JOURNAL OF
EMERGING INVESTIGATORS

Article

Assessing the association between developed surface area and land surface temperature of urban areas

Niko Ustin[1], Ethan Khorana[1], Jillian Ungar[2], Samuele Nusca-Dagon[3], Julia Hodges[4], Zachary Kordeleski[5], Derick Delloro[1], Claudia Tausz[1]

[1] Mastery School of Hawken, Cleveland, Ohio
[2] Vanderbilt University, Nashville, Tennessee
[3] John Cabot University, Rome, Italy
[4] Bennett Day School, Chicago, Illinois
[5] Explore Learning, Charlottesville, Virginia

SUMMARY

Urban heat islands are metropolitan areas with warmer land surface temperatures (LST) than the surrounding rural areas, and are associated with worse air and water quality, increased numbers of blackouts and power outages, and increased heat-related illnesses and deaths. The objective of this study was to investigate the contributions of infrastructure development to LST in a sample of neighborhoods in Cleveland, Ohio. Utilizing satellite images, we compared the performance of color analysis software to manual measurement in assessing the amount of developed geographic surface area in a sample. We investigated the correlation between surface area development and LST. Color analysis produced comparable results relative to the manual measurement of the extent of development in a sample (R^2 = 0.9, $p < 0.0001$). Results show a moderate degree of linearity between the proportion of development in an area and its LST (R^2 = 0.48, $p < 0.0001$). This suggests that development may be a contributor to rising temperatures in urban areas. Efforts toward sustainable development and increases in greenery could help slow the rising temperatures. Assessments of the proportion of areas that are developed could guide where these efforts are focused.

INTRODUCTION

Urban Heat Islands (UHIs) are a common environmental challenge that almost all urban areas face. UHIs are urban spaces with substantially hotter land surface temperatures (LST) than the surrounding rural areas (1). By 2050, the magnitude of UHI contribution to global warming is expected to be 40-70% of that from greenhouse gases (2). This is important because greenhouse gases are widely considered the most significant cause of climate change.

In the next 60 years, the average summer high in Cleveland, OH is expected to increase by 5°C (3). Cleveland's urban areas are, on average, 3.9°C hotter than the surrounding areas and as high as 8.3°C to 11.1°C hotter in some parts of the city. Cleveland ranks fifteenth in the United States in terms of the difference in its LST compared to surrounding areas (4). Cleveland is expected to finish out the century with the third greatest increase in heat-related mortality out of any major city in the United States (5). Additionally, with racial minorities being overrepresented in urban centers, they bear disproportionate consequences from UHIs (6). This all means that there are major health-related risks of UHIs, particularly in Cleveland, and those risks are disproportionately felt my racial minority populations.

The biggest contributor to the UHI phenomenon is increased amounts of infrastructure, in addition to a lack of greenery in cities (1). Recent studies have shown that larger amounts of infrastructure lead to an increase in heat absorption (1). Some common infrastructure materials, including roofing, pavement, and siding, tend to have low albedo values. Albedo is defined as the ratio of reflected solar radiation from a surface to the total amount of incident solar radiation received by the surface (7). In short, albedo describes the degree to which a surface absorbs energy from the sun. Materials with lower albedo values will absorb a larger proportion of heat from the sun, which leads to more heat retention in urban areas (7).

The purpose of this study was to determine the relationship between quantified development and LST. Previous studies have assessed the overall UHI effect but fail to determine its correlation with quantified development (8-10). We hypothesized that increased developed land percentages would correspond with increased LSTs in sampled areas. To efficiently quantify developed surface area, a novel method was created using color analysis. Since surfaces with low albedo values are likely associated with grey and black colors, we used the distribution of color to estimate developed surface area. We found that LSTs increased in sampled areas as their developed land percentages increased. This result helps to further establish the relationship between land development and UHIs.

