Research Paper

Greening in style: Urban form, architecture and the structure of front and backyard vegetation

Alessandro Ossola, Dexter Locke, Brenda Lin, Emily Minor

ABSTRACT

Residential yards comprise most land and green space across cities. Despite yards being ubiquitous, little comprehensive information exists on how vegetation varies between front and backyards. This hinders our ability to optimize greening interventions on private urban land.

We devised an accurate GIS algorithm to locate and classify front and backyards within residential landscapes. By applying this method to the greater Boston area, we measured vegetation structure (i.e., canopy cover, height and volume) of front and backyards with LiDAR and multispectral imagery. We further investigated relationships between urban form, architectural style, socio-economics, and the structure of front and backyard vegetation across Boston’s residential landscapes.

Among the 85,732 residential parcels that were not corner lots and had cadastral and architectural data available, backyards were twice as large as front yards on average and had significantly greater canopy cover, vegetation volume and taller trees. Parcel-level characteristics, including vegetation in the corresponding front or backyard, as well as morphological characteristics of parcels, were the best predictors of vegetation structure. House architectural style was related to vegetation structure. The neighborhood socio-economic characteristics were the least important factors in predicting yard vegetation structure.

Our study highlights that urban greening in yards depends on urban form and morphology at the parcel scale, and as such, it could be enhanced through urban to provide opportunities for additional vegetation. Architectural style might represent a further filter by which residents manage vegetation in their home environment, making it possible to devise strategies to green our cities – in style.

1. Introduction

Residential yards, the outdoor areas within residential parcels (Larson, Hoffman, & Riplinger, 2017), often contain most of the green space, vegetation, and trees in cities (Haase, Jänicke, & Wellmann, 2019; Lin, Meyers, & Barnett, 2015; Loram, Tratalos, Warren, & Gaston, 2007). Yards have become important refuges for people, plants, and wildlife within dense built landscapes and significant areas for the provision of ecosystem services (Cameron et al., 2012; Haaland & Konijnendijk, 2015; Lin et al., 2017).

Urban planning delineates urban form and morphology – the physical layout of built-up areas – by allocating physical space between front and backyards based on the relative position of a house within a parcel and its street access. Discriminating between front and backyards is not trivial, as residents’ yard use and aesthetic preferences often differ between these two spaces (Iles, 2008); the front yard represents a “visible symbol of self”, whereas the backyard a “personal pleasure ground” (Larsen & Harlan, 2006). Despite the individual dimension of residential yards, the social and cultural norms of a community, and even its regulations and informal rules, play an important role in affecting individual design preferences and attitudes toward residential landscaping (Locke, Roy Chowdhury, et al., 2018; Martin, Peterson, & Stabler, 2003; Nassauer, Wang, & Dayrell, 2009). This is particularly evident in front yards, the “more public” private spaces, where residents may feel more compelled to conform their landscaping to maintain an appropriate public image and care for the local community (Nassauer, 2011; Zmyslony & Gagnon, 1998). Often, front yard preferences align with the socio-cultural background of a community and change predictably across neighborhoods based on their socio-economic characteristics (Larsen & Harlan, 2006). Design and landscaping preferences of front and backyards are enacted through different management practices of these two urban spaces (Larsen & Harlan, 2006; Larson, Casagrande, Harlan, & Yabiku, 2009; Nassauer et al., 2014). Yard management mostly affects vegetation. It is perhaps not surprising that previous field studies found significant differences in vegetation characteristics in front and backyards within the same parcel (Dorney, Guntenspergen, Keough, & Stearns, 1984; Richards, Mallette, Simpson, & Macie, 1984; van Heezik, Freeman, Porter, & Dickinson, 2013).

To further complicate this picture, architectural style and housing type also define the character of a residential landscape, delineating
neighborhood boundaries and the physical geography of the city (Fusch & Ford, 1983). The architectural style of a house primarily engenders the symbol of a resident's social and economic status, and its desirability for others (Cherubin & Wilderman, 1986; Nasar, 1989). Lawns, gardens, and trees are often an integral part of house design and style, and as such, residents landscape their yard to highlight value and prestige of buildings and even discourage vandalism (Foster, Giles-Corti, & Knuiman, 2011; Harris & Brown, 1996). Yard vegetation defines the architectural space, and it has been linked to the ability of individuals to recall a building among many others through cognitive mapping (Evans, Smith, & Pezdek, 1982; Smardon, 1988). Thus, while it is reasonable to expect the house architectural style to be related to yard vegetation, limited support for this theory has been found to date (Nassauer et al., 2009).

