minimizing the risk of pollen allergy. Nonetheless, there remain other environmental features that could potentially influence human sentiments, such as building texture, acoustic frequency and quality, surface temperatures, and specific odorous gases (Maria & Hollander, 2025; Perrine et al., 2025; Zhang & Kang, 2022). Although this study does not incorporate an exhaustive set of environmental variables across multiple senses, and their omission may constrain a more comprehensive understanding of every sensory pathway, this study still provides a baseline analysis for capturing multi-sensory experiences at the population level.

4.2. Role of representativeness in human sentiment analysis

Although the intricate interplays between them are analyzed through multivariate linear regression, there are still some limitations and challenges that need to be addressed. The most prominent issue is spatial representativeness, which may introduce uncertainty and confounding factors into the analysis.

Spatial representativeness means that observations can reflect the overall human sentiment of a certain place, and it is crucial in human sentiment analysis. In this research, relatively low coefficients of determination are observed when all the grid cells are considered. By contrast, when we filter out grid cells containing tweets less than a certain threshold, the $adj.R^2$ of the model increases. Specifically, when the tweet count threshold exceeds 10, the $adj.R^2$ value remains satisfactory and stable, demonstrating that a sufficient volume of human sentiment observations is significant for representativeness. In general, we provide a rule-of-thumb that more than 10 unique sentiment observations from social media are relatively reliable in reflecting general sentiments at a place in the Bronx. Our findings have implications for generating representative results on human sentiments. For example, to reflect general sentiments within an approximately 100 m by 100 m area, at least responses from more than ten participants are needed (Wang et al., 2023). It also suggests that social media can serve as a relatively reliable tool for evaluating public space, enabling urban designers to make data-driven decisions about urban well-being and create more inclusive urban environments. While these findings may be limited to specific urban contexts with mild climates and relatively flat terrain, previous studies on mobile environmental sensing support the existence of the 'rule-of-thumb' and empirically demonstrate that 4–10 repeated measurements on distinct days at the same location are adequate to generate representative spatial patterns. Yet, the rule-of-thumb's spatial and seasonal transferability still needs careful investigation (Patton et al., 2015; Yuan et al., 2024; Zalzal et al., 2019). Different rules of thumb may emerge if research is conducted in other areas with distinct built environments or during other months of the year.

4.3. Effectiveness of cross-sensory interactions

To evaluate the role of cross-sensory interactions, we compare the results from three sets of models, including cross-sensory interaction models, non-interaction baseline models, and a Random Forest regression. The interaction models and baseline models use the same forward selection linear regression method, while the baseline omits all interaction terms. The Random Forest model takes in all explanatory and control variables and accounts for non-linearity and interactions automatically in training. Despite the relatively high explanatory performance of the interaction models, the baseline models, which include single representative indicators for each sense in the explanatory variables, achieve results comparable to those of the interaction models (see

Fig. S7 in Section 8 of the SI). Similar to the results of interaction models, the visual environmental indicator 'beautiful' emerges as the primary contributor to positive sentiments, while 'wealthy' for visual sense and 'sound level' for hearing consistently exhibit negative coefficients as the tweet count threshold increases. When the tweet count threshold exceeds ten, the $adj.R^2$ values of baseline models plateau above 0.80 for positivity and peak at approximately 0.60 for negativity. The robust explanatory power of the baseline models indicates that cross-sensory interactions contribute little to the model's explanatory power, suggesting that explicit interactions between different sensory pathways are not particularly prominent. This is further evidenced by Random Forest models, which yield similar R^2 values of approximately 0.80 and 0.70 for positivity and negativity, respectively (see Fig. S10 in Section 10 of the SI). Overall, while cross-sensory interactions naturally occur in multi-sensory experiences, place-based human sentiments are primarily driven by the effects of independent indicators for each sense rather than their complex combinations in our experimental settings.

