This manuscript has been published in final form at https://doi.org/10.1016/j.ufug.2019.01.008

Urban greenness extracted from pedestrian video and its relationship with surrounding air temperatures

Kris Y. Hong[a], Pak Keung Tsin[b], Matilda van den Bosch[b], Michael Brauer[b], Sarah B. Henderson[a,b,*]

[a] Environmental Health Services, BC Centre for Disease Control, Vancouver, BC, Canada
[b] School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada

Abstract

Urban greenness has been associated with a wide range of health benefits, partially due to local cooling. Several studies on these health benefits have assessed individual and population exposure to urban greenness using the Normalized Difference Vegetation Index (NDVI) from different satellite platforms. Recent comparisons between birds-eye NDVI and street-level measurements suggest that NDVI can severely misclassify individual exposure which, in turn, can bias epidemiologic effect estimates. Pedestrian video data may provide a novel source of individual, eye-level information on both indoor and outdoor exposure to vegetation. The objective of this pilot study was to examine the potential of pedestrian video data for assessing exposure to urban greenness using secondary data collected for a different study on microscale urban air temperatures. Image processing was used to extract green, yellow, and shaded pixels from ~10 million frames of video footage collected during 40 sampling runs of 20 urban routes measuring 8–10 km each. Resulting greenness values (combined total of green, yellow, and shaded pixels) were compared with concurrent air temperatures using correlations, time series plots, and maps. Correlations ranged from -0.61 to 0.34 and were in the expected direction for 31 of 40 runs. Time series plots and overlay maps showed clear inverse relationships in many cases. Most routes originally chosen to characterize higher temperature areas had negative correlations. Flat and weakly positive relationships tended to occur when conditions were overcast or routes were closer to large waterways. Secondary data are limited for such assessments, but the methods described here are promising for future exposure assessment.

1. Introduction

A rapidly-growing body of literature on urban greenness has found that it benefits a wide range of physical and mental health outcomes in surrounding populations [1,2]. One mechanism for these benefits is through mitigation of the urban heat island (UHI) effect [3] where heat-generating activities, low-albedo infrastructure, and sparse vegetation make more developed areas warmer than their suburban and rural surroundings [4]. Given projected increases in hot weather events [5] and heat-related morbidity and mortality [6] due to climate change, urban greening to mitigate UHI effects is a growing public health priority [7]. Such greening would also have co-benefits for mental and physical health in urban environments, but adequate measurements of greenness are required to optimize future interventions.

Urban greenness is a challenging exposure to measure because it varies over time at a relatively small spatial scale. Most epidemiologic studies have used remote sensing data, particularly the normalized difference vegetation index (NDVI), to evaluate greenness at the individual or population scale [8]. The NDVI is calculated using the different wavelengths of light absorbed by vegetation in comparison with hard surfaces, such as concrete. Values range from -1 to 1, with water bodies having values near -1, surfaces such as rock and snow having values below 0.1, grassy areas falling in the range of 0.2 to 0.3, and healthy forests having values over 0.6 [9]. Several recent studies have also highlighted the utility of eye-level data, such as imagery from Google Street View [10]. Comparisons between these bird's-eye and eye-level views suggest that remotely-measured greenness is variably and moderately correlated with vegetation observed at

the street-level [11,12]. This makes sense, given that satellite images cannot accurately measure vertical vegetation, the density of the foliage cover, or vegetation that is obstructed by taller features of the landscape. However, satellite images can capture changes in greenness over time, whereas static, eye-level images only show greenness at a single point in time.

When it comes to the relationship between urban greenness and surrounding temperature, most studies have compared measures such as NDVI with remotely-measured land surface temperature (LST) [13,14,15]. Given the limitations of satellite images for accurate assessment of vegetation, estimates derived from these methods are likely to have large errors. In addition, the health risks associated with extreme hot weather are typically attributed to actual or apparent air temperature, and the relationship between air temperature and LST can be complex and nonlinear, such that areas with higher LST may not feel hotter than those with lower LST [16]. To address the latter limitation, some studies have combined NDVI, LST, and other variables to model air temperatures at high spatial resolution across urban areas [17,18]. We used such methods to develop the greater Vancouver heat map following an extreme hot weather event in 2009 [19], and found it to be a better indicator of population mortality than LST alone [17]. In the summer of 2014 we conducted a large campaign to evaluate the map using mobile microscale measurements of air temperature along 20 routes around the city [20].

