Published in final edited form as:

J Expo Sci Environ Epidemiol. 2022 March ; 32(2): 213–222. doi:10.1038/s41370-022-00409-w.

Spatial resolution of Normalized Difference Vegetation Index and greenness exposure misclassification in an urban cohort

Raquel B. Jimenez[a,*], **Kevin J. Lane**[a], **Lucy R. Hutyra**[b], **M. Patricia Fabian**[a]

[a]Department of Environmental Health. School of Public Health, Boston University. 715 Albany Street, Boston, Massachusetts 02118. United States of America

[b]Department of Earth and Environment, Boston University. 685 Commonwealth Avenue, Boston, Massachusetts 02215. United States of America

Abstract

Background: The Normalized Difference Vegetation Index (NDVI) is a measure of greenness widely used in environmental health research. High spatial resolution NDVI has become increasingly available; however, the implications of its use in exposure assessment are not well understood.

Objective: To quantify the impact of NDVI spatial resolution on greenness exposure misclassification.

Methods: Greenness exposure was assessed for 31,328 children in the Greater Boston Area in 2016 using NDVI from MODIS ($250m^2$), Landsat 8 ($30m^2$), Sentinel-2 ($10m^2$), and the National Agricultural Imagery Program (NAIP, $1m^2$). We compared continuous and categorical greenness estimates for multiple buffer sizes under a reliability assessment framework. Exposure misclassification was evaluated using NAIP data as reference.

Results: Greenness estimates were greater for coarser resolution NDVI, but exposure distributions were similar. Continuous estimates showed poor agreement and high consistency, while agreement in categorical estimates ranged from poor to strong. Exposure misclassification was higher with greater differences in resolution, smaller buffers, and greater number of exposure quantiles. The proportion of participants changing greenness quantiles was higher for MODIS (11%–60%), followed by Landsat 8 (6%–44%), and Sentinel-2 (5%–33%).

Significance: Greenness exposure assessment is sensitive to spatial resolution of NDVI, aggregation area, and number of exposure quantiles. Greenness exposure decisions should ponder relevant pathways for specific health outcomes and operational considerations.

*Corresponding author: raque@bu.edu (Raquel B. Jimenez).

Author contributions

RBJ was responsible for research design, retrieving and processing data, conducting statistical analysis, and writing the manuscript. MPF contributed to research design, provided feedback on the manuscript, and contributed to writing it. KJL and LH provided feedback on research design and the manuscript.

Competing interests

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

Keywords

Normalized Difference Vegetation Index (NDVI); Spatial resolution; Greenness; Exposure assessment; Exposure misclassification; Epidemiological studies

Introduction

Research across multiple disciplines has drawn attention to the critical role of vegetation in maintaining environmental quality, facilitating social cohesion, and promoting healthy behaviors in urban areas (1–4). Vegetation can moderate exposures to environmental hazards by reducing ambient concentrations of air pollution (5,6), acting as a physical and psychological buffer for noise (7,8), and reducing heat exposure (9). Accumulating evidence suggests predominantly beneficial associations between exposure to vegetation and multiple mental and physical health outcomes (10–15). While there is growing consensus on the salutatory effects of vegetation, research findings on the greenness-health relationship are mixed for several outcomes, including birth outcomes (16–18), academic performance (19–22), and asthma and allergic respiratory diseases (23–25). These inconsistencies might arise from differences in study area and population, or differences in study design, including how exposure to green spaces is conceptualized and measured. In fact, the heterogeneity in green space data and spatial methods used to assess exposure have been pointed as potential driver behind the conflictive evidence observed in the literature (26,27).

Many environmental health studies rely on greenness data –a measure of relative abundance and spatial distribution of vegetation– to characterize exposure (28–30). Greenness is generally assessed with spectral indices that identify vegetation using reflectance measurements taken by instruments on satellites and planes. Among the many vegetation indices available, the Normalized Difference Vegetation Index (NDVI) is the most commonly used in urban green space assessment (31) and environmental health studies (26,27,32). NDVI applications have increased due to the availability of satellite data over long periods of time and geographic extents at increasingly higher spatial and temporal resolution (24,27,33). Coarser resolution data (pixel size of 250 m^2 or more) are more likely to include a mix of land cover classes (e.g., impervious, water, vegetation, etc.) within the extent of each pixel (34–37). The mix of signals from different surfaces dilutes the spectral signal of vegetation, hampering the identification of small green spaces such as backyards, pocket parks, and street trees, which account for a large share of greenness exposure in urban contexts (38). Oppositely, the smaller pixels in high-resolution NDVI data (pixel size of 10 m^2 or less) are more likely to capture pure land cover classes, and therefore can better identify small-scale green cover. This is especially relevant for exposure assessment in urban areas where land cover presents high variability (39). Moreover, NDVI presents strong spatial scale dependencies when assessing vegetation over mixed surfaces, suggesting that NDVI derived from data at different spatial resolutions might not be directly comparable in such settings (36). While this issue is less relevant for pixels of pure vegetation, such as forests (40), the majority of epidemiologic studies are set in urban areas, where land cover is highly heterogeneous. Together, these conditions suggest that spatial resolution of NDVI data might affect exposure estimates.

Greenness exposure is typically assigned to individuals by averaging NDVI within radial buffers centered on residential addresses or locations of interest (25,28,41–43). Buffers of different radii are used to represent a variety of spatial scales and exposure contexts where interactions with natural green space occur. Smaller buffer sizes capture vegetation in the immediate surroundings of individuals, such as private yards and street trees, and larger buffer sizes account for neighborhood-level greenness that might be encountered during commuting or recreational activities (26,27,44). Studies exploring the interactions between scale of analysis and NDVI resolution have observed changes in effect size and statistical significance of associations across buffer sizes for several health outcomes (45–47). While these differences might owe to the definition of aggregation areas or reflect the fact that different ecosystem services of vegetation occur at different scales, they also suggest that the associations between greenness and health might be susceptible to scale problems. However, few studies have explicitly analyzed how greenness data resolution and aggregation areas might impact estimates of greenness exposure (48), and more studies are needed to understand the extent to which scales of analysis and NDVI resolution impact greenness exposure assessment.

Overall, there is limited understanding on the impact of methodological decisions regarding aggregation scale and spatial resolution of NDVI data on greenness exposure estimates. NDVI derived from multiple sensors at different spatial resolutions are usually treated as interchangeable in their ability to characterize exposure to vegetation in the environmental health literature. Furthermore, the implications of using NDVI at a specific spatial resolution in relation to different scales of analysis are rarely addressed in the greenness-health literature. In this study, we illustrate the impact of NDVI resolution and scale of analysis on greenness exposure misclassification in a cohort of urban children.

1. Materials and methods

1.1. Study area and population.

The study area corresponds to the Greater Boston Area (GBA), a predominantly urban region comprising the easternmost third of the state of Massachusetts. The GBA presents humid continental climate with hot, humid summers and cold winters. We used addresses from a database of electronic health records of patients between 5 and 18 years who received medical attention at the Boston Medical Center (BMC), located in Boston, Massachusetts, USA. BMC is the largest urban safety net hospital in New England and primarily serves a low-income and vulnerable population. For this study, we selected a subset of the cohort corresponding to 31,328 children whose addresses were registered during visits to BMC in 2016. Addresses were geocoded using a cascading matching method using both the MassGIS Statewide Master Geocoder (49) and the ESRI ArcGIS World Geocoder in ArcMap version 10.7. (50). Protocols were approved by the Institutional Review Board at Boston University. Figure 1 shows the distribution of NDVI in the GBA and the City of Boston, where 63% of the study population is located. NDVI greater than zero is indicative of live vegetation, whereas negative values indicate highly reflective surfaces, such as water, ice, or impervious surfaces.

1.2. Greenness data sources and data processing

NDVI for the GBA was derived from four remote sensing datasets: Moderate Resolution Imaging Spectroradiometer (MODIS), Landsat 8, Sentinel-2, and the National Agricultural Imagery Program (NAIP). We selected remote sensing data collected during a five-week period during top-growth season in the Northern Hemisphere in 2016, as this is the most recent year in which NAIP data was available for the study area. Table 1 provides the technical specifications of remote sensing instruments associated to each dataset, overpass time in study area, and date range of selected images.

(i) MODIS (250 m^2): The Terra MODIS Vegetation Indices version 6 is a post-processed data product developed by the U.S. National Aeronautics and Space Administration. A dataset of NDVI estimates with global coverage is generated every 16 days at 250 m^2 spatial resolution using an algorithm that chooses the best available pixel value from all the acquisitions from the 16-day window, selecting low cloud cover and the highest NDVI value (51). This dataset is readily available for download or use in web-based GIS platforms and it is widely used in environmental health studies.

