Published in final edited form as:
Ann Epidemiol. 2011 December ; 21(12): 946–952. doi:10.1016/j.annepidem.2011.09.001.

Validation of the Normalized Difference Vegetation Index as a measure of neighborhood greenness

Isaac C. Rhew, PhD, MPH[a,*], **Ann Vander Stoep, PhD**[b,c], **Anne Kearney, PhD**[d,e], **Nicholas L. Smith, PhD, MPH**[b,f,g], and **Matthew D. Dunbar, PhD**[h]

[a]Social Development Research Group, University of Washington, Seattle, WA, USA

[b]Department of Epidemiology, University of Washington, Seattle, WA, USA

[c]Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, USA

[d]School of Forest Resources, University of Washington, Seattle, WA, USA

[e]Kearney Environmental LLC, Seattle, WA, USA

[f]Seattle Epidemiologic Research and Information Center of the Department of Veterans Affairs Office of Research and Development, Seattle, WA, USA

[g]Group Health Research Institute, Seattle, WA, USA

[h]Center for Studies in Demography and Ecology, University of Washington, Seattle, WA, USA

Abstract

Purpose—To assess the validity of a GIS measure, the Normalized Difference Vegetation Index (NDVI), as a measure of neighborhood greenness for epidemiologic research.

Methods—Using remote-sensing spectral data, NDVI was calculated for a 100-m radial distance around 124 residences in greater Seattle. The criterion standard was rating of greenness for corresponding residential areas by three environmental psychologists. Pearson correlations and regression models were used to assess the association between the psychologists' ratings of greenness and NDVI. Analyses were also stratified by residential density to assess whether the correlations differed between low and high density.

Results—Mean NDVI among this sample of residences was .27 (SD = 0.11; range: −.04 to .54), and the mean psychologist rating of greenness was 2.84 (SD = 0.98; range: 1 to 5). The correlation between NDVI and expert ratings of greenness was high (r = .69). The correlation was equivalently strong within each strata of residential density.

Conclusions—NDVI is a useful measure of neighborhood greenness. In addition to showing strong correlation with expert ratings, this measure has practical advantages including availability of data and ease of application to various boundaries which would aid in replication and comparability across studies.

*Corresponding Author. University of Washington, Social Development Research Group, 9725 3rd Avenue NE, #401, Seattle, WA 98115. Tel.: 206-221-1823; fax: 206-543-4507; rhew@u.washington.edu.

Background

Interest in neighborhood "greenness", or exposure to vegetation and green space, and its association with physical and mental health is increasing. Studies suggest that neighborhood greenness is positively associated with self-reported health (1, 2), physical activity (3–6), and mental health (1,8,9), as well as with reduced disparities in cardiovascular health (7). As research moves forward, valid measurement of neighborhood greenness is an issue warranting consideration. Objective or independently assessed measures are often preferred, since study participants may not be accurate reporters. General reporting inaccuracy could result in exposure misclassification (8). More problematic is that self-report of neighborhood characteristics may be differentially biased by a participant's health status particularly where the outcome involves psychopathology such as depression (9–11). There are also logistical considerations. Although independent ratings of greenness by research staff have been previously used and confer advantages, including the ability to ascertain particular qualities and details about greenness (12), this particular method is resource intensive and may not be feasible particularly in studies spanning many and/or large geographic areas. Further, it is important to identify the spatial scale at which a particular neighborhood characteristic affects human health (11, 13). This goal would be aided by the capacity to easily quantify the characteristic across various geographic boundaries and scales.

Geographic Information Systems (GIS) measures derived from independently collected data have often been utilized in studies of the built environment and health (11). Although objective, GIS measures also can be subject to biases and errors that could introduce exposure misclassification (14, 15). Validation studies provide a foundation upon which the accuracy of GIS measures used in research can be determined.

One GIS measure of area-level greenness, the Normalized Difference Vegetation Index (NDVI), shows promise for use in epidemiologic studies. This measure has been widely used to assess levels of vegetation in agriculture and forestry research (16). NDVI provides an objective metric for levels of healthy vegetation that can be easily summarized and applied across areas of different spatial scales of interest for virtually all land surfaces across the earth. NDVI data are collected using remote-sensing data that are publicly available. This could aid in replication and comparability of studies across different geographic regions.