Overall, most of the current empirical evidence about yard vegetation structural characteristics comes from a small number of field studies, performed in a limited number of parcels and neighborhoods. Furthermore, most vegetation studies of residential yards have been limited to front yards because of the restricted access to the back of the house (Threlfall et al., 2016; Zmyslony & Gagnon, 1998). Current limitations have made it difficult to quantify the full extent of urban greening (especially the design, management, and structure of front and backyards), the multi-scale factors that underpin urban greening, and how these translate into ecological functions and services (Cook, Hall, & Larson, 2012; Lin et al., 2017; Locke, Avolio, et al., 2018). Ultimately, there is a clear need to study residential yards at a resolution that matches individual management and design choices, which means differentiating between front and backyards and clarifying factors affecting vegetation at the parcel and yard scale (Locke, Avolio, et al., 2018; Locke, Roy Chowdhury, et al., 2018). This will allow urban planners to design urban space functional for people and biodiversity (Belaire, Westphal, & Minor, 2016; Beninde, Veith, & Hochkirch, 2015).

A better understanding of how urban vegetation structure changes across urban landscapes is needed, as this structure dictates ecological functioning (Ossola, Hahs, & Livesley, 2015; Ossola, Hahs, Nash, & Livesley, 2016), and benefits for people and biodiversity (Goddard, Dougill, & Benton, 2013; Ossola, Nash, Christie, Hahs, & Livesley, 2015; Smith, Warren, Thompson, & Gaston, 2006). In this study, we address this need by devising a new method to locate and classify front and backyards across residential neighborhoods within the greater Boston area. We further measure how vegetation structure varies across non-corner residential yards (n = 85,732 parcels, n = 171,464 yards) by addressing the following questions and hypotheses:

1. Are there differences in vegetation structure between front and backyards? We first hypothesize backyards to have greater vegetation structure than front yards (i.e., higher canopy cover, height and volume), as residents may prefer front yards with extensive lawns and simple vegetation.

2. How do differences in front and backyard vegetation relate to urban form, architectural style and neighborhood socio-economic characteristics? Here we test a second hypothesis that the house architectural style is related to the structure of yard vegetation, as it might affect residents’ planting and landscaping decisions.

We then discuss the implications of our results for urban planning and greening interventions.

2. Methods

2.1. Study area

The study area comprises the city of Boston and 41 adjacent municipalities in the greater Boston area (land area: 501 km²; population: 1,569,000 inhabitants). Founded in 1630, Boston is one of the oldest US cities. From 1880 to 1920, Boston experienced significant growth and expansion from an annexation of towns. In the early 20th century, professionals asserted the ideals of city planning in the urban landscape of Boston (Kennedy, 1992). High levels of immigration in the early 20th century led to an increased urban densification in the city center, with many native Bostonians leaving for the suburbs. The modern era of Boston’s planning and development history since the early 1980s has been characterized by revitalization to deal with changes in population growth, economic prosperity, and social problems. To bring back order and growth to the ailing city from the massive suburbanization, a vision for a “New Boston” was developed (O’Connor, 2001). As a result, Boston has one of greatest variety of architectural styles in the US. Compared to many other US cities, the greater Boston area has higher residential cover (47.33% of land) and urban vegetation cover (35.38%) (Ossola & Hopton, 2018a). However, attempts to further increase Boston’s vegetation cover, particularly in underserved areas downtown, had little success due to poor governance and limited funding (Foo, McCarthy, & Bebbington, 2018). Boston has a humid continental climate with cold winters, and hot and wet summers (mean annual temperature = 9.6 °C; mean annual precipitation = 1233 mm) (PRISM Climate Group, 2015).

2.2. Open data sources

We obtained residential parcel polygons and cadastral data, building footprints, and road centerlines from the open data portal of the Commonwealth of Massachusetts (2017) for 41 municipalities and from data portal of the City of Boston (2017). Residential parcels containing 4 or more residential buildings, multiple units, flats, deckers, apartment blocks, condominiums and mixed-use parcels were excluded from analyses as they lack defined front and backyards and may have multiple land managers. We obtained LiDAR data for the year 2014 from the US Geological Survey (“MA Post-Sandy CMG 2013-14” point cloud dataset, nominal point spacing 0.7 m, vertical accuracy 0.05 m, horizontal accuracy 0.35 m). High-resolution imagery (1 m) in the visible and near-infrared spectra for the year 2014 was obtained from the National Agriculture Imagery Program (NAIP) led by the US Department of Agriculture (USDA). We obtained socio-economic variables (n = 37) for 193 census tracts in the study area from the Census Bureau (2010). These variables are routinely used in studies on urban canopy cover and structure (Biggby, McHale, & Hess, 2014; Locke, Landry, Grove, & Roy Chowdhury, 2016; Ossola & Hopton, 2018a, 2018b) and relate to neighborhood demographics, urban historical development (e.g., number of buildings developed per decade), housing, income, education, employment and social inequality (Appendix A).