(ii) Landsat 8 (30 m^2): Surface reflectance data from the Operational Land Imager instrument on the Landsat 8 satellite is available every 16 days at 30 m^2 resolution with global coverage. We selected images for the study area with cloud cover less than 10%.

(iii) Sentinel-2 (10 m^2): The Multispectral Instrument on board of the Sentinel-2 satellite collects multispectral images at 10 m^2 resolution with global coverage every 5 days. To minimize the influence of atmospheric disturbances, the inclusion of images was restricted to those cloud cover less than 10%.

(iv) NAIP - National Agricultural Imagery Program (1 m^2): Multispectral aerial imagery acquired during the agricultural growing season in the continental United States. NAIP data are collected, managed, and validated by the U.S. Department of Agriculture (52) and it is used as the base for the development of high-resolution land cover maps in several states of the U.S. Reflectance measurements at 1 m^2 resolution for each state are available every 3 to 5 years (53). All images from the 2016 campaign in the study area were included, yielding minor overlap of images and complete spatial coverage of the study area.

MODIS, Landsat 8, Sentinel 2, and NAIP data were retrieved and processed in Google Earth Engine, a cloud-based computer platform that leverages publicly available remote sensing data repositories and Google's computational capabilities to perform large scale spatial analysis that typically challenge computational infrastructure due to processing requirements (54). First, NDVI for each Landsat 8, Sentinel-2, and NAIP scene was calculated from reflectance measurements in the red (R) and near infrared (NIR) bands using the following equation: $(NIR-R)/(NIR+R)$. Large water bodies and cloudy pixels were masked from NDVI calculations to reduce misclassification of non-vegetated pixels. Then, we created a greenest pixel composite (i.e., selecting the highest pixel value observed during the study period) for each NDVI dataset, integrating all available imagery into a single raster file with continuous coverage for the study area.

1.3. Greenness exposure assessment

Greenness exposure was assigned using Euclidean buffers of 50 m, 250 m, 500 m, and 1,000 m radii centered at participant's geocoded residential addresses, following exposure assessment methods traditionally used in the greenness-health literature (14,25,28,41–43). Residential greenness was estimated as the pixel area-weighted average NDVI for each buffer size. Negative NDVI values were included in calculations, as setting negative NDVI values to zero prior aggregation in buffers would bias exposure distributions by truncating its left tail, thus artificially reducing its variance. Nevertheless, we calculated greenness exposure estimates from the four NDVI datasets setting negative NDVI pixels to zero, to evaluate the sensitivity of our findings to this methodological decision.

1.4. Statistical analysis

Using a reliability assessment framework, we evaluated continuous and categorical (two, three, four, and five quantiles) estimates of NDVI, matching how greenness exposure is operationalized in environmental epidemiology studies. Reliability is defined as the extent to which greenness exposure estimates for the study population can be replicated using NDVI data at different spatial resolution. We considered each NDVI sensor as a rater, thus, the inter-rater reliability assessment measures how alike are exposure estimates from different sensors for the same set of subjects. Agreement and consistency in continuous NDVI exposure estimates were evaluated using Intra-Class Correlation Coefficients (ICC). The ICC provides a measure of the magnitude of the relationship between multiple assessments of the same variable across raters, while taking into account rater bias (55). ICC ranges between 0 and 1, where greater values indicate higher reliability. We use mixed models to derive the ICC (56), incorporating subjects' variability using a random effect, and between-rater variability using a fixed effect for raters. Following the model definitions by McGraw and Wong (57), we used two-way mixed models to estimate ICC for consistency (ICC(C,1)) and absolute agreement (ICC(A,1)).

Reliability in categorical estimates was assessed in terms of agreement in participants' exposure quantiles across combinations of buffer size and number of classes in categorical variables. Light's kappa (κ), a multi rater generalization of Cohen's κ (58,59), was used to assess agreement in exposure quantiles derived from all NDVI datasets. Cohen's κ was used to evaluate pairwise agreement in categorical estimates from Sentinel-2, Landsat 8, and MODIS data using NAIP data as reference, under the assumption that the most granular data is less prone to be affected by the mixed pixel problem and therefore provides the best representation of ground vegetation. Values of κ are interpreted as follows: 0–0.20 indicates no agreement, 0.21–0.39 minimal, 0.40–0.59 weak, 0.60–0.79 moderate, 0.80–0.90 strong, and >0.9 indicates almost perfect agreement (58). Finally, using NAIP data as reference, we evaluated the magnitude of misclassification by quantifying the proportion of the study population that shifts greenness quantiles across buffer size.

We replicated statistical analyses using exposure estimates from the four datasets where pixels with negative NDVI values were set to zero prior aggregation in buffers, in order to evaluate the sensitivity of our results to the decision of keeping negative values in our calculations. Sensitivity analyses are presented in Supplementary Information (SI).

2. Results

2.1. Continuous greenness exposure estimates

Figure 2 presents the final NDVI datasets in relation to a true color image in the study area, emphasizing the differences in spatial resolution of the data.

Figure 3 presents the distributions of greenness exposure estimates derived from the four NDVI datasets by buffer size. A comparison of the exposure distributions excluding and including negative NDVI values in calculations is provided in SI Figure 1. Summary statistics are presented in Table 2. Greenness exposure estimates were greater for coarser resolution data and decreased with higher spatial resolution of NDVI. Exposure distributions followed similar trends within and across buffer sizes, where estimates from NAIP were consistently lower than Sentinel-2, Landsat 8, and MODIS data, respectively. In general, narrower ranges and higher mean and median values were observed at coarser pixels and larger buffers. For all buffer sizes, the mean and median of NAIP estimates were negative, while for Sentinel-2, Landsat 8 and MODIS, were greater than zero.

Table 3 shows ICC coefficients of agreement and consistency for continuous NDVI exposure variables by buffer size. Overall, agreement in continuous exposure estimates was poor, with small, statistically significant coefficients and wide confidence intervals across buffer sizes. Consistency ranged from good to excellent, indicating that exposure estimates from the four NDVI data products yield similar distributions of exposure estimates in terms of participant's ranked order. Agreement and consistency increased with buffer size, indicating that exposure estimates are more alike in larger aggregation areas. Agreement and consistency between estimates where negative NDVI pixels were set to zero followed similar trends; however, agreement was higher and consistency was lower compared to estimates where negative NDVI values were included (SI Figure 2)

2.2. Categorical greenness exposure estimates

Figure 4 shows Light's κ for combinations of greenness exposure quantiles and buffer sizes. The results show agreement increases with buffer size and less number of quantiles, and that differences in agreement between categorical variables decrease with buffer size. The lowest level of agreement was $\kappa = 0.39$ (five quantiles, 50 m buffer), while the highest was $\kappa = 0.88$ (two quantiles, 1,000 m buffer). Agreement in categorical greenness estimates where negative NDVI values were set to zero followed similar trends (SI Figure 3).

2.3. Greenness exposure misclassification

Figure 5 shows agreement between NDVI exposure estimates from Sentinel-2, Landsat 8, and MODIS compared to NAIP across buffer size and NDVI quantiles. The highest levels of agreement were observed for Sentinel-2, followed by Landsat 8, and MODIS, suggesting that differences in exposure estimates relative to the high resolution NAIP data decrease with reduced differences in spatial resolution of NDVI data. The greatest discrepancies were observed at 50 m buffers, with differences in κ decreasing with greater buffer size, especially between Landsat 8 and Sentinel-2. Exposure estimates from MODIS data present the lowest agreement in relation to NAIP estimates for 50 m buffers, with κ values ranging from minimal (κ = 0.28, five quantiles) to weak (κ = 0.48, two quantiles).

The plot series in Figure 6 shows the proportion of participants assigned to a different greenness quantile using Sentinel-2, MODIS, and Landsat 8 NDVI compared to NAIP NDVI across buffer sizes and number of quantiles in categorical variables. The proportion of the study population who changed at least one position in their greenness category ranged between 11% and 60% for MODIS, between 6% and 44% for Landsat 8, and between 5% and 33% for Sentinel-2 for the combinations of 50 m buffer and five quantiles, and 1000 m buffer and two quantiles, respectively. The proportion of participants who changed greenness quantiles and the magnitude of change (i.e. the number of positions of the shift) increased with greater number of quantiles in categorical variables and smaller buffer size. Larger differences in spatial resolution of data led to greater movement in greenness categories for any given buffer size and quantiles. For example, when variables structured using quintiles are applied to 50 m buffers (top right plot in Figure 6), 60%, 44%, and 33% of participants changed at least one position; 22%, 7%, and 2% change at least two positions; 6%, 0.7%, and 0.07% change at least three positions in the categorical distribution of MODIS, Landsat 8, and Sentinel-2, respectively. In that same combination of 50 m buffer and quintiles, changes of four positions in the exposure categories relative to NAIP data were observed for Landsat 8 (0.04%) and MODIS (0.8%), but not in Sentinel-2. While misclassification for exposure quantiles where negative NDVI were set to zero was higher, the results followed similar trends across buffer size and number of exposure quantiles (SI Figure 4).