NDVI has been used recently in studies of neighborhood greenness and health (3, 17, 18). However, little is known about its validity. Validation studies are needed to determine whether this GIS measure accurately reflects the levels of greenness that can be observed from the ground. Further, because NDVI was created for use in agriculture and forestry, it is unclear how well it measures vegetation in urban settings. For example, spatial distribution of vegetation may be less concentrated and more difficult to detect in higher compared to lower density neighborhoods.

The aim of this study was to assess the inter-method reliability of NDVI by examining the correspondence between NDVI scores and observed residential greenness ratings of environmental psychology experts for the corresponding area. In addition, the study examined whether the correlation between NDVI and experts' greenness ratings differed across strata of residential density.

Methods

Residential areas for this study were selected from residences of participants in the Neighborhood Quality of Life Study (NQLS) (19). The NQLS is a large cross-sectional study of 2,199 adults living in the Greater Seattle or Baltimore/Washington D.C. regions.

The aim of this study is to examine the association between neighborhood built environment and physical activity. Sixteen urban or suburban neighborhoods from the Seattle and 17 neighborhoods from the Baltimore/D.C. regions were selected based on variability in walkability and neighborhood income (20). For this analysis, 131 residences were selected from only the 16 Seattle region neighborhoods. Depending on neighborhood geographic size, between six and ten residences were initially selected from each of the neighborhoods. Residences were chosen to represent spatially distinct locations within the neighborhoods. NDVI and other neighborhood characteristics were not known prior to selecting the sample of residences. Addresses were geocoded using ArcGIS (ESRI, Redlands, CA).

Measures

The principle underlying NDVI is that healthy green vegetation reflects more infrared radiation and absorbs more energy in the red wavelength when compared to unhealthy vegetation or to non-vegetated surfaces. NDVI is calculated according to the level of reflectance of near-infrared and visible red wavelength spectra detected by satellite. Spectral data captured on July 23, 2006 from Landsat 5, were downloaded from the United States Geological Survey Global Visualization Viewer website (21).

Using spectral data available at a 30-m by 30-m resolution, NDVI was calculated according to the following algorithm:

$$NDVI=(NIR-RED)/(NIR+RED)$$

where NIR is the amount of near-infrared wavelength reflectance, and RED is the amount of red wavelength reflectance detected. Scores can range from −1 to 1, where −1 indicates no presence of vegetation, and 1 indicates dense levels of healthy vegetation. Mean NDVI was calculated for a 100-m radial buffer around each residence using the ArcGIS zonal statistics tool. A 100-m buffer, rather than a larger spatial scale, was selected because there could be important effects of more proximal distance and/or views of nature from one’s residence. Further, we believed that if validity were established at a small spatial scale where issues of remote-sensing imaging resolution are more prominent, then accuracy of NDVI would also be acceptable at larger spatial scales.

The criterion standard for validating the NDVI rating of greenness was constructed by having environmental psychologists provide independent ratings of greenness on the basis of visual examination of photographs taken around each of the residences. To approximate the 100-m radial distance selected for the NDVI measure, areas along the immediate block and adjoining blocks of the participant residences were photographed from the street. Photos were taken during mid-September and early October 2008 which should minimize seasonal differences between image capture of the NDVI spectral data and the photographs used for rating. Because of the two year time difference in ratings, of the original 131 residences selected for this study, three (2.3%) were excluded because of obvious recent changes to the landscape (e.g. significant new development in close proximity to the residence which might indicate removal of trees or other vegetation), two (1.5%) were removed because of problems with clarity of photographs, and two (1.5%) were excluded because their addresses could not be located. A total of 124 residential areas were retained for the analysis.

For each selected residence, four photos were selected by one of the investigators (IR) to provide adequate representation of the vegetation within a 1–2 block radius. The sets of photos for each residential area were then presented to three experts (including co-author, AK) in environmental psychology. Each psychologist was asked to rate the sets of photographs for greenness on a scale from 1 (no or very little greenness) to 5 (very high

greenness). The mean of the three ratings was calculated for each residence. See examples of photographs for low- and high-rated greenness areas and their corresponding maps of NDVI online at the journal’s website.