2.3. Geospatial analyses

Front and backyards, or corner yards, for each residential parcel were geolocated and classified by using the following algorithm in ArcGIS Desktop 10.5 (ESRI, Redlands, CA). The centroid of the largest building in each residential parcel was identified. We assumed the largest structure was the main residential building, as opposed to other non-residential structures (e.g., garage, garden shed). For each of those centroids, a line perpendicular to the street centerline was drawn. Then a second perpendicular line to the newly created line was drawn, passing through the building centroid, and extending to the parcel’s lateral boundaries. The second line was used to split each parcel polygon into front and backyards. All building footprints within a parcel were further subtracted to define the geometry of yards. The yard strips between the lateral sides of the main residential building and the parcel lateral sides were split evenly between paired front and backyard. Residential parcels located within 15 m from road intersections were classified as corner parcels. As these parcels lay at the intersection between two or more streets, a clear distinction between front and backyards in these parcels is not possible (Fig. 1). Thus, corner parcels and yards therein (n = 44,336) were excluded from the study (Fig. 1, Appendix E).

Accuracy of the yard classification algorithm was calculated by
visually interpreting randomly-selected yards (n = 2000) and calculating the percentage of yards correctly geolocated and classified as front and backyards. For each parcel, total building area, parcel area, front and backyard area, building-street offset, distance from downtown, and geographic coordinates were calculated in ArcGIS Desktop 10.5 (ESRI, Redlands, CA).

Parcel-level data on the household type (i.e., 1, 2 or 3 families), architectural style, real estate value, and year of construction of the main residential building were extracted from cadastral datasets for each municipality in the study area (City of Boston, 2017; Commonwealth of Massachusetts, 2017). The 12 most common architectural style classes found across the Boston area, as classified from the cadaster, were considered in the study (Fig. 2). Complete cadastral data were not available for all residential buildings in the study area. After constraining our analyses to 1-, 2-, and 3-family homes not located on corner lots, 171,464 yards in 85,732 parcels were statistically analyzed (Appendix B).

We calculated four vegetation structure metrics for each front and back yard: (i) canopy cover (% yard cover), (ii) average canopy height (m), (iii) maximum canopy height (m), and (iv) projected volume of residential forest (m³ vegetation/m² yard area). Methods used are described and validated in previous studies (Ossola & Hopton, 2018a, 2018b). Vegetation height was interpolated from the LiDAR point cloud at 1.5 m horizontal resolution in ArcGIS Desktop 10.5 (ESRI, Redlands,
CA). Canopy cover was modelled at 1.5 m resolution by using a supervised classification approach based on data fusion of the vegetation height map, the visible and near-infrared imagery and Normalized Difference Vegetation Index (NDVI) map calculated from the same imagery. A minimum of 100,000 pixels for each of three cover classes (i.e., woody vegetation, herbaceous vegetation, non-vegetated) were manually classified to calculate a signature file used by the maximum likelihood supervised classifier that produced the canopy cover map (Singh, Vogler, Shoemaker, & Meentemeyer, 2012). The volume of residential forest was calculated by multiplying canopy cover and height, assuming this volume to be entirely occupied by vegetation. This modelling approach has been used to measure residential vegetation structure across entire urban landscapes and multiple cities (Ossola & Hopton, 2018a, 2018b), though it does not fully discriminate the ownership of vegetation and trees along property lines (e.g., trees overhanging from nearby parcels, rights-of-ways and easements). However, as Massachusetts law recognizes the right of residents to cut vegetation overhanging from nearby properties (e.g., Macero v. Busconi Corp., Civil law case N. 99-03577E (Middlesex Super.Ct.), 12 Mass. L. Rep. 521 (2000)), we assumed the management of vegetation at the property line to be independent from vegetation ownership.

2.4. Statistical analyses

For each of the four vegetation structure metrics described above, analyses were carried out in two phases. In the first phase, comparisons of vegetation structural metrics in relation to yard position (i.e., front and back) and categorical independent predictors (e.g., number of families) were performed by using pairwise Wilcoxon rank sum tests in R 3.4.1 (R Core Team, 2017). Average values are expressed with their standard error of the mean (SEM).

