This paper is peer-reviewed and accepted at the 34th Annual GIS Research UK Conference (GISRUK 2026).
https://zenodo.org/records/19387282
Doi: 10.5281/zenodo.19387282

Assessing Park Satisfaction from Google Maps Reviews: Novel Evidence from Multimodal Text–Image Analysis

Youlong Gu[1,2], Wenpei Li[1], Angelia Sia[3] and Filip Biljecki[1,4]

[1]Department of Architecture, National University of Singapore, Singapore
[2]School of Geography, University of Leeds, United Kingdom
[3]National Parks Board, Singapore
[4]Department of Real Estate, National University of Singapore, Singapore

Summary

Parks are essential to urban well-being, making park satisfaction crucial for sustainable city development. Traditional survey-based approaches to understand sentiment towards parks among residents are often costly, time-consuming, and limited in scale. Recent social media–based studies have scaled such research but predominantly focus on text and frequently overlook visual information and the joint effects of text–image representations. This study presents an automated multimodal framework using crowdsourced reviews from Google Maps to model park satisfaction by integrating textual and visual features. Using Singapore as a case study, we analysed 76,869 textual reviews and 184,322 images associated with them. The results show that multimodal models are more useful than text-only approaches, with textual sentiment, emotional attributes, and image temporal characteristics identified as the most influential factors. These findings highlight the importance of multimodal analysis for advancing park research and informing planning and policy practices.

KEYWORDS: Urban green spaces, Park perception, Crowdsourced data, Vision-language model.

1 Introduction

Parks are a crucial category of urban green spaces, providing essential recreational services that generate a wide range of physical, social, and psychological benefits for residents (Wolch et al., 2014; Koohsari et al., 2015; Akpinar et al., 2016; Remme et al., 2021; Petrunoff et al., 2021). These benefits emerge from residents' interactions with parks, making the relationship between parks and their users central to understanding patterns of park use and evaluating their broader societal value (Kaczynski et al., 2014; Edwards et al., 2015; Veitch et al., 2021). Therefore, how people perceive urban parks, and the extent to which these spaces meet their needs and expectations, is a critical factor linking park provision to broader goals of urban livability and social well-being.

Traditionally, park perception and satisfaction have been measured using qualitative and self-reported approaches, such as surveys and interviews (Kaczynski et al., 2014; Edwards et al., 2015; Petrunoff et al., 2021). While these methods provide valuable insights, they are often time-consuming, resource-intensive, and constrained by limited sample sizes, restricting their scalability and temporal coverage (Donahue et al., 2018; Chen et al., 2018). In recent years, the rapid growth of user-generated content has offered a promising alternative for studying park use and satisfaction, enabling the large-scale and cost-effective capture of user experiences. Existing studies have leveraged various forms of online content, including user density and contribution intensity (Levin et al., 2017; Chen et al., 2018; Huai et al., 2023), social media posts (Plunz et al., 2019; Lai and Deal, 2023) and place-based reviews (Liu and Xiao, 2021; Yang and Zhang, 2024; Zhao et al., 2024; Song et al., 2026), to examine a wide range of topics, such as spatial patterns of visitation, public sentiment, and perception.

Despite these advances, important research gaps remain. Existing studies leveraging user-generated data have largely focused on textual content, with limited consideration of the rich visual context information embedded in user-generated images (Song et al., 2026). In addition, existing image-based approaches are largely restricted to embedding-level representations, offering limited interpretability (Zhao et al., 2024). Overall, systematic multimodal approaches that extract and integrate more fine-grained representations from both textual and visual content to improve the understanding and modelling of park satisfaction remain insufficiently explored.

Motivated by these gaps, this study addresses two research questions:

- Can visual information from user-generated images complement textual information in modelling park satisfaction, and how can their joint effects be systematically evaluated?
- Which textual and visual representations can be quantitatively modelled to predict park satisfaction?

To answer these questions, we propose an automated framework that utilises Google Maps Reviews and multimodal foundation models to extract structured representations from review texts and images. We compare unimodal and multimodal models to assess the performance in predicting park satisfaction and to identify influential textual and visual features. The proposed framework provides a scalable and interpretable framework for advancing park analysis with implications for urban planning and policy.

2 Methodology

2.1 GMR data collection

We developed an automated pipeline to collect Google Maps Reviews (GMR) data using a custom-built tool. GMR was chosen as the primary data source because of its strong place specificity, enabling insights into local residents' place-based experiences, and its rapid growth in data volume (Veitch et al., 2021; Yang and Zhang, 2024; Song et al., 2026).

