Urban Forestry & Urban Greening 22 (2017) 124–135

Contents lists available at ScienceDirect

Urban Forestry & Urban Greening

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

Original article

# Data quality in citizen science urban tree inventories

Lara A. Roman[a,*], Bryant C. Scharenbroch[b,c], Johan P.A. Östberg[d], Lee S. Mueller[e,f], Jason G. Henning[g], Andrew K. Koeser[h], Jessica R. Sanders[i], Daniel R. Betz[a,j], Rebecca C. Jordan[j]

[a] US Forest Service, Philadelphia Field Station, 100 N. 20th St., Suite 205, Philadelphia, PA 19103, USA
[b] University of Wisconsin – Stevens Point, College of Natural Resources, 800 Reserve Street, Stevens Point, WI 54481, USA
[c] The Morton Arboretum, 4100 Illinois Route 53, Lisle IL 60532, USA
[d] Swedish University of Agricultural Sciences, Department of Landscape Architecture, Planning and Management, P.O. Box 66, SE-230 53 Alnarp, Sweden
[e] Friends of Grand Rapids Parks, PO Box 3199, Grand Rapids, MI 49501, USA
[f] Davey Resource Group, 295 S Water Street, #300, Kent, OH 44240, USA
[g] The Davey Institute and US Forest Service, Philadelphia Field Station, 100 N. 20th St., Suite 205, Philadelphia, PA 19103, USA
[h] Department of Environmental Horticulture, CLCE, IFAS, University of Florida – Gulf Coast Research and Education Center, 14625 County Road 672, Wimauma, FL, 33598, USA
[i] Technical Services and Research, Casey Trees, 3030 12th St. NE, Washington, DC 20017, USA
[j] Human Ecology, School of Environmental and Biological Sciences, Rutgers University, New Brunswick, NJ 08901, USA

ARTICLE INFO

*Article history:*
Received 1 July 2016
Received in revised form
18 November 2016
Accepted 1 February 2017
Available online 4 February 2017

*Keywords:*
Diameter at breast height
Observation error
Species misidentification
Tree monitoring
Urban forest
Volunteer monitoring

ABSTRACT

Citizen science has been gaining popularity in ecological research and resource management in general and in urban forestry specifically. As municipalities and nonprofits engage volunteers in tree data collection, it is critical to understand data quality. We investigated observation error by comparing street tree data collected by experts to data collected by less experienced field crews in Lombard, IL; Grand Rapids, MI; Philadelphia, PA; and Malmö, Sweden. Participants occasionally missed trees (1.2%) or counted extra trees (1.0%). Participants were approximately 90% consistent with experts for site type, land use, dieback, and genus identification. Within correct genera, participants recorded species consistent with experts for 84.8% of trees. Mortality status was highly consistent (99.8% of live trees correctly reported as such), however, there were few standing dead trees overall to evaluate this issue. Crown transparency and wood condition had the poorest performance and participants expressed concerns with these variables; we conclude that these variables should be dropped from future citizen science projects. In measuring diameter at breast height (DBH), participants had challenges with multi-stemmed trees. For single-stem trees, DBH measured by participants matched expert values exactly for 20.2% of trees, within 0.254 cm for 54.4%, and within 2.54 cm for 93.3%. Participants' DBH values were slightly larger than expert DBH on average (+0.33 cm), indicating systematic bias. Volunteer data collection may be a viable option for some urban forest management and research needs, particularly if genus-level identification and DBH at coarse precision are acceptable. To promote greater consistency among field crews, we suggest techniques to encourage consistent population counts, using simpler methods for multi-stemmed trees, providing more resources for species identification, and more photo examples for other variables. Citizen science urban forest inventory and monitoring projects should use data validation and quality assurance procedures to enhance and document data quality.

Published by Elsevier GmbH.

## 1. Introduction

Citizen scientists have been involved with ecological monitoring across a range of programs, expanding public engagement in research (Dickinson et al., 2012). In the ecological sciences, citizen science engages the public in authentic research, typically through volunteers collecting field data (Dickinson et al., 2012), which promotes environmental awareness, scientific literacy, and social capital (Cooper et al., 2007; Bonney et al., 2009; Conrad and Hilchey 2011; Crall et al., 2013). While the data generated by citizen scientists has been used for research and natural resource management (Dickinson et al., 2010; Tulloch et al., 2013), concerns have

* Corresponding author.
*E-mail address:* lroman@fs.fed.us (L.A. Roman).

http://dx.doi.org/10.1016/j.ufug.2017.02.001
1618-8667/Published by Elsevier GmbH.

been raised about data quality (Bird et al., 2013; Lewandowski and Specht 2015).

Assessments of observation error in citizen science have had mixed results concerning both the level of volunteer accuracy and implications of those findings for applying citizen science to research and management. Species misidentification and incomplete taxonomic resolution in citizen science projects can lead to interpretation problems, such as overestimation of species diversity (Gardiner et al., 2012), and limited research utility of volunteer-generated species lists beyond community-level assessments (Kremen et al., 2011). For example, Kremen et al. (2011) found that volunteers missed half the bee groups recorded by researchers. However, citizen science studies focused on coral reefs, crabs, and plants have concluded that data collected by volunteers was mostly accurate, and sometimes of comparable quality to data collected by professionals (Delaney et al., 2008; Edgar and Stuart-Smith 2009; Crall et al., 2011; Butt et al., 2013; Danielsen et al., 2013). For example, Crall et al. (2011) found that volunteer species accuracy for invasive plants was 72%, compared to 88% for professionals, with both groups having lower accuracy for difficult-to-identify species. These varied studies of citizen science data quality also had widely different task complexity, with species identification involving less than 10 to dozens or even hundreds of species. With the quality of volunteer data as well as task complexity varying by case, and each case having particular data quality needs, new implementations of citizen science should include pilot testing and accuracy evaluations.

