STANDARDIZATION OF STREET SAMPLING UNITS TO IMPROVE STREET TREE POPULATION ESTIMATES DERIVED BY I-TREE STREETS INVENTORY SOFTWARE

Mason F. Patterson

Thesis submitted to the faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
In
Forestry

P. Eric Wiseman, Chair

Susan D. Day

John A. McGee

Philip J. Radtke

May 2, 2012

Blacksburg, Virginia

Keywords: coefficient of variation, community forests, municipal forestry, sampling error, street tree inventory, urban forestry

STANDARDIZATION OF STREET SAMPLING UNITS TO IMPROVE STREET TREE POPULATION ESTIMATES DERIVED BY I-TREE STREETS INVENTORY SOFTWARE

Mason F. Patterson

ABSTRACT

Street trees are a subpopulation of the urban forest resource and exist in the rights-of-way adjacent to public roads in a municipality. Benefit-cost analyses have shown that the annual benefits provided by the average street tree far outweigh the costs of planting and maintenance. City and municipal foresters spend a majority of their time and resources managing street tree populations. Sample street tree inventories are a common method of estimating municipal street tree populations for the purposes of making urban forest policy, planning, and management decisions.

i-Tree Streets is a suite of software tools capable of producing estimates of street tree abundance and value from a sample of street trees taken along randomly selected sections (segments) of public streets. During sample street tree inventories conducted by Virginia Tech Urban Forestry, it was observed that the lengths of the sample streets recommended by i-Tree varied greatly within most municipalities leading to concern about the impact of street length variation on sampling precision.

This project was conducted to improve i-Tree Streets by changing the recommended sampling protocol without altering the software. Complete street tree censuses were obtained from 7 localities and standardized using GIS. The effects of standardizing street

segments to 3 different lengths prior to sampling on the accuracy and precision of i-Tree Streets estimates were investigated though computer simulations and analysis of changes in variation in number of trees per street segment as a basis for recommending procedural changes.

It was found that standardizing street segments significantly improved the precision of i-Tree Streets estimates. Based on the results of this investigation, it is generally recommended that street segments be standardized to 91m (300 ft) prior to conducting a sample inventory. Standardizing to 91m will significantly reduce the number of trees, the number of street segments, and the percentage of total street segments that must be sampled to achieve an estimate with a 10% relative standard error.

The effectiveness of standardization and the associated processing time can be computed from municipal attributes before standardization so practitioners can gauge the marginal gains in field time versus costs in processing time. Automating standardization procedures or conducting an optimization study of segment length would continue to increase the efficiency and marginal gains associated with street segment standardization.

ACKNOWLEDGEMENTS

First, I would like to extend my sincere gratitude to my committee chair and academic advisor, Dr. P. Eric Wiseman and the other members of my graduate committee Dr. Philip J. Radtke, Dr. John A. McGee, and Dr. Susan D. Day for the support, guidance, and experience they provided in this endeavor. I want to send a warm thank you to the rest of Virginia Tech Department of Forest Resources and Environmental Conservation faculty and staff for the nurturing atmosphere and learning environment they have provided since I came to Blacksburg in 2005. Thanks to Joe Gregory, Scott Maco, Al Zelaya, and the rest of The Davey Resource Group for brokering the exchange of street tree data with multiple US municipalities. I want to acknowledge Paula Peper of the Pacific Southwest Research Station for putting me into contact with the reference cities used in the development of i-Tree Streets software. The representatives from the 20+ localities that offered datasets for this endeavor, specifically Kerry Gray of Ann Arbor, MI; Betsy Reeves of Berkeley,CA; Maggie Riales of Dillon, SC; Brian Dick of Lakeland, FL; Domenic Staine of Lodi, NJ; Walter Warriner of Santa Monica, CA; and Joyce Hamburg of Southgate KY deserve acknowledgement for their cooperation and assistance with this endeavor. Finally, I would like to thank my family. Thanks to my mother and father, Nancy Foushee and Scott Patterson, and both sets of grandparents for their unconditional love and support. Thanks to my brother Dylan. The breadth of time and experiences we have shared together are irreplaceable to me. Finally, thanks to Julianne, for the special part of my life you have been since I showed up at Virginia Tech. This project would not have been possible without help and encouragement from everyone. I thank you sincerely and look forward to the future.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 – INTRODUCTION

1.1. Urban Forests

Urban forests are ecosystems existing in and around urbanized areas. They consist of woody vegetation (primarily trees) and associated natural resources. The urban forest in the continental US is comprised of approximately 3.8 billion trees, covering 27% of all urban lands (Dwyer et al. 2000). Roughly 75% of all US citizens live in urban areas and interact with urban forests on a day-to-day basis (Nowak et al. 2001). The compensatory value of US urban forests is about $2.4 trillion (Nowak et al. 2002). Valuations of urban forests, typically based on the social, environmental and economic benefits they provide, have been thoroughly documented in the literature (Dwyer et al. 1992; Miller 1997; Nowak et al. 2001).

Urban foresters manage urban forest structure (i.e., tree canopy cover, species composition, tree size distribution, and stocking) to influence biophysical function and increase urban forest value (McPherson et al. 1997). The initial step in the management of any resource is conducting an assessment of that resource (Miller 1997). Urban forest assessments provide municipal foresters with baseline information for planning the management of tree populations to maximize their benefits and minimize their costs. The customary source of data for an urban forest assessment is a tree inventory, which is a record of the location, characteristics, and assessment of individual trees within a well-defined population (Bond and Buchanan 2006). In the municipal forestry context, street tree inventories are a common method of collecting baseline data to conducting an urban forest assessment.

1.2. Street Trees

Street trees are a subpopulation of urban forests that exist in the public rights-of-way between the edges of roadways and the adjacent land parcels. Kielbaso (1990) estimated the total United States street tree population to be just shy of 62 million with enough crown cover to shade Delaware. Maco and McPherson (2003) estimated that the street tree resource in Modesto, CA provided $1.2 million in net annual benefits ($52 per tree; $23 per person) McPherson et al. (2005) found that the average tree in five US municipalities produced net annual benefits averaging between $21 and $38. From an assessment of street tree populations in fourteen Virginia municipalities (Wright et al. 2011), it was estimated that street trees provide an average of $1.1 million dollars in annual benefits ($74 per tree; $23 per person)[1].

Street trees, along with trees in parks and on other municipal lands, are the only urban trees under the direct management of local officials. Inspection, planting, maintenance, and removal of street trees are often major duties of municipal foresters and the crews they oversee. The proximity of street trees to a community increases their potential to influence and affect citizens of that community. Many citizens, especially those without private trees at their residences, interact with and are directly affected by street trees more often than most of the trees in the urban forest.

Street trees are often planted at sites that exhibit adverse growing conditions and thus must be highly tolerant of various environmental factors. Street trees commonly endure low soil volumes and soils with mineral deficiencies, high salinity, overabundance of heavy metals,

[1] The fourteen Virginia municipalities studied by Wright et al. 2011 were Abingdon, Arlington, Charlottesville, Fredericksburg, Harrisonburg, Leesburg, Lexington, Lynchburg, Martinsville, Radford, Richmond, Roanoke, Williamsburg, and Winchester.

extremely high or low pH, and heavy compaction (Day et al. 2010) . These conditions can increase the susceptibility of street trees to damage from pests and pathogens. It is common for municipal forestry programs to allocate a majority of the urban forest management efforts to selection and maintenance of street trees capable of thriving in urban environments.

Street trees are also critical from a functional standpoint. They cover and shade impervious surfaces, intercepting stormwater, mitigating runoff, and reducing the effects of the urban heat island. Street trees intercept particulates from automobile exhaust and other soil and airborne pollutants (Dwyer et al. 1992). It is important that street trees be proactively managed and maintained in good condition due to their impact on the public, the functional benefits they provide, and the resources required for their management.

1.3. Street Tree Inventory

Street tree inventories are an accounting of street trees and their attributes of interest such as species, size, condition, location, and management needs. Street tree populations are assessed using either a census (complete) inventory or a survey (sample) inventory. A complete inventory enumerates every street tree in a geographically defined population. These inventories can number in the tens-of-thousands of trees depending on the size of the locality.

A sample inventory uses statistical methods to characterize a geographically defined population of trees. Sample inventories are often preferred by municipal foresters because the composition of street tree populations are dynamic and the costs of creating and maintaining complete inventory databases is substantial. A survey inventory using statistical sampling can provide a representation of a street tree population at a high level of confidence (Jaenson et al.1992) and can provide sufficient accuracy to permit urban forest planning and management at considerable savings relative to complete inventories (Maco and McPherson 2003). Complete inventories give the most accurate representation of street tree populations; however, they are more costly than sample inventories and the degree of accuracy obtained from them may exceed the information needs of planning and management.

The financial constraints on many municipal forestry programs help justify the use of sample inventories. Increasing pressure is being placed on municipal forestry programs to demonstrate the value of street tree resources to justify funding. Municipal forestry programs receive the majority of their funding from general funds accrued from municipal taxes (Thompson and Ahern 2000). Cutbacks in municipal budgets and increased competition for available funding have led to a decreased allocation of funds to municipal forestry programs. When conducted properly, sample inventories are a cost effective tool for characterizing and managing street tree populations. However, when survey inventories are conducted without the proper sampling guidelines and attention to detail they can lead to inaccurate assessments and misinformed management decisions.

1.4. Sample Street Tree Inventory Using i-Tree Streets

The i-Tree Streets software application (hereafter referred to as STREETS) is a street tree inventory and assessment tool developed by the US Forest Service that is available free of charge to both private- and public-sector urban foresters. Using either complete or sample tree inventory data, the software is capable of estimating street tree abundance, composition, ecosystem services, and monetary value. Roughly two-thousand urban forest assessment projects were conducted using i-Tree software tools in 2010 (Jonnes 2011). STREETS determines street tree ecosystem services and socio-economic benefits using empirical models developed from geographic data and tree growth data for the twenty most populous street tree species in sixteen United States reference cities (McPherson 2010). Users select the most appropriate reference city for modeling street tree growth and benefits in their localities.

STREETS serves as more than a tool for characterizing street tree populations for the purpose of making management decisions. It is unique compared to other inventory tools because it provides scientifically valid estimates of street tree function and monetary worth at the click of a button. STREETS has quickly become the preferred tool of municipal foresters for conducting street tree assessments because it is free to the public and seamlessly integrates field inventory and computational assessment processes. As STREETS popularity amongst municipal foresters continues to grow, it is increasingly important to investigate potential ways of improving the reliability of the estimates generated by the software.

Specific sampling guidelines must be adhered to in order to attain an unbiased, accurate sample street tree inventory using STREETS. In STREETS, the sampling unit is a linear portion of a public roadway called a street segment. Conducting a sample inventory using STREETS first requires delineating a sampling frame of street segments within a geographic area of interest (typically the municipal boundary) and then randomly selecting street segments from the sampling frame for subsequent collection of tree inventory data. These steps in the sampling protocol are typically performed using GIS techniques. The conventional source of the street segment GIS data are polyline shapefiles called TIGER/Line, which is developed and distributed annually by the US Census Bureau (TIGER/Line Shapefiles 2011). The lengths of TIGER/Line street segments within localities exhibit non-normal distributions, with most streets having relatively short lengths from a few feet to hundreds of feet. Outlier segments typically exist in every locality with lengths ranging from hundreds to thousands of feet (Figure 1.1). Although many municipalities now create analogous street segment datasets independently, their datasets typically exhibit similar asymmetries in street segment length distribution.

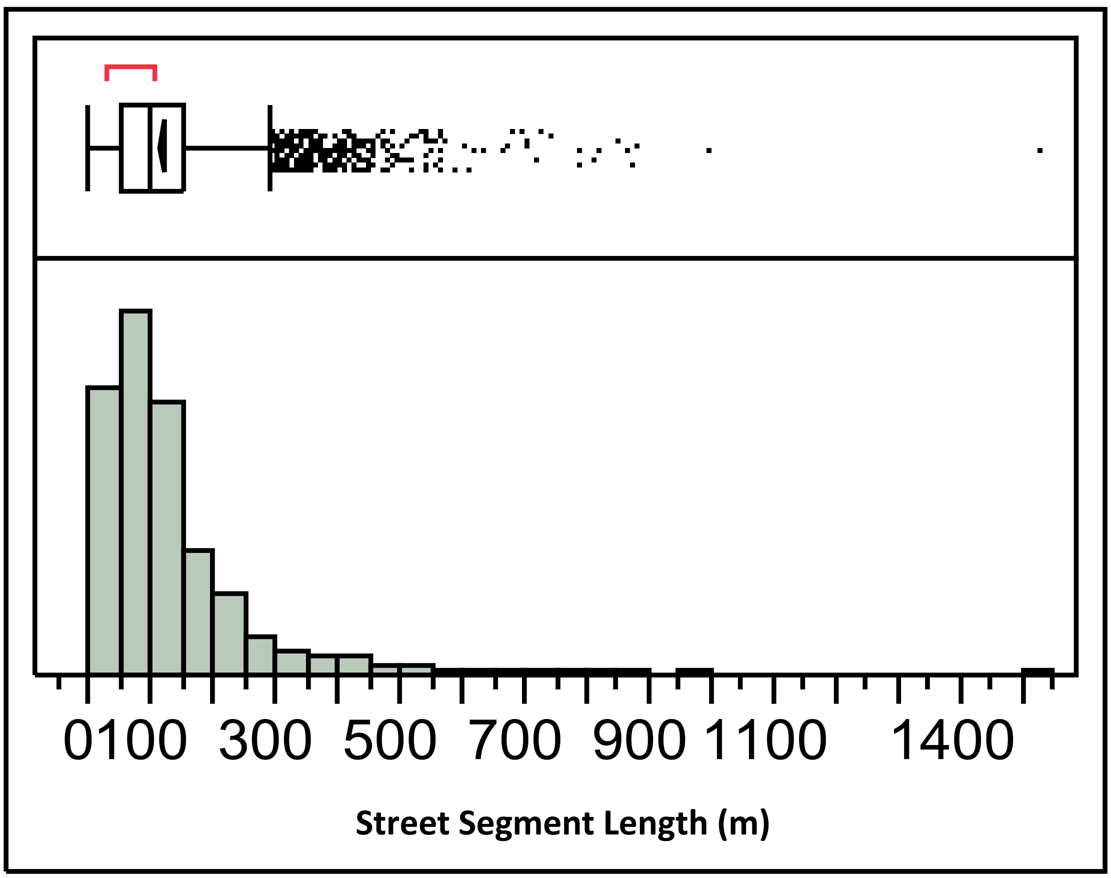


Figure 1.1. Histogram and boxplot of TIGER/Line street segment lengths (shown here in meters) for Ann Arbor, MI (*N* = 5,107). Distributions of TIGER/Line segments are typically right skewed, but the severity of skewness varies across localities.