3. Discussion

In this study, we quantified greenness exposure for a large urban cohort of children using NDVI data at varying spatial resolutions following exposure assessment methods used in epidemiological analysis, and compare them across multiple buffer sizes and quantile classifications to illustrate how the spatial scale of data might lead to greenness exposure misclassification. The use of coarser NDVI data led to substantial differences in categorical greenness exposure, and the severity of exposure misclassification depended on the spatial scale of analysis in relation to the quantile classification. The proportion of individuals who shifted exposure categories increased with coarser data, greater number of quantiles, and smaller buffers. These results indicate that exposure misclassification related to differences in NDVI spatial resolution is sensitive to buffer size and number of classes in categorical variables, adding to a limited body of literature analyzing the impact of NDVI data resolution and aggregation area in the context of environmental health studies. Our results are in line with previous studies which have, to varying extents, found that spatial resolution influences the magnitude or statistical significance of observed greenness-health associations (45–47).

The decision on which scale and resolution of NDVI to use when analyzing associations between greenness and health is not straightforward. First, researchers should identify the ecosystems services at play for the hypothesized exposure pathways in relation to the outcome of interest, in order to define the relevant area of exposure. For example, if greenness is thought to influence mental health through direct views of greenery from windows (60), or acting as a psychological or physical buffer from ambient noise at night (7,8), then small buffers should be used to represent residential exposure. If greenness is thought to influence mental health through increased physical activity (61), then larger buffers should be used to account for neighborhood-level exposure, as the assumption would be that greener neighborhoods are more inviting to perform physical activity. Once the aggregation area or buffer size is defined, the decision on spatial resolution of NDVI data should be driven primarily by the characteristics and spatial structure of vegetated spaces in the area of interest in relation to the relevant exposure pathways. In urban areas, a large share of greenness exposure comes from vegetation in backyards, pocket parks, large planters on sidewalks, and street trees. Given the small extent of vegetated features in urban settings, high resolution NDVI should be used to estimate exposure, as coarser resolution data might not be able to adequately characterize the distribution of greenness. This is particularly relevant for smaller buffers, as using MODIS or even Landsat 8 NDVI in a 50 m buffer might substantially increase exposure misclassification. However, coarser NDVI can be used for larger aggregation areas (buffers of 500 m, 100 m) if high-resolution data is not available, as the differences in the magnitude of misclassification between NDVI at coarser resolution is attenuated at larger buffers. In general, researchers should avoid using buffers that are smaller than pixels of the greenness data.

It is worth to note, however, that exposure pathways are complex, intertwined, and often unknown for various health outcomes, and that exposure areas will likely capture multiple pathways that might occur simultaneously. Furthermore, the decisions on aggregation area and spatial resolution of data have practical implications for statistical inference in epidemiological analyses. For example, it has been shown that the variance of greenness estimates declines with larger aggregation areas (48,62,63), which can negatively affect the ability of epidemiological models to detect statistical associations between greenness and health. Furthermore, several studies have illustrated the impacts of the Modifiable Areal Unit Problem when working with remote sensing data on greenness estimates and the observed associations between greenness and health (26,45,46,62). These challenges add to the complexity of selecting the appropriate scale for aggregation and resolution of greenness data and highlight the need of evaluating the sensitivity of analytical results towards methodological decisions regarding the units of analyses in greenness exposure assessment. The most common approach found in the epidemiological literature is performing sensitivity analysis using exposure estimates derived for multiple aggregation areas and data at different spatial resolutions to assess changes in the observed green space-health associations derived from statistical models (45,46,64,65). Alternatively, researchers should consider evaluating the sensitivity of buffer limits based on NDVI data resolution to better understand the implications of data decisions on exposure variance (see Labib et al. in (48)). Analyses where the effect of zoning and spatial resolution decisions on the variability of greenness exposure estimates are evaluated jointly can help researchers identify the appropriate buffer size for a given research question and analysis plan.

In addition to theoretical considerations, researchers should bear in mind the tradeoffs between ease of access and use, temporal and spatial availability, and accuracy of greenness data. High resolution NAIP data provides a more accurate appraisal of greenness, however, processing 1 m^2 data even for relatively small geographic extents is computationally intensive and thus processing requirements might constitute a practical constraint for researchers. Additionally, it is available only in the United States for the summer season, precluding its use to assess seasonal exposure or other parts of the world. Sentinel-2 and Landsat data are collected periodically, have global coverage, and are available starting 2015 and 1972, respectively; however, both require some degree of preprocessing to estimate NDVI. Precalculated MODIS NDVI data products have global coverage, are easily accessible for download, and are available since the year 2000, nevertheless, its coarse resolution increases the potential for exposure misclassification. Future studies should explicitly state the rationale for the selection of greenness data, provide complete references for NDVI data, and detailed methods used for NDVI data treatment in preparation to assess exposure (e.g.: cloud masking, compositing, treatment of missing and negative NDVI values) in order to improve comparability between studies, evaluate generalizability of findings, and enable replicability of methods.

The use of NDVI for greenness exposure assessment presents multiple advantages for environmental health research, such as broad spatial and temporal coverage across the globe, increasing availability at higher temporal and spatial resolutions, numerous NDVI data products being publicly accessible, and it being a standardized, validated measure of greenness exposure (31,66). However, NDVI does not provide information on vegetation species, green space structure, or green space use, and therefore is limited in its ability to characterize the personal contact or experience of greenness. Some studies have used other exposure metrics such as additional spectral indices and spectral unmixing (67), composite indices that incorporate multiple dimensions of natural spaces (68,69), or harmonize multiple green space metrics (48) to acknowledge the multiple scales and pathways by which vegetation impacts health. Studies have measured participants' perceived greenness exposure using surveys to characterize the subjective experience of greenness, observing significant associations with health outcomes and healthy behaviors and differences between perceived and objectively measured greenness using NDVI (70–72). Finally, there is a growing body of studies that used street-level imagery to characterize ground vegetation and views of greenery in public health studies (73). In general, these studies apply machine-learning methods to street view images to classify eye-level features of the built environment, including vegetation (33). Interestingly, studies have observed low correlations between exposure estimates derived from NDVI and street view data (74–78). Together, these findings suggest that the different measures of greenness might be capturing different attributes of the urban greenscape, and that the use of multiple metrics is necessary to comprehensively assess the relationship between urban greenness and health.

Our analyses are limited by our inability to control for factors that influence spectral measurements from different sensors when comparing NDVI exposure estimates, beyond differences in spatial resolution of the data. There is a set of factors that influence the reflectance measurements used to estimate NDVI, including differences in the technical specifications of instruments related to spectral resolution of bands, the platform in which sensors are mounted (aircraft versus satellites), and variations in overpass time, which might lead to differential influence of shadows from trees and tall buildings by sensor. Additionally, there are differences in data manipulation when calculating NDVI for the study area and period: a greenest pixel approach was used to composite overlapping images from satellites captured repeatedly over the study period, compared to a mosaic approach used for the aircraft data, where a single image captured on specific dates within the study period is available for each segment of the study area. However, selecting data for the same time window and masking cloudy pixels from our calculations addressed these limitations and renders the data comparable. Furthermore, we replicated methods for processing NDVI used to assess greenness exposure in epidemiologic and environmental health studies, framing our results in the context of current applications of NDVI data. Finally, as greenness exposure estimates were not linked to a health outcome, we were not able to appraise the effects of exposure misclassification on the observed effect estimates or to assess whether misclassification due to resolution and scale is differential or non-differential. Nevertheless, we hypothesize that exposure misclassification is non-differential, as nothing indicates that the measurement error stemming from technical specifications of a sensor that determine spatial resolution of NDVI data would be associated with a health outcome. Additional research is needed in other populations, geographic settings, and health studies to evaluate the generalizability of our results and better understand the impact of NDVI data on greenness misclassification.