5. Conclusion and future work

In conclusion, the study reveals how urban physical environments shape nuanced human sentiments through multiple senses. We propose a framework that integrates multi-sensory experiences, visual, auditory, tactile, and olfactory, using multiple data sources, including drive-by sensing and street view images. Then, we apply multivariate linear regression to explain sentiments extracted from social media using environmental indicators derived from the abovementioned data sources. By incorporating multiple senses, this approach offers a more holistic understanding of how people perceive urban spaces. Results show that, in addition to visual senses, auditory and tactile senses can also influence human sentiments. Furthermore, a sensitivity analysis is performed to confirm the robustness of the proposed framework. Besides, this research also has practical implications for urban planning and design, as it suggests potential targeted strategies to optimize urban soundscapes and thermal comfort.

In the future, we would like to further expand the exploration in three directions. Firstly, the number of tweets in grid cells is relatively small, leaving a lot of locations uncovered. It is attributed to the difficulty in measuring urban environments and human sentiments in good spatio-temporal accordance. To address this issue, we will use other data sources to access human sentiments. For example, RedNote, a popular Chinese social media platform that is similar to Instagram, could provide novel insights into how people feel from a different perspective. Besides, reviews on Google Maps directly reflect people's judgment toward a specific place, which would be more explicit than those on social media. Secondly, this study only considers PM2.5 concentration but does not account for other airborne substances that may affect human sentiments in locations, such as VOCs. VOCs play a crucial part in olfactory sense and contribute to the formation of secondary particulate matter (PM) aerosols. They react with oxidants such as nitrogen oxides and then produce chemicals that can condense into PM. Meanwhile, ozone produced in these reactions can further accelerate the formation of PM aerosols under strong sunlight. The impacts of PM aerosols are twofold: odorous particles affect people's sense of smell, while drastically reduced visibility makes the city appear hazy and dusky, in turn diminishing the perceived beauty and vibrancy of the urban environment. In view of this, we suggest measuring various odorous substances, including total VOCs, H_2S, and NH_3, alongside $PM_{2.5}$ in future studies for a more comprehensive assessment of urban smell sense. Lastly, how the senses investigated in this research affect human sentiments is highly entangled with the local context. For example, people from a tropical climate might find the temperature in New York's summer quite pleasant, or overall, New York's noise levels are absolutely insufferable. Future studies are needed in various urban settings across cultures to draw generalizable conclusions.

CRediT authorship contribution statement

Xiaotong Zhang: Writing – review & editing, Writing – original draft, Formal analysis, Visualization, Validation, Methodology, Conceptualization. **Yuhao Kang:** Writing – review & editing, Data curation, Conceptualization. **Simone Mora:** Writing – review & editing. **Fábio Duarte:** Writing – review & editing. **Carlo Ratti:** Writing – review & editing, Resources. **An Wang:** Writing – review & editing, Validation, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization.

Acknowledgements

This study is funded by the Smart City Research Institute (SCRI) at the Hong Kong Polytechnic University (Project ID: P0058094).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compenvurbsys.2026.102489.

Data availability

Data will be made available on request.

References

Anjomshoaa, A., Mora, S., Schmitt, P., & Ratti, C. (2018). Challenges of Drive-By IoT sensing for smart cities: City scanner case study. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers, Singapore, Singapore*. https://doi.org/10.1145/3267305.3274167

Apte, J. S., Messier, K. P., Gani, S., Brauer, M., Kirchstetter, T. W., Lunden, M. M., … Hamburg, S. P. (2017). High-resolution air pollution mapping with Google street view cars: Exploiting big data. *Environmental Science & Technology, 51*(12), 6999–7008. https://doi.org/10.1021/acs.est.7b00891

Barbieri, F., Camacho-Collados, J., Neves, L., & Espinosa-Anke, L. T. (2020). Unified benchmark and comparative evaluation for tweet classification. In *arXiv 2020. arXiv preprint arXiv:2010.12421*.

Bricka, S., Reuscher, T., Schroeder, P., Fisher, M., Beard, J., & Sun, X. L. (2024). *Summary of travel trends: 2022 national household travel survey*.

Buxton, R. T., Pearson, A. L., Allou, C., Fristrup, K., & Wittemyer, G. (2021). A synthesis of health benefits of natural sounds and their distribution in national parks. *Proceedings of the National Academy of Sciences, 118*(14), Article e2013097118.

Cai, Y., Chen, H., Zhang, X., Bu, W., & Ning, W. (2025). Benefits of Gardenia jasminoides Ellis's floral volatile components on human emotions and moods. *Scientific Reports, 15*(1), 4194.