Residential density was defined as the number of residential units per acre of residential land use based on King County tax parcel data for a 1-km road network buffer around each residence.

Data analysis

To examine the consistency of the ratings of greenness across the three experts, one-way ANOVA was used to calculate the intraclass correlation coefficient (ICC). To estimate the correlation between NDVI and the mean greenness rating as well as other neighborhood characteristics, Pearson correlation coefficients were calculated. To account for clustering of observations within study neighborhoods, mixed effects regression models with a random intercept for neighborhood were also used to estimate the slope of the association, with NDVI modeled as the independent variable and the greenness rating as the dependent variable. In regression analyses, NDVI was multiplied by a factor of 10 for ease of interpretation.

Additional analyses were conducted with stratification by residential density to examine whether the correlation between NDVI and greenness rating differed by levels of density. In these analyses, residential density was dichotomized at the median cutoff for this particular sample of Seattle neighborhoods ($<50^{th}$ percentile, $>50^{th}$ percentile). In a separate regression model, an NDVI-density interaction term was included, and Wald’s tests were used to assess the statistical significance.

Results

The mean NDVI for this sample of residential areas was 0.27 (SD = 0.11; range: −0.04 to 0.54). The mean greenness ratings by the three environmental psychologists were 3.05, 3.04, and 2.41, and the overall mean across the three raters was 2.84 (SD = 0.98; range 1 to 5). Inter-rater reliability for psychologists’ greenness ratings was high (ICC = 0.82). The mean residential density in the sampled areas was 10.6 residential units per acre (SD = 10.4; range: 0.8 to 65.8).

Correlation between NDVI and expert-rated greenness was high ($r = 0.69$, $p<0.001$). Using mixed effects regression models, a 0.10 increase in mean NDVI was associated with a 0.52 increase in psychologist-rated greenness (95% CI: 0.39, 0.65). There was a moderate inverse correlation between NDVI and neighborhood density ($r = -0.42$, $p<0.001$). In stratified analyses, the correlation between NDVI and expert-rated greenness remained strong in both low ($r = .70$, $p<0.001$) and high density ($r = 0.71$, $p<0.001$) areas. In regression models, there was no moderating effect of density on the association between psychologist greenness rating and NDVI (interaction-p = .98).

Discussion

Based on findings from this analysis, NDVI appears to be a valid measure for quantifying levels of residential area greenness that can be viewed from the ground. NDVI was highly correlated with expert ratings of greenness, the associations were linear, and associations remained strong in both low- and high-density areas.

It is important to recognize ways in which this measure of greenness is distinct from other measures used in research. Some studies have used independently observed measures of

greenness where particular qualities of vegetation such as type, diversity or organization can be observed, but obviously NDVI cannot provide these details. Further, other studies from the Netherlands assessed the proportion of land classified from public records as public green space surrounding individuals' residences (1). NDVI, however, measures overall level of vegetation. This distinction is important in that different aspects of greenness may operate through different mechanisms to promote health. Whereas neighborhood green space might make its contribution through providing publicly accessible spaces that individuals travel to and utilize for recreational physical activity and social interaction (6, 22, 23), overall levels of vegetation may promote health through passive viewing of nature from one's own home or in transporting to or from one's home. Kaplan and Kaplan proposed that viewing nature captures individuals' involuntary attention—a type of passive, effortless attention that contrasts with directed attention which requires effort and focus—which promotes restoration from mental fatigue (24). In support of this theory, studies have shown that residential views of nature, which are not included as areas of exposure in studies of public green space, are associated with general psychological well-being (25), as well as with fewer acts of aggression in inner city settings (12). It is also possible that general greenness could improve the aesthetics of one's neighborhood which may encourage local physical activity (24, 26), reduce stress and improve mental health (27).