In addition to the thermometer and global positioning system (GPS) carried by the pedestrian who collected the mobile monitoring data, a shoulder-mounted video camera was used to record every route. Although these data were intended to allow retrospective assessment of anomalies in the air temperature measurements, they also provide a novel source of information on urban greenness. This is a pilot study to assess whether useful information about surrounding greenness can be extracted from pedestrian video, which would provide a new method for greenness exposure assessment. First, an estimate of greenness percentage is extracted from ~10million frames of video footage. Second, the utility of the greenness values is assessed through quantitative comparison with simultaneous information on air temperature and qualitative comparison with satellite images and land use maps.

2. Methods

2.1 Study area

The greater Vancouver urban area is situated on the southwest coast of British Columbia, Canada. The summer climate is typically mild, with an average daily maximum temperature of 21.3 °C [21]. It is one of the greenest cities in the world [22], partly due to its location within the Pacific temperate rainforest. There are also 144,702 documented street trees in the City of Vancouver (roughly 1259 street trees per km^2), which excludes park trees and trees within private property lines [23].

2.1 Data collection

The methods for the 2014 study on microscale urban temperature variation are described in detail elsewhere [20]. In brief, 20 walking routes were selected around greater Vancouver with approximate lengths of 8–10 km per route, such that they could be covered at an ambulatory pace within 2–3 hours (Figure S1). All routes covered predominantly residential areas, and were chosen to characterize areas with high heat exposure according to the previously developed greater Vancouver heat map [17], high population density, or high social vulnerability [19,20]. Air temperature, position, and video data were collected using a Met One 064-2 thermometer, a Garmin GPSMAP 78s GPS, and GoPro Hero 3 video camera, respectively (Figure S2). The thermometer was chosen for its accuracy of ± 0.1 °C and a response time of 10s [24], and the GPS was selected for its accuracy of 10m [25]. The thermometer was sheltered in a radiation shield and mounted to a plastic pipe that extended 50 cm from the body of the pedestrian to prevent transfer of body heat (Figure S2). The camera was mounted on the left shoulder of the pedestrian, recording

continuous footage at 29.97 frames per second. All instrument times were synchronized prior to each run, with temperature and GPS data logged at 10-second intervals. Forty sets of data were collected by the same individual for the 20 routes, where each route was monitored twice (two triplicate runs were omitted). Replicate runs were walked on (1) the opposite sides of the street and (2) in opposite directions from each other. Data collection began in late May and ended in early September, and all routes were sampled between 15:00 and 18:00 to ensure that the hottest hours of the day were captured. Overcast days were not sampled.

2.3 Video analysis

Video data were analyzed in Python version 3.5 using the OpenCV library version 3.1.0 [26]. Each video frame was processed in the Hue-Saturation-Value (HSV) colour space. In brief, hue represents the colour, saturation represents the colour intensity, and value represents the brightness. In OpenCV, hue ranges from 0 to 180, saturation ranges from 0 to 255, and value ranges from 0–255. Each video frame had a resolution of 1280 × 738 pixels. Pixels that were green, yellow, or in shade were individually extracted from each frame by searching for HSV values that fell within the relevant ranges (Table 1). Both green and yellow pixels required two HSV ranges to exclude narrow bands of values that were not characteristic of the vegetation in the video, such as green utility poles (Fig. 1). These ranges were chosen based on standard OpenCV parameters and visual comparisons of different parameterizations. Yellows were included because of the yellowing of vegetation due to dry summer conditions (Fig. 1). Shaded pixels were included because vegetation can generate large areas of shade on sunny days, and because separating shade from low saturation and low value vegetation was too computationally challenging for this pilot study. The green, yellow, and shaded pixels were then combined into a single image to represent the overall influence of vegetation, and the greenness variable was calculated by dividing the number of vegetation pixels by the total number of pixels in the frame. The result was a greenness value ranging from 0 to 1 for every frame of recorded video, allowing us to average the values over different time intervals.

Table 1. Hue-Saturation-Value (HSV) ranges for green, yellow, and shaded pixels.

Pixel Colour	Hue	Saturation	Value
Green	26–90 26–90	51–255 78–255	103–255 38–255
Yellow	15–25 15–25	51–255 78–255	103–255 38–255
Shaded	0–179	0–77	0–102

2.4 Data processing

The results of the video analysis were imported into R version 3.3.2 for subsequent data processing and visualization. First, the frame-by-frame greenness values were averaged into 1-s intervals and linked to corresponding 10-s GPS and temperature data. Next, the greenness values were lagged by five seconds because the camera captured the entire area in front of the pedestrian, whereas the thermometer and GPS device took measurements at the position of the pedestrian. To further compensate for this discrepancy, a 45-second right-aligned moving average was used to smooth the greenness values. This corresponds to an integrated distance of approximately 60 m given a walking speed of approximately 1.4 m/s. Because the background air temperature changed during each 2–3 hour measurement period, the raw mobile air temperature measurements were adjusted to account for these shortterm temporal trends using data from Vancouver International Airport, as previously described [20].

