Urban Forestry & Urban Greening 21 (2017) 11–15

Contents lists available at ScienceDirect

Urban Forestry & Urban Greening

journal homepage: www.elsevier.com/locate/ufug

Short communication

Google Street View shows promise for virtual street tree surveys

Adam Berland*, Daniel A. Lange

Department of Geography, Ball State University, Muncie, IN, USA

ARTICLE INFO

Article history:
Received 8 August 2016
Received in revised form 13 October 2016
Accepted 6 November 2016
Available online 13 November 2016

Keywords:
Municipal forestry
Species identification
Street-level photography
Tree inventories

ABSTRACT

Geospatial technologies are increasingly relevant to urban forestry, but their use may be limited by cost and technical expertise. Technologies like Google Street View^{TM} are appealing because they are free and easy to use. We used Street View to conduct a virtual survey of street trees in three municipalities, and compared our results to existing field data from the same locations. The virtual survey analyst recorded the locations of street trees, identified trees to the species level, and estimated diameter at breast height. Over 93% of the 597 trees documented in the field survey were also observed in the virtual survey. Tree identification in the virtual survey agreed with the field data for 90% of trees at the genus level and 66% of trees at the species level. Identification was less reliable for small trees, rare taxa, and for trees with multiple species in the same genus. In general, tree diameter was underestimated in the virtual survey, but estimates improved as the analyst became more experienced. This study is the first to report on manual interpretation of street tree characteristics using Street View. Our results suggest that virtual surveys in Street View may be suitable for generating some types of street tree data or updating existing data sets more efficiently than field surveys.

1. Introduction

Street trees—trees in the public right-of-way along streets—are a prominent component of the urban forest (Berland and Hopton, 2014; McPherson et al., 2016). Street trees provide valuable environmental, social, and economic benefits (Mullaney et al., 2015). Given the importance of street trees, field-based tree surveys are carried out to facilitate informed management of this municipal resource. Street tree surveys are used to characterize attributes of street tree assemblages such as the number of trees, tree sizes, and species composition (Pedlar et al., 2013). These surveys may be used to identify management needs associated with planting, pruning, tree removal, mitigating hazards and infrastructure conflicts, and understanding the potential and realized impacts of tree pests and pathogens. Furthermore, street tree survey data are frequently used to estimate the environmental and aesthetic benefits of street trees using models like i-Tree Streets (i-Tree, 2016).

Field-based street tree surveys provide invaluable information for municipal forest managers, but they are costly, labor-intensive, time-consuming, and pose safety risks (e.g., automobile crashes) to field crews (Alonzo et al., 2016). These problems can be mitigated through the use of geospatial technologies, which are emerging with broad applications for characterizing and managing the urban forest (Ward and Johnson, 2007). Remote sensing has driven major advances in quantifying urban forest abundance (McGee et al., 2012; O'Neil-Dunne et al., 2014) and quality (Alonzo et al., 2014) using airborne and satellite imagery. Remote sensing techniques have permitted powerful analyses such as early detection of tree pests (San Souci et al., 2009), comparison of tree distributions to population characteristics (Landry and Chakraborty, 2009; Berland et al., 2015), and estimation of tree benefits (Alonzo et al., 2016). However, despite the capabilities and consistent advances in the use of remote sensing to characterize the urban forest, remote sensing techniques remain largely inaccessible to non-experts including most urban forest managers. In addition, high-quality remote sensing products are generally expensive because they require costly data sets and a paid geospatial analyst. As such, these products are out of reach for most municipalities due to inadequate funding and in-house expertise.

Geospatial technologies that are simple to use and free or inexpensive may be more useful to a larger number of municipal forest managers as compared to sophisticated approaches requiring expensive data. Google Street View^{TM} (GSV) is a geospatial platform that provides ground-based panoramic photographs captured along streets (Anguelov et al., 2010). GSV is available online (http://www.google.com/streetview) or through Google Earth^{TM} (https://www.google.com/earth/), and it provides excellent cov-

* Corresponding author at: Department of Geography, 2000 W University Ave, Muncie, IN 47306, USA.
E-mail address: amberland@bsu.edu (A. Berland).

http://dx.doi.org/10.1016/j.ufug.2016.11.006
1618-8667/

erage of streets in the US and many other countries around the world. GSV's street-level perspective of a community is a powerful tool for virtual exploration of neighborhoods, and this technology has been used in several research applications. Li et al. (2015a,b) used GSV to characterize greenery in urban streetscapes by quantifying green pixels in Street View images. Rousselet et al. (2013) used GSV to study the distribution of a forest pest by identifying the pest's distinctive nests. GSV has also been used in public health studies to conduct virtual audits of the built environment (Badland et al., 2010; Clarke et al., 2010). In these diverse applications, GSV has shown promise for providing useful virtual information that reduces the need to visit field sites.