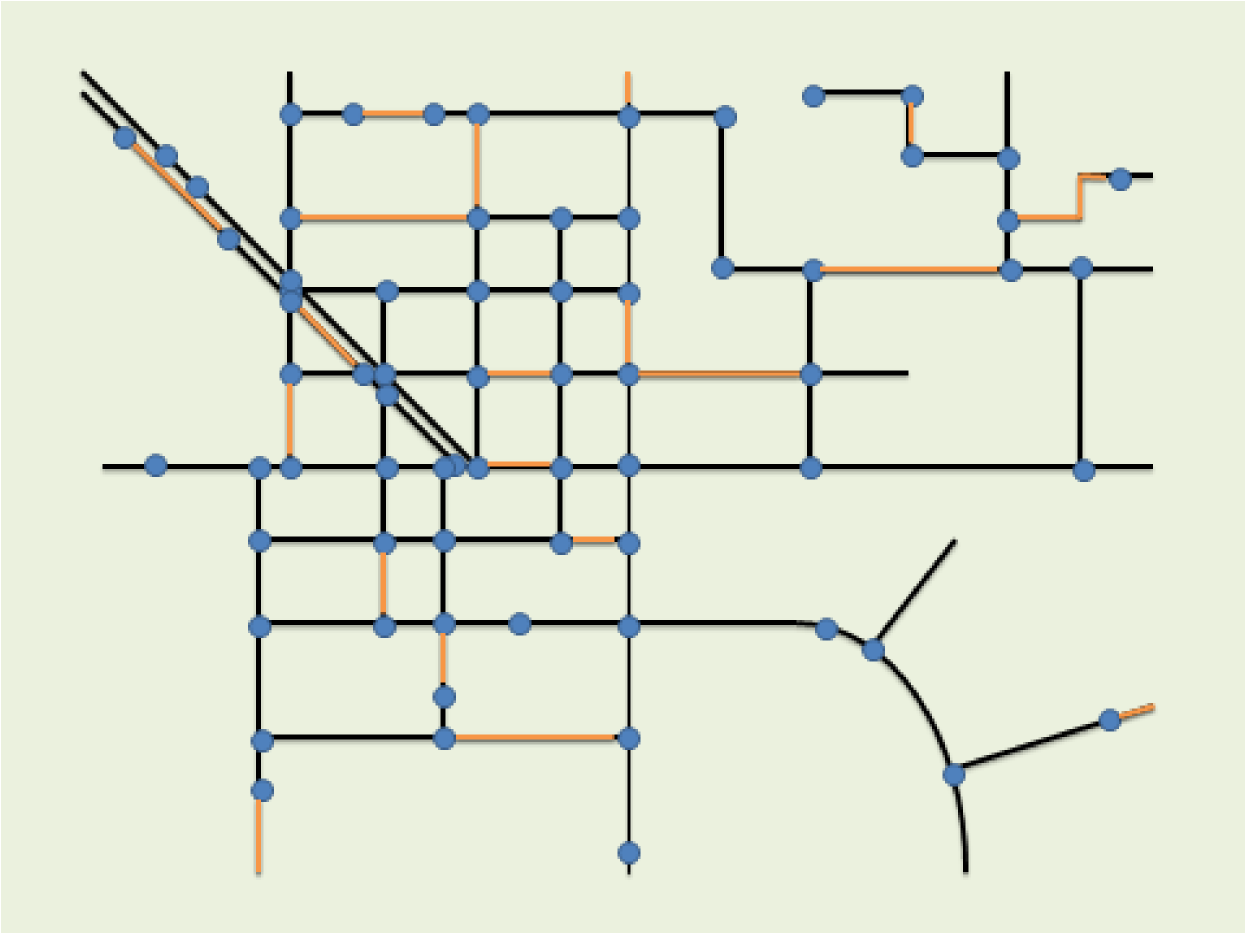


Figure 1.2. Schematic of a simple random sample of twenty street segments (highlighted in orange) from a sampling frame of approximately 100 street segments (black). Endpoints of street segments are depicted by blue dots.

Figure 1.2 is a schematic of a random sample of twenty street segments from a sampling frame comprising roughly 100 street segments. The schematic demonstrates two key characteristics of conventional street segments: variation in their length and in their termination points. Length of street segments can vary considerably depending on the geography of the locality. Segments also often span two street intersections, but may also appear and terminate arbitrarily along streets with few points of intersection.

Once a sample of street segments has been drawn from the sampling frame, trees existing in the right-of-way adjacent to selected segments are tallied and their attributes are measured. The STREETS software is then used to calculate an estimate of the total street tree population for the area of interest along with a standard error for the total, both of which are based on the average number of trees per segment and the variation in the number of trees from segment to segment. Using the population estimate (along with data on tree size and species) as input data, STREETS computes street tree benefits from empirical models developed for biophysical functions such as stormwater interception, energy conservation, and air pollution absorption. According to STREETS developers, a simple random sample of 3-6% of the street segments in a municipality should provide a standard error of the total population estimate that is no larger than 10% of the estimate itself (i-Tree 2011). This error rate is generally regarded as a sufficient level of precision to permit confident inferences from street tree assessments; however, this guideline has not been thoroughly investigated to determine whether it fulfills the stated outcome.

1.5. Problem Statement

Since 2008, researchers in Virginia Tech's Department of Forest Resources and Environmental Conservation have conducted street tree assessments using STREETS in about a dozen Virginia municipalities. In these assessments, it has been observed that the sampling intensity guidelines provided by STREETS (i.e., sampling 3-6% of street segments) rarely affords a relative standard error that is less than 10% of the population estimate (i-Tree 2011). Large standard errors produce greater uncertainty in street tree population estimates and thus uncertainty in the estimation of street tree benefits from these estimates. If a high level of confidence cannot be placed in a street tree assessment, then it is useless as a basis for policy, planning, and management of street trees.

A secondary consequence of high standard error is the additional cost for supplemental sampling that must be conducted in order to reduce the standard error. Sample size must be increased by a factor of four to reduce relative standard error by half because standard error is in squared units and is inversely related to sample size (STRATUM 2012). This leads to excessive costs and limits the effectiveness of performing supplemental sampling.

Closer investigation of this issue has revealed that in many of the assessed localities there is considerable variation in the length of sampled street segments. This variation in sampling units likely contributes to sampling error and raises concerns about the accuracy and precision of street tree population estimates obtained using the conventional STREETS sample inventory protocol. Increasing sampling unit uniformity (i.e., reducing variability in street segment length) prior to drawing an inventory sample should decrease the variation in tree count across sampled segments. Reducing variability in tree count between segments reduces the minimum number of segments that must be sampled (n) to achieve a desired level of error. Standardizing street segment length may also lead to gains in the accuracy of total street tree population estimates produced by STREETS.

1.6. Project Scope and Objectives

McPherson (2010) called for a "systematic analysis of the sensitivity of i-Tree Streets' output to the probabilistic range of input values." He identified six types of errors that must be examined, the most significant of which is sampling error; "how well a tree sample reflects the actual tree population." The scope of this study was an investigation of the sample street tree inventory protocol utilized by i-Tree Streets assessment software. Specifically, the study focused on the response of street tree sampling error to modifications in the program's conventional sampling protocol, which entailed standardization of street segment lengths prior to sample selection and inventory data collection. The goal of this study was to improve the conventional STREETS sampling protocol to provide a better tool for street tree inventory and assessment. To achieve this goal, a GIS procedure for standardizing street segment length was developed and tested using complete tree inventory datasets obtained from several US municipalities. The research objectives and hypotheses were:

1. Examine the effects of standardizing street segment length on the sampling error of STREETS estimates for total tree populations.

   Hypothesis: Standardizing street segments to a uniform length prior to sampling will provide total tree population estimates with lower relative standard errors than samples taken with unstandardized street segments.

2. Examine the effects of street segment length and sampling intensity on the accuracy and precision of STREETS estimates for total tree populations.

   Hypothesis: Standardizing street segments to a uniform length prior to sampling at the recommended intensity will provide more accurate and precise estimates of street tree populations than unstandardized street segments.

CHAPTER 2 – LITERATURE REVIEW

The following review provides a description of street tree sampling and i-Tree Streets software (hereafter referred to as STREETS). It also provides background on the conventional street segment geospatial data source (TIGER/Line Shapefiles) and concludes with a discussion of plot size optimization with respect to improvements in sampling precision.

2.1. Street Tree Sampling

Since the computerization of street tree inventory became popular in the United States in the 1970s (Smiley and Baker 1988), researchers and practitioners have made incremental efforts to increase the efficiency of street tree sampling. Mohia et al. (1978) developed a methodology for sampling to determine the population distribution of common street tree species. The investigation was conducted using primitive computers to simulate the sampling of the known population of street trees in Poughkeepsie, NY. The most common trees were found to be distributed in clusters, thus the sampling protocol identified called for cluster sampling of street segments and inventorying the most prevalent species adjacent to every segment in the cluster (sometimes at intervals). The recommended protocol required that the practitioner identify and reasonably estimate the most prevalent tree species before conducting a street tree inventory. The findings indicated that one can achieve an acceptable level of variance in estimates of stem height and diameter by sampling >50 clusters of streets. Sampling at intervals was found to decrease the total sample required for an adequate estimate but increases the standard deviation of the sample.

Jaenson et al. (1992) identified a statistical method of estimating street tree populations from a stratified sample of blocks and sections of linear street. This method required no preexisting information on the street tree population of interest and took steps to ensure that the sample units were relatively uniform. It was concluded that by conducting a presample and subsequently sampling ~2,000 trees, one could accurately (± 10%) estimate the total number of street trees in the municipality, the species composition, and the size of DBH classes. The results were validated by comparing estimates obtained using the recommended procedure against known population counts in 4 New York municipalities. Maco and McPherson (2003) combined the sampling methodology used by Jaenson et al. (1992) with the benefit cost analysis used to assess urban trees in Modesto, CA (McPherson et al. 1999). The procedures used by Maco and McPherson (2003) evolved to become the Street Tree Assessment Tool for Urban Forest Managers (STRATUM), the predecessor and underlying mechanics of STREETS (STRATUM 2012). However, the controls in sampling unit uniformity recommended in Jaenson et al. (1992) were not applied to the STREETS sampling protocol (i-Tree 2011).

2.2. Overview of i-Tree Streets Software

STREETS is one available tool in the i-Tree software suite. In order for STREETS to report on abundance, composition, and value, a minimum of three attributes-street segment identifier, species, and DBH-must be collected for each inventoried tree. STREETS uses simple random sampling statistics to estimate total street tree abundance. The mean number of trees tallied on sampled segments is multiplied by the total number of street segments in the sampling frame to calculate the tree population estimate. The variance in the number of trees on each sampled segment is used to calculate the standard error of the tree population estimate

(STRATUM 2012). The estimated abundance of each species comprising the street tree population is proportionate to the number of trees of that species in the sample inventory (Maco and McPherson 2003). The abundance of trees with various attributes (maintenance needs, wire conflicts, pruning needs, hazardous defects, etc.) can also be calculated in proportion to their presence in the sample. Using STREETS, one can produce reports describing street tree composition by species, size, and condition.

Street tree data from sixteen reference cities are used to model street tree benefits in different climate zones in the United States (Fig 2.1). A reference city is selected using the STRATUM Climate Zone Map or by matching an area of interest with the reference city that is most similar in species composition, precipitation, average annual heating degree days, and average annual cooling degree days (McPherson 2010). For each reference city the biophysical benefits were modeled based on structural attributes for 40 specimens of the 20 most popular street tree species. The benefits of a street tree population are modeled by categorizing each tree by linking it to one of the 20 most popular street trees based on physiology and leaf morphology. If none of the 20 most popular species are similar enough to be selected, the tree can be linked to a broader species model that is based on leaf morphology and stature (Leaf Morphologies: Broadleaf Evergreen, Broadleaf Deciduous, Conifer Evergreen, Palm Evergreen; Sizes: Small, Medium, Large). Each tree is also categorized by its DBH class.

Benefits are modeled based on population fragments (Maco and McPherson 2003). For example, if i-Tree estimates that there are 20 eastern hemlocks with a DBH of 8-16 cm within a locality and each hemlock conserves 10 kWh of energy annually from the model, then the total annual energy saving produced by this portion of the street tree population is 200 kWh. A dollar value is then calculated by multiplying this value by the local price of a kWh of energy.

Figure 2.1. Map of the United States showing the delineation of climate zones and reference cities used by i-Tree Streets for modeling tree growth and ecosystem processes (Climate 2012).

This method of benefit estimation has important implications for this research project. Improvements in the accuracy and/or precision of overall street tree populations estimates will necessitate further research to test the impacts on estimates of population characteristics such as species and size compositions. Refined methods identified as being more accurate or precise predictors of total population may improve the representativeness of estimates of species and size distributions or population fragments and the estimation of benefits provided.