Conclusion

Our results demonstrate the impact of NDVI spatial resolution, aggregation area, and quantile classification on estimates of greenness exposure in an urban cohort. We recommend using fine spatial resolution NDVI data to estimate greenness exposure in small aggregation areas in order to minimize exposure error, especially in urban areas where land cover is highly heterogeneous at small scale. NDVI product selection and processing decisions in future work should take into account hypothesized exposure pathways, physiological mechanisms, the spatial structure of vegetated areas driving the exposure in the area of study, as well as operational constrains, such as data processing demands.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors gratefully acknowledge Dr. Howard Cabral for his advice on statistical analysis. We sincerely thank the three anonymous reviewers whose comments and suggestions helped improve this manuscript.

Funding

This work was supported by a National Science Foundation Research Traineeship (NRT) grant to Boston University (DGE 1735087), and grant R01ES027816 from the National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH).

Data availability

Greenness data generated in this study are available from the corresponding author upon reasonable request.

References

1. Twohig-Bennett C, Jones A. The health benefits of the great outdoors: A systematic review and meta-analysis of greenspace exposure and health outcomes. Environ Res. 2018 Oct 1;166:628–37. [PubMed: 29982151]
2. La Fortezza R, Carrus G, Sanesi G, Davies C. Benefits and well-being perceived by people visiting green spaces in periods of heat stress. Urban For Urban Green. 2009 Jan 1;8(2):97–108.
3. Escobedo FJ, Kroeger T, Wagner JE. Urban forests and pollution mitigation: Analyzing ecosystem services and disservices. Environ Pollut. 2011 Aug 1;159(8–9):2078–87. [PubMed: 21316130]
4. Douglas I Urban ecology and urban ecosystems: understanding the links to human health and well-being. Curr Opin Environ Sustain. 2012 Oct 1;4(4):385–92.
5. Diener A, Mudu P. How can vegetation protect us from air pollution? A critical review on green spaces' mitigation abilities for air-borne particles from a public health perspective - with implications for urban planning. Sci Total Environ. 2021 Nov 20;796:148605. [PubMed: 34271387]
6. Grote R, Samson R, Alonso R, Amorim JH, Cariñanos P, Churkina G, et al. Functional traits of urban trees: air pollution mitigation potential. Front Ecol Environ. 2016 Dec 1;14(10):543–50.
7. Dzhambov A, Dimitrova D. Urban green spaces′ effectiveness as a psychological buffer for the negative health impact of noise pollution: A systematic review. Noise Heal. 2014;16(70):157.
8. Bolund P, Hunhammar S. Ecosystem services in urban areas. Ecol Econ. 1999 May 1;29(2):293–301.
9. Kong L, Lau KK-L, Yuan C, Chen Y, Xu Y, Ren C, et al. Regulation of outdoor thermal comfort by trees in Hong Kong. Sustain Cities Soc. 2017 May 1;31:12–25.
10. van den Berg M, Wendel-Vos W, van Poppel M, Kemper H, van Mechelen W, Maas J. Health benefits of green spaces in the living environment: A systematic review of epidemiological studies. Vol. 14, Urban Forestry and Urban Greening. Elsevier GmbH; 2015. p. 806–16.
11. Gascon M, Triguero-Mas M, Martínez D, Dadvand P, Forns J, Plasència A, et al. Mental Health Benefits of Long-Term Exposure to Residential Green and Blue Spaces: A Systematic Review. Int J Environ Res Public Health. 2015 Apr 22;12(4):4354–79. [PubMed: 25913182]
12. Crouse DL, Pinault L, Balram A, Hystad P, Peters PA, Chen H, et al. Urban greenness and mortality in Canada's largest cities: a national cohort study. Lancet Planet Heal. 2017 Oct 1;1(7):e289–97.
13. Orioli R, Antonucci C, Scortichini M, Cerza F, Marando F, Ancona C, et al. Exposure to Residential Greenness as a Predictor of Cause-Specific Mortality and Stroke Incidence in the Rome Longitudinal Study. Environ Health Perspect. 2019 Feb;127(2):027002. [PubMed: 30775931]
14. Vienneau D, de Hoogh K, Faeh D, Kaufmann M, Wunderli JM, Röösli M. More than clean air and tranquillity: Residential green is independently associated with decreasing mortality. Environ Int. 2017 Nov 1;108:176–84. [PubMed: 28863390]
15. McCormick R Does Access to Green Space Impact the Mental Well-being of Children: A Systematic Review. J Pediatr Nurs. 2017 Nov 1;37:3–7. [PubMed: 28882650]
16. Cusack L, Larkin A, Carozza SE, Hystad P. Associations between multiple green space measures and birth weight across two US cities. Health Place. 2017 Sep 1;47:36–43. [PubMed: 28711859]
17. Hu CY, Yang XJ, Gui SY, Ding K, Huang K, Fang Y, et al. Residential greenness and birth outcomes: A systematic review and meta-analysis of observational studies. Vol. 193, Environmental Research. Academic Press Inc.; 2021. p. 110599. [PubMed: 33307084]
18. Dadvand P, Sunyer J, Basagaña X, Ballester F, Lertxundi A, Fernández-Somoano A, et al. Surrounding greenness and pregnancy outcomes in four Spanish birth cohorts. Environ Health Perspect. 2012 Oct;120(10):1481–7. [PubMed: 22899599]
19. Browning MHEM, Kuo M, Sachdeva S, Lee K, Westphal L. Greenness and school-wide test scores are not always positively associated – A replication of "linking student performance in Massachusetts elementary schools with the 'greenness' of school surroundings using remote sensing." Landsc Urban Plan. 2018 Oct;178:69–72.

20. Wing TVL, Tuen YTT, Wen-Chi P, Wu C-D, Shih-Chun CL, Spengler JD. How is environmental greenness related to students' academic performance in English and Mathematics? Landsc Urban Plan. 2019 Jan 1;181:118–24.
21. Thiering E, Markevych I, Brüske I, Fuertes E, Kratzsch J, Sugiri D, et al. Associations of Residential Long-Term Air Pollution Exposures and Satellite-Derived Greenness with Insulin Resistance in German Adolescents. Environ Health Perspect. 2016 Aug;124(8):1291–8. [PubMed: 26863688]
22. Hodson CB, Sander HA. Relationships between vegetation in student environments and academic achievement across the continental U.S. Landsc Urban Plan. 2019 Sep 1;189:212–24.
23. Lambert KA, Bowatte G, Tham R, Lodge C, Prendergast L, Heinrich J, et al. Residential greenness and allergic respiratory diseases in children and adolescents – A systematic review and meta-analysis. Vol. 159, Environmental Research. Academic Press Inc.; 2017. p. 212–21. [PubMed: 28803150]
24. Gernes R, Brokamp C, Rice GE, Wright JM, Kondo MC, Michael YL, et al. Using high-resolution residential greenspace measures in an urban environment to assess risks of allergy outcomes in children. Sci Total Environ. 2019 Jun 10;668:760–7. [PubMed: 30865906]
25. Andrusaityte S, Grazuleviciene R, Kudzyte J, Bernotiene A, Dedele A, Nieuwenhuijsen MJ. Associations between neighbourhood greenness and asthma in preschool children in Kaunas, Lithuania: A case-control study. BMJ Open. 2016;6(4).
26. Labib SM, Lindley S, Huck JJ. Spatial dimensions of the influence of urban green-blue spaces on human health: A systematic review. Vol. 180, Environmental Research. Academic Press Inc.; 2020. p. 108869. [PubMed: 31722804]
27. Markevych I, Schoierer J, Hartig T, Chudnovsky A, Hystad P, Dzhambov AM, et al. Exploring pathways linking greenspace to health: Theoretical and methodological guidance. Environ Res. 2017 Oct 1;158:301–17. [PubMed: 28672128]
28. Cusack L, Larkin A, Carozza S, Hystad P. Associations between residential greenness and birth outcomes across Texas. Environ Res. 2017 Jan 1;152:88–95. [PubMed: 27743971]
29. Fong KC, Hart JE, James P. A Review of Epidemiologic Studies on Greenness and Health: Updated Literature Through 2017. Curr Environ Heal reports. 2018;5(1):77–87.
30. Zhan Y, Liu J, Lu Z, Yue H, Zhang J, Jiang Y. Influence of residential greenness on adverse pregnancy outcomes: A systematic review and dose-response meta-analysis. Vol. 718, Science of the Total Environment. Elsevier B.V.; 2020. p. 137420. [PubMed: 32325616]
31. Rhew IC, Vander Stoep A, Kearney A, Smith NL, Dunbar MD. Validation of the Normalized Difference Vegetation Index as a Measure of Neighborhood Greenness. Ann Epidemiol. 2011 Dec 1;21(12):946–52. [PubMed: 21982129]
32. Gascon M, Triguero-Mas M, Martínez D, Dadvand P, Rojas-Rueda D, Plasència A, et al. Residential green spaces and mortality: A systematic review. Environ Int. 2016 Jan;86:60–7. [PubMed: 26540085]
33. Shahtahmassebi AR, Li C, Fan Y, Wu Y, lin Y, Gan M, et al. Remote sensing of urban green spaces: A review. Urban For Urban Green. 2021 Jan 1;57:126946.
34. Badgley G, Field CB, Berry JA. Canopy near-infrared reflectance and terrestrial photosynthesis. Sci Adv. 2017 Mar 1;3(3).
35. Jiang Z, Chen Y, Li J, Dou W. The impact of spatial resolution on NDVI over heterogeneous surface. In: International Geoscience and Remote Sensing Symposium (IGARSS). 2005. p. 1310–3.
36. Jiang Z, Huete AR, Chen J, Chen Y, Li J, Yan G, et al. Analysis of NDVI and scaled difference vegetation index retrievals of vegetation fraction. Remote Sens Environ. 2006 Apr 15;101(3):366–78.
37. Chen X, Wang D, Chen J, Wang C, Shen M. The mixed pixel effect in land surface phenology: A simulation study. Remote Sens Environ. 2018 Jun 15;211:338–44.
38. Sun Y, Meng Q, Sun Z, Zhang J, Zhang L. Assessing the impacts of grain sizes on landscape pattern of urban green space. https://doi.org/101117/122285177. 2017 Oct 24;10462:934–40.
39. Small C A global analysis of urban reflectance. Int J Remote Sens. 2005 Feb 16;26(4):661–81.