Given the widespread use of GSV, we conducted this exploratory study to understand whether this technology can be employed to conduct robust virtual surveys of street trees. In other words, can we generate reliable data about street trees based solely on GSV imagery? We assessed the quality of street tree data generated using GSV imagery by comparing it to existing field data. Specifically, we evaluated data regarding the numbers, sizes, and species of trees. As compared to field surveys, we expected that this virtual approach could save time, money, and reduce safety hazards associated with field work. As this study is the first of its kind, this article provides valuable information about the capabilities, limitations, and possible applications of GSV image interpretation in urban forestry research and practice.

2. Methods

2.1. Study area

This study was conducted in the following three municipalities in suburban Cincinnati, OH, USA: Mt. Healthy (39.23°N, 84.55° W; population 6061), Reading (39.23°N, 84.44° W; population 10,357), and Wyoming (39.23°N, 84.47° W; population 8404). These municipalities are primarily comprised of low- to medium-density residential land. In each municipality, over 70% of the houses were built before 1970 (US Census Bureau, 2016). Wyoming has an active planting program for street trees; the other two have existing street trees but no current planting program.

2.2. Field data collection

In 2013, we randomly sampled street segments in the study area, surveying >10% of the total length of local, public streets in each municipality. On each street segment, we followed the i-Tree Streets sampling protocol (i-Tree, 2016). We recorded the geographic location of every tree encountered in the public right-of-way, with trees defined as woody vegetation >2.5 cm (1 in.) diameter at breast height (dbh). Trees were identified to the species level, and dbh was measured to the nearest 0.1 cm. A geographic information system (GIS) was used to plot the locations of trees inventoried in the field.

2.3. Virtual survey of street trees

We used GSV to virtually survey the same street segments that were sampled in the field. The virtual survey was conducted from October 2015 to March 2016, roughly 2.5 years after the field survey. The analysis was conducted by a single individual who held a bachelor's degree in biology with a focus on field botany. The analyst had never visited the study area, so he had no prior knowledge of which tree species to expect.

GSV was accessed using Google Earth^{TM} Pro, which was freely available and supported the import of GIS shapefiles to view street segment lines within the GSV interface. This feature was useful for ensuring that the field crew and virtual survey analyst both inventoried the same street segments. Along each street segment, the analyst noted each tree's geographic location, identified the tree to the species level, and manually estimated dbh to the nearest cm based on the tree's appearance in the photograph within the landscape context. The analyst also noted the date of the GSV image. We did not collect data regarding tree health, evidence of pests, or infrastructure conflicts in this exploratory study.

The analyst calibrated his virtual survey estimates at multiple stages of the study. Prior to the virtual survey, 20 trees of various species and size classes were selected in our local community; the analyst estimated the dbh and species of these trees using GSV, and visited them in the field to assess his performance. Then during the virtual survey, the analyst completed one municipality at a time, and assessed his performance in that municipality before moving on to the next. After generating unsatisfactory dbh estimates in Mt. Healthy, the analyst used trees of known dbh to make a pictorial reference guide to assist in estimating dbh based on a tree's appearance in GSV. We assumed this midstream calibration would lead to varying levels of data quality among the study municipalities, but would improve the overall performance of the virtual survey for the study area as a whole.

2.4. Comparison of field data and virtual survey data

We compared the virtual survey to the field data to assess the performance of the virtual survey. We approached this comparison with the assumption that the field data were correct, and that disagreement between the two data sets was introduced by the virtual survey (but see Discussion points below about species identification and temporal mismatch between field and virtual data collection). To compare observations, trees from the two surveys were first matched by geographic location using a spatial join in ArcGIS 10.3 (ESRI, 2014) and subsequent manual adjustments. Trees were matched across the two surveys if they were located at the same street address, similarly sized, and represented the same genus or genera that could reasonably be confused.