2.3. TIGER/Line Shapefiles

TIGER (Topologically Integrated Geographic Encoding and Referencing) is a system developed and used by the US Census Bureau for mapping and studying census geography (US Census Bureau 2011). TIGER files contain cartographic information and characteristic attributes describing US roads, utilities, boundaries, landmarks, hydrologic features, and geography. TIGER/Line files were first released in 1990 in the format of ASCII text files. The Census Bureau continued to release TIGER/Line files in this format through 2006 (TIGER/Line Shapefiles 2011). Prior to 2007, these features were identified using two-digit Census Feature Class Codes (CFCC).

In 2002, the comprehensive Master Address File/TIGER Accuracy Improvement Project (MTAIP) was undertaken to improve the accuracy of TIGER/Line data. MTAIP focused on improving the accuracy of feature locations by obtaining and utilizing the most accurate data provided by non-federal governmental agencies. These agencies included state, local, or tribal governing bodies. MTAIP was conceived under the principal that by "enhancing geographic partnerships" with state and local governments, the US Census Bureau could provide the general public with free, high accuracy data describing census geography (US Census Bureau 2011). However, in order to provide the general public with the benefits of high accuracy data, the US Census Bureau was forced to sacrifice some consistency between data provided by different sources.

In 2007, the Census Bureau began producing TIGER/Line data in the form of shapefiles for use in GIS. Feature classes in TIGER/Line Shapefiles are coded using the Master Address

File/ TIGER Feature Class Code (MTFCC), a one letter, four-digit code (US Census Bureau 2011). Appendix 1 of i-Tree Streets User's Manual recommends TIGER/Line Shapefiles as a source of geospatial street segment data for preparing sample street tree inventories. Because TIGER/Line shapefiles have national coverage and are the recommended street segment data source for the STREETS sampling protocol, they were chosen as the most suitable test data for this study.

2.4. Plot Size Optimization and Sampling Improvements in Other Natural Resource Contexts

Plot size optimization refers to the identification and establishment of a size or spatial uniformity in sampling units to achieve the minimum time necessary to locate and sample a population of interest at a desired accuracy (Zeide 1980). The principal determinant of optimum plot size are relative variability and relative cost (Wiegert 1962). The cost of field sampling is linked to both plot size and sample size. Plot size is generally inversely related to the number of plots that must be sampled to achieve a desired precision (Gerik and Rush 1985; Poultney et al. 1997; Xiao et al. 2004; Nowak et al. 2008) and must be increased by a factor of four to reduce the sample variance by half (Zeide 1980). This leads to a tradeoff during plot size optimization between the gains of reducing variation and the costs associated with sampling larger plots.

While attempts to control for segment length have been executed during prior street tree sampling refinements (Jaenson et al. 1992; Maco and McPherson 2003), an investigation in pursuit of an optimum segment length for street tree sampling has not yet been undertaken. However, reducing sample variation as a means of investigating precision during plot size optimization has been successful using fixed area sampling units, linear transects, and variable area plots in the sampling of related natural resources. Should standardization prove to be an effective means of reducing sample variance, it will be logical next step to pursue and identify optimal segment length for street tree samplings. A strong understanding of the results of past efforts at identifying optimum plot sizes and other improvements in natural resources sampling will allow for more useful interpretation of the results of this investigation.

Fixed area sampling units are commonly used to sample and estimate natural resources abundance and characteristics and can be geometrically categorized as plots and strip plots (Avery and Burkhart 2001). Standardized street segments are essentially two dimensional plots with one fixed dimension, so it is reasonable to expect similar trends in changes in variance with standardized length as have been found in optimization studies fixed area plots. Nowak et al. (2008) tested the effects of the size of fixed radius plots on the precision of urban forest population estimates in Syracuse, NY. The time required to conduct samples of 26 fixed radius plots (0.017, 0.04, and 0.067 ha) was weighed against the precision of calculated estimates. A sample of two hundred 0.04 ha plots was identified as the optimum for urban forests in general. They estimated that conducting such a sample would take a two-person crew roughly 14 weeks and would yield a relative standard error of about 12%. While relative standard error was found to decrease with plot size, the optimum plot size was determined to be the median size of 0.04 ha because of the additional time required for the sampling of larger plots.

Xiao et al. (2004) conducted a fixed area optimization study of 6 m wide strip plots used to measure vegetation cover. Plots lengths of 10-100 m were tested and it was concluded that a sample of 176, 45-m strip plots was optimum for mapping vegetation cover at regional and local scales. McCormick and Choat (1987) compared visual strip plot sampling of reef fish and identified 37.5 m × 10 m plots to be the optimum size with an approximate 70% gain in the time required for sampling regardless of desired precision. Plot sizes, sample sizes, relative variability, and relative costs varied greatly between the fixed area plot size optimization studies described above. However, these studies demonstrate the tradeoff between reduction in variability and increase in cost as plot size and sample size increase. While optimizing standardized length is beyond the scope of this project, understanding changes in relative cost and relative variability based on plot size and sample size will allow for a more sound interpretation of the results of this investigation and will help to guide recommendations for which standardized length is best suited for certain types of field situations.

Linear transect sampling is a form of distance sampling (Buckland and Laake 1993; Buckland et al. 2001) commonly used in biometrics (Patil et al. 1979; Zahl 1989; Marques and Buckland 2003) to estimate total population or density. Subjects are sampled if they are within a designated distance of a line (Marques and Buckland 2003). STREETS sampling protocol differs from linear transect sampling because the width of the detection distance is not fixed from plot to plot and they are not randomly placed but randomly selected. However, street segments are linear in their nature and a tree is detected based on it being within a certain distance of a segment. Marques and Buckland (2003) applied a conditional likelihood approach to linear transect sampling to incorporate covariates. If variables other than segment length affect the presence and densities of street trees they could be incorporated into STREETS sampling protocol using the Marques and Buckland (2003) procedure.

Variable area transect (VAT) sampling and estimation (Parker 1979) are used to measure stationary subjects in fields related to ecology and biology. During VAT sampling; a fixed width strip is sampled until a predetermined number of subjects are sampled (Engeman et al. 2005). Variable area transect sampling differs from fixed area and linear transect sampling because the sampling unit is defined by the subject count instead of a spatial relationship. Engeman and Sugihara (1998) conducted Monte Carlo optimization simulations of VAT sampling of 64 spatial patterns and concluded that the sample of 3-6 subjects was considered optimum for the majority of field situations. A follow up simulation study of VAT sampling led to the identification of sampling 5-7 subjects with $20 \leq n \leq 40$ (Engeman et al. 2005). While STREETS calculations and methods do not lend themselves to VAT sampling, it is an intriguing possibility for street tree sampling because municipal street segments are essentially a network of strips with their boundaries variable yet set based on the public right-of-way.

While the studies described above are not focused on street tree sampling and do not use street segments for their sampling units, the methods and findings of each study can be applied to this project. Fixed are plots, linear transects, and variable area transects all exhibit similarities to the sampling units used in STREETS inventories but also exhibit distinct differences (resource sampled, population distribution, plot dimensions, relative costs, etc.). However, understanding the tradeoff between cost and precision based on sample and plot size, the incorporation of covariates into sampling of natural resources, and adaptive sampling of resources with non-normal distributions will be useful interpreting the results of this project and determining how to pursue further improvements in street tree sampling.

CHAPTER 3 – MATERIALS AND METHODS

3.1. Obtaining Datasets and Selecting Research Municipalities

The following sections describe how complete street tree inventory data and street segment data were obtained, formatted, and used to select the seven research municipalities used in this investigation. The purpose of these steps was to create and select datasets for testing research hypotheses.

3.1.1. Obtaining Candidate Tree Inventory Datasets

Complete street tree inventory datasets were requested from thirty municipalities in nineteen states in the continental United States (Table 3.1). Fifteen of the municipalities were reference cities used for i-Tree Streets development and modeling. The fifteen other localities had conducted street tree inventories through contracts with the Davey Resource Group (Kent, OH). These thirty localities were targeted because they had all contracted or coordinated with the Davey Resources Group and were consistent from a methodological standpoint. Fifteen of the thirty localities responded to the data request and provided candidate datasets that met the preliminary selection criteria (i.e., complete inventories conducted within the last ten years with geographic coordinates for each tree (Figure 3.1)).

Table 3.1. Thirty US municipalities that received requests for street tree inventory datasets for use in this study.

Glendale, AZ	Boise, ID	Auburn Hills, MI	Queens, NY
Berkeley, CA	Elgin, IL	Minneapolis, MN	Lower Merion, PA
Claremont, CA	Avon, IN	Des Peres, MO	Charleston, SC
Modesto, CA	Indianapolis, IN	Wildwood, MO	Dillon, SC
Santa Monica, CA	Valparaiso, IN	Charlotte, NC	Longview, WA
Fort Collins, CO	La Grange, KY	Lodi, NJ	Sun Prairie, WI
Lakeland, FL	Southgate, KY	Albuquerque, NM	
Orlando, FL	Ann Arbor, MI	Roswell, NM	

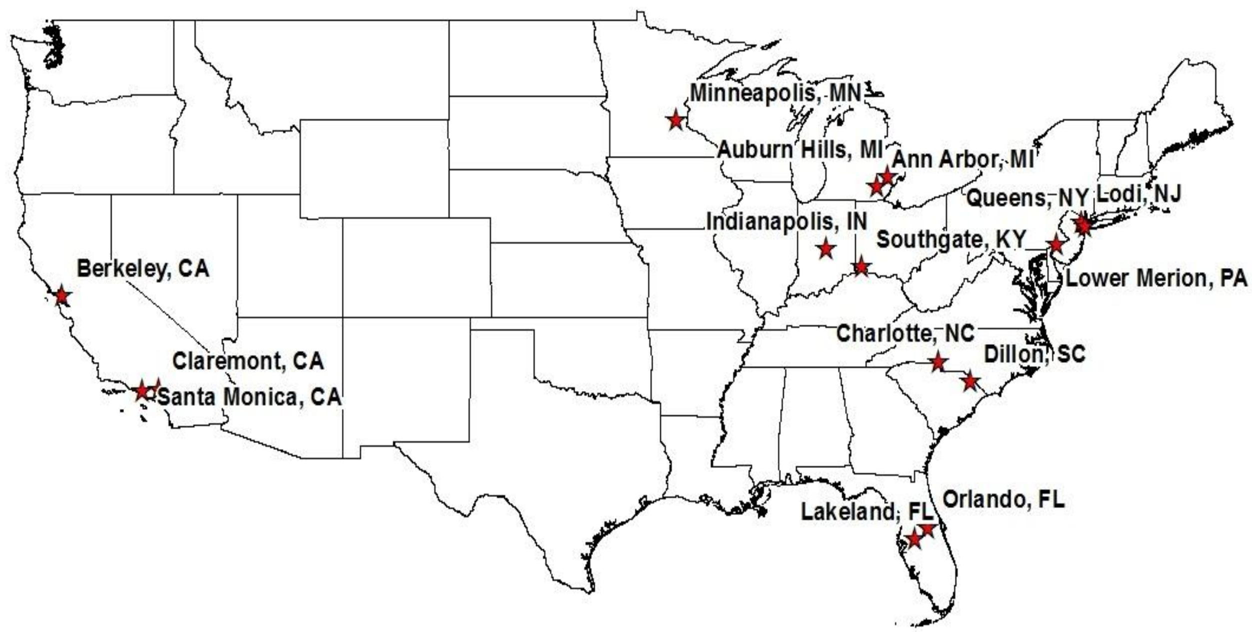


Figure 3.1. Locations of fifteen municipalities that provided candidate tree inventory datasets that met the preliminary data screening criteria for this study.

3.1.2. Obtaining TIGER/Line Street Segments

A sampling frame is defined as the source or population that a sample is selected from (Särndal et al. 2003). In this study, TIGER/Line street segment datasets were chosen as the source for experimental sampling frames because they have national coverage and are the recommended data source in the STREETS protocol. For each locality, TIGER/Line shapefiles from the year most proximate to the completion date of the corresponding tree inventory were downloaded from the US Census Bureau (TIGER/Line Shapefiles 2011). The completion date of each tree inventory was determined by sorting date records in the GIS attribute table of each dataset. If the inventory was conducted by the Davey Resource Group and had no date information in the GIS attribute table, then inventory completion date was determined from the metadata provided by the Davey Resource Group (Table 3.2). If the inventory was conducted for a STREETS reference city and contained no date information, then date was determined from the locality's Municipal Forest Resources Analysis report prepared by US Forest Service (US Forest Service 2010). The Orlando and Berkeley TIGER/Line shapefiles most proximate in time to the inventory completion date were noticeably out of line when overlaid with street tree points in GIS. In both instances, the TIGER/Line shapefile from the following year was used because they matched public streets better. Every dataset and shapefile was either received or projected in the state plane zone of its respective municipality.

Table 3.2. Characteristics of seven US municipalities in the final selection for this study. More elaborate metadata are provided in Appendix B. [1]Source: (US Census Bureau 2012).

Municipality	**Tree Inventory Date**	**Human Pop.**[1]	**Land Area (sq. mi.)**[1]	**Pop. Density (#/sq. mi.)**	**Street Trees**	**Street Miles in Inventory**
Ann Arbor	2011	115,204	27.8	4,140	40,321	375
Berkeley	2001	109,408	10.5	10,450	30,378	262
Dillon	2007	6,740	5.2	1,290	1,840	41
Lakeland	2007	96,623	65.3	1,480	26,193	596
Lodi	2007	23,970	2.3	10,583	5,153	49
Santa Monica	2011	88,679	8.4	10,538	31,694	240
Southgate	2006	3,704	1.4	2,594	199	12

Street polylines categorized as private streets, alleys, interstates and accompanying ramps, and trails were excluded from the study because trees residing there are typically not under municipal jurisdiction and are commonly excluded from street tree inventories. Thus only TIGER/Line shapefiles with MTFCC (Master Address File/Tiger Feature Class Code) of S1200 and S1400 (US Census Bureau 2011) were included in the experimental sampling frames. The final step in delineating the experimental sampling frames was the acquisition of municipal boundary shapefiles from local officials or from web-based GIS data warehouses at the local or state level (Appendix A).