Data collection is conducted in two stages. First, using 2000 m radius queries and 'park' as the search term, we retrieve metadata for all park-related points of interest (POIs). The average park rating is subsequently used as the ground-truth measure of user satisfaction, following previous practices (Huai et al., 2023). Reviews are then automatically collected for each POI, including user-generated texts and images, with all user identifiers anonymised.

2.2 Multimodal alignment and attribute extraction

As user-generated content, real-world GMR data often exhibit substantial multimodal noise, such as weak semantic correspondence between review text and images, or conflicting emotional cues across modalities. To mitigate these issues, we implement a multimodal alignment procedure as a preprocessing step.

Using OpenCLIP ViT-H/14, (Schuhmann et al., 2022; Radford et al., 2021), text and images are embedded into a shared vector space, and pairwise text–image similarity scores are computed to identify images and to weight image-level contributions. Following multimodal alignment, we further transform both modalities into structured representations to investigate how specific indicators contribute to quantitative outcomes. For the text modality, we extract sentiment, emotions, and topics, while image indicators include scene types, visual topics, and temporal information. Review texts are processed using the Qwen3 Instruct 4B model, and review images with Qwen3-VL Instruct 8B model (Bai et al., 2025; Yang et al., 2025), using designed prompts to output attributes in a unified JSON schema we conceived. Detailed definitions and attributes are provided in Table 1.

Table 1: Unified JSON schema for structured information extraction from text and image modalities.

Modality	**Field**	**Description**	**Example**
Text	sentiment	Overall sentiment polarity in $[-1, 1]$	`0.7`
Text	emotions	Top-2 emotions with confidence (0–1)	`Joy:0.9, Neutral:0.1`
Text	topics	High-level semantic concepts discussed	`[water, nature, pond]`
Image	scene type	Top-2 Places365 labels with confidence	`park:0.85, pond:0.78`
Image	topics	High-level summary of visual content	`[urban_park, water_lilies]`
Image	time	Temporal cue: day, night, or unknown	`day`

2.3 Feature set design and modelling

Based on these structured outputs, we construct three progressively extended feature sets for park satisfaction prediction, as summarised in Table 2. Given that sentiment has been validated as an effective proxy for park perception and satisfaction (Huai et al., 2023; Mohamed and Kronenberg, 2025), the first feature set (Model A: Text Baseline) includes only the mean sentiment score of textual reviews and serves as a baseline for assessing the explanatory power of sentiment polarity alone.

The second feature set (Model B: Text Structural) extends this baseline by incorporating measures of structural complexity. Beyond numerical sentiment indicators, we compute entropy-based metrics across multiple dimensions such as emotional distributions and thematic topics, to capture whether reviews associated with a given location exhibit concentrated or diverse expressive patterns. Accordingly, all entropy-based indicators in the feature sets are named using the suffix diversity.

Table 2: Summary of feature sets used for park satisfaction prediction.

Model	Feature family	Feature
A: Text Baseline	Text sentiment	Text mean sentiment
B: Text Structural	Text structure	Text mean sentiment, Text emotion diversity, Text topic diversity
C: Multimodal Structural	Text + Image structure	Text mean sentiment, Text emotion diversity, Text topic diversity, Image scene diversity, Image topic diversity, Image daytime share

The third feature set (Model C: Multimodal Structural) further extends the same structural perspective to the visual modality while keeping all textual features unchanged. Image-level features include the entropy-based diversity of scene category distributions and image topic distributions, as well as the proportion of daytime images. All image features are aggregated at the location level using text-image semantic similarity scores as weights, such that images with higher semantic relevance to their associated reviews contribute more strongly to the final multimodal representation.

All three feature sets are evaluated using linear Ridge regression with cross-validation. A linear model is adopted to ensure that changes in predictive performance reflect genuine information gains from additional features rather than increased model flexibility. Ridge regression estimates coefficients by minimising:

$$\min_{\boldsymbol{\beta}} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda\|\boldsymbol{\beta}\|_2^2, \tag{1}$$

where $\mathbf{X}$ denotes the standardized feature matrix, $\mathbf{y}$ represents the observed park satisfaction score, $\boldsymbol{\beta}$ are the regression coefficients, and λ controls the strength of L2 Regularisation. Regularisation stabilises coefficient estimation under correlated predictors, enabling robust and interpretable comparisons across feature sets. To reduce data noise, the analysis is restricted to POIs with more than 20 reviews. Models are trained independently for each location, and performance is summarised using the mean and standard deviation of R^2 and the mean absolute error (MAE) and its standard deviation across cross-validation folds.