In the second phase, Distributed Random Forest models (DRFs) were used to predict front and backyard vegetation metrics. DRFs is a regression technique whereby multiple regression trees are built on sub-sets of observations and predictors to allow a final prediction based on all regression trees generated (i.e., random forest). The random selection of predictors in random forest modelling reduces bias and data overfit, and it has been widely used in ecological prediction (Prasad, Iverson, & Liaw, 2006). This modelling approach has the following attributes: (i) it handles multicollinear variables, (ii) it handles predictor variables regardless of their distribution, (iii) it is suitable for modeling large and high dimensional datasets, and (iv) it handles variables with high number of categories (e.g., 12 architectural styles). DRF modelling was performed in H2O, as this artificial intelligence analytical platform can efficiently run complex cloud-based machine learning DRF algorithms (H2O.ai, 2018, version 3.20.0.2) and be interfaced through R 3.4.1 (R Core Team, 2017). The yard vegetation dataset was randomly split into training, validation, and test datasets containing, 80%, 10% and 10% of observations, respectively, as these sample sizes are routinely used in DRF models (H2O.ai). DRF models were created by fitting a maximum of 200 statistical trees (max depth = 30) for each of 5-fold cross validations and by calculating a final DRF model for each vegetation metric in both front and backyards. The predictor variables for each DRF model fit into one of three categories: architectural style, yard position (i.e., front and back), and categorical independent predictors (e.g., number of families)
categories: (i) parcel-level vegetation metrics for the opposite yard within each parcel (i.e., back and front), (ii) parcel-level morphological variables related to building and urban form (e.g., yard area, house footprint area, house-street offset, parcel area, year construction of house, house value, distance from downtown and number of families), and (iii) socio-economic variables at census tract level (e.g., income, education, employment, household characteristics, Appendix A). Model fit was assessed by using standard machine learning metrics as Mean Squared Error (MSE), Root Mean Squared Logarithmic Error (RMSLE) and residual deviance in relation to number of DRF trees and leaves fitted (H2O.ai, 2018).

3. Results

3.1. Vegetation structure in front and backyards

The accuracy of the GIS algorithm devised to geolocate and identify front and backyards exceeded 98.35%. Overall, backyards were almost twice as large as front yards (423.27 ± 0.006 m² and 232.02 ± 0.003 m², respectively). Single family homes had significantly larger yards (256.06 ± 1.01 m² and 465.77 ± 1.97 m² for front and backyards, respectively) compared to two-family (140.78 ± 0.84 m² and 261.70 ± 1.76 m²) and three-family homes (125.28 ± 2.24 m² and 236.04 ± 4.07 m²) (Fig. 3A). Estates and mansions had yards seven times larger than conventional houses (Fig. 4A). On average, the proportional allocation between front and backyards was similar among households with different number of families and ranged between 63.32 and 63.97% for backyards, and 36.02 and 36.67% for front yards (Fig. 3B). Similarly, the average proportional allocation between yards in relation to building architectural styles was limited to a 7.75% variation range (Fig. 4B). Of the 85,732 properties considered, the average distance between the residential building and the street centerline was 19.06 ± 0.02 m. Bungalows had the lowest average house to street centerline distance (17.98 ± 0.08 m) and estate/mansion properties had the highest (28.67 ± 0.75 m).

Across the study area, backyards had significantly greater canopy cover than front yards (64.03 ± 0.09% of yard area and 43.37 ± 0.11%, respectively), as well as significantly greater average canopy height (6.66 ± 0.02 m; 4.07 ± 0.01 m), maximum canopy height (15.98 ± 0.02 m; 10.28 ± 0.03 m) and vegetation volume (4.95 ± 0.01 m³/m²; 2.51 ± 0.01 m³/m²). Vegetation metrics in front and backyards were significantly related to the number of families (Fig. 3C-G), as well as building architectural style (Fig. 5, Appendix C).

3.2. Urban form, style and social effects on yard vegetation

Distributed Random Forest models (DRFs) predicting front and backyard vegetation variables were able to explain about half of the initial deviance, on average (Table 1). Canopy cover, average

Fig. 3. Relationship between household type and average yard area (A), percent yard area (B), percent canopy cover (C), average canopy height (D), maximum canopy height (E), and average vegetation volume (F) in front and backyards. Lower case letters above each boxplot represents statistically equal mean values for front yards (x, y, z) and backyards (a, b, c) based on pairwise Wilcoxon tests.
vegetation height, and vegetation volume were better explained by vegetation characteristics of the opposite yard within each parcel (Fig. 6, Appendix D), whereas maximum canopy height was better explained by parcel and yard size (Fig. 6). Overall, vegetation metrics of the opposite yard within each parcel and parcel-level variables related to urban form and architecture were more important in predicting yard vegetation structure compared to socio-economic variables (Fig. 6, Appendix D).