3.1.3. Formatting Inventory Datasets, Municipal Boundaries, and TIGER/Line Shapefiles

In some instances, a large cluster of street segments within the municipal boundary lacked trees, stumps, or planting spaces when viewed in the GIS display. If these clusters of streets could be identified as locations not likely to have been inventoried (e.g. golf courses, airports, graveyards) using satellite imagery in Google Earth (Google Inc.; Mountain View, CA) a revised municipal boundary layer was manually digitized in ArcGIS (ESRI Inc.; Redlands, CA) to exclude these locations. These revised boundary layers were also manually digitized around tree inventory datasets that were incomplete (i.e., not all public streets within the municipality were inventoried). In all cases, the original boundary layer was clipped (shrunk) to the dimensions of the revised boundary layer. This procedure was conducted on 5 of the 15 datasets considered and 3[2] of the 7 selected research localities.

[2] Berkeley, CA and Southgate, KY had secondary boundary layers digitized due to the presence of an un-inventoried campus and golf course. A second boundary layer had to be digitized for Dillon, SC because it was a complete inventory of part of the municipality, or a partial inventory.

For each locality, the TIGER/Line shapefile depicting public streets was clipped to the finalized municipal boundary layer to produce a finalized experimental sampling frame. An attribute field was added to each experimental sampling frame and the length of each segment in the sampling frame was calculated using the ArcGIS Calculate Geometry Function. GIS points for inventoried street trees were then clipped to the finalized municipal boundary, discarding any points for trees residing outside the finalized boundary. In addition, the GIS attribute table of each tree inventory was filtered to discard any records that were not actual trees (e.g., stumps, vacant planting sites, records with no trunk diameter entry) and the filtered dataset was exported to a new tree points layer. If the attribute table of a dataset contained a field describing the locational nature of the trees (e.g., street, park, median, borderline, government), then any tree that could be a street tree (street, median, borderline, blank) was selected and exported to a potential street trees layer.

Tree points had to be uniquely assigned to specific street segments in order to conduct the simulations and statistical analyses. To do this, the distance of each tree point to the nearest street segment in each experimental sampling frame was computed using the Near Analysis function of ArcGIS 3D Analyst. The frequency distributions of these distances were then calculated using JMP 9 (SAS Institute Inc.; Cary, NC). Based on this information, it was determined that trees greater than 61 meters from a street segment could not be spatially associated with a particular segment and were subsequently discarded from the tree point datasets. Not more than 2.5% of the trees were discarded from any of the datasets based on their distance from street centerlines.

The number of trees adjacent to each unstandardized street was tallied using a spatial join in ArcGIS. In sampling frames other than those in Dillon, SC and Southgate, KY the joined tree count was greater than inventory points count by up to 75. It was hypothesized that single trees were joined to multiple segments in the larger datasets due to them being the same distance or rounded to the same distance from two separate street segments. Finally, the coefficient of variation (CV) for the mean tree count per segment and the coefficient of determination (R^2) for the linear relationship between segment length and tree count were calculated to aid in the selection of municipalities for the inventory simulations (see section 3.4 for further discussion of statistical analysis).

3.1.4. Selection of Research Municipalities

Seven localities were selected from the fifteen candidates for use in the inventory simulations and statistical analyses. These localities were selected to provide broad geographic representation of the continental US and to span the range of CV and R^2 values computed for the candidate localities. The median CV for per-segment tree count in research localities was 1.53 in Dillon, SC and ranged from 1.25 in Ann Arbor, MI to 2.64 in Southgate, KY (Table 3.3). The median R^2 for segment length-tree count was 0.2848 in Ann Arbor and ranged from 0.0581 in Dillon to 0.4426 in Lodi. Metadata on census geography and urban forest environment in research municipalities is provided in Appendix B.

3.2. Procedures for Standardizing Street Segment Length

The following sections describe procedures for segment length standardization and compared to unstandardized datasets as a test of the hypothesis that standardization improves the accuracy and precision of STREETS inventories. The initial step in standardization was to splice segments together to create a population of segments with the maximum possible lengths. These long segments were then divided to approximately 150, 300, and 450ft, hereafter referred to as (46, 91, and 137m).

Table 3.3. Report of the segment count, coefficient of variation (CV) of trees per segment, and coefficient of determination R^2 of the strength of the relationship between street length and tree count for unstandardized sampling frames in municipalities who submitted usable datasets. Data are sorted alphabetically by municipality. R^2 values were all highly significant ($P \leq 0.0001$). Selected research localities are denoted with an asterisk (*). Queens, NY was removed from consideration due to large dataset size.

Municipality	**Street Segment Sampling Frame (N)**	**Coefficient of Variation (%)**	**Coefficient of Determination (R^2)**
Ann Arbor, MI*	5,107	125	0.2848
Auburn Hills, MI	1,334	326	0.2419
Berkeley, CA*	4,435	128	0.3123
Charlotte, NC	26,129	198	0.2052
Claremont, CA	2,596	142	0.1128
Dillon, SC*	624	153	0.0581
Indianapolis, IN	11,071	236	0.1156
Lakeland, FL*	7,362	209	0.1047
Lodi, NJ*	822	146	0.4426
Lower Merion, PA	2,862	160	0.3936
Minneapolis, MN	14,434	110	0.3087
Orlando, FL	13,202	227	0.1856
Santa Monica, CA*	3,730	154	0.3042
Southgate, KY*	172	264	0.1177

Spliced segments smaller than these target lengths were classified as “snippets” and either discarded or respliced to adjacent segments. Six dataset were generated for each combination of target length and snippet fate. Appendix C outlines the GIS procedure for street segment standardization.

3.2.1. Splicing Street Segments Prior to Standardization

Street segments were spliced together to maximize their overall length prior to their division. Segments were spliced to intersecting segments with the same name using the Dissolve function in ArcGIS (ESRI, Redlands CA). Street name segments were subsequently spliced across intersections until they dead-ended into open space or another continuous length of street (Appendix C; Presentation 1).

3.2.2. Selecting Standardized Street Segment Lengths

Standardized lengths of 46, 91, and 137 m were selected because they roughly corresponded to the median values of the frequency distribution quartiles of segment length in the unstandardized sampling frames (Table 3.4). A difference of no more than 12 meters existed between the selected target lengths and median values for percentiles of interest. Past experiences with sample street tree inventories conducted by Virginia Tech Urban Forestry suggest that sampling units of these sizes would be practical for field operations.

Table 3.4. Quartiles of the length of unstandardized street segments in research municipalities. Median street length within each quartile is denoted with an asterisk (*).

Municipality	25th Percentile Conventional Segment Length (m)	50th Percentile Conventional Segment Length (m)	75th Percentile Conventional Segment Length (m)
Ann Arbor	39.21	70.62	156.77
Berkeley	44.92	90.18*	120.11
Dillon	46.01	77.30	120.98
Lakeland	52.45*	63.13	148.53
Lodi	52.82	95.73	147.87*
Santa Monica	54.75	97.14	161.59
Southgate	88.15	111.92	116.94

3.2.3. Generating Standardized Street Segment Datasets

Some short spliced segments intersected the municipal boundary and were artifacts of larger segments that were clipped to the boundary layer. Spliced street segments less than the largest standardized length of 137 meters that intersected municipal boundary layers along with their associated trees were selected and removed from spliced sampling frames to refine the population of streets being sampled to include those most likely to be managed by the municipality. The spliced segments in each layer were then divided in a number of segments with lengths in the range of the identified target length plus or minus ten percent of the target length (i.e. 41-50m, 82-101m, and 123-151m) using the ArcGIS split function.

Two alternate procedures were pilot tested to identify an appropriate method for handling spliced segments shorter than the standardized ranges. In the Splice-Dice-Discard procedure, short splice segments and their tallied trees were removed from the sampling frame and street tree population. In the Splice-Dice-Resplice procedure, short spliced segments longer than half the lower endpoint of the target length range were considered to be street segments; segments shorter than half the lower endpoint were spliced with the nearest adjacent segment based on the rules outline in Appendix C: Presentation 2. This procedure relaxed the target length range for segment standardization to ± half of the target length range (21-71m, 41-142m, 62-213m).

3.3. Simulation Studies of Conventional and Refined Sampling Protocols

Street tree simulations were conducted to test the hypothesis that standardized segment sampling frames produce more accurate estimates of street tree population that unstandardized segments. Street tree inventory simulations were conducted on unstandardized and Splice-Dice-

Resplice (due to the results of the pilot study 4.1) sampling frames in Ann Arbor, Santa Monica, and Southgate using the statistical software R version 2.14.1 (R Development Corp Team; Vienna, Austria). These localities were selected for the simulation study because they provided broad geographic representation and spanned the range of CV values for the mean tree count per segment observed across all research localities (see Table 3.3 in section 3.1.4). Thirty replicated inventory simulations were conducted for each municipality by drawing random samples (5% sampling intensity) from the conventional segment sampling frame and the three standardized segment sampling frames. A computational script (Appendix D) was run on each sampling frame to calculate an estimated street tree population and standard error for each replicated simulation in each locality. The equations used for this script were obtained from the i-Tree Streets literature (STRATUM 2012).

3.4. Statistical Analyses

The effect of standardization on the distribution of segment lengths and correlation between street length and tree count compared to unstandardized segments was assessed preliminarily. Variation in labor time based on standardized segment length was analyzed using paired t-tests. Unstandardized segment count and total municipal street length were assessed as potential predictors of the amount of labor time required to standandardize segments in a municipality. The relationship of both parameters to labor time was assessed using linear regression.

For the simulation study of segment standardization effects on tree population estimates, one sample t-tests were conducted ($\alpha = 0.05$) to identify significant differences between the 30 mean population estimates from each simulated sampling frame dataset and the true population. One-way ANOVA was conducted within each locality and a Tukey-Kramer means comparison was used to test for differences in the population estimates amongst unstandardized and standardized segment sampling protocols. A 95% confidence interval was used to evaluate population estimator precision. One-tailed paired t-tests were conducted on the average standard error of the simulated estimates to test for changes in precision between unstandardized and standardized sampling frames.

The coefficient of variation (CV) was computing by dividing the standard deviation by the mean trees per segment for conventional and standardized sampling frames in all research localities. The minimum number of street segments (n_{i}) that must be sampled to achieve a confidence interval half width of 20% of the population estimate was calculated from the coefficient of variation and adjusted using a finite population correction (n_{min}). One-tailed paired t-tests were conducted to test for reductions in minimum sample size as afforded by street segment standardization.

$$CV = \frac{s}{\bar{x}} \times 100$$

where *s* = the standard deviation of the number of trees per segment
$\bar{x}$ = the mean tree count per segment

$$n_i = \left(\frac{CV \times t}{AE}\right)^2 \qquad \text{(Ziede 1980)}$$

where *CV* = coefficient of variation for tree count per segment
AE = desired ratio of confidence interval half-width to population estimate
t = Student's t-value ($\alpha = 0.05$)

$n_{min} = n_i \times \frac{N}{(N+n_i-1)}$ (Berenson 2004)

where n_i = uncorrected minimum n

N = total segments in the sampling frame

A common criterion for precision of street tree population estimates is a relative standard error not larger than 10% of the estimate. The minimnum number of sampled street segments was divided by the total segment count in the sampling frame to calculate percent sampling intensity. Conversely the standard error that could be achieved using a 5% segment sampling intensity (STREETS recommends 3-6% intensity) was calculated for unstandardized and standardized sampling frames (STRATUM 2012).

In order to further assess the accuracy of unstandardized and standardized sampling frames, the percent bias was calculated using the complete datasets and all frames in all localities. The mean population estimate was computed by multiplying the mean tree count per street segment by total sampling frame segments. The true population value was subtracted from the mean population estimate to calculate the bias. Relative biased was then calculated by dividing the bias by the true population total. The mean population estimate for Splice-Dice-Discard sampling frames was adjusted to correct for negative bias due to the removal of snippets. Adjusted mean population estimates were calculated as the product of mean trees per segment and the total segments in the sampling frame before undersized street segments were removed. Differences in percent bias, CV, and required labor time between Splice-Dice-Discard and Splice-Dice-Resplice sampling procedures were assessed during the pilot study using paired t-tests to determine which technique would be more preferable.

A post hoc investigation of variables that might be used to estimate the effects of standardization was conducted based on the hypothesis that the effects of standardization were reflective of available growing space and the level of uniformity of street tree distribution along linear segments. Impervious surface percentage, street tree density, and maximum unstandardized street length were analyzed as potential variables in predicting the effects of standardization to 91 m using linear regression. Ninety-one meter sampling frames were selected for this analysis based on the results of the CV analysis (Section 4.4). Street tree density was calculated by dividing the number of street trees in each municipality by the total miles of street in each sampling frame. Percent impervious surface values were found using downloadable GIS shapefiles provided by the US Forest Service (US Forest Service 2010).

CHAPTER 4 – RESULTS

4.1. Preliminary Results of Splice-Dice-Discard and Splice-Dice-Resplice Procedures

Because there were concerns that the street segment standardization procedure might bias the street tree inventory sampling method, the bias was evaluated by multiplying the mean per-segment tree count by the total segments in each sampling frame. The bias in estimates of tree count derived from street segment sampling frames created using the Splice-Dice-Discard standardization procedure was highly significant ($P \leq 0.0001$) and substantially greater than those of the Splice-Dice-Resplice procedure (Table 4.1). When paired by segment standardization length, Splice-Dice-Resplice procedures had minimal yet significant increases in tree count CV (1-4%) compared to Splice-Dice-Discard procedures in the same municipality (46m: $P=0.0465$; 91m: $P=0.0072$; 137m: $P=0.0375$). The Splice-Dice-Discard standardization procedures were abandoned following this preliminary analysis because to their relatively high biases and low level of difference in CV compared to Splice-Dice-Resplice frames. Unstandardized and Splice-Dice-Resplice sampling frames were unbiased but had experimental error resulting from rounding and hypothesized joining of single tree records to multiple streets (Sections 2.3 and 5.1.2). This error was reflected in bias calculations at <1% (Table 4.2) and does not apply to recommended field procedures.