40. Nouri H, Anderson S, Sutton P, Beecham S, Nagler P, Jarchow CJ, et al. NDVI, scale invariance and the modifiable areal unit problem: An assessment of vegetation in the Adelaide Parklands. Sci Total Environ. 2017 Apr 15;584–585:11–8.
41. Grazuleviciene R, Danileviciute A, Dedele A, Vencloviene J, Andrusaityte S, Uždanaviciute I, et al. Surrounding greenness, proximity to city parks and pregnancy outcomes in Kaunas cohort study. Int J Hyg Environ Health. 2015 May 1;218(3):358–65. [PubMed: 25757723]
42. Dadvand P, de Nazelle A, Figueras F, Basagaña X, Su J, Amoly E, et al. Green space, health inequality and pregnancy. Environ Int. 2012 Apr 1;40:110–5. [PubMed: 21824657]
43. Nordbø ECA, Nordh H, Raanaas RK, Aamodt G. GIS-derived measures of the built environment determinants of mental health and activity participation in childhood and adolescence: A systematic review. Vol. 177, Landscape and Urban Planning. Elsevier B.V.; 2018. p. 19–37.
44. Browning MHEM, Lee K. Within What Distance Does Greenness Best Predict Physical Health? A Systematic Review of Articles with GIS Buffer Analyses across the Lifespan. Int J Environ Res Public Health. 2017;14(7).
45. Su JG, Dadvand P, Nieuwenhuijsen MJ, Bartoll X, Jerrett M. Associations of green space metrics with health and behavior outcomes at different buffer sizes and remote sensing sensor resolutions. Environ Int. 2019 May 1;126:162–70. [PubMed: 30798197]
46. Reid CE, Kubzansky LD, Li J, Shmool JL, Clougherty JE. It's not easy assessing greenness: A comparison of NDVI datasets and neighborhood types and their associations with self-rated health in New York City. Health Place. 2018 Nov 1;54:92–101. [PubMed: 30248597]
47. Browning MHEM Locke DH. The greenspace-academic performance link varies by remote sensing measure and urbanicity around Maryland public schools. Landsc Urban Plan. 2020 Mar 1;195.
48. Labib SM, Lindley S, Huck JJ. Scale effects in remotely sensed greenspace metrics and how to mitigate them for environmental health exposure assessment. Comput Environ Urban Syst. 2020 Jul 1;82:101501.
49. MassGIS (Massachusetts Bureau of Geographic Information). MassGIS Data: Master Address Data - Basic Address Points [Internet]. 2021 [cited 2021 Jan 13]. Available from: https://docs.digital.mass.gov/dataset/massgis-data-master-address-data-basic-address-points
50. ESRI. ArcMap v10.7. Redlands, CA; 2019.
51. Didan K MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006. 2015.
52. U.S. Department of Agriculture. NAIP Imagery [Internet]. [cited 2021 Oct 19]. Available from: https://www.fsa.usda.gov/programs-and-services/aerial-photography/imagery-programs/naip-imagery/index
53. National Agricultural Imagery Program (NAIP). 2016 Massachusetts NAIP Imagery. USDA FSA Aerial Photography Field Office; 2019.
54. Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, Moore R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens Environ. 2017 Dec;202:18–27.
55. Liu J, Tang W, Chen G, Lu Y, Feng C, Ty XM. Correlation and agreement: overview and clarification of competing concepts and measures. Shanghai Arch Psychiatry. 2016;28(2):115. [PubMed: 27605869]
56. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychol Bull. 1979;86(2):420–8. [PubMed: 18839484]
57. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychol Methods. 1996;1(1):30–46.
58. McHugh ML. Interrater reliability: The kappa statistic. Biochem Medica. 2012;22(3):276–82.
59. Cohen J A Coefficient of Agreement for Nominal Scales. Educ Psychol Meas. 1960 Apr 2;20(1):37–46.
60. Li D, Sullivan WC. Impact of views to school landscapes on recovery from stress and mental fatigue. Landsc Urban Plan. 2016 Apr 1;148:149–58.
61. Pun VC, Manjourides J, Suh HH. Association of neighborhood greenness with self-perceived stress, depression and anxiety symptoms in older U.S adults. Environ Heal 2018 171. 2018 Apr 16;17(1):1–11.

62. Jelinski DE, Wu J. The modifiable areal unit problem and implications for landscape ecology. Landsc Ecol 1996 113. 1996;11(3):129–40.
63. Helbich M Spatiotemporal Contextual Uncertainties in Green Space Exposure Measures: Exploring a Time Series of the Normalized Difference Vegetation Indices. Int J Environ Res Public Heal 2019, Vol 16, Page 852. 2019 Mar 8;16(5):852.
64. Houston D Implications of the modifiable areal unit problem for assessing built environment correlates of moderate and vigorous physical activity. Appl Geogr. 2014 Jun 1;50:40–7.
65. Tsai W-L, Leung Y-F, McHale MR, Floyd MF, Reich BJ. Relationships between urban green land cover and human health at different spatial resolutions. Urban Ecosyst. 2019 Apr 12;22(2):315–24.
66. Gascon M, Cirach M, Martínez D, Dadvand P, Valentín A, Plasència A, et al. Normalized difference vegetation index (NDVI) as a marker of surrounding greenness in epidemiological studies: The case of Barcelona city. Urban For Urban Green. 2016 Sep 1;19:88–94.
67. Sadeh M, Brauer M, Dankner R, Fulman N, Chudnovsky A. Remote sensing metrics to assess exposure to residential greenness in epidemiological studies: A population case study from the Eastern Mediterranean. Environ Int. 2021 Jan 1;146:106270. [PubMed: 33276312]
68. Rugel EJ, Henderson SB, Carpiano RM, Brauer M. Beyond the Normalized Difference Vegetation Index (NDVI): Developing a Natural Space Index for population-level health research. Environ Res. 2017 Nov 1;159:474–83. [PubMed: 28863302]
69. Mavoa S, Lucassen M, Denny S, Utter J, Clark T, Smith M. Natural neighbourhood environments and the emotional health of urban New Zealand adolescents. Landsc Urban Plan. 2019 Nov 1;191:103638.
70. Hipp JA, Gulwadi GB, Alves S, Sequeira S. The Relationship Between Perceived Greenness and Perceived Restorativeness of University Campuses and Student-Reported Quality of Life: http://dx.doi.org/101177/0013916515598200. 2015 Aug 16;48(10):1292–308.
71. Tilt JH, Unfried TM, Roca B. Using Objective and Subjective Measures of Neighborhood Greenness and Accessible Destinations for Understanding Walking Trips and BMI in Seattle, Washington: http://dx.doi.org/104278/0890-1171-214s371. 2016 Nov 28;21(4 SUPPL.):371–9.
72. Loder AKF, Gspurning J, Paier C, Poppel MNM van. Objective and Perceived Neighborhood Greenness of Students Differ in Their Agreement in Home and Study Environments. Int J Environ Res Public Heal 2020, Vol 17, Page 3427. 2020 May 14;17(10):3427.
73. Kang Y, Zhang F, Gao S, Lin H, Liu Y. A review of urban physical environment sensing using street view imagery in public health studies. https://doi.org/101080/1947568320201791954. 2020 Jul 2;26(3):261–75.
74. Larkin A, Hystad P. Evaluating street view exposure measures of visible green space for health research. J Expo Sci Environ Epidemiol 2018 294. 2018 Jan 19;29(4):447–56.
75. Lu Y, Yang Y, Sun G, Gou Z. Associations between overhead-view and eye-level urban greenness and cycling behaviors. Cities. 2019 May 1;88:10–8.
76. Helbich M, Poppe R, Oberski D, Zeylmans van Emmichoven M, Schram R. Can't see the wood for the trees? An assessment of street view- and satellite-derived greenness measures in relation to mental health. Landsc Urban Plan. 2021 Oct 1;214:104181.
77. Wang R, Feng Z, Pearce J, Yao Y, Li X, Liu Y. The distribution of greenspace quantity and quality and their association with neighbourhood socioeconomic conditions in Guangzhou, China: A new approach using deep learning method and street view images. Sustain Cities Soc. 2021 Mar 1;66:102664.
78. Villeneuve PJ, Ysseldyk RL, Root A, Ambrose S, DiMuzio J, Kumar N, et al. Comparing the Normalized Difference Vegetation Index with the Google Street View Measure of Vegetation to Assess Associations between Greenness, Walkability, Recreational Physical Activity, and Health in Ottawa, Canada. Int J Environ Res Public Heal 2018, Vol 15, Page 1719. 2018 Aug 10;15(8):1719.