2.5 Comparison with air temperatures

Correlations between surrounding air temperatures and greenness values were used to evaluate the direction, strength, and consistency of the relationship across all 40 runs. Scatterplots were used to show how the sets of data points compared, and time series plots of the normalized (z-score)

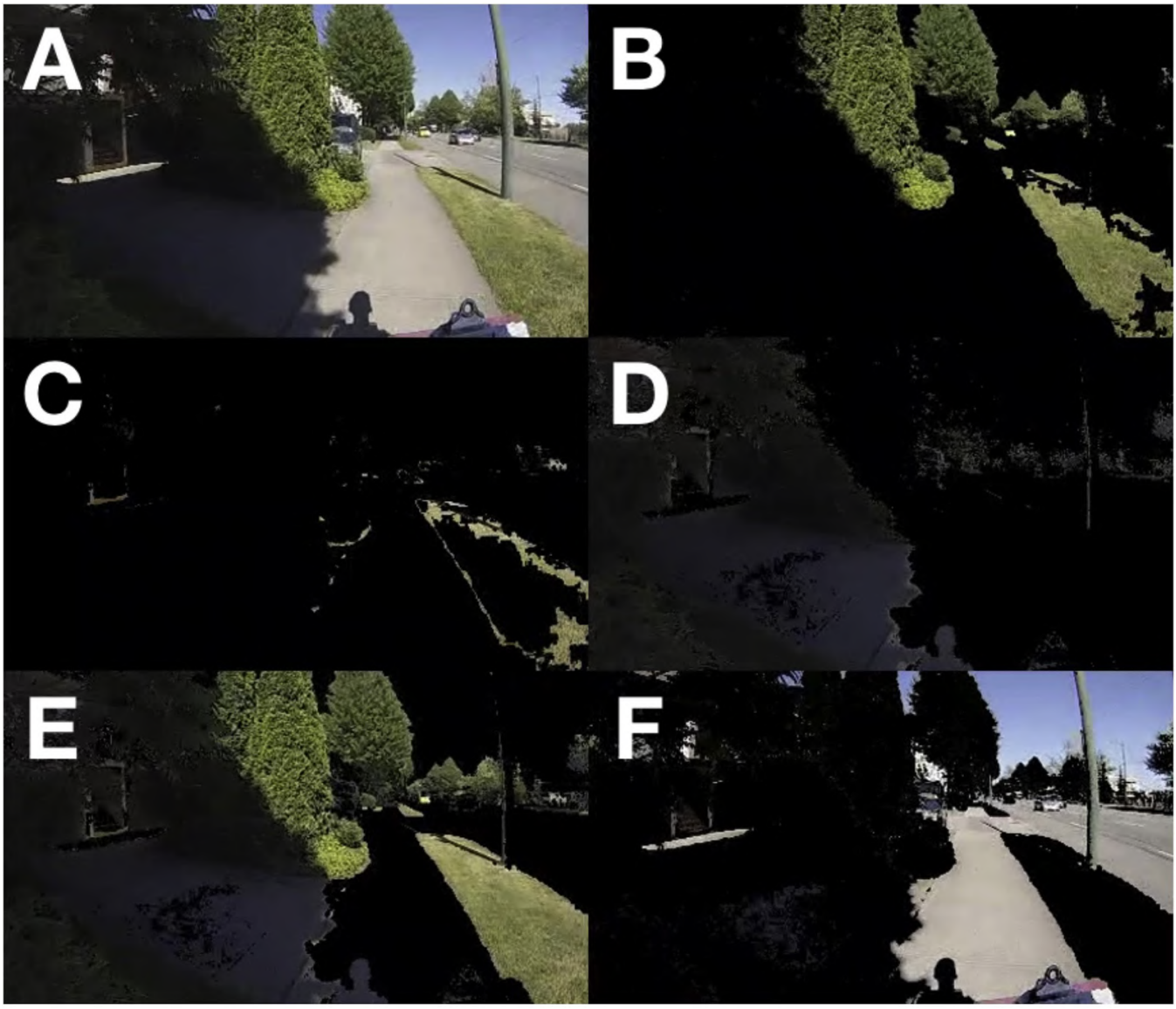

Fig. 1. Pixel extraction example for a single frame in route 8A showing: the original image (A); the extracted green pixels (B); the extracted yellow pixels (C); the extracted shaded pixels (D); all extracted pixels combined (E); and the inverse (F).

values were used to show the comparison on a temporal scale. Maps of the air temperature observations and greenness values were overlaid on satellite and land use imagery to visualize the data relative to their spatial surroundings.