3 Results

3.1 GMR of park-related public green space in Singapore

Singapore was selected as the study area due to its strong national commitment to urban greening (Tan, 2015; Friess, 2017; Sia et al., 2023), providing a rich park context and policy-relevant implications of park satisfaction at the national scale.

Using our automated web-scraping framework and the search keyword ‘park’, we collected a comprehensive set of park-related POIs across Singapore, resulting in a total of 1,938 POIs. While the primary category corresponds to parks, the retrieved POIs also include related public green spaces such as ‘gardens’, ‘playgrounds’, ‘hiking areas’ and so on. These categories were retained because they exhibit similar descriptive patterns with parks and provide valuable textual and visual user-generated content. Figure 1 presents a typical example of a place review and its associated user review.

Figure 1: A typical park review example. (A) POI-level summary information. (B) An individual user's textual review. (C) Images associated with the same user review.

In total, 1,140 POIs were found to have been reviewed, from which we collected 76,869 textual reviews and 184,322 images, spanning a temporal range from October 2012 to October 2025. For each POI, we extracted the average user rating and used it as the ground-truth indicator of park satisfaction. The mean rating across all selected POIs is 4.42, out of a maximum score of 5. In terms of review volume, a POI has an average of 569 reviews, with a median of 10 reviews, reflecting the highly skewed distribution of user engagement. Figure 2 maps the spatial distribution of park satisfaction across Singapore, based on ratings and the number of reviews.

3.2 Model performance evaluation

A minimum threshold of 20 textual reviews was applied, resulting in 654 POI locations. Within this subset, 174,111 text-image pairs were (average semantic similarity = 0.247) identified, and structured multimodal features were generated using the Qwen3 series models for subsequent Ridge regression analysis.

As shown in Table 3, the baseline sentiment model (Model A) achieves an average R^2 of 0.344 and $R^2_{\text{std}} = 0.061$, indicating that textual sentiment alone explains a substantial proportion of the variation in park satisfaction and yields stable performance across folds. Incorporating textual structural features in Model B leads to a marked performance improvement, with R^2 increasing to 0.549, suggesting the added value of text structure such as emotional diversity and thematic richness. Further incorporating visual structural features in Model C leads to a modest additional increase in explanatory power ($R^2 = 0.602$)., while reducing performance variability relative to Model B ($R^2_{\text{std}} = 0.071$ vs. 0.074), This result indicates that image-based features, while offering limited gains in mean performance within a linear framework, contribute to more stable and consistent predictions across locations.