4. Discussion

4.1. Vegetation structure in front and backyards

In our analysis of over 85,000 parcels in the Boston metropolitan area, we observed significant differences in the vegetation in front and backyards. As hypothesized, backyards had significantly greater horizontal and vertical vegetation structure than front yards. This result held true for single and multi-family houses, as well as for houses with different architectural styles. These results appear consistent with previous evidence suggesting that residents prefer manicured front lawns over more complex woody vegetation, shrubs and large trees (Belaire et al., 2016; Feagan & Ripmeester, 2001).

In Boston, the lower percent canopy cover in front yards, as compared to backyards, might be partially explained by the lower absolute area available for planting woody vegetation with greater cover, like broad canopy trees. Tree cover, in fact, appears to increase with yard size, particularly in enclosed backyards, as noticed in exurban Michigan (Nassauer et al., 2014). Further, front yards are often “staged” as open landscapes with the house in the background representing a symbol of wealth, social status and power (Cherulnik & Wilderman, 1986; Nasar, 1989). In this context, lawns in the front of the house are often designed and landscaped to highlight, and not hide, the building, its architecture and value (Evans et al., 1982). Front yards engender personal pride, social relatedness, and connection to the neighboring community (Harris & Brown, 1996; Quayle & van der Lieck, 1997). In this way, landscaping manicured lawns and simple vegetation in front yards can facilitate property upkeep, tidiness, order and conformity, demonstrating so called “cues to care” (Nassauer, 1988, 1995) that residents place toward their community (Foster et al., 2011; Hess, 2008; Nassauer, 2011). More complex front yard vegetation might have been preferred by residents valuing privacy and protection from noise or pollution (Smardon, 1988; Zagorski, Kirkpatrick, & Stratford, 2004). On the other hand, landscaping preference in backyards is “more likely to follow individual fantasy” (Larsen & Harlan, 2006), thus allowing the creation of more complex vegetation structure in backyards (Belaire et al., 2016).

Boston’s backyard trees were taller than those in front yards. Maximum tree canopy height in both front and backyards was better predicted by the yard size, rather than the vegetation characteristics of the opposite yard within each parcel (i.e., back and front). A study of 107 yards in 10 suburbs in Hobart, Australia found no differences in tree canopy height between front and backyards, but also no differences in yard size (Daniels & Kirkpatrick, 2006). This suggests that residents of the Boston area may prefer planting and managing taller trees in larger yard spaces and further away from perceived risks (e.g., tree falling on the house or street) (Dilley & Wolf, 2013).

4.2. Urban form, style and social effects on yard vegetation

Except for maximum tree canopy height, vegetation characteristics of the opposite yard within a parcel were the best predictors for the vegetation metrics investigated. In other words, parcels vary in their canopy cover, even though there are differences between yards. This suggests that, at a city scale, resident landscaping preference and vegetation management are best predicted at the level of the individual parcel. Residents with a relatively high (or low) amount of vegetation in their front yard are more likely to have a relatively high (or low) amount of vegetation in their backyard. This contrasts with what found in smaller and more localized vegetation surveys of landscaping preferences in Phoenix, AZ (Larsen and Harlan (2006) and garden styles between front and backyards in Hobart, Australia (Daniels and Kirkpatrick (2006), where significant differences in vegetation...
characteristics were observed in about half of the yards investigated (232 and 107 parcels, respectively). In this way, the effects of resident landscaping preference and vegetation management of front and backyards could well vary from property to property at the parcel scale, to be then increasingly homogenized at larger landscape and city scales (Cook et al., 2012).

Parcel-level metrics related to urban form and architectural style ranked as the second most important class of predictors of the overall yard vegetation structure. This is not surprising as urban form and physical space availability are significant predictors of vegetation cover at landscape, city and urban macro-scales (Biggsby et al., 2014; Nassauer et al., 2014; Ossola & Hopton, 2018a, 2018b; Troy, Grove, O’Neil-Dunne, Pickett, & Cadenasso, 2007). The shape and size of yards have been found to be inextricably linked to the history and types of urban development, which could in turn affect the availability of space for vegetation (Conway & Hackworth, 2007; Gill et al., 2008; Lin et al., 2017; Smith et al., 2006). In this way, urban form and available space could well be limiting our ability to achieve higher canopy cover and taller vegetation and trees, particularly as human perception of vegetation is highly related to density, space and field of view of a landscape scene (Smardon, 1988; Ulrich, 1986).