3. Results

The mean air temperature measured across the 40 runs ranged from 19.8 to 31.9 °C, with standard deviations ranging from 0.26 to 1.21 °C. In comparison, the mean greenness ranged from 52% to 65%, with standard deviations ranging from 6% to 13%. The correlations between air temperature and greenness ranged from −0.61 to 0.34, with 31 of 40 runs having correlations in the expected negative direction (Fig. 3, Table S1). There was no clear spatial pattern in the correlation values, though route 9 was the only route with two positive values (Fig. 3) and the only route with part of its path directly traversing coastal waters (Figure S3H). Of the remaining seven routes with positive correlations, four had one weakly positive correlation and one strongly negative correlation, two had one moderately positive correlation and one moderately negative correlation, and one had weak correlations in both cases. There were several routes with two moderately or strongly

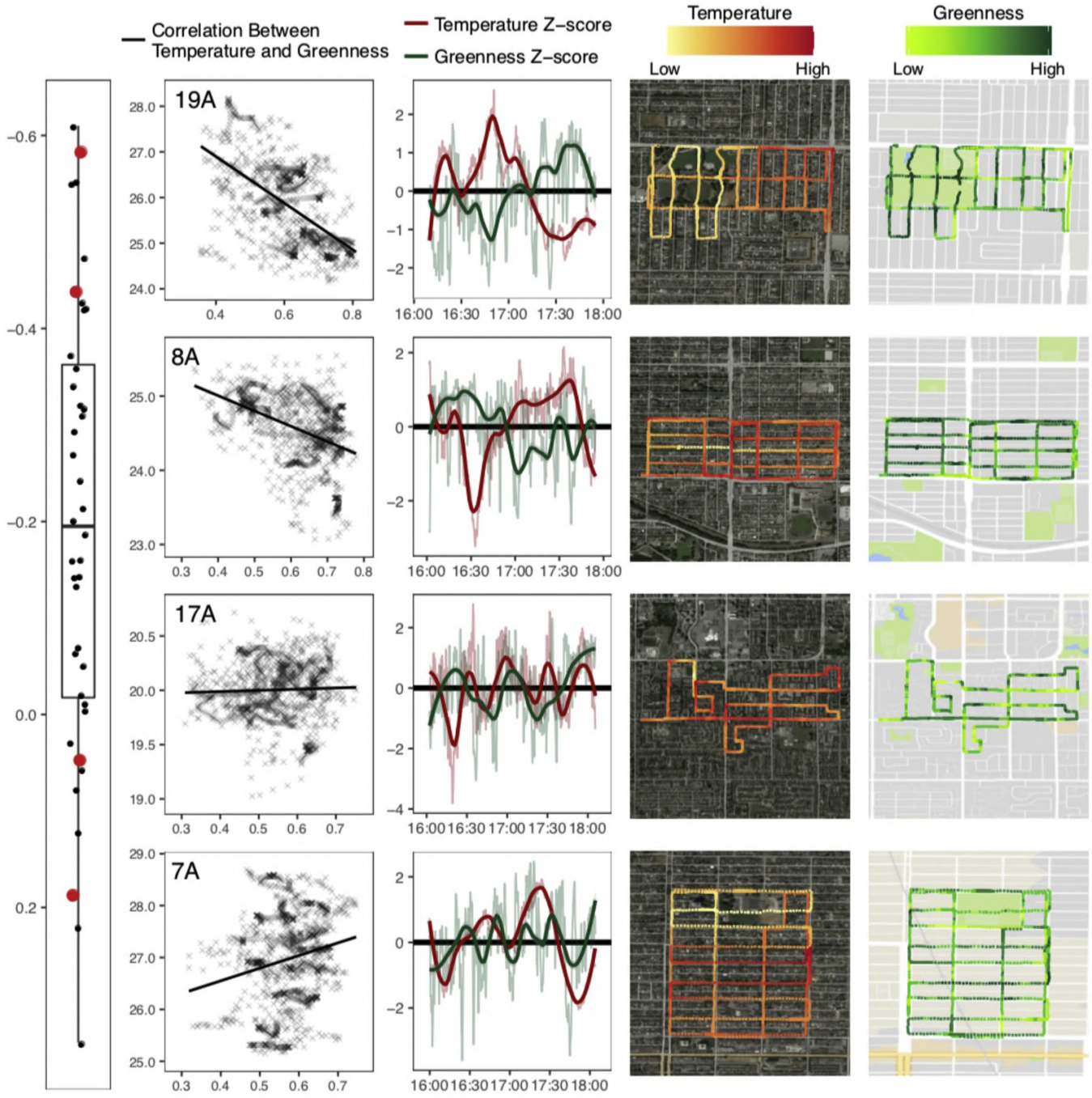

Fig. 2. The column on the far left is a boxplot showing the distribution of correlation coefficients between greenness and temperature. The white box indicates the interquartile range (IQR), the black line indicates the median, and the whiskers indicate the most extreme measurement that is no more than 1.5 times the IQR away from the white box in both directions. The enlarged red points indicate the four routes that are further visualized in the corresponding plots to the right. These routes provide examples of strongly negative, weakly negative, flat, and positive correlations, but all routes are shown in the Supplemental Materials (Figures S3A-S3H). For each of these routes there is: (1) a scatterplot showing the relationship between adjusted air temperature (y-axis) and greenness (x-axis); (2) a time-series plot showing the temporal relationship between raw and smoothed z-scores for the temperature measurements and greenness values; (3) a map showing the spatial distribution of temperature measurements overlaid on a satellite image; and (4) a map showing the spatial distribution of greenness values overlaid on a map of basic land use, including smaller roads (narrow white), larger roads (wide white), arterial roads (yellow), and parks (green).