Table 4.1. Comparison of rounding and experimental error of Splice-Dice-Resplice sampling frames with bias of Splice-Dice-Discard and Splice-Dice-Discard (adjusted) sampling frames. Data are sorted by locality and standardized length.

Municipality	Standardized Length (m)	Splice-Dice-Resplice	Splice-Dice-Discard	Splice-Dice-Discard (adjusted)
		Experimental Error	% Bias	
Ann Arbor	46	0.12	-1.04	1.16
	91	0.02	-3.70	4.04
	137	0.05	-6.88	8.43
Berkeley	46	0.18	-0.29	0.96
	91	0.08	-1.49	3.40
	137	0.09	-2.80	5.69
Dillon	46	0.11	0.34	1.10
	91	0.08	-0.25	2.55
	137	0.06	-1.28	3.82
Lakeland	46	0.18	-0.72	0.92
	91	0.09	-3.41	2.93
	137	0.10	-5.58	6.57
Lodi	46	-0.04	-0.37	2.28
	91	0.06	-1.67	7.40
	137	-0.03	-5.53	14.72
Santa Monica	46	0.00	-0.33	1.14
	91	0.12	-1.06	4.21
	137	0.06	-2.07	8.39
Southgate	46	0.93	-0.81	1.67
	91	-0.38	-4.55	5.24
	137	0.23	-9.46	10.57

Table 4.2. Experimental error of unstandardized and Splice-Dice-Resplice standardized segments.

Municipality	Unstandardized	46 m	91 m	137 m
	— % Experimental Error —			
Ann Arbor	0.06	0.12	0.02	0.05
Berkeley	0.15	0.18	0.08	0.09
Dillon	0.04	0.11	0.08	0.06
Lakeland	0.06	0.18	0.09	0.10
Lodi	0.02	-0.04	0.06	-0.03
Santa Monica	0.27	0.00	0.12	0.06
Southgate	0.26	0.93	-0.38	0.23

Standardizing segment length effectively narrowed and normalized the range of segment lengths in every sampling frame (Figure 4.1). Distributions of unstandardized segment length had similar interquartile ranges but had much more extreme minimum and maximum values (Table 4.3).

Table 4.3. Distributions of unstandardized segment lengths in each research locality.

Municipality	Sampling Frame (N)	Min	25th Percentile	50th Percentile	75th Percentile	Max
		— Street Segment Length (m) —				
Ann Arbor	5,107	0.01	52.82	95.73	147.87	1527.86
Berkeley	4435	0.24	44.92	90.18	120.11	1867.51
Dillon	624	1.56	88.15	111.92	116.94	399.08
Lakeland	7,362	0.14	54.75	97.14	161.59	2240.30
Lodi	822	0.02	46.01	77.30	120.98	449.54
Santa Monica	3,730	0.01	52.45	63.13	148.53	888.23
Southgate	172	1.01	39.21	70.62	156.77	432.01

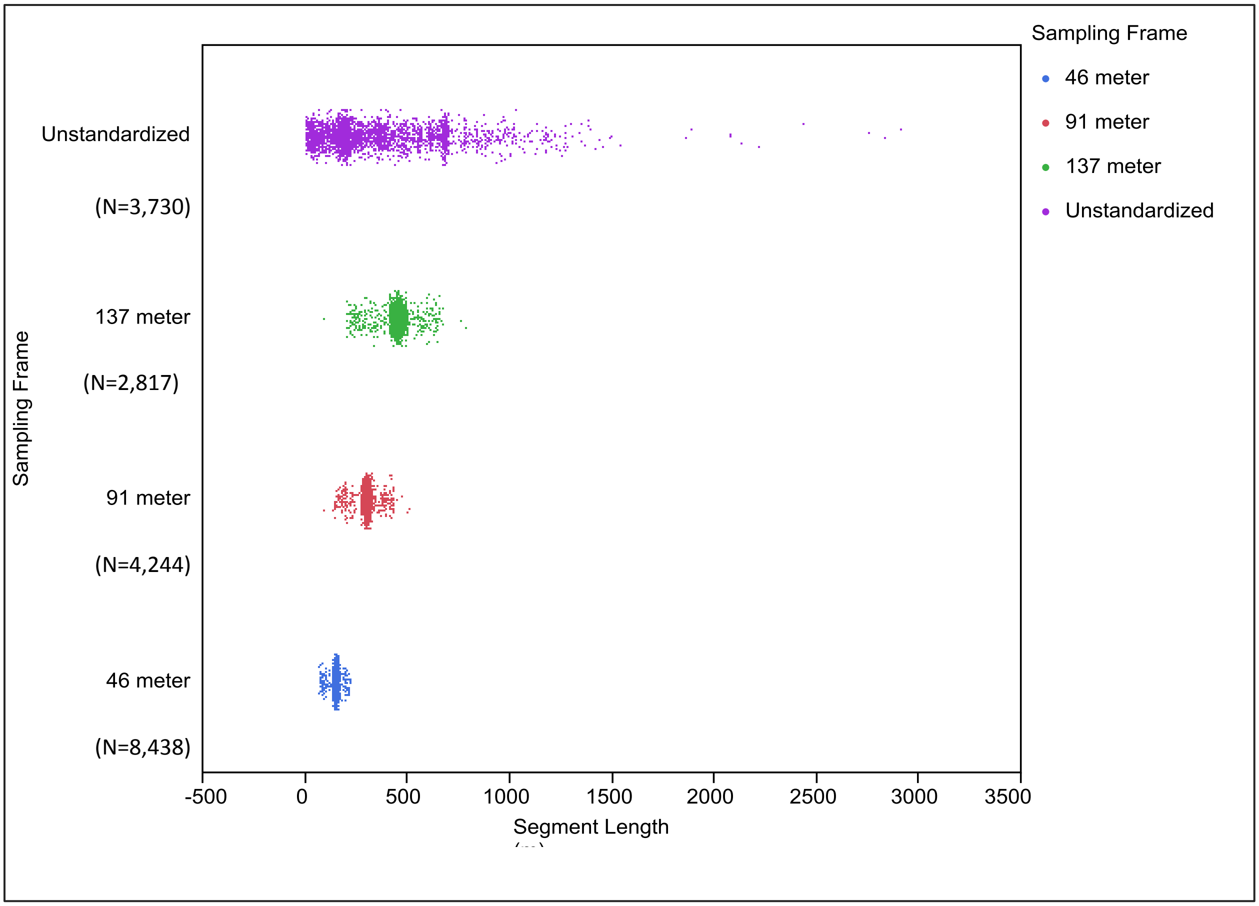


Figure 4.1. An example of standardization effects on street segment length distribution in Santa Monica, CA, which had the median maximum unstandardized sampling length among the seven research localities.

Across localities, the relationship between segment length and tree count in unstandardized sampling frames was highly significant ($P \leq 0.0001$) and the strength of relationship ranged from weak to moderate ($0.0581 \leq R^2 \leq 0.4426$). This potential source of sampling bias was substantially mitigated through segment length standardization as evidenced by the consistently weak relationship (all $R^2 \leq 0.0336$) between segment length and tree count in the standardized sampling frames (Table 4.4).

Table 4.4. Coefficient of determination (R^2) for linear relationship between street length and tree count for unstandardized and standardized sampling frames in each research municipality. Asterisks(*) denote significant relationships at α=0.05.

	Street Segment Sampling Frame			
	Unstandardized	**46 meter**	**91 meter**	**137 meter**
Municipality	**——— Coefficient of Determination (R^2) ———**			
Ann Arbor	0.2848*	0.0037*	0.0116*	0.0336*
Berkeley	0.3123*	0.0001	0.0161*	0.022*
Dillon	0.0581*	0.0018	0.0003	0.0091*
Lakeland	0.1047*	0.0003*	0.002*	0.0058*
Lodi	0.4426*	0.0053*	0.0138*	0.0209*
Santa Monica	0.3042*	0.0025*	0.0019*	0.0023*
Southgate	0.1177*	0.0004	0.004	0.0009

4.2. Processing Time Analysis

The amount of time spent processing unstandardized street segment sampling frames using the Splice-Dice-Resplice standardization procedure was tracked for each combination of locality and segment standardization length. Total processing time ranged from 1.08 hours in Dillon, SC to 17.5 hours in Lakeland, FL (Figure 4.2). The pace of standardization ranged from 129 to 624 segments per hour and 14 to 66 kilometers per hour. Across localities and standard lengths, the median pace of standardization was 417 segments per hour or 47.5 kilometers per hour.

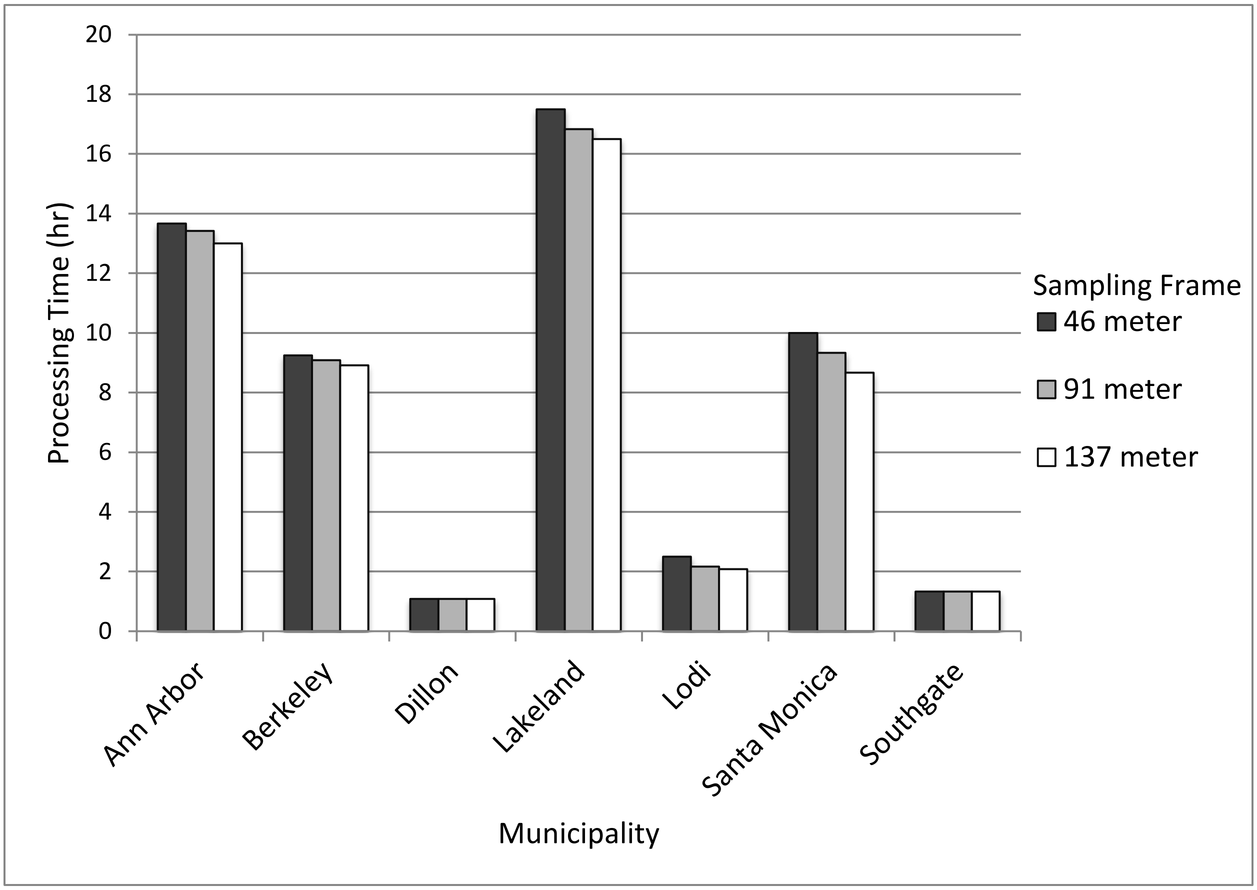


Figure 4.2. Processing time required to standardize street segment sampling frames to lengths of 46, 91, and 137 meter in seven US municipalities.

A significant positive relationship was found between processing time and both unstandardized segment count (R^2=0.98, P<0.0001) and street mileage (R^2=0.97, P<0.0001). When paired by locality, standardizing 137 m sampling frames took significantly less time than standardizing 91 m (P=0.0441) and 46 m (P=0.0297) sampling frames. Likewise, the 91 m sampling frames took significantly less time to standardize than 46 m sampling frames(P=0.0304). However, these differences in time were minor, with maximum difference in labor time between sampling frames in the same locality of 1.33 hours.

4.3. Street Tree Inventory Simulations

Replicated simulations of street tree inventories were conducted in a sub-set of the investigated localities: Ann Arbor, Dillon, Lodi, and Santa Monica. Simulations using both unstandardized and standardized sampling frames produced accurate estimates of the street tree population total in all localities. The true street tree population was within the 95% confidence interval bounds of the simulated tree population estimates for all sampling frames (Table 4.5). However, standardization reduced the standard errors of population estimates (46m: P=0.0294; 91m: P=0.0540; 137m: P=0.0439), effectively narrowing the confidence interval bounds (Table 4.5). No statistically significant differences were found between population estimates and true population values in any sampling frame. Likewise, no significant difference was found amongst the estimates generated by the various sampling frames when compared within each locality.