While data quality from volunteers is sometimes questioned, field data collected by researchers and their paid crews is not free of errors. When forest monitoring is conducted by researchers, examining the extent and sources of error helps to identify best practices for training crews, conducting field work and managing data (van Doorn 2014). Whether data are produced by paid or unpaid field crews, observation errors can be documented and potentially minimized through quality assurance and data validation (Ferretti 2009; Wiggins et al., 2011), and quantified error can be accounted for in statistical models (Chave et al., 2004). Evaluations of citizen science data quality can therefore be viewed in the larger context of best management practices for ecological monitoring (Lindenmayer and Likens 2010). As with any ecological research, assessing observation error in citizen science is critical to both designing effective programs and determining appropriate uses of the data.

In this paper, we present a pilot study about data quality in urban tree inventories collected by volunteers. We focused on street tree inventories, as street trees are on the front lines of engagement and management in municipal forestry. Street tree inventories record the locations and particular attributes of trees in sidewalks and other street-side environments. Such inventories are used for a wide range of purposes, including managing tree risk, prioritizing maintenance, mapping storm-damaged trees, charting species diversity and size class distribution, and estimating ecosystem services (Jim and Liu 2001; Harris et al., 2004; McPherson et al., 2005; Sjöman et al., 2012; Bond 2013; McPherson and Kotow 2013; Östberg and Sjögren 2016). Researchers and managers also use repeated inventories and systematic monitoring to explore trends in street tree populations, such as composition changes and mortality rates (Dawson and Khawaja 1985; Roman et al., 2013; Roman et al., 2014). Depending on the particular objectives of urban forest inventories, the data quality necessary and qualifications of those collecting the data may differ.

While street tree inventories are traditionally carried out by professional arborists, citizen scientists are now used in many cities. Examples of citizen science in urban forest management include the street tree census in New York City, NY (Silva et al., 2013; Campbell 2015), the Tree Inventory Project in Portland, OR (St. John 2011), the OpenTreeMap software, which has been used in cities in the United States, Canada and the United Kingdom (www.opentreemap.com, Kocher 2012), and survival monitoring for planting programs across the United States (Roman et al., 2013; Silva and Krasny 2014). Citizen science can improve volunteer knowledge about trees (Cozad 2005) and some authors have suggested that engaging volunteers in data collection can build support for municipal and nonprofit programs (McPherson 1993; Bloniarz and Ryan 1996). The application of citizen science in urban forestry builds on a rich tradition of volunteerism in urban forest management, with volunteers engaging in tree planting and other forms of stewardship (Romolini et al., 2012). Such activities can deepen participants' civic engagement and cultivate a sense of empowerment (Westphal 2003; Fisher et al., 2015; Ryan 2015).

Yet, even with urban foresters already using volunteers to collect data, integrating volunteer data into urban forest research and management has been met with skepticism due to lack of information about observation errors in the urban forest context (Roman et al., 2013). The complexity of tasks in urban tree inventories may make such work particularly challenging for citizen scientists who lack prior experience. Specifically, field crews must contend with high species diversity, with, on average, 77 tree species across 38 cities world-wide (Yang et al., 2015), and substantially higher in some municipalities, such as 161 species in Chicago, IL (Nowak et al., 2013). This includes native and exotic species, and identification guides for novices are not widely available. Urban tree inventories also typically involve measuring diameter at breast height (DBH), and if monitoring DBH change is desired, this requires field crews to make consistent measurements that allow for longitudinal tracking of individual tree growth over years or even decades.

There are only two previous studies about volunteer data quality in urban tree inventories. Although both studies conclude that there is potential for relying on field data collected by volunteers, accuracy rates for certain variables do not seem tenable for research and management applications. Cozad (2005) studied volunteer accuracy for a street tree inventory in Minneapolis, MN, and found 76% accuracy for DBH tree size class reported by volunteers, and 80% accuracy for species identification. Bloniarz and Ryan (1996) studied volunteer accuracy in Brookline, MA, and found that 94% of volunteers agreed with arborists for genus identification, and 80% for species identification; although that study considered only the most common species. Both studies found relatively low data quality for volunteers reporting maintenance needs (49% in Cozad 2005; 75% in Bloniarz and Ryan 1996), indicating that such evaluations should be performed to professionals. As public participation in urban tree inventories and monitoring expands, it is essential to build upon these studies with evaluations of data quality in more locations, and to make explicit connections between observed data quality and appropriate data uses.

Our study compared street tree data collected by experts to data collected by field crews with novice and intermediate levels of prior experience as a pilot test of new tree monitoring protocols. The goals of our study were to (1) identify the magnitude and frequency of inconsistencies in urban forestry field data; (2) determine whether novice and intermediate crews differ in their performance for genus and DBH; and (3) generate suggestions to revise training and data collection procedures in ways that may enhance data quality. We then draw lessons learned for volunteer data collection in urban forestry, with comparisons to prior studies (Bloniarz and Ryan 1996; Cozad 2005), and provide recommendations for designing field methods suitable to volunteers as well as appropriate applications of citizen science in urban forestry.