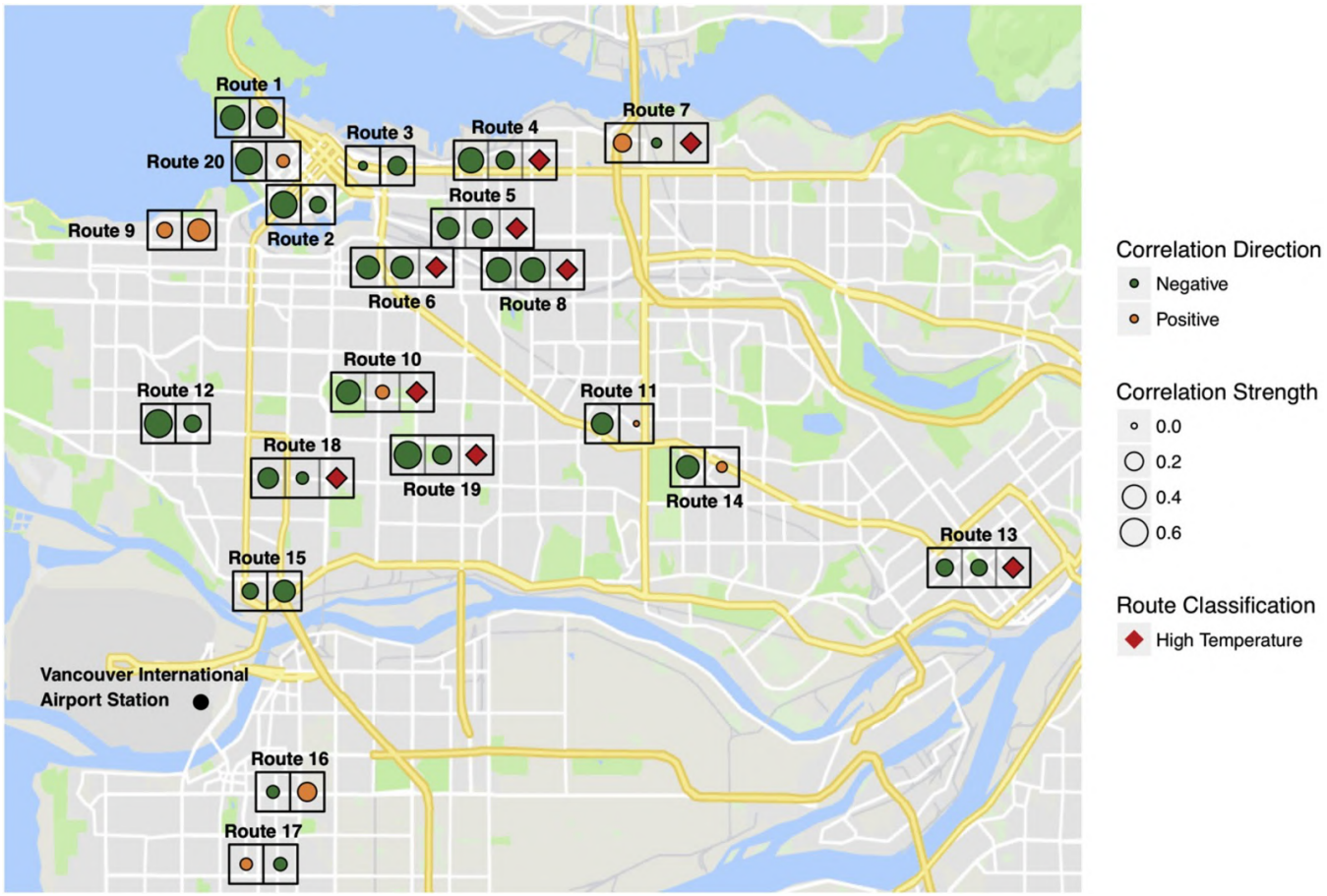

Fig. 3. Map of the study area in greater Vancouver, Canada, summarizing the observed correlations between greenness values extracted from pedestrian video and concurrent air temperatures for two replicates of 20 routes sampled in the summer of 2014. The size of each dot indicates the strength of the correlation, while the colour indicates the direction with the expected direction indicated in green. Red diamonds indicate routes that were originally chosen to reflect higher heat exposures according to the previously-developed greater Vancouver heat map[17].

negative correlations, many covering inland areas originally chosen to reflect higher heat exposures according to the previously-developed greater Vancouver heat map [17] (Fig. 3).

Routes 19A, 8A, 17A, and 7A, representing strongly negative, weakly negative, flat, and positive correlations, respectively, were selected for detailed visualization here with scatterplots, time series plots, and maps (Fig. 2). Routes 19A and 8A were located inland from the coastal and riverine areas while 17A and 7A were located closer to the water (Fig. 3). These visualizations for all routes can be found in the Supplemental Materials (Figures S3A-S3H).

Scatterplots showed marked variability in the relationship between air temperature and greenness, even where the correlation was relatively strong (Fig. 2). All four routes had clustering of observations in some areas, particularly route 7A where air temperatures were often constant over a relatively wide range of greenness values. This may suggest a problem with the operation of the thermometer during this run, which is further supported by the time-series plots. Both routes 19A and 7A spanned a difference of 4 °C between the lowest and highest measurements, but the time-series plot for 7A showed much less variation in air temperature than the plot for 19A (red lines). Another possible explanation for the weaker correlation is cloud cover, given that video footage from 7A showed that the route was overcast in some places.

For the time series plots (Fig. 2), localized smoothing of the z-score values showed a clear inverse relationship between air temperature and greenness for routes 19A and 8A, with lines often