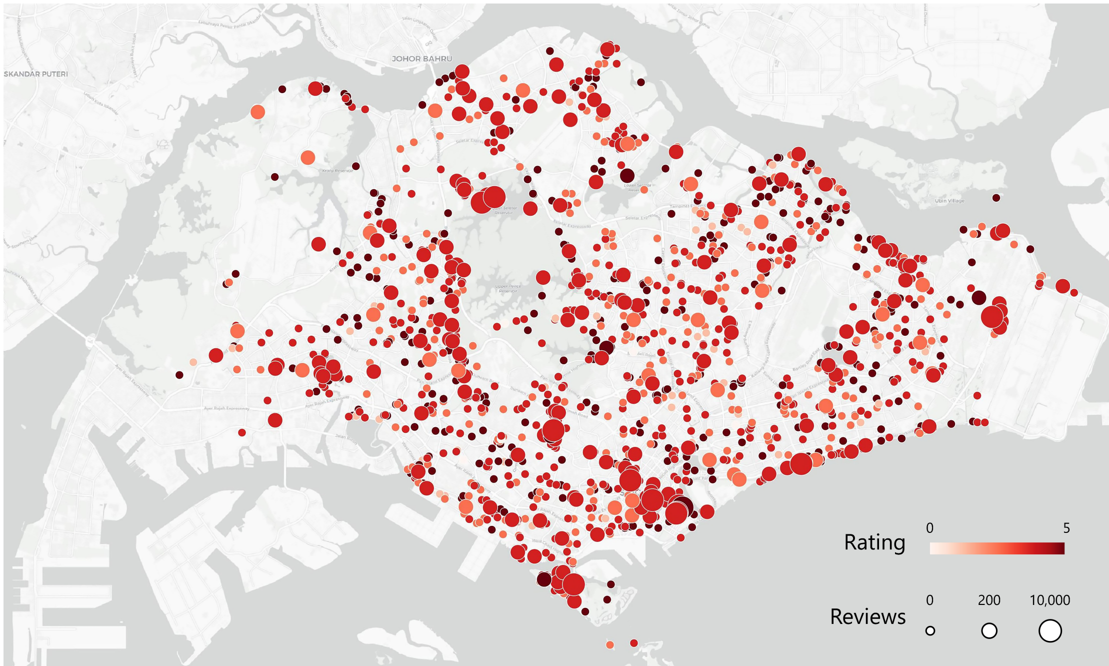

Figure 2: Spatial distribution of park-related public green space POIs across Singapore. Basemap: (c) OpenStreetMap contributors.

Table 3: Cross-validated performance of Models A to C for park satisfaction prediction.

Model	**n**	$\mathbf{R^2_{mean}}$	$\mathbf{R^2_{std}}$	$\mathbf{MAE_{mean}}$	$\mathbf{MAE_{std}}$
Model A: Text Baseline	1	0.344	0.039	0.179	0.008
Model B: Text Structural	3	0.549	0.074	0.151	0.007
Model C: Multimodal Structural	6	0.602	0.071	0.140	0.011

3.3 Feature contributions

We selected the best-performing linear model in terms of predictive accuracy, Model C (Multimodal Structural), and applied SHAP (SHapley Additive exPlanations) to interpret feature contributions at the sample level. Figure 3 illustrates the distribution of SHAP values for each feature and contribution patterns.

The results indicate that **Text mean sentiment** is the most dominant predictor, exhibiting the largest mean SHAP value (0.26) and positive effects, indicating that more positive emotional expressions in reviews are strongly associated with higher predicted park satisfaction. This finding is consistent with the conclusions drawn from the baseline models. In contrast, **Text emotion diversity** shows a substantial negative contribution (-0.17), implying that reviews with more complex and diverse emotions are associated with lower satisfaction. Such reviews often combine positive

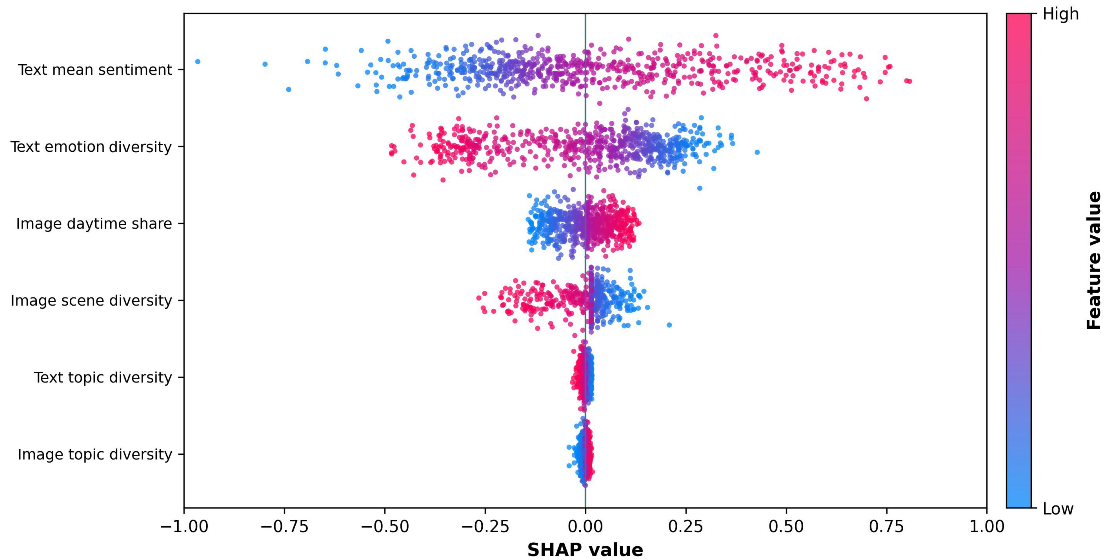

Figure 3: SHAP summary of multimodal feature contributions.

and negative aspects, reflecting more critical and nuanced user experiences, and may therefore be informative for diagnostic analyses.

Regarding visual features, **Image daytime share** exhibits a modest but consistent positive contribution (0.05), indicating that the temporal visual characteristics are stably associated with park satisfaction. Meanwhile, **Image scene diversity** demonstrates a slight negative effect (−0.05), suggesting that visually complex scenes may correspond to the identification of multiple issues or concerns within the park environment. By comparison, both **Text topic diversity** and **Image topic diversity** contribute negligibly, suggesting that the number of topics is not directly associated with predicted satisfaction.