RESULTS

A total of 24 blocks from six Cleveland neighborhoods were sampled totaling 3,888,000 m^2 of surface area (**Figures 1 and 2**). The proportion of development in blocks within the two neighborhoods used for method validation ranged from 5.6% to 99.8% when assessed with the manual method and from 0.9% to 100% when assessed with the automated method (**Figures 3 and 4**). The manual method involved the perimeter of development being outlined by hand in order to determine the developed area, while the automated method relied on color detection software to determine the extent of developed area. The proportion of development in sampled blocks assessed via the automated method in

https://doi.org/10.59720/23-235

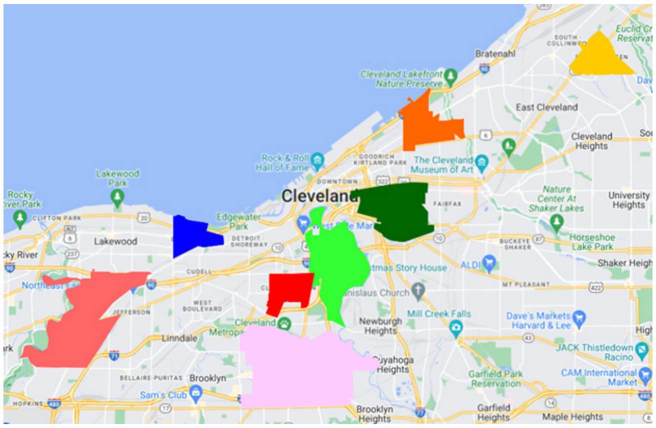

Figure 1. Neighborhoods in Cleveland used for sampling. A map showing the six neighborhoods sampled for comparison of LST to developed area percentage and the two neighborhoods (light and dark green) used to validate the color analysis method for automated calculation of percent development. The three majority non-white neighborhoods are shown in orange, red, and yellow.

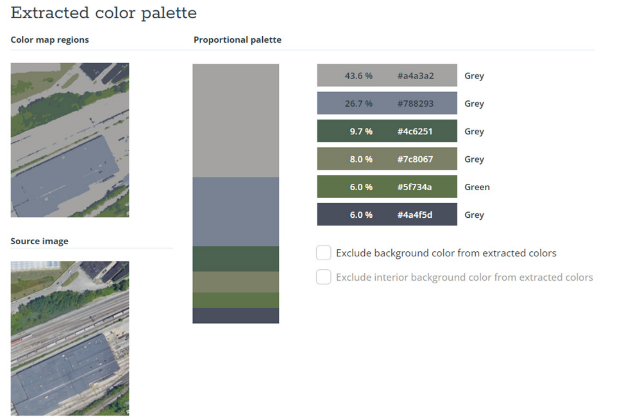

Figure 3. Color analysis of sample block. Screenshot from TinEye Color Extraction Lab software showing a source image (lower left) and the corresponding color mapped regions generated by the software (upper left). The portion and percentage of the processed image that is each color are displayed on the right.

the six neighborhoods that were then tested for correlation with LST ranged from 6.2% to 86.2% (**Figure 5**). Median LST in those six neighborhoods ranged from 24.39 °C to 32.8 °C. The three neighborhoods with the highest average temperatures of sampled blocks all have a majority non-white population (**Figure 5**). The three neighborhoods with the lowest average temperatures of sampled blocks all have a majority white population. The three sampled majority non-white neighborhoods ranked 1st, 3rd, and 4th in terms of highest average developed surface percentage across sampled blocks in the neighborhood (**Figure 6**). The three sampled majority white neighborhoods ranked 2nd, 5th, and 6th in terms of highest average developed surface percentage across sampled blocks in the neighborhood (**Figure 6**).

Validity testing of the color analysis method showed that it had a high degree of linearity and parity (slope ≈ 1) when compared to the manual development quantification method ($R^2 = 0.9$, $p < 0.0001$) (**Figure 7**). This means that the automated method was relatively accurate at estimating developed area, as a percentage of total area, through color analysis. There is a moderate degree of linearity between developed surface percentage and LST in any given sampled block ($R^2 = 0.48$, $p < 0.0001$) (**Figure 8**).

DISCUSSION

This study measured the association between the proportion of development in an urban area and its LST. The results show a moderate degree of linearity between developed surface area and LST ($R^2 = 0.48$, $p < 0.0001$) (**Figure 8**). This suggests that development may be an important contributor to higher LSTs in metropolitan areas. Future investigations of other potential contributors to this variability are warranted and could include comparisons of correlation within and between neighborhoods, and exploration of other natural geographic factors, vegetation, the built environment, traffic, industry, and neighborhood demographics. Future research could also investigate whether similar results are found in