Aspects of spectral data sources should be considered when using NDVI to characterize neighborhood greenness. For this study, we used remote sensing data captured from the Landsat5 satellite that have a relatively fine 30-m by 30-m resolution. There are other publicly available spectral data that have image resolutions that are more coarse (e.g. 1 .1-km Advanced Very High Resolution Radiometer (AVHRR) (28) or Moderate Resolution Imaging Spectroradiometer (MODIS) (29)). The data with more coarse resolution have the advantage of being publicly available for many dates in a given year (e.g. AVHRR has revisit time of 12 hours and MODIS of 2 days), while Landsat 5 data are currently available for only a few specific dates. For assessments of neighborhood greenness at a more localized scale, like the 100-m buffer used in this study, it is preferable to use finer spatial resolution data because using lower resolution data (e.g. 1 km or 250 m) would include vegetation that exists outside the 100-m buffer. However, if the spatial scale selected is much larger (e.g. city or county level) and there is particular interest in the change in neighborhood greenness over time, or if temporal consistency with an outcome measure is a high priority, then the use of lower resolution data may be appropriate.

There were limitations to the current study. First, NDVI was based on satellite images captured in 2006. Thus, levels of vegetation may have changed by the time photographs were taken in 2008. However, to address this possibility, residential areas where there was obvious change to the landscape (e.g. significant new construction around the residence) were not included in the analysis. Second, the ratings were based on photos taken from the street, and thus part of the foliage found in back yards was not taken into account for greenness ratings.

Because of the high correlation with experts' ratings of greenness, it appears that NDVI accurately reflects the amount of neighborhood greenness that can be observed directly by humans from the ground. As an objective measure, NDVI removes the possibility of same source bias that can occur when study participants report both their perceptions of their surroundings and their health outcome. NDVI also has practical advantages. It can be easily derived using no-cost publicly available data and can be applied to various geographic boundaries. This could facilitate common use across studies in different geographic regions which would aid in replication and comparison of findings. Thus, NDVI appears to be a valid and practical metric for use in epidemiologic research of associations between neighborhood-level greenness and health.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Acknowledgments

While working on this study, Dr. Rhew was supported by NIH grant T32 HD052462-01. Additional support for this research came from a *Eunice Kennedy Shriver* National Institute of Child Health and Human Development research infrastructure grant, 5R24HD042828, to the Center for Studies in Demography & Ecology at the University of Washington. The authors wish to thank Drs. Jenna Tilt, Gordon Bradley, Brian Saelens, and Jon Wakefield for their consultation on this study. Further, we are grateful to Drs. Saelens, Larry Frank, and Jim Sallis for allowing us to use data from the Neighborhood Quality of Life Study.

List of abbreviations

GIS	Geographic Information Systems
NDVI	Normalized Difference Vegetation Index
NQLS	Neighborhood Quality of Life Study
ICC	Intraclass correlation coefficient