The year of house construction was a better predictor of yard vegetation structure than the number of buildings built within each decade in the neighborhood. Similarly, the type of household (i.e., 1, 2, 3 families) recorded at parcel-level was a better predictor than household variables recorded at neighborhood level. Overall, this suggests that more accurate accounts of urban vegetation and trees would benefit from detailed assessments at the yard level, thus allowing a better evaluation of the multiscale effects and interactions occurring across residential landscapes (Cook et al., 2012; Harris et al., 2012). It also suggests that even with some neighborhood influences, there is a clear need to measure urban vegetation at the scale of management, which is the parcel or sub-parcel area. Although prior researchers have conceptualized residential land management as multi-scalar, with homes nested in neighborhoods, and neighborhoods within municipalities (Chowdhury et al., 2011), our approach was able to quantify the relative importance of each scale across our study area.

Our second hypothesis, that the house architectural style is related to the structure of yard vegetation, was confirmed. While different Boston neighborhoods have different building style compositions, as observed in other US cities (Bastian, 1980), vegetation structure varied significantly across both front and backyards and by architectural style. Mansions and estates hosted the tallest and most prominent trees as these are likely preferred in prestigious residential scenes (Ulrich,
Table 1
Summary statistics for the training and testing of the Distributed Random Forest (DRF) models predicting yard vegetation metrics.

<table>
<thead>
<tr>
<th>Vegetation metric predicted</th>
<th>Yard type</th>
<th>Number of DRF trees</th>
<th>MSE (training)</th>
<th>RMSE (training)</th>
<th>Mean residual deviance (training)</th>
<th>MSE (testing)</th>
<th>RMSE (testing)</th>
<th>Mean residual deviance (testing)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canopy cover</td>
<td>Backyard</td>
<td>109</td>
<td>0.043</td>
<td>0.044</td>
<td>0.043</td>
<td>0.044</td>
<td>0.044</td>
<td>0.044</td>
</tr>
<tr>
<td>Canopy cover</td>
<td>Front yard</td>
<td>87</td>
<td>0.048</td>
<td>0.048</td>
<td>0.048</td>
<td>0.048</td>
<td>0.048</td>
<td>0.048</td>
</tr>
<tr>
<td>Maximum canopy height</td>
<td>Backyard</td>
<td>100</td>
<td>22.683</td>
<td>25.482</td>
<td>25.092</td>
<td>5.092</td>
<td>5.048</td>
<td>5.009</td>
</tr>
<tr>
<td>Average canopy height</td>
<td>Backyard</td>
<td>102</td>
<td>11.332</td>
<td>11.332</td>
<td>11.332</td>
<td>11.185</td>
<td>11.185</td>
<td>11.185</td>
</tr>
<tr>
<td>Forest volume</td>
<td>Front yard</td>
<td>92</td>
<td>5.680</td>
<td>5.680</td>
<td>5.680</td>
<td>5.680</td>
<td>5.680</td>
<td>5.680</td>
</tr>
</tbody>
</table>

Future efforts could further test whether architectural style can be used as a robust proxy to downscale census socio-economic data from neighborhood-level to parcel- and yard-level. In this regard, neighborhood-level socio-economic characteristics were the least important variables in predicting front and backyard vegetation structure. This finding adds new evidence to the debate about the importance of socio-economic characteristics to yard vegetation. On one hand, studies in several cities have found characteristics such as age, income, or race to be important predictors of yard vegetation in general (Belaire et al., 2016; Minor, Belaire, Davis, Franco, & Lin, 2016) and the diversity of flowering plants in particular (Lowenstein & Minor, 2016). On the other hand, income and socio-demographic composition was not related to urban and residential tree cover (Duncan et al., 2014; Ossola & Hopton, 2018a, 2018b; Pham, Apparicio, Landry, Séguin, & Gagnon, 2013) or yard landscaping preference (Larsen & Harlan, 2006) in other cities. Similarly, in our study, median household income from the census was not a strong predictor for most measures of vegetation structure, and was similar in importance to real estate value at parcel level.

4.3. Where are all the yards?

The algorithm devised to locate and classify yards achieved very high accuracy. Yard classification was consistent across diverse neighborhoods characterized by heterogeneous urban form, planning schemes and street networks (e.g., regular vs sinuous) (Fig. 1). The few misclassifications of yards were plausibly attributable to imprecisions in the original street geospatial layer or to complex parcel morphologies (Fig. 1, Appendix E). A previous attempt to locate private gardens across Dunedin, New Zealand, based on vegetation classification of satellite imagery, achieved lower accuracy (90.7%) than our algorithm (Mathieu, Freeman, & Aryal, 2007). In a recent study, Haase et al. (2019) managed to quantify vegetation cover across residential yards in Leipzig, Germany, from satellite imagery but with no distinction between front and backyards. By making use of freely available cadastral and urban form data (i.e., parcel, building and street geometry), our algorithm allows researchers to more precisely locate and classify all yards across entire urban landscapes, regardless of vegetation composition or other yard attributes not related to urban morphology and planning. Since the advent of high-resolution, high-accuracy tree canopy mapping in urban areas, parcel-scale studies have become more popular (Ossola & Hopton, 2018b). Thus, the GIS algorithm presented here offers a new tool to inform greening efforts on private residential land. Further, this tool allows the scale of research on urban greening to better match the scale of decision-making and management: the parcel and its yards. Ultimately, this methodological approach allows researchers to (i) expand beyond the limitations of physical access to properties and small to moderate field-based studies of yards, (ii) examine variable importance across scales, and (iii) investigate how yard...
Fig. 6. Importance of variables predicting yard canopy cover (A, B) and maximum canopy height (C, D) calculated from DRF modelling. Parcel-level vegetation variables are highlighted in green, parcel-level morphological variables related to urban form in grey and census tract-level variables in black. Only the 30 most important predictors are plotted. Graphs for other vegetation metrics are reported in Appendix D. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
and sub-parcel level characteristics may moderate the front/back differences in vegetation structure.