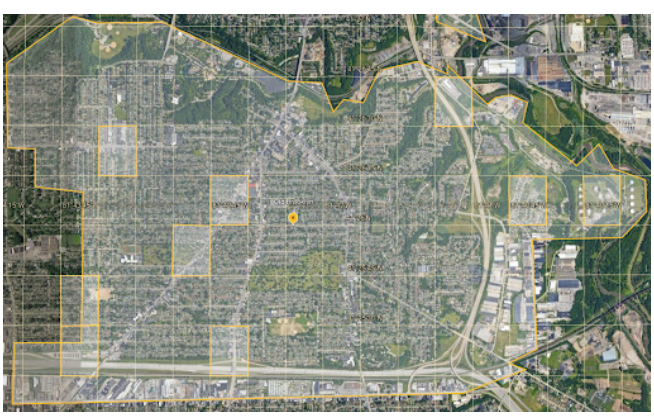

Figure 2. A representative image of the random sampling process employed to determine the development percentage in each selected block. Nine sampled blocks are gridded-off in a neighborhood in Cleveland with ~90 blocks total and an area of ~14,580,000 m².

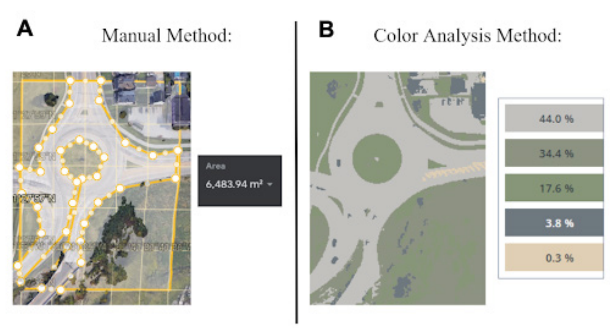

Figure 4. Comparison of manual versus automated color development analysis methods. A: the manual method of surveying the developed surface area in a block. The dots show the perimeter of development being outlined in order to determine the developed area, and then from that the development percentage. Structures, driveways, and streets were considered development. B: the color analysis method. Greys and blacks were considered infrastructure and were consequently summed to determine the development percentage of a sampled block.

https://doi.org/10.59720/23-235

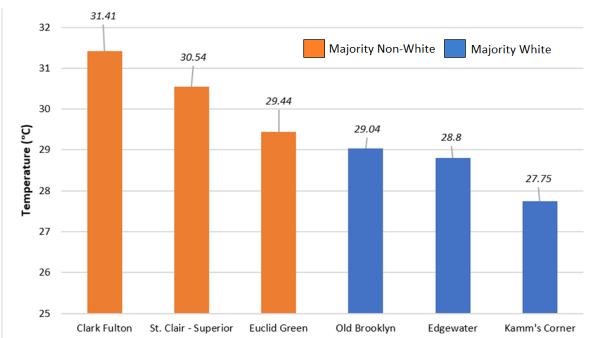

Figure 5. Average LST across all sampled blocks in each neighborhood. Orange colored bars denote a majority non-white neighborhood, and blue bars denote a majority white neighborhood. Racial demographics were sourced from City-Data (12).

other cities.

An automated development analysis method was employed, using color as a proxy for development, to efficiently quantify the amount of development within an area. Results show that the novel method produced very similar results to the manual assessment method, with the linear regression having a high degree of linearity and a slope of ~1 ($R^2 = 0.9$, $p < 0.0001$) (**Figure 7**). The use of color as a proxy for development was formed through consideration of the visual presentation of albedo since grey and black surfaces likely correspond with lower solar reflectiveness, and color analysis proved effective at corresponding with albedo. With the automated method being a viable option to use over the manual method, future researchers could employ it to save time and analyze larger geographic areas. One limitation of this experiment is the sample size, which could be addressed in future studies by sampling several cities and increasing the number of neighborhoods sampled in each city. In addition, LST data were values from summer 2016, while developed surface area was assessed from images from 2021. The portion of developed surface area in sampled blocks may have changed in the intervening years. Ideally, the LST data and satellite images would be from the same year.