To compare tree counts, we noted the total number of trees in each survey, the number of matched trees, and the percent of street segments on which the tree counts agreed between the two surveys. For matched trees, we compared the level of agreement in tree identification and size class estimates between the field and virtual surveys using raw percent agreement and Cohen's kappa. Kappa tests whether the observed level of agreement exceeds agreement arising through random chance; this is valuable in situations where species diversity is poor and the virtual analyst could produce fairly high percent agreement by simply guessing the most common species every time. We calculated kappa using the irr package in R version 3.2.2 (R Core Team, 2015). To assess agreement on tree identification, kappa was computed separately for genus and species. To assess size class agreement, trees were first binned into the following dbh classes commonly employed in forestry: 0–7.6 cm (0–3 in.), 7.6–15.2 cm (3–6 in.), 15.2–30.5 cm (6–12 in.), 30.5–45.7 cm (12–18 in.), 45.7–61.0 cm (18–24 in.), 61.0–76.2 cm (24–30 in.), and >76.2 cm (>30 in.). Then we calculated a weighted kappa on these ordinal data, wherein virtual dbh estimates were penalized more harshly as their distance from the observed size class increased.

In addition to kappa, we used linear regression analysis in R version 3.2.2 (R Core Team, 2015) to compare dbh estimates from the virtual survey to field measurements. In the regression equation, a slope of 1 would indicate that the set of virtual dbh estimates was proportional to field measurements, a slope <1 would indicate dbh was underestimated in the virtual survey, and a slope >1 would indicate dbh was overestimated in the virtual survey. The coefficient of determination (R^2) was used to evaluate the consistency of

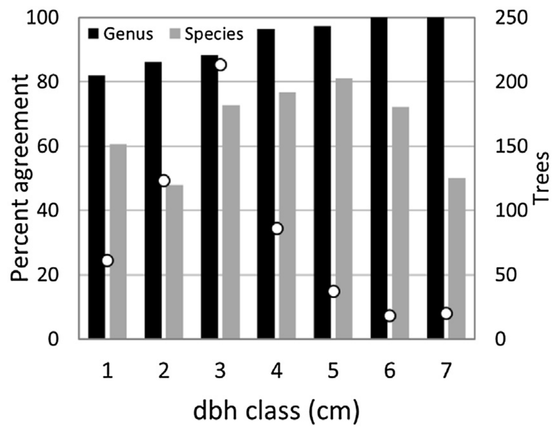

Fig. 1. Agreement between the field and virtual surveys in terms of tree identification at the genus and species levels. Dbh classes are: 1 (0–7.6 cm), 2 (7.6–15.2 cm), 3 (15.2–30.5 cm), 4 (30.5–45.7 cm), 5 (45.7–61.0 cm), 6 (61.0–76.2 cm), and 7 (>76.2 cm). White circles indicate the number of trees in each dbh class.

virtual dbh estimates, where a higher R^2 value indicated a more consistent relationship between virtual dbh estimates and field measurements, and a lower R^2 value indicated a weaker pattern between virtual estimates and field measurements. Regressions were computed for each municipality individually and for all three municipalities combined.

3. Results

We conducted both field and virtual surveys for 115 street segments across the three study municipalities (Table 1). The virtual survey of Reading and Wyoming took 11.2 h for one analyst to complete (time was not recorded for Mt. Healthy). This was decidedly faster than the 28.5 h needed for a two-person field team to collect the same tree variables in Reading and Wyoming. Although we inventoried the same street segments in the two surveys, there was a temporal mismatch in data capture for every street segment. The field survey took place September to October 2013, while GSV imagery on the study street segments was acquired from July through September (leaf-on season) across the following years: 2009 (2% of imagery), 2011 (38%), 2014 (56%), 2015 (3%), and 2016 (<1%). We encountered 597 trees in the field survey and 606 trees in the GSV virtual survey (Table 1). Tree counts were the same between the two surveys for 67% of street segments. The average difference in tree counts between the field and virtual surveys was 0.07 trees per street segment, and the single largest discrepancy for a segment was seven trees. Over 93% of trees encountered in the field survey were matched between the two surveys based on their location, dbh, and genus (Table 1).

For matched trees, the genus identification was in agreement between the field and virtual surveys for 90% of trees (kappa = 0.88, $p < 0.001$), and agreement was higher for larger trees (Fig. 1). At the species level, the level of agreement for tree identification was 66% of trees (kappa = 0.64, $p < 0.001$). Agreement was very high for common species, particularly when that species was the only member of its genus represented in the study area. For example, *Pyrus calleryana* was the most common species encountered in the field (14% of matched trees), and the field and virtual surveys agreed for nearly 99% of individuals. Similarly, *Gleditsia triacanthos* and *Zelkova serrata* were both common (6% and 5% of matched trees, respectively) and identified with 100% agreement between the two surveys. On the other hand, agreement was much lower for genera with many species represented in the field survey. For instance, six *Ulmus* species represented 7% of the field survey, but there was 0% agreement at the species level between the two surveys. Likewise, agreement was low (37%) for the eight *Quercus* species comprising 7% of the field survey.