Camacho-Collados, J., Rezaee, K., Riahi, T., Ushio, A., Loureiro, D., Antypas, D., … Martínez Cámara, E. (2022). TweetNLP: Cutting-edge natural language processing for social media. In W. Che, & E. Shutova (Eds.), *Proceedings of the 2022 Conference on empirical methods in natural language processing: System demonstrations Abu Dhabi, UAE*.

Chen, T., Zhang, P., Chu, B., Ma, Q., Ge, Y., Liu, J., & He, H. (2022). Secondary organic aerosol formation from mixed volatile organic compounds: Effect of RO2 chemistry and precursor concentration. *npj Climate and Atmospheric Science, 5*(1), Article 95. https://doi.org/10.1038/s41612-022-00321-y

Dai, L., Zheng, C., Dong, Z., Yao, Y., Wang, R., Zhang, X., Ren, S., Zhang, J., Song, X., & Guan, Q. (2021). Analyzing the correlation between visual space and residents' psychology in Wuhan, China using street-view images and deep-learning technique. *City and Environment Interactions, 11*, Article 100069.

Gu, X., Beuster, L., Liu, X., van Leeuwen, E., Venverloo, T., & Duarte, F. (2026). Global patterns of inequality in pedestrian shade provision. *Nature Communications, 17*(1), Article 2563. https://doi.org/10.1038/s41467-026-69190-w

Guo, R., Qi, Y., Zhao, B., Pei, Z., Wen, F., Wu, S., & Zhang, Q. (2022). High-resolution urban air quality mapping for multiple pollutants based on dense monitoring data and machine learning. *International Journal of Environmental Research and Public Health, 19*(13), 8005. https://doi.org/10.3390/ijerph19138005

Hatzopoulou, M., Valois, M. F., Levy, I., Mihele, C., Lu, G., Bagg, S., … Brook, J. (2017). Robustness of land-use regression models developed from mobile air pollutant measurements. *Environmental Science & Technology, 51*(7), 3938–3947. https://doi.org/10.1021/acs.est.7b00366

Herrera, C., & Cabrera-Barona, P. (2022). Impact of perceptions of air pollution and noise on subjective well-being and health. *Earth, 3*(3), 825. https://doi.org/10.3390/earth3030047

Hu, Y., Deng, C., & Zhou, Z. (2019). A Semantic and Sentiment Analysis on Online Neighborhood Reviews for Understanding the Perceptions of People toward Their Living Environments. *Annals of the American Association of Geographers, 109*(4), 1052–1073. https://doi.org/10.1080/24694452.2018.1535886

Hu, Y., Wu, Y., Tantian, Z., & Sun, G. (2024). Capturing urban green view with mobile crowd sensing. *Ecological Informatics, 81*, Article 102640. https://doi.org/10.1016/j.ecoinf.2024.102640

Huang, J., Fei, T., Kang, Y., Li, J., Liu, Z., & Wu, G. (2024). Estimating urban noise along road network from street view imagery. *International Journal of Geographical Information Science, 38*(1), 128–155.

Ito, K., Kang, Y., Zhang, Y., Zhang, F., & Biljecki, F. (2024). Understanding urban perception with visual data: A systematic review. *Cities, 152*, Article 105169.

Kwan, M.-P. (2018). The neighborhood effect averaging problem (NEAP): An elusive confounder of the neighborhood effect. *International Journal of Environmental Research and Public Health, 15*(9), 1841. https://www.mdpi.com/1660-4601/15/9/1841.

Lam, C. K. C., Yang, H., Yang, X., Liu, J., Ou, C., Cui, S., … Hang, J. (2020). Cross-modal effects of thermal and visual conditions on outdoor thermal and visual comfort perception. *Building and Environment, 186*, Article 107297. https://doi.org/10.1016/j.buildenv.2020.107297

Lanotte, M., & Siracusa, G. (2022). Analysis and visualization of the sense of place from twitter data. In *Proceedings of the 7th International Workshop on Social Media World Sensors, Barcelona, Spain*. https://doi.org/10.1145/3544795.3544854

Li, W., Zhai, J., & Zhu, M. (2022). Characteristics and perception evaluation of the soundscapes of public spaces on both sides of the elevated road: A case study in Suzhou, China. *Sustainable Cities and Society, 84*, Article 103996.