Some foreseeable applications of these results are to use analysis of developed land, potentially accomplished with the

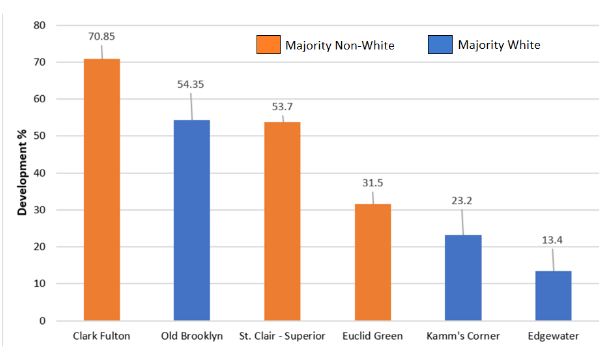

Figure 6. Average percentage of developed surface area across all sampled blocks in each neighborhood. Orange colored bars denote a majority non-white neighborhood, and blue bars denotes a majority white neighborhood. Racial demographics were sourced from City-Data.

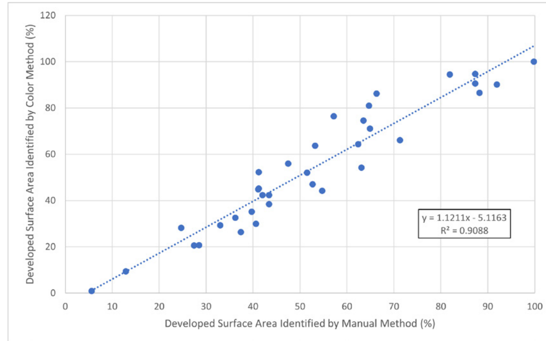

Figure 7. Correlation between manual development analysis method and the color development analysis method in the two neighborhoods used for method validation. Scatter plot showing the percentage of developed surface area calculated by the manual analysis method compared to the percentage of developed surface area calculated by the automated color development analysis method with a positive linear regression ($R^2 = 0.9$, $p < 0.0001$).

methodology developed for this study, in an area to support tree canopy assessment, strategic planting initiatives, or other heat mitigation efforts, as higher developed land portions may be a predictor of higher LST. City planners could also use assessments of developed land portions to target future infrastructure placement or to estimate the potential mitigating impact of adding tree canopy within a developed area on albedo and median temperatures. Expanding the scope of UHI investigations to further quantify its effects on racial minority populations would also be a valuable future research focus. Such a study may provide the necessary insight for overcoming the demographic and socioeconomic challenges related to this complex environmental justice issue.

MATERIALS AND METHODS

Six of 34 neighborhoods in Cleveland, Ohio, representing ~18% of Cleveland's total surface area, were selected for inclusion in this study (**Figure 1**). Neighborhoods were selected non-randomly in pairs to represent both majority racial-minority and majority white neighborhoods and matched for geographic similarity, such as proximity to bodies of water and industrial centers. Geographic factors were

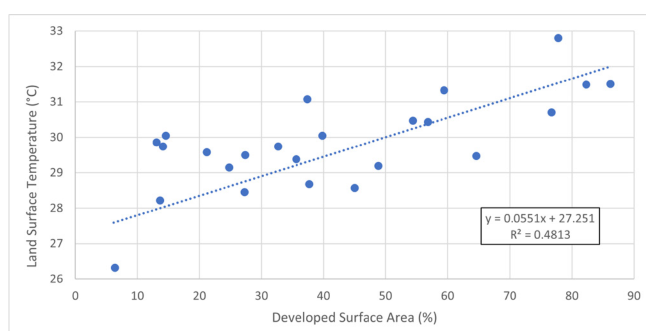

Figure 8. Correlation between developed surface area and land surface temperature in six Cleveland neighborhoods. A scatter plot comparing developed surface area, assessed with the automated color analysis method, to LST with a positive linear regression (R2 = 0.48, $p < 0.0001$).

https://doi.org/10.59720/23-235

matched to balance their potential contribution to LST.

To assess development, Google Earth images of the sampled neighborhoods from 2021 were used. A random sampling process was employed to acquire data that was representative of each neighborhood. Neighborhoods were gridded off into 162,000 m^2 blocks. Ten percent of the total blocks were surveyed (**Figure 2**). The sampled blocks were determined by a random number generator generating block coordinates. If a randomly sampled block had been selected previously, new random grid coordinates were generated.