Overall, the SHAP analysis reveals that emotional intensity and emotional structure in text are the primary drivers of satisfaction predictions, while visual features play a secondary but complementary role by refining predictions along temporal and scene-related dimensions.

4 Conclusion

This study presents a novelty in analysing park satisfaction using place reviews by incorporating visual information. Through quantitative modelling, we show that multimodal models incorporating visual information are consistently more useful than text-only models in predicting park satisfaction. Beyond conventional textual sentiment, our findings highlight the value of additional multimodal indicators, such as emotion entropy derived and image-based temporal attributes, as informative predictors for park satisfaction. Incorporating these dimensions enables more nuanced observations and diagnostic analyses of park quality and perception, going beyond what can be captured by aggregated rating scores alone.

Despite these contributions, the current set of indicators remains primarily at the level of modal associations. Future research could integrate review data with fine-grained geographic information to identify which specific physical elements or environmental characteristics drive variations in particular indicators within user evaluations. Moreover, the automated framework developed in this study provides a scalable foundation for extending the approach to broader geographic contexts, thereby supporting more comparative investigations of park satisfaction across cities and regions.

Acknowledgements

This research is part of the project Spatial-temporal trends in usage & perception of parks and green spaces, which is supported by the National Parks Board, Singapore.

Biography

Youlong Gu is a Research Engineer at the National University of Singapore. He focuses on using computer vision and statistics to evaluate urban visual data, leveraging it to map urban patterns and analyse urban dynamics.

Wenpei Li is a PhD student at the National University of Singapore, focusing on the application of big data, particularly textual data, in the study of urban green spaces.

Angelia Sia is the Deputy Director of Research at the National Parks Board, Singapore. She received her PhD degree from the National University of Singapore. Her research investigates the interactions between people and nature in cities, and the associated well-being benefits.

Filip Biljecki is an assistant professor at the National University of Singapore and the founder of the NUS Urban Analytics Lab. He holds a MSc and PhD degree from the Delft University of Technology in the Netherlands.

References

Akpinar, A., Barbosa-Leiker, C., & Brooks, K. R. (2016). Does green space matter? Exploring relationships between green space type and health indicators. *Urban Forestry & Urban Greening*, *20*, 407–418, `https://doi.org/10.1016/j.ufug.2016.10.013`.

Bai, S., et al. (2025). Qwen3-VL Technical Report.

Chen, Y., Liu, X., Gao, W., Wang, R. Y., Li, Y., & Tu, W. (2018). Emerging social media data on measuring urban park use. *Urban Forestry & Urban Greening*, *31*, 130–141, `https://doi.org/10.1016/j.ufug.2018.02.005`.

Donahue, M. L., Keeler, B. L., Wood, S. A., Fisher, D. M., Hamstead, Z. A., & McPhearson, T. (2018). Using social media to understand drivers of urban park visitation in the Twin Cities, MN. *Landscape and Urban Planning*, *175*, 1–10, `https://doi.org/10.1016/j.landurbplan.2018.02.006`.

Edwards, N., Hooper, P., Knuiman, M., Foster, S., & Giles-Corti, B. (2015). Associations between park features and adolescent park use for physical activity. *International Journal of Behavioral Nutrition and Physical Activity*, *12*(1), 21, `https://doi.org/10.1186/s12966-015-0178-4`.

Friess, D. A. (2017). Singapore as a long-term case study for tropical urban ecosystem services. *Urban Ecosystems*, *20*(2), 277–291, `https://doi.org/10.1007/s11252-016-0592-7`.

Huai, S., Liu, S., Zheng, T., & Van de Voorde, T. (2023). Are social media data and survey data consistent in measuring park visitation, park satisfaction, and their influencing factors? A case study in Shanghai. *Urban Forestry & Urban Greening*, *81*, 127869, `https://doi.org/10.1016/j.ufug.2023.127869`.

Kaczynski, A. T., et al. (2014). Are park proximity and park features related to park use and park-based physical activity among adults? Variations by multiple socio-demographic characteristics. *International Journal of Behavioral Nutrition and Physical Activity*, *11*(1), 146, `https://doi.org/10.1186/s12966-014-0146-4`.