Virtual survey estimates of dbh placed 67% of matched trees into the same size bin (weighted kappa = 0.73, $p < 0.001$). Regression slopes were <1 in each municipality, indicating that the analyst underestimated dbh in the virtual survey (Fig. 2). However, the analyst improved his estimates over the course of the study by evaluating his performance after each municipality before moving on to the next, as indicated by increasing R^2 values and slopes increasing toward 1 as he progressed from Mt. Healthy to Reading to Wyoming (Fig. 2).

4. Discussion

GSV imagery shows promise as a tool for conducting virtual surveys of street trees. As expected, the virtual survey was conducted by a single person more quickly, at lower cost, and with reduced safety risks as compared to the field survey. The intended use of street tree data ultimately dictates what is considered acceptable data quality, so we hesitate to make judgments about the ability to generate usable data with this approach. In general, the virtual survey produced similar data to the field survey, which is promising for the continued implementation of GSV as a geospatial tool to aid research and practice in urban forestry. For example, the similar number of total trees observed in the field and virtual surveys falls within the range of error expected when conducting a sample inventory for an i-Tree Streets study (i-Tree, 2016), and most trees were observed in both surveys (Table 1). When trees were not identified in both surveys, it could be attributable to two main issues. First, GSV imagery for the study area was acquired from 2009 to 2016, but no imagery was acquired in 2013 when the field survey took place. Thus, a number of trees were planted or removed in the time between the two surveys. Second, in locations without a sidewalk to delineate the public right-of-way, the virtual survey analyst occasionally had difficulty determining whether a tree should be counted as a street tree or not.

While our findings provide the first gauge of how well GSV can be used to conduct virtual street tree surveys, several limitations and recommendations should be considered. High agreement in genus identification suggests that analysts can use GSV to reliably identify most trees to the genus level, especially for larger trees (Fig. 1). Larger trees have more leaves and larger boles with defining bark characteristics to aid identification, and they extend closer to the GSV panorama center. Small trees with very few leaves are difficult to see in GSV, especially if the image quality is marginal at that location. While genus identification was fairly reliable, species identification was less successful, particularly for genera like *Acer, Quercus,* and *Ulmus* that had multiple species represented in our sample. In our study, the analyst had never visited the study area, but a local urban forester with knowledge of the species planted within a municipality would have an advantage in identifying trees using GSV imagery. Providing analysts with training materials including photos and lists of distinguishing characteristics for local street trees could improve species identification. Note that some of the observed disagreement for species identification likely stems from the field crews and virtual analyst recognizing the same type of tree but assigning it two different names; for example, a hybrid Freeman maple may have been recorded as any one of *Acer x freemanii, A. rubrum,* or *A. saccharinum.*

The analyst's estimates of dbh improved markedly as he learned from his performance in one municipality and applied that insight to the next (Fig. 2). For future studies, we recommend developing a photographic guide showing trees of various diameters in GSV imagery so analysts can compare trees in their studies to reference trees of known dbh. While larger trees were more readily

Table 1
Summary of trees inventoried in the field and virtual surveys, and number of trees matched between the two surveys based on their geographic location, species, and dbh.

Municipality	Street segments (total km)	Trees (field)	Trees (virtual)	Matched trees (% of field total)
Mt. Healthy	31 (4.4 km)	63	69	56 (89%)
Reading	49 (5.8 km)	64	55	54 (84%)
Wyoming	35 (5.7 km)	470	482	448 (95%)
All 3 combined	115 (15.9 km)	597	606	558 (94%)