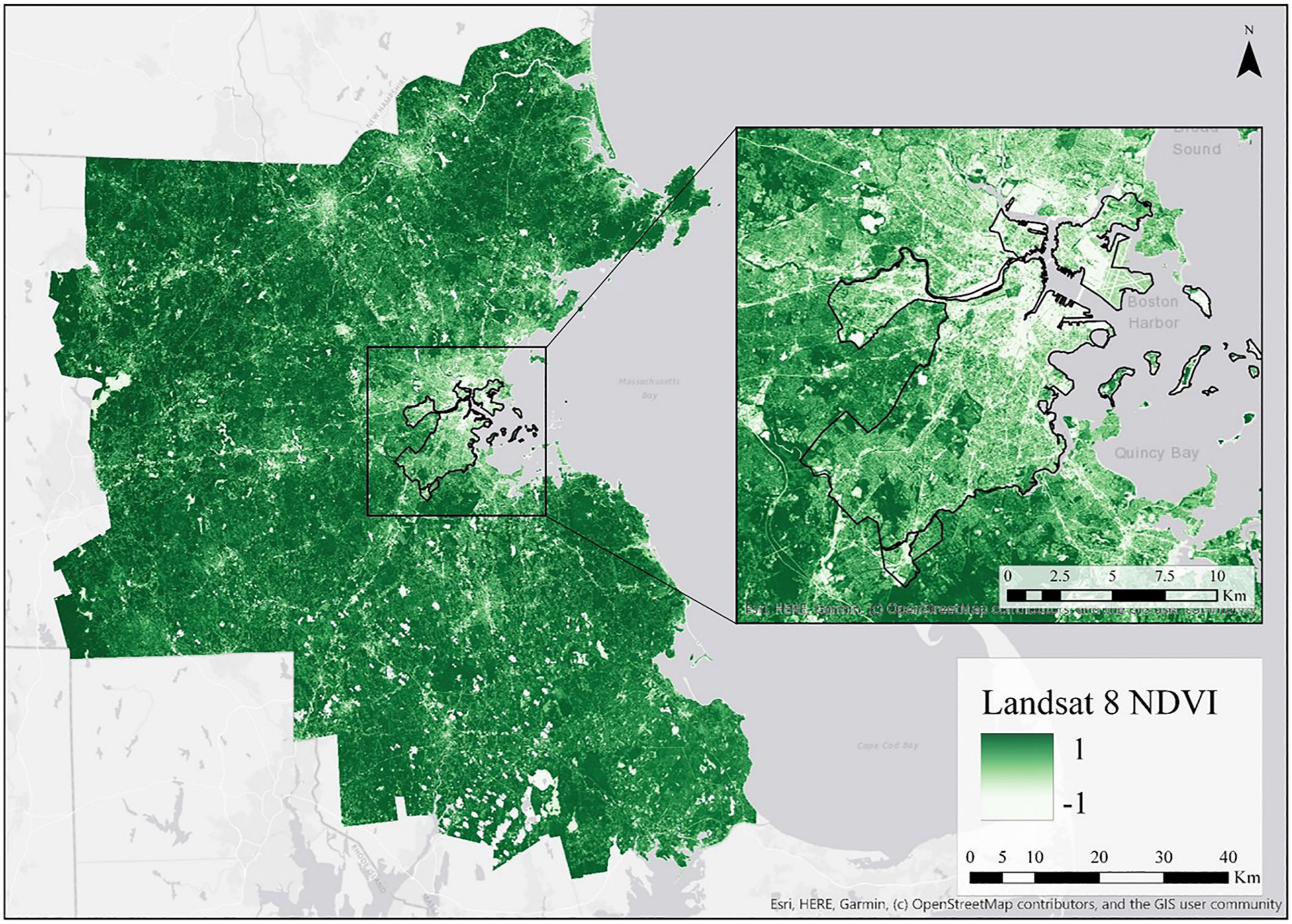

Fig. 1.

Greenness in the Greater Boston Area and city of Boston, Massachusetts in summer 2016. NDVI was derived from Landsat 8 surface reflectance data.

Fig 2.
Comparison of spatial resolution of NDVI from NAIP, Sentinel-2, Landsat 8, and MODIS. The figure shows NDVI from the four data products next to a true color image of the Boston Common and Public Garden in Massachusetts during the summer season of 2016.