crossing the zero threshold at the same time point. This pattern was apparent at the beginning of route 17A, but not afterwards, and the converse was apparent for route 7A where air temperature and greenness were most positively correlated. Of the four selected routes, background temperatures for 17A were lowest (20°C, Table S1) and the range of temperatures measured was the smallest (1.5 °C). Given that vegetation acts to cool the surrounding environment through shading and evapotranspiration, we expect that the relationship between greenness and air temperature would be attenuated on overcast days.

For the satellite and land use maps (Fig. 2), route 19A showed higher greenness values in the large park on the west side of the route compared with the residential area on the east side of the route, and especially along the major road at the far east. Likewise, the air temperatures were lowest in the park, and highest along the major road. There was no park area included in route 8A, but high greenness values were observed along all of the tree-lined residential streets. In comparison, low air temperature values were most apparent along the central street running from west to east. Route 17A covered a newer and more suburban residential area on the flat delta between the northern and southern branches of the Fraser River (Fig. 3). There is no clear relationship between temperature and greenness along this route, though temperatures were lowest in the northwest corner, around the largest park area.

Finally, route 7A showed that temperatures were clearly lowest in the large park on the north side of the route, and that greenness was also low in this area (Fig. 2). Review of the video for 7A showed two factors that likely affected the greenness values in the park: (1) slightly overcast conditions and (2) the camera angle included more sky than usual. In most cases the camera was angled such that 1/3 of the frame was above the horizon and 2/3 of the frame was below (Fig. 1). However, the park section of route 7A was closer to 1/2 above the horizon and 1/2 below, thereby driving down the greenness percentage of each frame. This could be standardized in future work. In addition to factors affecting the greenness extraction and possible problems with the thermometer, the northern part of 7A also slopes steeply down towards Burrard Inlet (Fig. 3), which often carries cooling westerly breezes in from the Pacific Ocean [27,28]. As such, meteorological conditions in this area may have a larger effect on air temperature than local greenness. We could not adjust our analyses for local wind effects, but they could be addressed in future work by adding a wind meter to the sampling equipment.