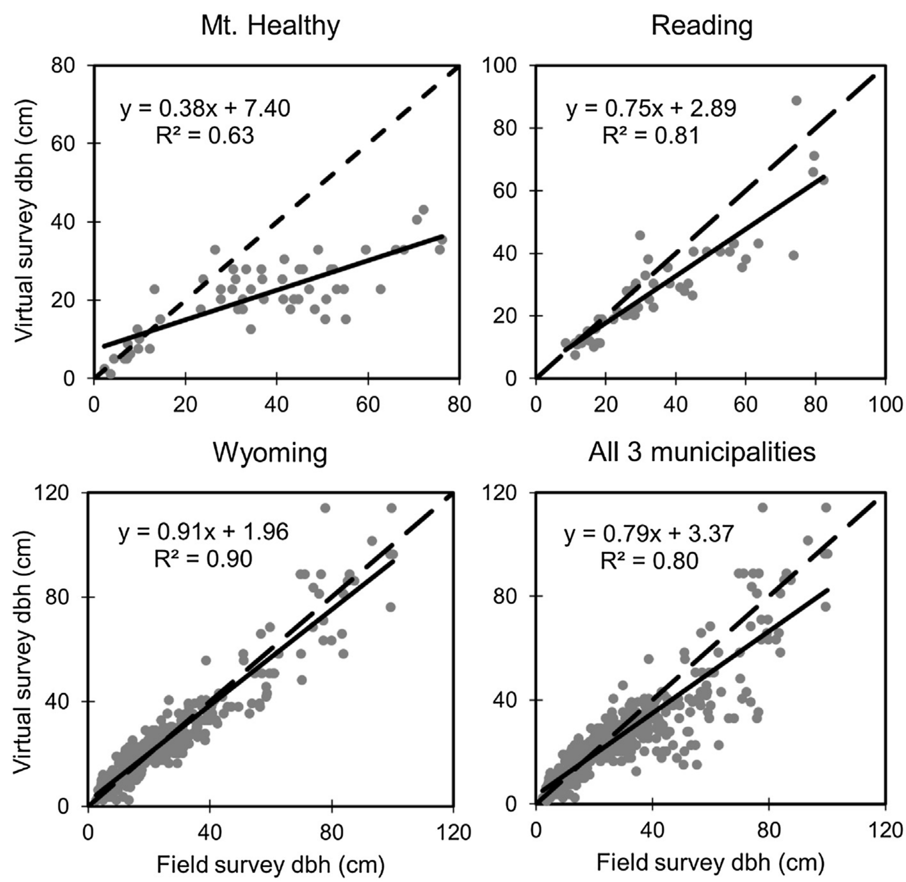

Fig. 2. Linear regressions assessing the relationship between field measurements and virtual survey estimates of dbh. The regression (solid) and 1-to-1 (dashed) lines are shown for reference. The analyst completed estimates in the order of Mt. Healthy, Reading, and then Wyoming, and evaluated his performance in each municipality before moving on to the next.

identified at the genus level, it was more difficult to accurately estimate dbh for larger trees, as evidenced by less consistent estimates for larger trees (Fig. 2). As dbh estimates were generally more accurate for smaller trees, the relatively high proportion of smaller trees in Wyoming may partly explain the improved success estimating dbh in that community (Fig. 2). We used manual estimates of dbh because we wanted to conduct a free analysis without introducing advanced computing requirements, but it may be possible to use image measurement software to automate dbh estimation. This automation could improve consistency across multiple analysts and save time, and may thus be valuable even if software-based estimates of dbh are somewhat inaccurate. More broadly, we are aware of ongoing efforts to automate the use of GSV imagery to generate a suite of tree data including tree counts, health, size, and species, but these techniques are still under development and require advanced computer skills beyond the reach of most communities. The manual approach proposed here provides a lower-tech alternative that is widely accessible to those with limited technological capabilities.

We used GSV to replicate an existing field survey for the sake of comparison, but practical applications could include updating outdated street tree inventories and generating new street tree data. In such cases, groundtruthing should be performed to assess the quality of tree data generated using GSV. We worked in lower density urban areas, so we do not know how well our approach would work in denser urban centers where street trees may be obscured by traffic or infrastructure in GSV imagery. This analysis was conducted by a single individual who has a background in field botany. As such, we cannot estimate inter-operator consistency or describe how reliably this approach could be implemented by volunteers or other non-experts. It is possible that our approach could be used by an urban forester to validate field data collected by volunteers.

Virtual street tree surveys in GSV are appealing for several reasons. They can be done with less time, labor, money, and safety risks than field surveys. Virtual surveys can be conducted year-round regardless of season or weather conditions. As opposed to other geospatial technologies used in urban forestry, GSV is freely available and easy to use. In our comparison of virtual survey data to field-based data, we found high levels of agreement in tree counts and genus identification, and identification was easiest for large trees. Identifying trees to the species level was more difficult. The analyst's ability to manually estimate dbh improved markedly over time, demonstrating a learning curve that could be eliminated in the future by implementing image measurement software. In light of these results, GSV shows strong potential for complementing field data collection as a means of generating information about street trees.