An automated method of assessing the percentage of development in an area was designed and then tested in two neighborhoods not included in the rest of the data (**Figure 1**). For the automated method, imported images from Google Earth were processed with the TinEye Color Extraction Lab software that quantifies the percentage of each color present in the image (**Figure 3**). Grays and blacks were assumed to correspond with structures, driveways, and streets. All other colors were assumed to represent undeveloped surfaces. Color areas that were associated with developed surfaces were summed to generate a final developed surface percentage within a sampled block. The manual development quantification method was then compared to the automated color analysis method for verification of the accuracy of the color analysis method by analyzing how they performed in identical sampled blocks from two additional neighborhoods not included in the main analysis (**Figure 4**). The 26 sampled blocks in these two neighborhoods each had an area of 16,000 m^2.

LSTs for each sampled block in the six neighborhoods were obtained using NASA and U.S. Geological Survey satellite data from Landsat 8, which collects data at a 30-meter resolution (11). For this study, the median LST (of all official summer days from June 21, 2016 to September 22, 2016) at the midpoint of the sampled block was used in the analysis. The data on LSTs and the images used for developed area analysis were from different date ranges due to limitations around the availability of the data.

A simple linear regression model was used to calculate correlation coefficients between the manual method and the automated color method in the two neighborhoods used for method validation. A simple linear regression model was also used to calculate correlation coefficients between the percentage of developed surface area identified by the automated method and the median LST at midpoint of each block in the six neighborhoods not used for validating the automated method.

ACKNOWLEDGMENTS

Thank you to Dr. Pauline Terebuh for feedback during the revision process and Blake Reese for assisting with data collection.

Received: August 24, 2023
Accepted: February 28, 2024
Published: November 11, 2024

REFERENCES

1. Memon, Rizwan A., et al. "A review on the generation, determination and mitigation of Urban Heat Island." *Journal of Environmental Sciences*, vol. 20, no. 1, pp. 120-128, Jan. 2008, https://doi.org/10.1016/S1001-0742(08)60019-4
2. Huang, Kangning, et al. "Projecting global urban land expansion and Heat Island intensification through 2050." *Environmental Research Letters*, vol. 14, no. 11, p. 114037, Nov. 2019, https://doi.org/10.1088/1748-9326/ab4b71
3. Fitzpatrick, Matthew C., et al. "Contemporary climatic analogs for 540 North American urban areas in the late 21st Century." *Nature Communications*, vol. 10, no. 1, Feb. 2019, https://doi.org/10.1038/s41467-019-08540-3
4. Sangiorgio, Valentino, et al. "Development of a holistic urban heat island evaluation methodology." *Scientific Reports*, vol. 10, no. 1, Oct. 2020, https://doi.org/10.1038/s41598-020-75018-4
5. "Killer Summer Heat: Projected Death Toll from Rising Temperatures in America due to Climate Change." *NRDC*. www.nrdc.org/resources/killer-summer-heat-projected-death-toll-rising-temperatures-america-due-climate-change. Accessed 3 Oct. 2022.
6. Hsu, Angel, et al. "Disproportionate exposure to urban heat island intensity across major US cities." *Nature Communications*, vol. 12, no. 1, May 2021, https://doi.org/10.1038/s41467-021-22799-5
7. Taha, Haider. "Urban climates and heat islands: Albedo, evapotranspiration, and anthropogenic heat." *Energy and Buildings*, vol. 25, no. 2, pp. 99-103, May 1998, https://doi.org/10.1016/S0378-7788(96)00999-1
8. Doulos, L, et al. "Passive Cooling of Outdoor Urban Spaces. The Role of Materials." *Solar Energy*, vol. 77, no. 2, 2004, pp. 231-249, May 2004, https://doi.org/10.1016/j.solener.2004.04.005
9. Akbari, H, et al. "Cool Surfaces and Shade Trees to Reduce Energy Use and Improve Air Quality in Urban Areas." *Solar Energy*, vol. 70, no. 3, 2001, pp. 295-310, Feb. 2001,https://doi.org/10.1016/S0038-092X(00)00089-X
10. Rossi, Federico, et al. "Analysis of Retro-Reflective Surfaces for Urban Heat Island Mitigation: A New Analytical Model." *Applied Energy*, vol. 114, Feb. 2014, pp. 621-631, Nov. 2013, https://doi.org/10.1016/j.apenergy.2013.10.038
11. EarthExplorer, 2016. [Online]. Available: earthexplorer.usgs.gov/.
12. City-Data, 2017. [Online]. Available: https://www.city-data.com/indexes/neighborhoods/OH/1/.