4.4. Implications for urban planning and greening

Our study highlights that urban greening in private yards is related to urban form and architecture, and as such, it could be enhanced through urban planning. This is important as modern residents express the will to plant less trees in the future (Dilley & Wolf, 2013), unnecessarily remove trees (Ossola & Hopton, 2018b), and even completely pave their yards (Perry & Nawaz, 2008). We argue that future urban planning could devise better urban forms, parcel and yard morphologies able to influence residents’ landscaping behaviors, thus enhancing vegetation structure and its numerous ecosystem services (e.g., biodiversity, urban heat mitigation, human health and wellbeing). Planning residential parcels with larger front and backyards would be the most logical way to increase vegetation structure, but this could exacerbate urban sprawl and its negative impacts on urban
sustainability. Additional evidence is however needed to clarify whether the surrounding land use (e.g., residential parcels facing a peri-urban forest vs parcels in a dense urban environment) and the climate in which a city is located might affect the patterns observed in our study. Finally, greening increments related to parcel size will not be unlimited, as suggested by a study of 1.4 million residential parcels in 1503 neighborhoods spread across nine US cities, whereby increases in residential vegetation cover would only be substantial for parcel areas up to 1000 m² (Ossola & Hopton, 2018b).

Where planning schemes do not allow larger residential parcels, significant gains in vegetation structure could be obtained by reducing building-street offsets and designing smaller front yards and larger backyards. Interestingly, regardless of parcel size, household occupancy, or architectural style, the proportional allocation between front and backyards was similar across Boston. A different yard proportional allocation could increase residential greening during urban renewal and land redevelopement and protect vegetation and trees during urban densification (Cheng, Ryan, Warren, & Nicolson, 2017; Lin & Fuller, 2013). Striking a balance between optimal urban form and private greening would also need to carefully consider trade-offs related to visual and physical access to green spaces and resource inequality (Danford et al., 2014). Innovative urban planning schemes could foster the individual and social creation of urban greening from residential land through neighborhood approaches geared toward shared yard governance and management (Dewaelheyns, Kerselaers, & Rogge, 2016; Ossola et al., 2018; Steenberg, Duinker, & Charles, 2013). Despite being uncommon throughout Boston (Harris et al., 2012), residents’ associations could also be leveraged to increment yard vegetation structure through more structurally elaborate gardening styles and landscaping rules (Harris et al., 2012). Urban planning could further increase urban greening by incorporating marketing strategies to influence human behavior, the human need for style, and individual preferences towards more complex yard vegetation and landscaping types (Grove et al., 2006).

In a study on the geography of architectural styles in San Diego, CA and Columbus, OH, Fusch and Ford (1983) stated: “By mapping and monitoring, [sic] urban house types we can gain a better understanding of the role of the three-dimensional landscape in shaping the morphology and socio-economic structure of cities”. Thirty-five years later, our study not only confirms this vision, but it sets a novel way to holistically understand how people interact with the vegetation and trees in their home environment and the larger residential landscape.

Acknowledgements

Authors kindly acknowledge USDA, USGS, MASS-GIS and the City of Boston for providing geospatial datasets. This work was supported by the National Socio-Environmental Synthesis Center (SESYNC) under funding received from the National Science Foundation DBI-1639145.