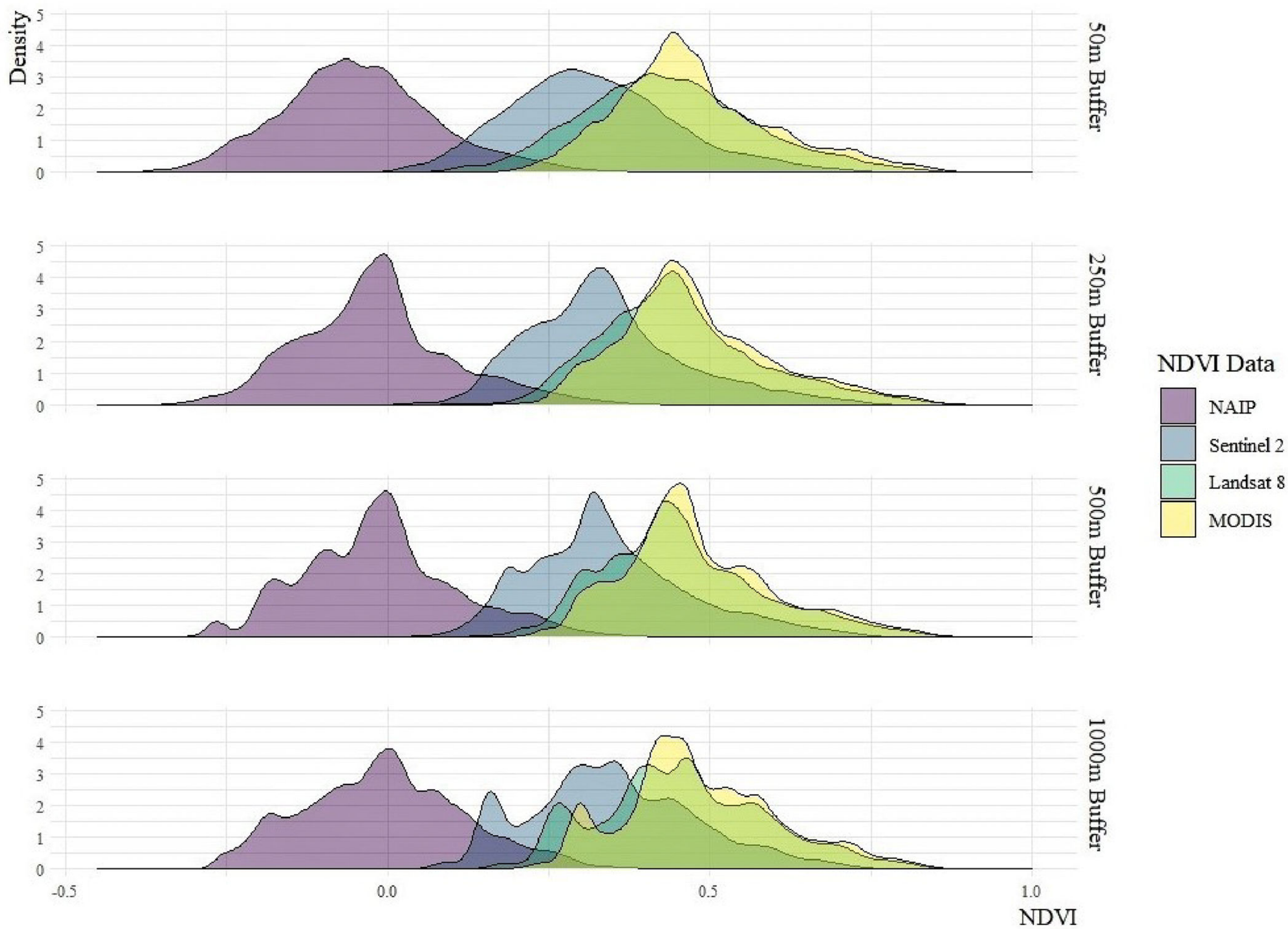

Fig 3.

Greenness exposure distributions for the study population (an = 31,328) across different buffer sizes. Panels show the distributions of continuous greenness exposure estimates derived from NAIP, Sentinel-2, Landsat 8, and MODIS in residential buffers of 50 m, 250 m, 500 m, and 1000 m radii.

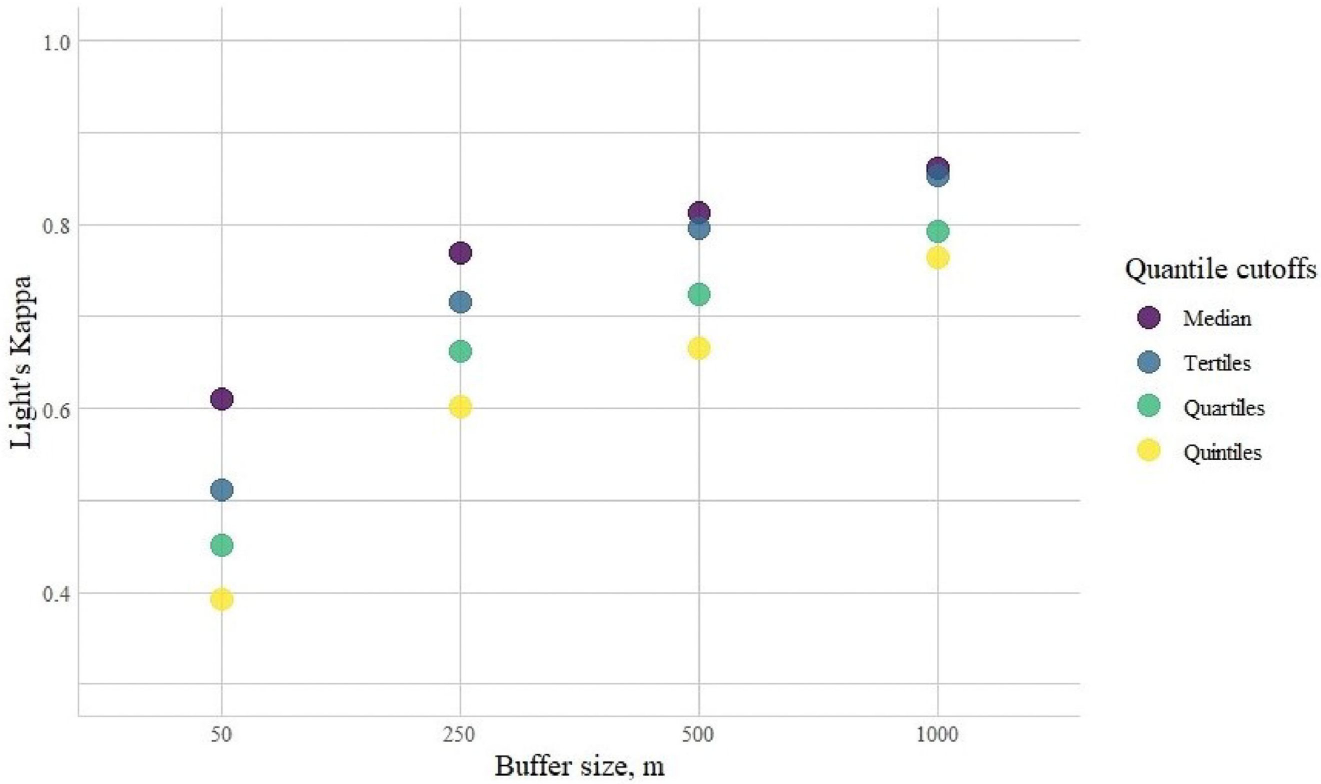

Fig 4.
Agreement between greenness exposure quantiles from NAIP, Sentinel-2, Landsat 8, and MODIS. The plot shows values of Light's ϰ for multi rater agreement in categorical greenness exposure estimates across buffer sizes and quantile cutoffs.

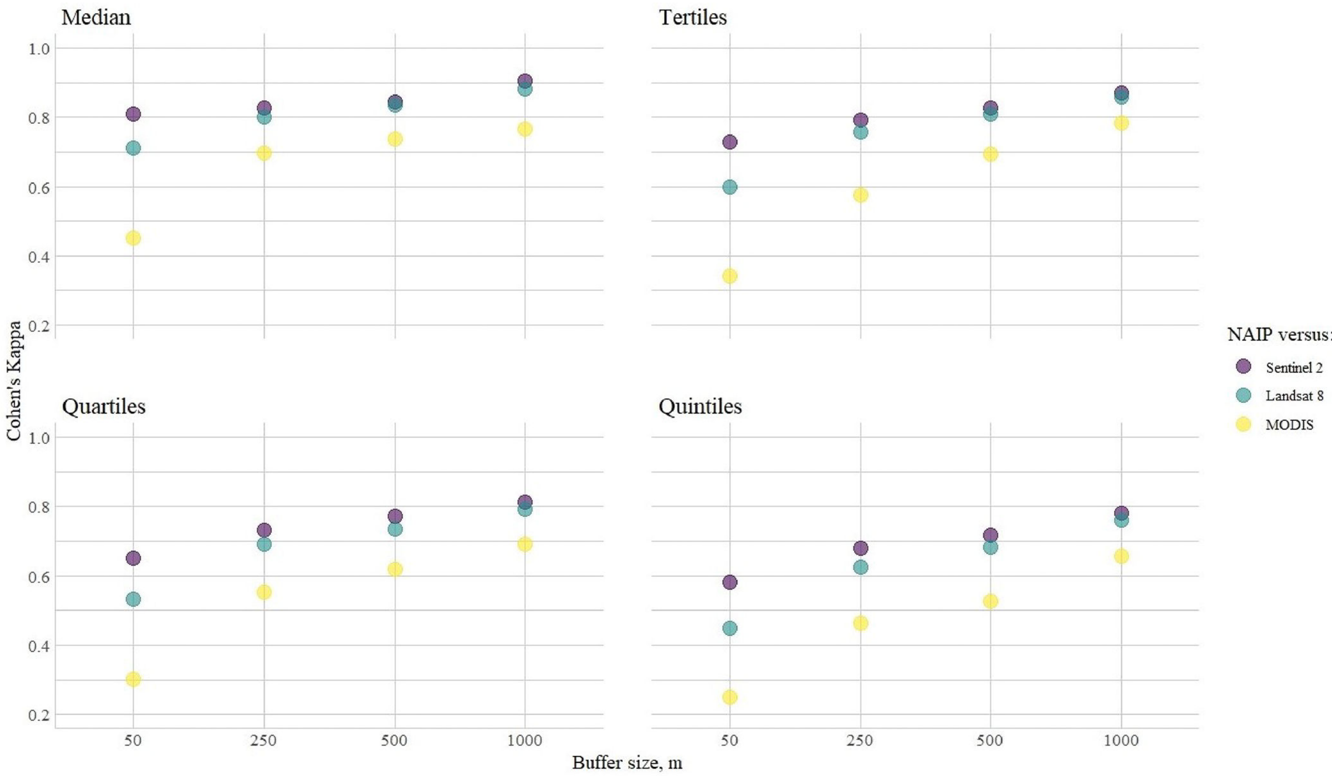

Fig. 5.