Liu, M., Luo, X., & Lu, W.-Z. (2023). Public perceptions of environmental, social, and governance (ESG) based on social media data: Evidence from China. *Journal of Cleaner Production, 387*, Article 135840.

Loureiro, D., Barbieri, F., Neves, L., Anke, L. E., & Camacho-Collados, J. (2022). Timelms: Diachronic language models from Twitter. In *arXiv preprint arXiv:2202.03829*.

Ma, W., & Thompson, W. F. (2015). Human emotions track changes in the acoustic environment. *Proceedings of the National Academy of Sciences, 112*(47), 14563–14568.

Maria, C., & Hollander, J. B. (2025). Urban façades and human stress: Evaluating façades' texture and layout. *Building Research and Information: The International Journal of Research, Development and Demonstration, 53*(8), 1039–1057. https://doi.org/10.1080/09613218.2025.2506065

Niu, K., & Xing, Y. (2024). *Opinion distribution: Spatial sentiment analysis of online restaurant reviews through BERT model and GIS. Fourth international conference on signal processing and machine learning (CONF-SPML 2024)*.

O'Keeffe, K. P., Anjomshoaa, A., Strogatz, S. H., Santi, P., & Ratti, C. (2019). Quantifying the sensing power of vehicle fleets. *Proceedings of the National Academy of Sciences, 116*(26), 12752–12757.

Park, Y., Zhao, Q., Guldmann, J.-M., & Wentz, E. A. (2023). Quantifying the cumulative cooling effects of 3D building and tree shade with high resolution thermal imagery in a hot arid urban climate. *Landscape and Urban Planning, 240*, Article 104874. https://doi.org/10.1016/j.landurbplan.2023.104874

Patton, A. P., Zamore, W., Naumova, E. N., Levy, J. I., Brugge, D., & Durant, J. L. (2015). Transferability and generalizability of regression models of ultrafine particles in urban neighborhoods in the Boston area. *Environmental Science & Technology, 49*(10), 6051–6060. https://doi.org/10.1021/es5061676

Perrine, A.-L., Pédrono, G., Morel, P., Empereur-Bissonnet, P., Richard, J.-B., Golliot, F., … Motreff, Y. (2025). Mental health impact of a large-scale industrial fire on exposed residents in France: The post fire 76 health survey. *Archives of Public Health = Archives belges de santé publique, 83*(1). https://doi.org/10.1186/s13690-025-01602-7, 125–111.

Shin, B., Ryu, S., Kim, Y., & Kim, D. (2022). Analysis on review data of restaurants in Google maps through text mining: Focusing on sentiment analysis. *Journal of Multimedia Information System, 9*(1), 61–68.

Simion, S., Kovacs, M., & Toth, L. (2024). *Urban noise, a source of discomfort. 389*. Les Ulis: EDP Sciences.

Siragusa, G., & Leone, V. (2018). Such a wonderful place: Extracting sense of place from twitter. In a. Gangemi, A. L. Gentile, A. G. Nuzzolese, S. Rudolph, M. Maleshkova, H. Paulheim, … M. Alam (Eds.), *the semantic web: ESWC 2018 satellite events Cham*.

Stone, A. A., Schneider, S., & Harter, J. K. (2012). Day-of-week mood patterns in the United States: On the existence of 'blue Monday', 'thank god it's Friday' and weekend effects. *The Journal of Positive Psychology, 7*(4), 306–314. https://doi.org/10.1080/17439760.2012.691980

Suk, H. W., Choi, E., Na, J., Choi, J., & Choi, I. (2021). Within-person day-of-week effects on affective and evaluative/cognitive well-being among Koreans. *Emotion, 21*(5), 1114–1118. https://doi.org/10.1037/emo0000930

U.S. Census Bureau. (2024). *QuickFacts: Bronx County, New York*. U.S. Department of Commerce. Retrieved June 3rd, 2025, from https://www.census.gov/quickfacts/fact/table/bronxcountynewyork/PST045224.