4. Discusion

The primary objective of this study was to evaluate whether pedestrian exposure to surrounding greenness could be successfully extracted from pedestrian video. Concurrent air temperatures were correlated with these values primarily to assess this objective, and secondarily to provide some street-level insight into the real-time greenness-temperature relationship observed in greater Vancouver during the summer of 2014. Although the correlations and visualizations suggest that our methods can generally identify vegetation from pedestrian video, there is room for improvement. Most significantly, we could not separate shade from dark and low saturation greenery because they share similar HSV ranges. Additionally, we classified all pixels by colour alone, meaning that some objects in the urban environment (e.g. cars and street signs) were misclassified as vegetation despite excluding the green and yellow HSV ranges that were characteristic of these objects. A similar study that extracted urban greenness from Google Street View stills used different methods [29], but encountered the same problems. Such limitations could be addressed using more refined methods such as advanced computer vision and deep learning.

Greenness is spatially and temporally variable over both the short-term and the long-term, and our measurements along each route were taken during the hottest hours of the day over the course of three months. As such, we did not attempt to evaluate the results against any other methods for measuring surrounding vegetation, such as the NDVI or

information from Google Street View. The most finely resolved NDVI data are measured by instruments such as Landsat, which only overpasses the study area at approximately 10:00 local time once every 16 days. Google Street View images are tagged with their capture dates, but Google strives to keep its data current and all 20 routes were last imaged in the summer of 2015 or later, whereas our data were collected in 2014.

We stress that this is an exploratory study using data collected for other purposes. Despite the limitations, our results suggest that street-level greenness can be extracted from pedestrian video using relatively simple tools, and that the extracted data are generally correlated with surrounding air temperatures in the expected direction. With further refinement, our methods could provide unprecedented spatial and temporal resolution for greenness exposure in individual-level studies. The first step would be to design a study with the specific objective of comparing video-derived greenness with other concurrent and widely used measures of greenness. For example, the timing of pedestrian video could be synchronized with Landsat overpasses on clear days, or with Google Street View plans to image a specific area. These routinely collected data could then be used to optimize the parameters used for the video processing (Table 1), ensuring more validity of the methods in future work.

Given that greenness may impact health via visually-stimulated stress reduction, eye-level video data may be more relevant than other measures. Monitoring studies that outfit participants with instruments such as GPS, pedometers, and air quality sensors could easily include an inexpensive video camera to capture total personal exposure (in both indoor and outdoor environments) to greenness for more comprehensive exposure estimates. Based on our experiences, we would recommend mounting cameras with wide-angle lenses in a location where the angle is expected to be most stable. While measuring street-level greenness can be time-consuming, it serves as an important complement to other measures. These data could also be useful for evaluating and validating the wide range of population-scale greenness models that are currently in use [8,30].

5. Conclusion

Urban greenness has a wide range of public health benefits, including the potential to mitigate the UHI effect. However, the exact relationship between vegetation and health benefits remains unclear, partly due to challenges in measuring greenness. We have proposed a colour-based image processing approach to estimate local greenness and vegetation from eye-level pedestrian video. Extracted greenness values were variable across each of the 20 routes and were negatively correlated with surrounding air temperatures in most cases. These results suggest that our methods can provide the basis for an efficient and effective approach to measuring greenness for future studies investigating the population health benefits of greenness.