References

1. Maas J, van Dillen SM, Verheij RA, Groenewegen PP. Social contacts as a possible mechanism behind the relation between green space and health. Health Place. 2009; 15:586–95. [PubMed: 19022699]
2. Maas J, Verheij RA, Groenewegen PP, de Vries S, Spreeuwenberg P. Green space, urbanity, and health: how strong is the relation? J Epidemiol Community Health. 2006; 60:587–92. [PubMed: 16790830]
3. Bell JF, Wilson JS, Liu GC. Neighborhood greenness and 2-year changes in body mass index of children and youth. Am J Prev Med. 2008; 35:547–53. [PubMed: 19000844]
4. Rodriguez DA, Brown AL, Troped PJ. Portable global positioning units to complement accelerometry-based physical activity monitors. Med Sci Sports Exerc. 2005; 37:S572–81. [PubMed: 16294120]
5. McGinn AP, Evenson KR, Herring AH, Huston SL, Rodriguez DA. Exploring associations between physical activity and perceived and objective measures of the built environment. J Urban Health. 2007; 84:162–84. [PubMed: 17273926]
6. Cohen DA, McKenzie TL, Sehgal A, Williamson S, Golinelli D, Lurie N. Contribution of public parks to physical activity. Am J Public Health. 2007; 97:509–14. [PubMed: 17267728]
7. Mitchell R, Popham F. Effect of exposure to natural environment on health inequalities: an observational population study. Lancet. 2008; 372:1655–60. [PubMed: 18994663]
8. Macintyre S, Macdonald L, Ellaway A. Lack of agreement between measured and self-reported distance from public green parks in Glasgow, Scotland. Int J Behav Nutr Phys Act. 2008; 5:26. [PubMed: 18454876]
9. Kawachi, I.; Berkman, LF. Introduction. In: Kawachi, I.; Berkman, LF., editors. Neighborhoods and Health. New York, NY: Oxford University Press; 2003.
10. O'Campo, P.; Caughy, MO. Measures of Residential Community Contexts. In: Oakes, JM.; Kaufman, JS., editors. Methods in Social Epidemiology. San Francisco, CA: Jossey-Bass; 2006.
11. Mair C, Diez Roux AV, Galea S. Are neighbourhood characteristics associated with depressive symptoms? A review of evidence. J Epidemiol Community Health. 2008; 62:940–6. 8 p following 946. [PubMed: 18775943]
12. Kuo FE, Sullivan WC. Aggression and violence in the inner city: Effects of environment via mental fatigue. Environment and Behavior. 2001; 33:543–571.
13. Messer LC. Invited commentary: Beyond the metrics for measuring neighborhood effects. Am J Epidemiol. 2007; 165:868–71. discussion 872–3. [PubMed: 17329714]
14. Ricketts TC. Geographic information systems and public health. Annu Rev Public Health. 2003; 24:1–6. [PubMed: 12471275]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

15. Boone JE, Gordon-Larsen P, Stewart JD, Popkin BM. Validation of a GIS facilities database: quantification and implications of error. Ann Epidemiol. 2008; 18:371–7. [PubMed: 18261922]
16. NASA. Landsat 7 science data user's handbook. http://landsathandbook.gsfc.nasa.gov/pdfs/Landsat7_Handbook.pdf
17. Liu GC, Wilson JS, Qi R, Ying J. Green neighborhoods, food retail and childhood overweight: differences by population density. Am J Health Promot. 2007; 21:317–25. [PubMed: 17465177]
18. 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. Am J Health Promot. 2007; 21:371–9. [PubMed: 17465183]
19. Sallis JF, Saelens BE, Frank LD, et al. Neighborhood built environment and income: examining multiple health outcomes. Soc Sci Med. 2009; 68:1285–93. [PubMed: 19232809]
20. Frank LD, Sallis JF, Saelens BE, et al. The Development of a Walkability Index: Application To the Neighborhood Quality of Life Study. Br J Sports Med. 2009
21. United States Geological Survey. USGS Global Visualization Viewer. 2011. http://glovis.usgs.gov/
22. Kuo FE, Sullivan WC, Coley RL, Brunson L. Fertile ground for community: Inner-city neighborhood common spaces. American Journal of Community Psychology. 1998; 26:823–851.
23. Cohen DA, Inagami S, Finch B. The built environment and collective efficacy. Health Place. 2008; 14:198–208. [PubMed: 17644395]
24. Kaplan, R.; Kaplan, S. The Experience of Nature: A Psychological Perspective. Cambridge: Cambridge University Press; 1989.
25. Kaplan R. The nature of the view from home: psychological benefits. Environment and Behavior. 2001; 33:507–542.
26. Humpel N, Owen N, Leslie E. Environmental factors associated with adults' participation in physical activity: a review. Am J Prev Med. 2002; 22:188–99. [PubMed: 11897464]
27. Mair C, Roux AVD, Shen M, et al. Cross-Sectional and Longitudinal Associations of Neighborhood Cohesion and Stressors with Depressive Symptoms in the Multiethnic Study of Atherosclerosis. Annals of Epidemiology. 2009; 19:49–57. [PubMed: 19064189]
28. National Ocean and Atmospheric Administration. Advanced Very High Resolution Radiometer - AVHRR. Camp Springs; MD: 2011. http://noaasis.noaa.gov/NOAASIS/ml/avhrr.html
29. National Aeronautics and Space Administration. MODIS - Web. 2011. http://modis.gsfc.nasa.gov/