Wang, A., Machida, Y., deSouza, P., Mora, S., Duhl, T., Hudda, N., … Ratti, C. (2023). Leveraging machine learning algorithms to advance low-cost air sensor calibration in stationary and mobile settings. *Atmospheric Environment, 301*, Article 119692. https://doi.org/10.1016/j.atmosenv.2023.119692

Wang, A., Mora, S., Machida, Y., deSouza, P., Paul, S., Oyinlola, O., Duarte, F., & Ratti, C. (2023). Hyperlocal environmental data with a mobile platform in urban environments. *Scientific Data, 10*(1), 524. https://doi.org/10.1038/s41597-023-02425-3

Wang, A., Paul, S., deSouza, P., Machida, Y., Mora, S., Duarte, F., & Ratti, C. (2023). Key themes, trends, and drivers of mobile ambient air quality monitoring: A systematic review and meta-analysis. *Environmental Science & Technology, 57*(26), 9427–9444. https://doi.org/10.1021/acs.est.2c06310

Wang, C., Bai, Y.-X., Li, X.-W., & Lin, L.-T. (2024). Effects of extreme temperatures on public sentiment in 49 Chinese cities. *Scientific Reports, 14*(1), 9954.

Wang, J., Guetta-Jeanrenaud, N., Palacios, J., Fan, Y., Kakkar, D., Obradovich, N., & Zheng, S. (2025). Unequal impacts of rising temperatures on global human sentiment. *One Earth, 8*(10), 101422.

Yan, J., Huang, X., Wang, S., He, Y., Li, X., Hohl, A., … Lin, B. (2023). Toward a comprehensive understanding of eye-level urban greenness: A systematic review. *International Journal of Digital Earth, 16*(2), 4769–4789.

Yuan, Z., Kerckhoffs, J., Li, H., Khan, J., Hoek, G., & Vermeulen, R. (2024). Hyperlocal air pollution mapping: A scalable transfer learning LUR approach for mobile monitoring. *Environmental Science & Technology, 58*(32), 14372–14383. https://doi.org/10.1021/acs.est.4c06144

Zalzal, J., Alameddine, I., El Khoury, C., Minet, L., Shekarrizfard, M., Weichenthal, S., & Hatzopoulou, M. (2019). Assessing the transferability of landuse regression models for ultrafine particles across two Canadian cities. *Science of the Total Environment, 662*, 722–734. https://doi.org/10.1016/j.scitotenv.2019.01.123

Zhang, B., & Kang, J. (2022). Effect of environmental contexts pertaining to different sound sources on the mood states. *Building and Environment, 207*, Article 108456.

Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H. H., Lin, H., & Ratti, C. (2018). Measuring human perceptions of a large-scale urban region using machine learning. *Landscape and Urban Planning, 180*, 148–160.

Zhang, M., Zhang, X., Guo, S., Xu, X., Chen, J., & Wang, W. (2021). Urban micro-climate prediction through long short-term memory network with long-term monitoring for on-site building energy estimation. *Sustainable Cities and Society, 74*, Article 103227.

Zhang, Y., Kwan, M.-P., & Ma, H. (2024). Sensing noise exposure and its inequality based on noise complaint data through vision-language hybrid method. *Applied Geography, 171*, Article 103369.

Zhao, N., Prieur, J.-F., Liu, Y., Kneeshaw, D., Lapointe, E. M., Paquette, A., … Rainham, D. G. (2021). Tree characteristics and environmental noise in complex urban settings–A case study from Montreal, Canada. *Environmental Research, 202*, Article 111887.

Zhuang, Y., Kang, Y., Fei, T., Bian, M., & Du, Y. (2024). From hearing to seeing: Linking auditory and visual place perceptions with soundscape-to-image generative artificial intelligence. *Computers, Environment and Urban Systems, 110*, Article 102122.

Zou, K., Yu, X., Kwok, C. Y. T., Wong, M. S., Kwan, M.-P., & Hou, H. (2025). Simulation and exposure assessment of hourly traffic noise in Hong Kong using a minimal error iterative model based on diversion strategies. *Computers, Environment and Urban Systems, 120*, Article 102300. https://doi.org/10.1016/j.compenvurbsys.2025.102300