References

1. James, P., Banay, R.F., Hart, J.E., Laden, F., 2015. A review of the health benefits of greenness. Curr. Epidemiol. Rep. 2, 131–142.
2. Nieuwenhuijsen, M.J., Khreis, H., Triguero-Mas, M., Gascon, M., Dadvand, P., 2017. Fifty shades of green: pathway to healthy urban living. Epidemiology 28, 63–71.
3. van den Bosch, M., Ode Sang, A., 2017. Urban natural environments as nature-based solutions for improved public health - a systematic review of reviews. Environ. Res. 158, 373–384.
4. Silva, H.R., Phelan, P.E., Golden, J.S., 2010. Modeling effects of urban heat island mitigation strategies on heat-related morbidity: a case study for Phoenix, Arizona, USA. Int. J. Biometeorol. 54, 13–22.
5. Meehl, G.A., Tebaldi, C., 2004. More intense, more frequent, and longer lasting heat waves in the 21st century. Science 305, 994–997.
6. Gasparrini, A., Guo, Y., Sera, F., Vicedo-Cabrera, A.M., Huber, V., Tong, S., de Sousa Zanotti Stagliorio Coelho, M., Nascimento Saldiva, P.H., Lavigne, E., Matus Correa, P., et al., 2017. Projections of temperature-related excess mortality under climate change scenarios. Lancet Planet. Health 1, e360–e367.
7. Bowler, D.E., Buyung-Ali, L., Knight, T.M., Pullin, A.S., 2010. Urban greening to cool towns and cities: a systematic review of the empirical evidence. Landsc. Urban Plan. 97, 147–155.
8. Rugel, E.J., Henderson, S.B., Carpiano, R.M., Brauer, M., 2017. Beyond the Normalized Difference Vegetation Index (NDVI): Developing a natural space index for population-level health research. Environ. Res. 159, 474–483.
9. Weier, J., Herring, D., 2000. Measuring Vegetation (NDVI and EVI). accessed in 2019. https://earthobservatory.nasa.gov/features/MeasuringVegetation.
10. Rundle, A.G., Bader, M.D.M., Richards, C.A., Neckerman, K.M., Teitler, J.O., 2011. Using Google Street View to audit neighborhood environments. Am. J. Prev. Med. 40, 94–100.
11. Jiang, B., Deal, B., Pan, H., Larsen, L., Hsieh, C.-H., Chang, C.-Y., Sullivan, W.C., 2017. Remotely-sensed imagery vs. eye-level photography: evaluating associations among measurements of tree cover density. Landsc. Urban Plan. 157, 270–281.
12. Larkin, A., Hystad, P., 2018. Evaluating street view exposure measures of visible green space for health research. J. Expo. Sci. Environ. Epidemiol.
13. Weng, Q., Lu, D., Schubring, J., 2004. Estimation of land surface temperature–vegetation abundance relationship for urban heat island studies. Remote Sens. Environ. 89, 467–483.
14. Zhang, X., Zhong, T., Feng, X., Wang, K., 2009. Estimation of the relationship between vegetation patches and urban land surface temperature with remote sensing. Int. J. Remote Sens. 30, 2105–2118.
15. Imhoff, M.L., Zhang, P., Wolfe, R.E., Bounoua, L., 2010. Remote sensing of the urban heat island effect across biomes in the continental USA. Remote Sens. Environ. 114, 504–513.
16. Vancutsem, C., Ceccato, P., Dinku, T., Connor, S.J., 2010. Evaluation of MODIS land surface temperature data to estimate air temperature in different ecosystems over Africa. Remote Sens. Environ. 114, 449–465.
17. Ho, H.C., Knudby, A., Sirovyak, P., Xu, Y., Hodul, M., Henderson, S.B., 2014. Mapping maximum urban air temperature on hot summer days. Remote Sens. Environ. 154, 38–45.
18. Kloog, I., Nordio, F., Coull, B.A., Schwartz, J., 2014. Predicting spatiotemporal mean air temperature using MODIS satellite surface temperature measurements across the northeastern USA. Remote Sens. Environ. 150, 132–139.

19. Kosatsky, T., Henderson, S.B., Pollock, S.L., 2012. Shifts in mortality during a hot weather event in Vancouver, British Columbia: rapid assessment with case-only analysis. Am. J. Public Health 102, 2367–2371.
20. Tsin, P.K., Knudby, A., Krayenhoff, E.S., Ho, H.C., Brauer, M., Henderson, S.B., 2016. Microscale mobile monitoring of urban air temperature. Urban Clim. 18, 58–72.
21. Government of Canada, 2016. Canadian Climate Normals 1981-2010 Station Data. October 27, 2016.
22. MIT Senseable City Lab. Treepedia, 2016. Exploring the Green Canopy in Cities Around the World. October 27, 2016.
23. City of Vancouver, 2016. Vancouver Street Trees. October 27, 2016.
24. Met One Instruments Inc, 2005. Model 064-1, 064-2 Temperature Sensor Operation Manual Document No 064-9800.
25. Garmin, 2019. Gpsmap 78 series.
26. Bradski, G., Kaehler, A., 2008. Learning Opencv: Computer Vision With the Opencv Library. O'Reilly Media, Inc.
27. Oke, T.R., Hay, J.E., 1994. The Climate of Vancouver. University of British Columbia.
28. Völker, S.B., Baumeister, H., Claßen, T., Hornberg, C., Kistemann, T., 2013. Evidence for the temperature-mitigating capacity of urban blue space – a health geographic perspective. Erdkunde 67, 355–371.
29. Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., Zhang, W., 2015. Assessing street-level urban greenery using Google Street View and a modified green view index. Urban For. Urban Green. 14, 675–685.
30. Markevych, I., Schoierer, J., Hartig, T., Chudnovsky, A., Hystad, P., Dzhambov, A.M., de Vries, S., Triguero-Mas, M., Brauer, M., Nieuwenhuijsen, M.J., et al., 2017. Exploring pathways linking greenspace to health: theoretical and methodological guidance. Environ. Res. 158, 301–317.