4.4. Analysis of Coefficient of Variation

Coefficients of variation of tree count per street segment for unstandardized segment sampling frames ranged 125-264% and averaged 164% (Table 4.6). Standardization of street segments reduced the coefficient of variation for all sampling frames except the 46 m standardization in Dillon and Lakeland. Standardization of street segments to lengths of 91 m and 137 m significantly reduced tree count CV relative to the unstandardized sampling frames (P=0.0006 and P=0.0002 respectively). Reductions in coefficient of variation were inversely related to standardized segments length (P=0.0174, R^2=0.2630).

Table 4.5. Results of replicated street tree inventory simulations (*n* = 30) using an unstandardized street segment sampling frame and three standardized street segment sampling frames in four US municipalities. P-values are reported for tow-tailed, one-sample *t*-test comparing the estimated tree population to the true tree population. Discrepancies in the true population reported for unstandardized and standardized sampling frames resulted from removal of residual street segments that intersected the municipal boundary and their associated street trees during dataset processing.

Municipality	P-value	True Population	Mean Estimate	Standard Deviation	Lower CI Bound (95%)	Upper CI Bound (95%)
Ann Arbor						
46 m	0.09	40,254	40,718	1,442	40,180	41,256
91 m	0.12	40,254	40,942	2,333	40,071	41,813
137 m	0.86	40,254	40,325	2,152	39,522	41,129
Unstandardized	0.62	40,321	40,046	3,015	38,920	41,171
Dillon						
46 m	0.08	1,837	1,705	394	1,558	1,852
91 m	0.36	1,837	1,910	431	1,749	2,071
137 m	0.80	1,837	1,861	506	1,672	2,050
Unstandardized	0.78	1,840	1,813	519	1,619	2,007
Lodi						
46 m	0.64	5,143	5,176	384	5,033	5,320
91 m	0.15	5,143	5,344	751	5,064	5,625
137 m	0.60	5,143	5,221	795	4,924	5,518
Unstandardized	0.33	5,153	5,404	1,384	4,887	5,920
Santa Monica						
46 m	0.38	31,643	31,930	1,770	31,269	32,591
91 m	0.74	31,643	31,531	1,893	30,824	32,238
137 m	0.53	31,643	31,382	2,259	30,538	32,225
Unstandardized	0.97	31,694	31,667	3,637	30,309	33,025

Table 4.6. The coefficient of variation (CV) of the number of trees per street segment for the unstandardized and three standardized sampling frames in the seven research municipalities. Three metrics calculated based on CV and used to judge the effects of standardization are reported.

Municipality	**Sampling Frame**	***N***	**Coefficient of Variation (CV)**	**Minimum *n* Required for 10% Relative Standard Error**	**Sampling Intensity for 10% Relative Standard Error**	**Relative Standard Error at 5% Sampling Intensity**
Ann Arbor	Unstandardized	5,107	125%	146	2.9%	7.6%
	46 m	13,171	97%	90	0.7%	3.7%
	91 m	6,633	87%	72	1.1%	4.7%
	137 m	4,450	84%	67	1.5%	5.5%
Berkeley	Unstandardized	4,435	128%	152	3.4%	8.4%
	46 m	9,195	100%	95	1.0%	4.5%
	91 m	4,614	85%	68	1.5%	5.5%
	137 m	3,081	79%	59	1.9%	6.2%
Dillon	Unstandardized	624	153%	165	26.5%	26.8%
	46 m	1,448	175%	245	16.9%	20.0%
	91 m	721	146%	160	22.1%	23.7%
	137 m	485	134%	127	26.3%	26.6%
Lakeland	Unstandardized	7,362	209%	397	5.4%	10.6%
	46 m	20,974	211%	419	2.0%	6.4%
	91 m	10,519	183%	312	3.0%	7.8%
	137 m	7,061	169%	264	3.7%	8.8%
Lodi	Unstandardized	822	146%	164	20.0%	22.1%
	46 m	1,708	93%	79	4.6%	9.8%
	91 m	862	86%	66	7.6%	12.7%
	137 m	570	81%	57	10.0%	14.9%
Santa Monica	Unstandardized	3,730	154%	215	5.8%	11.0%
	46 m	8,438	109%	113	1.3%	5.2%
	91 m	4,224	100%	94	2.2%	6.7%
	137 m	2,817	94%	82	2.9%	7.8%
Southgate	Unstandardized	172	264%	137	79.7%	87.7%
	46 m	408	261%	247	61.6%	56.0%
	91 m	208	214%	134	68.0%	65.1%
	137 m	139	204%	94	74.3%	75.6%

The minimum street segment sample size (*n*) required from each sampling frame to achieve a standard error not exceeding 10% of the total tree population estimate was reduced for all standardizations except 46 m in Lakeland, FL; Dillon, SC; and Southgate, KY (Figure 4.3). Averaged across all localities, standardizing segments to 91 m and 137 m significantly reduced the minimum sample size relative to the unstandardized sampling frame (Figure 4.3) (91 m: P=0.0101; 137 m: P=0.0012). There were roughly one-third and one-half as many 137 m segments and 91 m segments as there were 46 m segment in any municipal sampling frame. The magnitude of difference in total segment count in each sampling frame was greater than that of minimum sampling size required (*n*). Thus trends in the effects of standardization on minimum *n* are inverted with respect to sampling intensity.

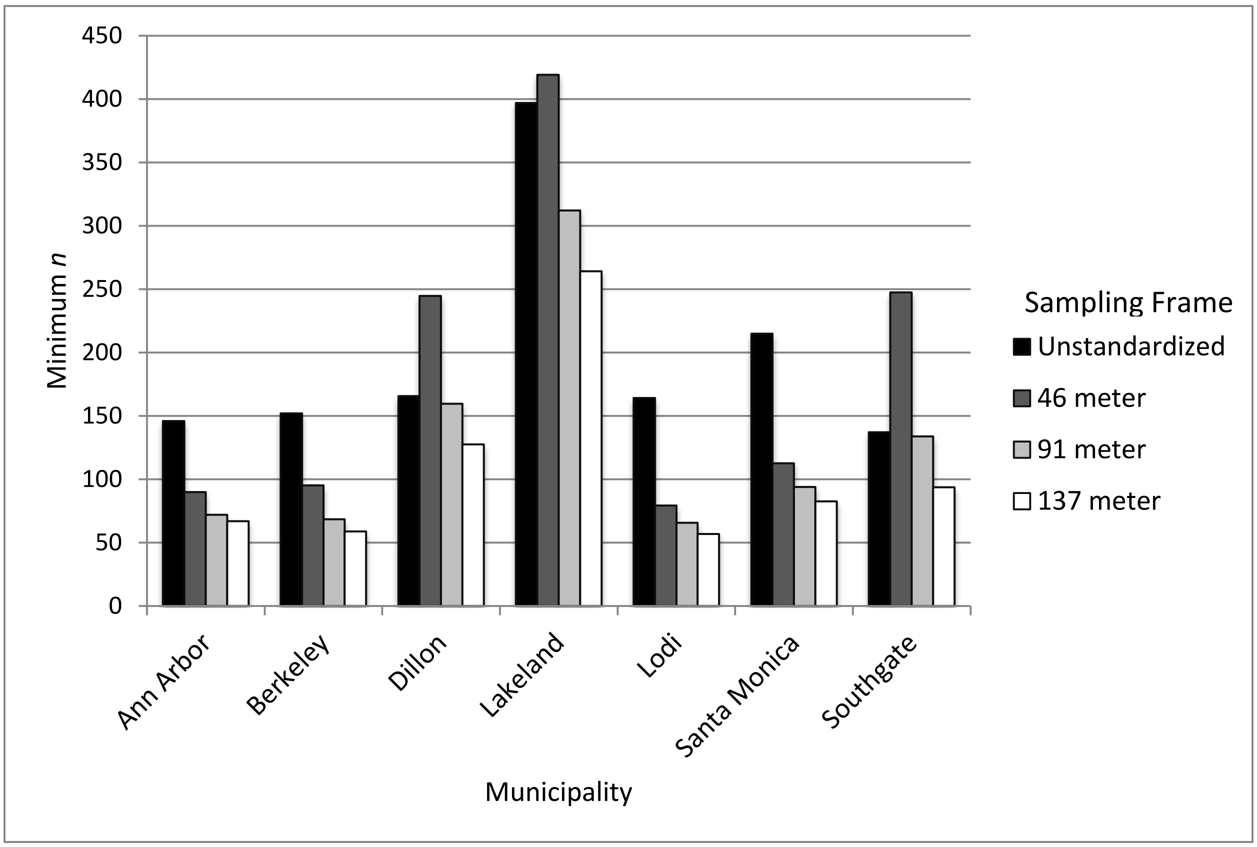


Figure 4.3. Minimum street segment sample size (n) required from each sampling frame to achieve a relative standard error not exceeding 10% of the total tree population estimate in seven US municipalities.

The sampling intensity (percent of street segments sampled) necessary to achieve a standard error not exceeding 10% of the total tree population estimate when sampling from unstandardized frames ranged from 3% in Ann Arbor to 80% in Southgate (Figure 4.4). In every instance, standardizing street segment sampling frames significantly reduced the minimum sample intensity compared to the unstandardized sampling frames (46 m: P=0.0093; 91 m: P=0.0100; 137 m: P=0.0211). Minimum sampling intensities were reduced by an average of 60% when sampling frames were standardized to 46 m, 45% at 91 m, and 33% at 137 m (Table 4.7). These percentages also reflect reductions in the total street length sampled because segment lengths in a standardized frame are close to the same size.

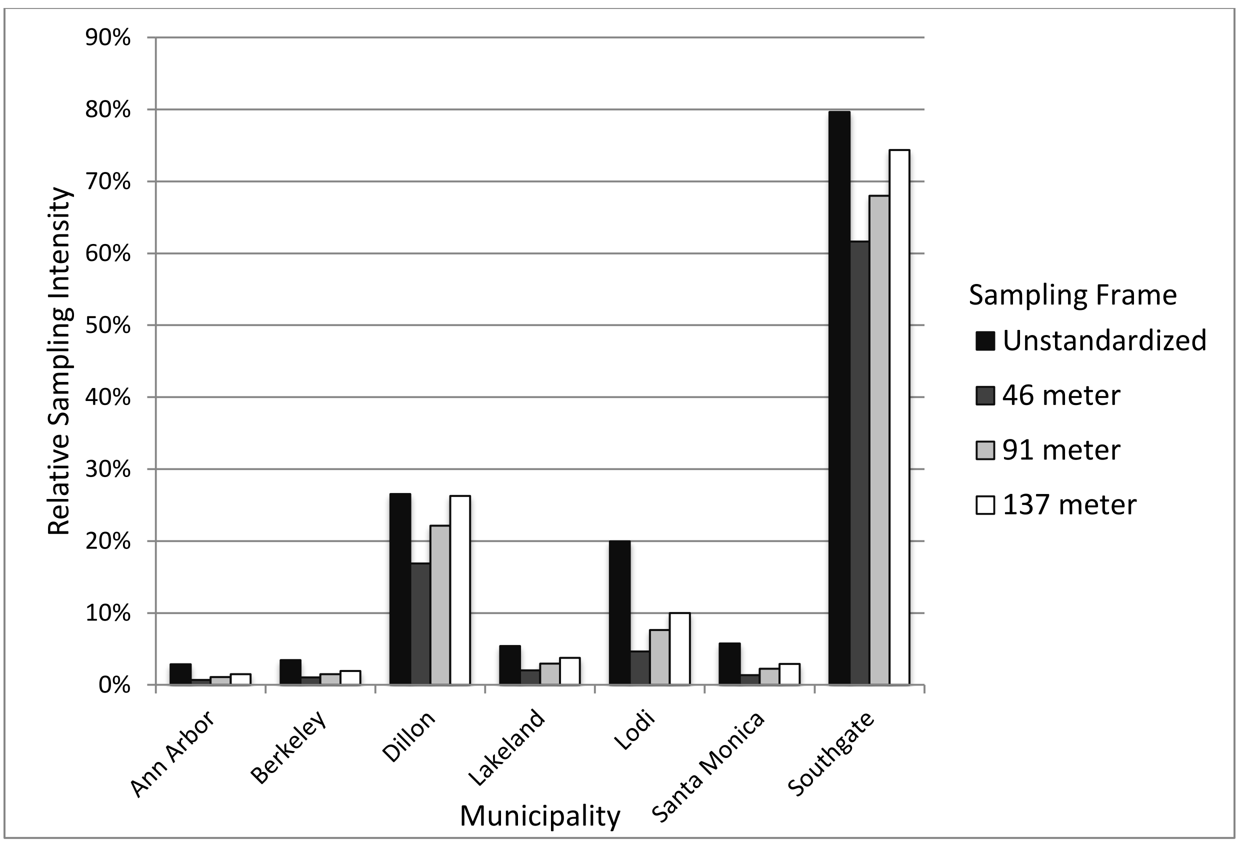


Figure 4.4. Relative sampling intensity (percent of street segments sampled) necessary to achieve a standard error not exceeding 10% of the total tree population estimate in seven US municipalities.

Table 4.7. Overall reduction in sampling intensity (percent of street segments sampled) necessary to achieve a relative standard error not exceeding 10% of the total tree population estimate for three standardized sampling frames compared to the unstandardized frames across seven US municipalities. P-values are reported for one-tailed, paired *t*-tests comparing the unstandardized sampling frame to each of the three standardized sampling frames**.**

	Street Segment Sampling Frame		
	46 meter	**91 meter**	**137 meter**
Municipality	**——— Reduction in Street Segment Sampling Intensity (%) ———**		
Ann Arbor	76	62	47
Berkeley	70	57	44
Dillon	36	17	1
Lakeland	63	45	31
Lodi	77	62	50
Santa Monica	77	61	49
Southgate	23	15	7
P-value	0.0093	0.0100	0.0211

The relative standard error resulting from a 5% sampling intensity of unstandardized street segment sampling frames ranged from 8% in Ann Arbor to 88% in Southgate (Figure 4.5). In every instance, standardizing the sampling frames significantly reduced the relative standard error compared to unstandardized sampling frames (46 m: P=0.021; 91 m: P=0.0236; 137 m: P=0.0189). On average, relative standard errors were reduced by 44% when sampling frames were standardized to 46 m, 31% at 91 m, and 21% at 137 m (Table 4.8). Sampling frames standardized to 91 m and 137 m also had significant reductions in CV, minimum sample size (*n*), minimum sampling intensity, and relative standard error compared to unstandardized sampling frames.