Agreement in greenness exposure quantiles from Sentinel 2, Landsat 8 and MODIS relative to NAIP NDVI. Panels show values of Cohen's κ for pairwise agreement in categorical exposure estimates across buffer sizes and greenness quantiles.

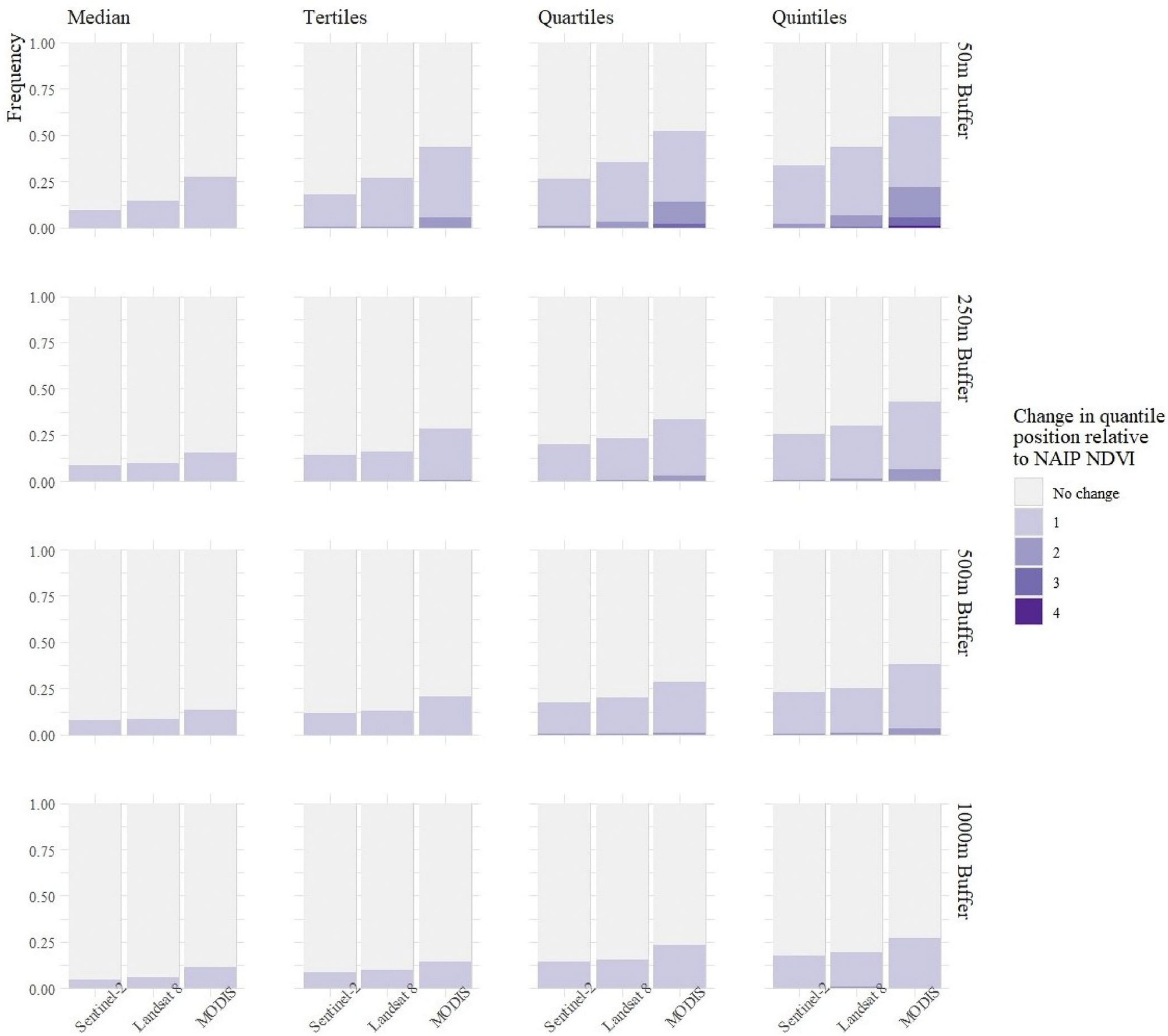

Fig. 6.

Greenness exposure misclassification relative to NAIP NDVI across buffer size and exposure quantiles. This plot series shows the proportion of participants that change positions in the categorical distribution of greenness exposure derived from Sentinel-2, Landsat 8, and MODIS data in relation to NAIP data. Plots are organized by buffer size in rows and number of classes of categorical variables in columns. The color legend indicates the magnitude of change in quantile positions relative to NAIP NDVI.

Table 1.

Technical specifications of remote sensing instruments

Remote sensing data	Instrument	Platform	Band wavelength, µm		Ground pixel resolution, m	Overpass time[2]	Temporal range of images	
			Red	Near Infrared			Start date	End date
NAIP[1]	Leica Geosystem's ADS100/SH100 digital sensors	Aircraft	0.62–0.65	0.81–0.88	1	15:00–19:00	3/7/2016	3/8/2016
Sentinel-2	Multispectral instrument	Satellite	0.63–0.70	0.73–0.93	10	15:41	7/7/2016	3/8/2016
Landsat 8	Operational Land Imager	Satellite	0.64–0.67	0.85–0.88	30	15:30	4/7/2016	22/7/2016
MODIS	Moderate Resolution Imaging Spectroradiometer	Satellite	0.62–0.67	0.84–0.88	250	15:44	11/7/2016	27/7/2016

[1] National Agricultural Imagery Program (NAIP).

[2] Overpass time for Sentinel 2, Landsat 8, and Terra satellite obtained from Spectator.earth. Available at https://app.spectator.earth/. Overpass time for NAIP imagery derived from metadata.

Table 2.

Summary statistics of greenness exposure distributions.

Buffer size	NDVI data	Min	Max	Median	Mean	SD	Variance	Skewness	Kurtosis
50 m	NAIP	−0.47	0.46	−0.05	−0.04	0.12	0.015	0.30	3.27
	Sentinel-2	−0.05	0.79	0.27	0.28	0.13	0.017	0.42	3.14
	Landsat 8	−0.42	0.89	0.42	0.43	0.14	0.019	0.26	3.05
	MODIS	0.06	0.89	0.43	0.45	0.13	0.016	0.60	3.24
250 m	NAIP	−0.37	0.46	−0.02	−0.02	0.12	0.014	0.45	3.41
	Sentinel-2	−0.03	0.77	0.29	0.30	0.12	0.015	0.67	3.47
	Landsat 8	0.08	0.87	0.44	0.45	0.13	0.015	0.58	3.32
	MODIS	0.10	0.89	0.43	0.45	0.12	0.014	0.64	3.27
500 m	NAIP	−0.35	0.46	−0.02	−0.01	0.12	0.014	0.41	3.17
	Sentinel-2	0.01	0.76	0.29	0.31	0.12	0.015	0.62	3.24
	Landsat 8	0.08	0.87	0.43	0.45	0.13	0.016	0.56	3.16
	MODIS	0.11	0.88	0.44	0.45	0.12	0.014	0.61	3.19
1,000 m	NAIP	−0.28	0.43	−0.01	0.00	0.12	0.015	0.24	2.73
	Sentinel-2	0.03	0.75	0.31	0.31	0.13	0.017	0.37	2.78
	Landsat 8	0.11	0.86	0.45	0.45	0.14	0.017	0.31	2.74
	MODIS	0.14	0.86	0.44	0.46	0.12	0.015	0.42	2.89

Author Manuscript
Author Manuscript
Author Manuscript

Table 3.

Intra-class correlation coefficients (ICC) for continuous NDVI exposure variables

	Absolute Agreement		Consistency	
Buffer size (m)	**ICC (A,1)** [¤]	**95% CI**	**ICC(C,1)** [¤]	**95% CI**
50	0.181 [**]	0.025 – 0.356	0.799 [*]	0.796 – 0.802
250	0.206 [**]	0.029 – 0.401	0.942 [*]	0.941 – 0.942
500	0.214 [**]	0.031 – 0.414	0.964 [**]	0.964 – 0.965
1,000	0.235 [**]	0.035 – 0.443	0.978 [**]	0.977 – 0.978

[¤] Model specifications by McGraw and Wong, 1996

[**] Significant at $p < 0.001$.

[*] Significant at $p<0.01$.

ICC <0.5 poor, 0.5–0.75 fair, 0.75–0.9 good, >0.9 excellent. Criteria from Koo and Li, 2016