Koohsari, M. J., et al. (2015). Public open space, physical activity, urban design and public health: Concepts, methods and research agenda. *Health & Place*, *33*, 75–82, `https://doi.org/10.1016/j.healthplace.2015.02.009`.

Lai, S. & Deal, B. (2023). Parks, Green Space, and Happiness: A Spatially Specific Sentiment Analysis Using Microblogs in Shanghai, China. *Sustainability*, *15*(1), 146, `https://doi.org/10.3390/su15010146`.

Levin, N., Lechner, A. M., & Brown, G. (2017). An evaluation of crowdsourced information for assessing the visitation and perceived importance of protected areas. *Applied Geography*, *79*, 115–126, `https://doi.org/10.1016/j.apgeog.2016.12.009`.

Liu, R. & Xiao, J. (2021). Factors Affecting Users' Satisfaction with Urban Parks through Online Comments Data: Evidence from Shenzhen, China. *International Journal of Environmental Research and Public Health*, *18*(1), 253, `https://doi.org/10.3390/ijerph18010253`.

Mohamed, A. A. & Kronenberg, J. (2025). Users' experiences of park accessibility and attractiveness based on online review analytics. *Scientific Reports*, *15*(1), 4268, `https://doi.org/10.1038/s41598-025-88500-8`. Publisher: Nature Publishing Group.

Petrunoff, N. A., et al. (2021). Associations of park access, park use and physical activity in parks with wellbeing in an Asian urban environment: a cross-sectional study. *International Journal of Behavioral Nutrition and Physical Activity*, *18*(1), 87, `https://doi.org/10.1186/s12966-021-01147-2`.

Plunz, R. A., Zhou, Y., Carrasco Vintimilla, M. I., Mckeown, K., Yu, T., Uguccioni, L., & Sutto, M. P. (2019). Twitter sentiment in New York City parks as measure of well-being. *Landscape and Urban Planning*, *189*, 235–246, `https://doi.org/10.1016/j.landurbplan.2019.04.024`.

Radford, A., et al. (2021). Learning Transferable Visual Models From Natural Language Supervision.

Remme, R. P., et al. (2021). An ecosystem service perspective on urban nature, physical activity, and health. *Proceedings of the National Academy of Sciences*, *118*(22), e2018472118, https://doi.org/10.1073/pnas.2018472118.

Schuhmann, C., et al. (2022). LAION-5B: An open large-scale dataset for training next generation image-text models.

Sia, A., Tan, P. Y., Kim, Y. J., & Er, K. B. H. (2023). Use and non-use of parks are dictated by nature orientation, perceived accessibility and social norm which manifest in a continuum. *Landscape and Urban Planning*, *235*, 104758, https://doi.org/10.1016/j.landurbplan.2023.104758.

Song, Q., Tian, S., Zheng, L., Zheng, Y., Qiu, L., Huang, B., & Ameijde, J. v. (2026). Beyond sentiment: Using large language models to decode multidimensional urban park perceptions for enhanced equality. *Landscape and Urban Planning*, *268*, 105571, https://doi.org/10.1016/j.landurbplan.2025.105571.

Tan, P. Y. (2015). Greening Singapore: Past Achievements, Emerging Challenges. In *50 Years of Urban Planning in Singapore*, World Scientific Series on Singapore's 50 Years of Nation-Building (pp. 177–195).

Veitch, J., Rodwell, L., Abbott, G., Carver, A., Flowers, E., & Crawford, D. (2021). Are park availability and satisfaction with neighbourhood parks associated with physical activity and time spent outdoors? *BMC Public Health*, *21*(1), 306, https://doi.org/10.1186/s12889-021-10339-1.

Wolch, J. R., Byrne, J., & Newell, J. P. (2014). Urban green space, public health, and environmental justice: The challenge of making cities 'just green enough'. *Landscape and Urban Planning*, *125*, 234–244, https://doi.org/10.1016/j.landurbplan.2014.01.017.

Yang, A., et al. (2025). Qwen3 Technical Report.

Yang, C. & Zhang, Y. (2024). Public emotions and visual perception of the East Coast Park in Singapore: A deep learning method using social media data. *Urban Forestry & Urban Greening*, *94*, 128285, https://doi.org/10.1016/j.ufug.2024.128285.

Zhao, X., Lu, Y., Huang, W., & Lin, G. (2024). Assessing and interpreting perceived park accessibility, usability and attractiveness through texts and images from social media. *Sustainable Cities and Society*, *112*, 105619, https://doi.org/10.1016/j.scs.2024.105619.