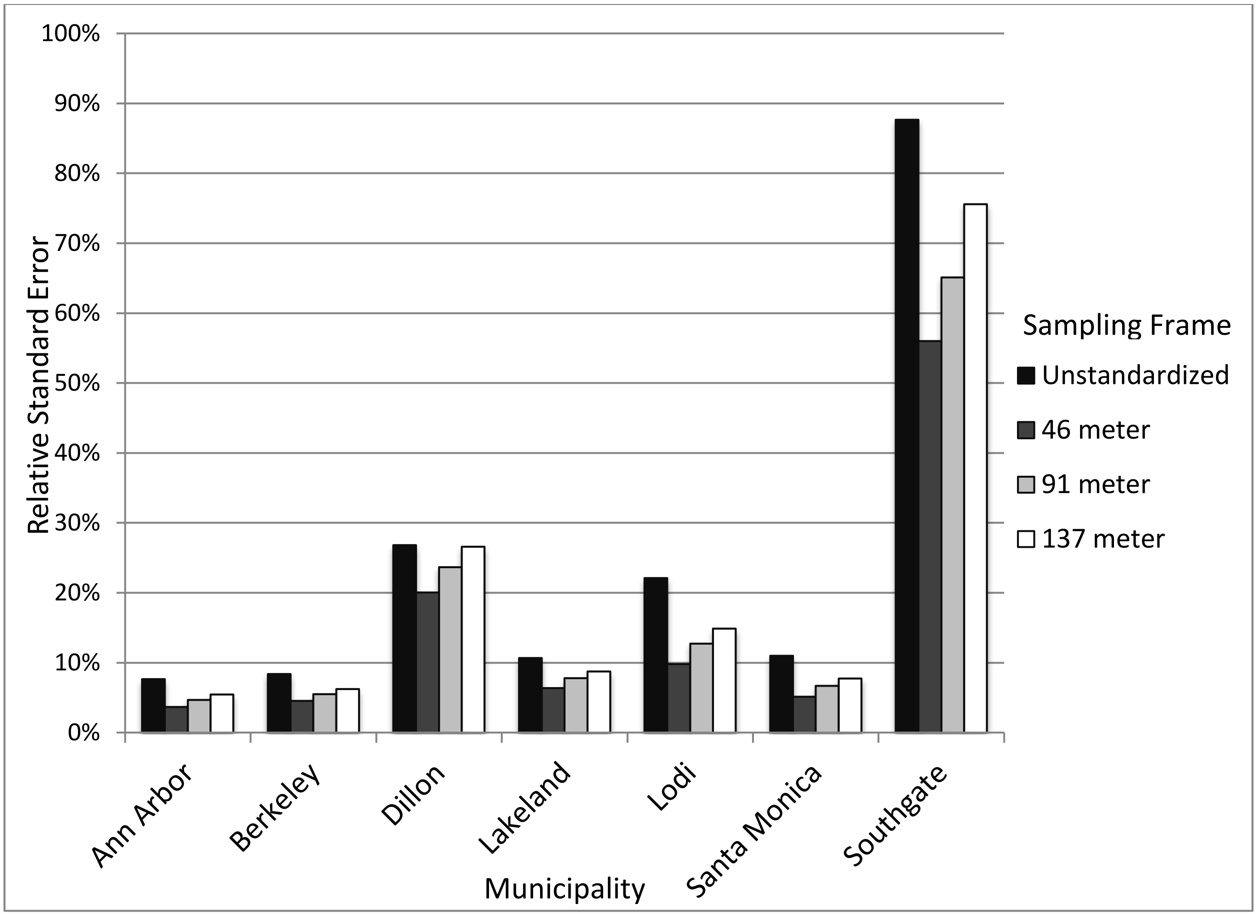


Figure 4.5. Relative standard error of the total tree population estimate achievable when conducting a sample street tree inventory using a 5% street segment sampling intensity in seven US municipalities.

Table 4.8. Overall reduction in relative standard error resulting from standardization of street segment length when conducting a street tree inventory using a 5% street segment sampling intensity in seven US municipalities. P-value is from two-tailed, paired t-test comparing each standardized sampling frame to the unstandardized sampling frame across all municipalities.

Sampling Frame	Ann Arbor	Berkeley	Dillon	Lakeland	Lodi	Santa Monica	Southgate
46 meters (P=0.0210)	52%	46%	25%	40%	56%	53%	36%
91 meters (P=0.0236)	39%	35%	12%	27%	43%	39%	26%
137 meter (P=0.0189)	28%	26%	1%	18%	33%	29%	14%

4.5. Prediction of the Magnitude of the Effects of Standardization

The effects of citywide impervious surface and street tree stocking (trees per unit street length) on overall reductions in sample intensity at 10% relative standard error were evaluated using linear regression (Figures 4.6.a and 4.6.b). Citywide impervious surface coverage and street tree stocking were shown to be significant parameters for predicting the effect of standardizing sampling frames to 91m. Percent impervious surface and street tree density were positively correlated to the variation in percent reduction in sampling intensity. Percent impervious surface and street tree density explained 67% (P=0.0249) and 80% (P=0.0064) of the variation in reductions in sampling intensity.

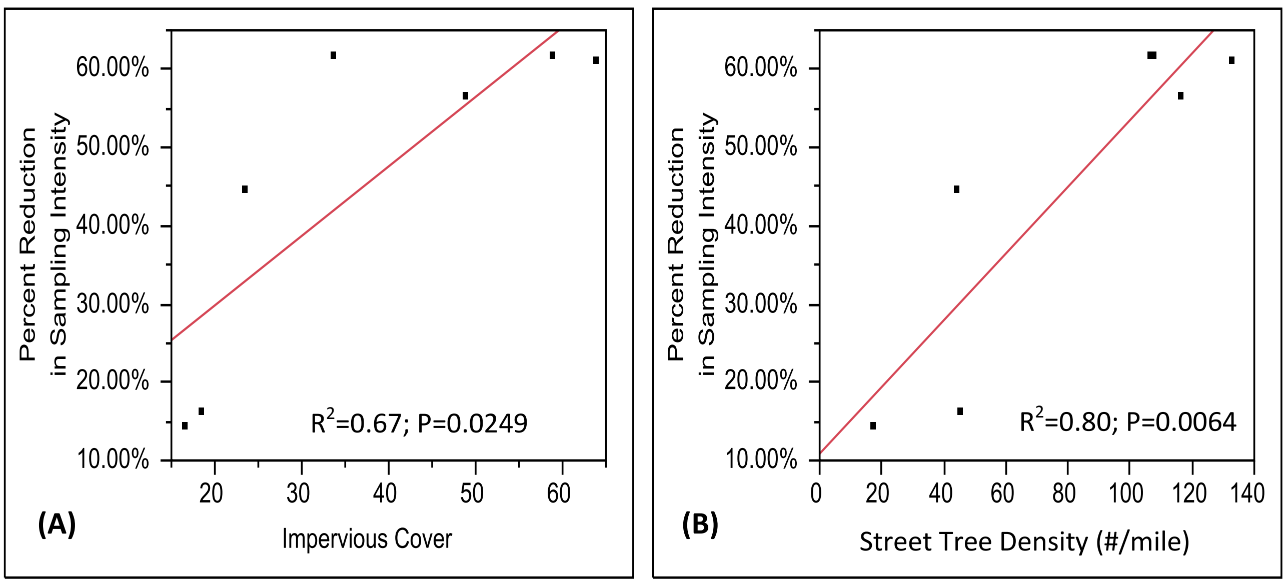


Figure 4.6. Relationship of (A) impervious cover and (B) street tree density to reduction in sampling intensity required to achieve a 10% relative standard error for the total street tree estimate using a sampling frame standardized to 91 m. The seven points in each graph represent the seven research municipalities used in this investigation.

CHAPTER 5 – DISCUSSION

This chapter synthesizes and evaluates the results and limitations of this investigation, with particular emphasis on applications for street tree managers. It also identifies further research questions that have come to light over the course of this effort that should be pursued in an effort to further improve i-Tree Streets and the science and practice of street tree sampling.

5.1. Recommendations for Street Tree Managers

Based on the reduction in the strength of the relationship between trees tallied and street length from unstandardized to standardized segments, it is concluded that variation in trees per segment due to street length was effectively controlled through segment length standardization. Standardization to lengths of 91 m and 137 m led to significant improvements in every metric for assessing estimator precision (CV, minimum sample size (n), and sampling intensity) and relative standard error. Sampling frames using 91 m standardization length outperformed 137 m sampling frames in two of three practical metrics for assessing estimator precision (sample intensity and relative standard error). Replicated simulations of street tree inventories along with analysis of estimator bias revealed that accurate estimates of street tree population are obtainable using both standardized and unstandardized sampling frames. However, simulations also indicated that standardization reduces the variability between population estimates compared to estimates generated from unstandardized frames.

Based on these findings it is generally recommended that a sampling frame be standardized to 91 m prior to conducting a STREETS inventory. This will substantially improve the precision of STREETS estimates and their relative standard errors. By implementing a 91 m street segment, practitioners wishing to minimize costs will have to conduct less fieldwork to achieve a population estimate they deem acceptable. Alternatively, practitioners that have a set level of funding would standardize to improve the precision obtainable from a certain level of field effort.

Under certain circumstances, it may be justifiable to standardize segments to 46 m or 137 m. Standardization to 46 m or 137 m is only advisable if a substantial discrepancy is anticipated between travel and measurement time at 91 m. Travel and measurement time will vary based on plot size and sampling intensity. Mode of transportation will influence travel and measurement time in the field. If transportation and/or sampling are to be done by automobile, traffic regulations and chances for congestion have major impacts on the time traveled. If transportation by automobile will be minimized to travel between clusters of proximate segments to be visited on foot, consideration should be given to total distance travelled and number of trips at each combination standardized length and sample size. Aside from transportation factors, the count of trees per segment length in a municipality, number of inventory attributes tallied, intensity of physical inventory measurements, and other customary methodology aspects should be considered in measurement time estimations.

Standardizing sampling frames to 46 m street segment length leads to reductions in sampling intensity and relative standard error, but increases in minimum sample size (n) compared to 91 m frames. Standardizing to 46 m would be most advisable when mobility from plot to plot is high and estimated measurement time is high. Standardizing frames to 137 m will reduce the minimum sample size but increase sampling intensity and relative standard error compared to standardizing to 91 m. This would only be recommended when mobility from plot to plot is very low (foot travel) and attribute data can be collected with ease. Standardization of sampling frames to lengths outside of the 46 – 137 m range is not recommended as segments shorter than 46 m are not practical for field measurements and results indicate that standardizing to lengths greater than 137 m may increases sampling intensity and relative standard error in some cases as shown in this study.

The percent reduction in relative standard error obtainable from a 10% sampling intensity of a sampling frame standardized to 91 m segments can be estimated from the impervious surface coverage (US Forest Service 2010) using the following equations:

% Reduction in Sampling Intensity (%10 RSE) = 0.1206 + (0.0089% impervious cover)*

(R^2=0.67, P=0.0249)

% Reduction in Sampling Intensity (%10 RSE) = 0.1087 + (0.0043 mean trees per mile)*

(R^2=0.80, P=0.0064)

A reliable estimate of processing time required to standardize a sampling frame to the segment lengths studied in this investigation can be obtained using the equations below. Processing time will vary based on the competency of the GIS technician processing the sampling frame dataset. During this investigation, segments were standardized at an average pace of 417 segments per hour or 47.5 kilometers per hour.

*Processing Time (hrs) = 0.3851 + (0.0023*unstandardized segment count)*

(R^2=0.98, P≤0.0001)

*Processing Time (hrs) = 1.1099 + (0.029*unstandardized segment length (miles))*

(R^2=0.97, P≤0.0001)

Using these equations, municipal foresters can weigh the gains in reduced field time against the cost of office time spent standardizing.

5.2. Caveats and Limitations of This Investigation

Practitioners must understand the caveats and limitations of this investigation to properly interpret and apply the results. The scope of this investigation was to improve the STREETS sampling protocol without altering the inner working of the software so that practitioners can achieve accurate and precise estimates of total street tree populations more efficiently. There are also limitations associated with procedural learning curve, software used, the geospatial data, locality size, geographic representation, and the method used to associate trees to streets.

The processing time was recorded for the standardization of 46, 91, and 137 m sampling frames in every locality and used in the processing time analysis (see Section 4.2). No attempt was made to quantify or account for the effects of GIS processor competency on standardization times. The procedure was conducted a number of times before generating the standardized datasets, so the GIS technician was familiar with the protocol. The localities were standardized in ascending order of total street segment length in the unstandardized sampling frame, so this could be a potential source of bias in the processing time analysis.

Calculation of STREETS benefits estimates was not practical because of the relative complexities of replicating benefits modeling using third party software. All of the street tree inventory data was collected by the same organization using roughly the same definition of a street tree and a public street segment. In general, street segments were defined as TIGER/Lines not classified as private streets, alleys, interstates, trails, and other limited access rights-of-way, and street trees were records classified as such in the inventory database. However, it is expected that some degree of inconsistency exists between each inventory due to the commercial nature of the Davey Resources Group and its ability to cater to municipal goals and objectives in their particular tree inventory contracts.