Appendix A: Socio-economic variables at census tract level (US Census Bureau, 2010)

<table>
<thead>
<tr>
<th>Description</th>
<th>Socio-economic variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median age (both sexes)</td>
<td>DP0020001</td>
</tr>
<tr>
<td>White alone or in combination with one or more other races</td>
<td>DP0090001</td>
</tr>
<tr>
<td>Black or African American alone or in combination with one or more other races</td>
<td>DP0090002</td>
</tr>
<tr>
<td>American Indian and Alaska Native alone or in combination with one or more other races</td>
<td>DP0090003</td>
</tr>
<tr>
<td>Asian alone or in combination with one or more other races</td>
<td>DP0090004</td>
</tr>
<tr>
<td>Native Hawaiian and Other Pacific Islander alone or in combination with one or more other races</td>
<td>DP0090005</td>
</tr>
<tr>
<td>Some Other Race alone or in combination with one or more other races</td>
<td>DP0090006</td>
</tr>
<tr>
<td>Population in households</td>
<td>DP0120000</td>
</tr>
<tr>
<td>Population in group quarters</td>
<td>DP0120014</td>
</tr>
<tr>
<td>Family households</td>
<td>DP0130002</td>
</tr>
<tr>
<td>Nonfamily households</td>
<td>DP0130010</td>
</tr>
<tr>
<td>Average household size</td>
<td>DP0160001</td>
</tr>
<tr>
<td>Average family size</td>
<td>DP0170001</td>
</tr>
<tr>
<td>Occupied housing units</td>
<td>DP0180002</td>
</tr>
<tr>
<td>Vacant housing units</td>
<td>DP0180003</td>
</tr>
<tr>
<td>Owner-occupied housing units</td>
<td>DP0210002</td>
</tr>
<tr>
<td>Renter-occupied housing units</td>
<td>DP0210003</td>
</tr>
<tr>
<td>Percent high school graduate or higher (table S1501)</td>
<td>HC01_EST_VC16</td>
</tr>
<tr>
<td>Percent bachelor's degree or higher (table S1501)</td>
<td>HC01_EST_VC17</td>
</tr>
<tr>
<td>All families (table S1702)</td>
<td>HC01_EST_VC01</td>
</tr>
<tr>
<td>Percent families below poverty level (table S1702)</td>
<td>HC02_EST_VC01</td>
</tr>
<tr>
<td>Gini Index (table B19083)</td>
<td>HD01_VD01</td>
</tr>
<tr>
<td>Median income (dollars) per household (table S1903)</td>
<td>HC02_EST_VC02</td>
</tr>
<tr>
<td>Working-age population 20–64 years (table S2301)</td>
<td>HC01_EST_VC24</td>
</tr>
<tr>
<td>In labor population 20–64 years (table S2301)</td>
<td>HC02_EST_VC24</td>
</tr>
<tr>
<td>Employed population 20–64 years (table S2301)</td>
<td>HC03_EST_VC24</td>
</tr>
<tr>
<td>Unemployment rate 20–64 years (table S2301)</td>
<td>HC04_EST_VC24</td>
</tr>
<tr>
<td>Median contract rent (table B25058)</td>
<td>HD01_VD01</td>
</tr>
<tr>
<td>Buildings built in 2005 or later (table B25034)</td>
<td>HD01_VD02</td>
</tr>
<tr>
<td>Buildings built 2000–2004</td>
<td>HD01_VD03</td>
</tr>
<tr>
<td>Buildings built 1990–1999</td>
<td>HD01_VD04</td>
</tr>
<tr>
<td>Buildings built 1980–1989</td>
<td>HD01_VD05</td>
</tr>
<tr>
<td>Buildings built 1970–1979</td>
<td>HD01_VD06</td>
</tr>
<tr>
<td>Buildings built 1960–1969</td>
<td>HD01_VD07</td>
</tr>
<tr>
<td>Buildings built 1950–1959</td>
<td>HD01_VD08</td>
</tr>
<tr>
<td>Buildings built 1940–1949</td>
<td>HD01_VD09</td>
</tr>
<tr>
<td>Buildings built 1939 or earlier</td>
<td>HD01_VD10</td>
</tr>
</tbody>
</table>
Appendix B: Map of the 85,732 parcels (in red) in the greater Boston area investigated in this study for which complete cadaster data for residential buildings is available (i.e., architecture style, real estate value, year of construction)
Appendix C. Average canopy height (1) and forest volume (2) in Boston’s front and backyards in relation to the house architectural style.
Appendix D: Variable importance predicting yard canopy volume (1, 2) and average canopy height (3, 4) calculated from the distributed random forest modelling. Parcel-level vegetation variables are highlighted in green, parcel-level morphological variables related to urban form in grey and census tract-level variables in black. Only the 30 most important predictors are plotted.
Appendix E. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.landurbplan.2019.02.014. These data include Google maps of the most important areas described in this article.
References


Nassauer, J. I., Wang, Z., & Dayrell, E. (2009). What will the neighbors think? Cultural and ecological drivers, motivations and barriers for biodiversity management in re-


Ossola, A., & Hopton, M. E. (2018a). Climate di...