Acknowledgement

We thank Dustin Herrmann and two reviewers for constructive comments that helped improve the manuscript. This work was supported by the Department of Geography at Ball State University. Field data referenced in this study were collected while AB held a National Research Council Research Associateship Award at the US EPA.

References

Alonzo, M., Bookhagen, B., Roberts, D.A., 2014. Urban tree species mapping using hyperspectral and lidar data fusion. Remote Sens. Environ. 148, 70–83.
Alonzo, M., McFadden, J.P., Nowak, D.J., Roberts, D.A., 2016. Mapping urban forest structure and function using hyperspectral imagery and lidar data. Urban For. Urban Green. 17, 135–147.
Anguelov, D., Dulong, C., Filip, D., Frueh, C., Lafon, S., Lyon, R., Ogale, A., Vincent, L., Weaver, J., 2010. Google Street View: capturing the world at street level. Computer 6, 32–38.
Badland, H.M., Opit, S., Witten, K., Kearns, R.A., Mavoa, S., 2010. Can virtual streetscape audits reliably replace physical streetscape audits? J. Urban Health 87, 1007–1016.
Berland, A., Hopton, M.E., 2014. Comparing street tree assemblages and associated stormwater benefits among communities in metropolitan Cincinnati, Ohio, USA. Urban For. Urban Green. 13, 734–741.
Berland, A., Schwarz, K., Herrmann, D.L., Hopton, M.E., 2015. How environmental justice patterns are shaped by place: terrain and tree canopy in Cincinnati, Ohio, USA. Cities Environ. (CATE) 8, Article 1.
US Census Bureau, 2016. American FactFinder, http://factfinder.census.gov/ (Accessed August 2016).
Clarke, P., Ailshire, J., Melendez, R., Bader, M., Morenoff, J., 2010. Using Google Earth to conduct a neighborhood audit: reliability of a virtual audit instrument. Health Place 16, 1224–1229.
ESRI, 2014. ArcGIS for Desktop Version 10.3. Redlands, CA.
i-Tree, 2016. i-Tree Streets, http://www.itreetools.org/streets/index.php (Accessed August 2016).
Landry, S.M., Chakraborty, J., 2009. Street trees and equity: evaluating the spatial distribution of an urban amenity. Environ. Plann. A 41, 2651–2670.
Li, X., Zhang, C., Li, W., Kuzovkina, Y.A., Weiner, D., 2015a. Who lives in greener neighborhoods? The distribution of street greenery and its association with residents' socioeconomic conditions in Hartford, Connecticut, USA. Urban For. Urban Green. 14, 751–759.
Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., Zhang, W., 2015b. Assessing street-level urban greenery using Google Street View and a modified green view index. Urban For. Urban Green. 14, 675–685.
McGee III, J.A., Day, S.D., Wynne, R.H., White, M.B., 2012. Using geospatial tools to assess the urban tree canopy: decision support for local governments. J. For. 110, 275–286.
McPherson, E.G., van Doorn, N., de Goede, J., 2016. Structure, function and value of street trees in California, USA. Urban For. Urban Green. 17, 104–115.
Mullaney, J., Lucke, T., Trueman, S.J., 2015. A review of benefits and challenges in growing street trees in paved urban environments. Landsc. Urban Plann. 134, 157–166.
O'Neil-Dunne, J., MacFaden, S., Royar, A., 2014. A versatile, production-oriented approach to high-resolution tree-canopy mapping in urban and suburban landscapes using GEOBIA and data fusion. Remote Sens. 6, 12837–12865.
Pedlar, J.H., McKenney, D.W., Allen, D., Lawrence, K., Lawrence, G., Campbell, K., 2013. A street tree survey for Canadian communities: protocol and early results. For. Chron. 89, 753–758.
R Core Team, 2015. R: A Language and Environment for Statistical Computing, https://www.R-project.org (Accessed August 2016).
Rousselet, J., Imbert, C.-E., Dekri, A., Garcia, J., Goussard, F., Vincent, B., Denux, O., Robinet, C., Dorkeld, F., Roques, A., Rossi, J.-P., 2013. Assessing species distribution using Google Street View: a pilot study with the pine processionary moth. PLoS One 8, e74918.
San Souci, J., Hanou, I., Puchalski, D., 2009. High-resolution remote sensing image analysis for early detection and response planning for emerald ash borer. Photogramm. Eng. Remote Sens. 75, 905–909.
Ward, K.T., Johnson, G.R., 2007. Geospatial methods provide timely and comprehensive urban forest information. Urban For. Urban Green. 6, 15–22.