Steps were taken to ensure that the datasets were as comparable as possible before beginning the investigation into street segment standardization. TIGER/Lines were obtained from the nearest available year to the year of inventory and those classified as limited access rights-of-way were removed from the datasets. Tree records classified as private, park, alley, or planting space and those ~61m or more from the nearest public centerline were removed from populations. However, it should be noted that field data entry and the methodology used to filter large datasets could have led to some level of discrepancy between the true population of inventoried street trees and segments and the depictions of those populations in finalized GIS datasets.

Local geospatial street centerline data could be used instead of TIGER/Line Shapelifes to produce standardized sampling frames. Practitioners should select and customize the sampling frames prior to standardization so public streets are accurately represented in the sampling frame.

As long as the TIGER/Line sampling frame and the local sampling frame represent roughly the same population of street segments the final product will essentially be the same regardless of which source data is used. If the two perspective sampling frames have major differences in the streets that are represented; practitioners are advised to select the sampling frame that is most representative of streets existing in the public right-of-way. Based on the results of the processing time analysis (Section 4.2) there could be an increase in associated processing time if on sampling frame has substantially more segments. Priority should be given to the sampling frame with the fewest segments when sampling frames are similar in representation and coverage.

It is also possible that some of the TIGER/Lines within the area of interest were used in the investigation but were not sampled in the field because they did not exist at the time or were otherwise deemed inappropriate. This would lead to segments that were not sampled being tabulated in this investigation as segments with no trees. It is expected that such error would lead to a dilution in the relationship between street length and tree count and thus would also dilute the magnitude of the effects of standardization. Based on this assumption, it is reasonable to conclude that the effects of standardization in a real world situation would be greater than those identified in this investigation.

Some limitations exist in the size range of localities represented in these inventories. Large datasets such as the obtained from Charlotte, NC and Minneapolis, MN could not be used in this investigation due to time constraints associated with performing six standardizations on each dataset. The largest locality that could be examined was Lakeland with a land area of 168 square kilometers and 959 kilometers of public streets. Southgate had the smallest land area at 3.5 square kilometers and the shortest total length of public streets at nearly 19 kilometers.

The seven localities used in this investigation span the Continental US, but there were large geographic expanses that did not contain a research locality. There was reasonable coverage of the Atlantic seaboard, Midwest, and southern Pacific seaboard. However, areas in the Pacific Northwest, desert Southwest, Great Plains, and New England were underrepresented. These areas should be considered in future research since the drivers of street tree density and distribution in these areas may differ somewhat from research localities based on geographic differences and environmental and human impacts.

During data formatting, street segments were spatially assigned to the closest street trees and a total count of the existing street trees was generated in the GIS attribute table of the sampling frame. Using this technique in the methodology may have resulted in some inconsistencies between geospatial representation and ground truth. In a small number of instances, when segments were spatially assigned to the nearest trees, the tree counts appearing in the street segment attribute tables were slightly higher (≤ 75) than the true number of trees (see Section 3.1.3). This phenomenon did not occur in any of the segment populations in Dillon or Southgate; or the standardized populations in Lodi. It is hypothesized that the additional trees are the result of the same tree being spatially related to multiple segments due to it being exactly or very close to the same distance, depending on rounding, from two segments. This would help explain its occurrence in the larger datasets. There is also some chance that trees that were joined to single segments were joined to segments other than those they would have been inventoried on in the field. This could have occurred as the result of GPS error, abnormal rights-of-way, road centerline geographic error, medians, or other discrepancies between what segment the tree is closest to and what segment it actually fronts on. Due to steps taken in the dataset formatting phase (see Section 3.1.3), these occurrences were probably rare. These effects can be assumed to be minor, but must be considered when interpreting the results.

5.3. Future Inquiries to Improve i-Tree Streets and Street Tree Sampling

The discovery of improvements in STREETS precision through the standardization of street segment sampling frames has prompted additional research questions and highlighted the need to continue improving and developing pre-sampling protocols. This section will discuss avenues for continuing to study and improve street tree sampling.

When sampling at the same intensity, a higher level of confidence can be placed in street tree population estimates, benefits estimates, and value estimates from standardized street segment sampling frames because benefits calculations are based on the species and sizes of trees in a more precisely estimated population. However, it is important to point out that samples of standardized street segments can produce equal or more accuracy in estimating predominant species and size classes when compared to sampling unstandardized street segments at the same intensity. If a future investigation yielded such results, it would further validate the advantages of sampling with segments of standardized length. Such an assessment could be conducted using ArcGIS and STREETS subpopulation estimators (i-Tree Cooperative STRATUM 2012). Standardization could affect STREETS reports on other tree population attributes such as maintenance tasks and invasive pest vulnerabilities. However, an investigation of these effects is not considered as high a priority because they do not contribute to the appraised value of street tree populations.

While this study does generally recommend standardizing segment length to 91 m and provides suggestions for determining if another standardized length is more suitable, it is not an optimization study because the cost of gains in precision are not measured in units of field time. If it is confirmed that standardization does not negatively impact streets benefits estimation, it would be desirable to conduct a segment length optimization study using the formatted datasets from this project. An optimal routing and networking project could be conducted in ArcGIS to simulate sample inventories at varying intensities and segment sizes to determine optimum street segment length based on the marginal gain in precision per unit field work. Simulated inventories could assess the effects of varying delays in travel time and measurement time on optimal plot size and would be a more detailed assessment of which length is best suited for a certain array of sampling conditions.

Street segment standardization procedures require an intermediate knowledge of GIS software at a minimum and can take roughly two workdays for localities the size of Lakeland, FL. Automation of sampling frame standardization procedures within ArcGIS would increase the marginal gain of standardizing street segments. An ArcGIS toolbox module could be produced with effective tools for splicing segments (Appendix C, Step 1.3), dicing spliced sampling frames (Appendix C, Step 2), and re-splicing snippets (Appendix C, Step 3). An automated module for splicing, dicing, and resplicing segments would allow users to standardize centerline datasets by manually uploading the dataset and working through three automated steps using an installable toolbar. Practitioners with a basic knowledge of GIS principles would be able to produce standardized sampling frames in much shorter time spans using. Such an innovation would ensure that municipal foresters would not have to gauge the potential statistical benefits of sampling frame standardization against the time required to conduct the standardization procedures.

Standardization controlled the 6% to 44% variation in trees per segment in unstandardized sampling frames explained by street length, meaning unidentified variables account for over half of the variation in trees per segment. The results of this investigation indicate that street tree density and percent impervious surface are statistically significant estimators of the effects of standardization. A more thorough analysis of the drivers of street tree stocking (tree count per segment or segment length) would be valuable to optimizing street tree sampling. Assessing the influences of available growing space, municipal regulation, property level management, and environmental characteristics on stocking levels within the public rights-of-way would serve as a viable starting point for such a study. If other influential variables were identified, a procedure similar to the one identified by Marques and Buckland (2003) could be used to incorporate into STREETS as an expansion tool.

Potential methods of improving street tree sampling exist beyond the realm of i-Tree Streets. Tree count per street segment generally exhibits a right-skewed distribution with the mode tree count being zero. From this it is hypothesized that there may be value in attempting to estimate street tree populations using the median, a binomial distribution, or cluster sampling.

Such a method, if effective and economical, could be incorporated into more complex models of street tree populations such as those used by STREETS.

As i-Tree Streets continues gain popularity among municipal foresters it will become increasingly important that the results of street tree inventories be accurate, precise, and comparable from locality to locality. The adoption of standardization techniques by municipal forestry practitioners will allow for savings in time and resources allocated to field sampling and allow for a more consistent comparison of inventories conducted in different locations. Standardizing street segments is the first step towards improving i-Tree Streets inventories and many opportunities exist to build upon the findings of this study. However, street tree sampling is a relatively young practice when compared to the science sampling of traditional forest populations and much of the variation in street tree distributions and population is still attributable to unknown factors. Sweeping changes in street tree inventory protocol will only become more difficult to implement as the current procedures are adopted and established. Researchers and scientists from various disciplines of natural resources are encouraged to pursue improvements in street tree sampling while the realm of possibilities and opportunity for improvement is at its greatest. Street tree sampling practitioners are encouraged to remain open to studying and potentially adopting sampling procedures from disciplines with highly developed sampling strategies, regardless of their compatibility current protocol.

LITERATURE CITED

Avery, T. E., and Burkhart, H.E. (2001). Forest measurements 5th ed. Boston, MA, McGraw Hill. .

Berenson, M. L., D.M. Levine, T.C. Krehbiel (2004). Basic Business Statistics. Upper Saddle River, NJ, Prentice Hall.

Bond, J., and B. Buchanan (2006). Best Management Practices-Tree Inventories

Champaign, IL, International Society of Arboriculture.

Buckland, S. T., Anderson, D.R., Burnham, K.P. and and J. L. Laake (1993). Distance sampling: Estimating abundance of biological populations. Encyclopedia of Biostatistics. London, Chapman and Hall.

Buckland, S. T., D.R. Anderson, K.P Burnham, J.L. Laake, D.L. Borchers, and L. Thomas (2001). Introduction to Distance Sampling. Oxford, Oxford University Press.

Climate. (2012). "i-Tree Streets Climate Zone Map." i-Tree Archives Retrieved August 2010, from http://www.itreetools.org/resources/archives.php.

Day, S. D., P.E. Wiseman, S.B. Dickinson, and J.R. Harris (2010). "Tree root ecology in the urban environemnt and implications for a sustainable rhizosphere " Arboriculture and Urban Forestry **36**(5): 193-205.

Dwyer, J. F., D.J. Nowak, M.H. Noble, and S.M. Sisinni (2000). Assessing our nation's urban forests: Connecting people with ecosystems in the 21st century. General technical Report PNW-490. Portland, OR, USDA Forest Service.

Dwyer, J. F., E.G. McPherson, H.W. Schroeder, and R.A. Rowntree (1992). "Assessing the benefits and costs of the urban forest." Journal of Arboriculture **18**(5): 227-234.

Engeman, R. M., and Sugihara, R.T. (1998). "Optimization of variable area transect sampling using Monte Carlo simulation." Ecology **79**(4): 1425-1434.

Engeman, R. M., R.M. Nielson, and R.T. Sugihara (2005). "Evaluation of optimized variable area transect sampling using totally enumerated field data sets." Environmentrics **16**: 767-772.

ESRI Inc. (2010). "ArcGIS 10.0." Retrieved August 2011.

Gerik, T. J., and Rush, C.M. (1985). "Optimizing plot size for field studies of Phymatotrichum root rot of cotton " Phytopathology **75**(2): 240-243.

i-Tree (2011). i-Tree Streets User's Manual V. 4.0, i-Tree Cooperative Initiative.

Jaenson, R., N. Bassuk, S. Schwager, and D. Headley (1992). "A statistical method for the accurate and rapid sampling of urban street tree populations." Journal of Arboriculture **18**(4): 171-183.

Jonnes, J. (2011) "What is a tree worth?" The Wilson Quarterly.

Kielbaso, J. J. (1990). "Trends and issues in city forests." Journal of Arboriculture **16**(3): 69-76.

Maco, S. E., and E.G. McPherson (2003). "A practical approach to assessing structure, function, and value of street tree populations in small communities." Journal of Arboriculture **29**(2): 84-97.

Marques, F. F. C., and S.T. Buckland (2003). "Incorporating covariates into standard line transect analyses." Biometrics **59**: 924-935.

McCormick, M. I., and J.H. Choat (1987). "Estimating total abundance of a large temperate-reef fish using visual strip transects." Marine Biology **96**(4): 469-478.

McPherson, E. G. (2010). "Selecting reference cities for i-Tree Streets." Arboriculture and Urban Forestry **36**(5): 230-240.

McPherson, E. G., D.J. Nowak, G. Heisler, S. Grimmond, C. Souch, R. Grant, and R.A. Rowntree (1997). "Quantifying urban forest sturcture, function, and value: the Chicago Urban Forest Climate Project." Urban Ecosystems **1**: 49-61.

McPherson, E. G., J.R. Simpson, P.J. Peper, and Q. Xiao (1999). "Benefit-cost analysis of Modesto's municipal urban forest." Journal of Arboriculture **25**(5): 235-247.

McPherson, E. G., J.R. Simpson, P.J. Peper, S.E. Maco and Q. Xiao (2005). "Municipal forest benefits and costs in five US cities." Journal of Forestry **103**: 411-416.
Miller, R. W. (1997). Urban Forestry: Planning and Managing Urban Greenspaces. Upper Saddle River, NJ, Prentice Hall.
Mohia, P., L. Smith, F. Valentine, W. Stiteler, T. Elias, and R. Westfall (1978). Structure of urban street tree populations and designs for estimating their parameter. Metropolitan Tree Improvement Alliance (METRIA) Proc., 1**:** 28-43.
Nowak, D. J., D.E. Crane, and J.F. Dwyer (2002). "Compensatory values of urban trees in the United States." Journal of Arboriculture **28**(4): 194-199.
Nowak, D. J., J.T. Walton, J.C. Steven, D.E. Crane, and R.E. Hoehn (2008). "Effect of plot and sample size on timing and precision of urban forest assessment." Arboriculture and Urban Forestry **34**(6): 386-390.
Nowak, D. J., M.H. Noble, S.M. Sisinni, and J.F. Dwyer (2001). "Assessing the US urban forest resource." Journal of Forestry **99**(3): 37-42.
Parker, K. R. (1979). "Density estimation by variable area transect." Journal of Wildlife Management **43**: 484-492.
Patil, S. A., K.P. Burnham, and J.L. Kovner (1979). "Nonparametric estimation of plant density by the distance method." Biometrics **35**(3): 597-604.
Poultney, R., Riley, J., and Webster, R. (1997). "Optimizing plot size and shape for field experiments on terraces. ." Experimental Agriculture **33**(1): 51-64.
Särndal, C., B. Swensson, and Jan Wretman (2003). Model Assisted Survey Sampling. New York, NY, Springer-Verlag New York Inc. **9-12 pp.**
Smiley, E. T. a. F. A. B. (1988). "Options in street trees inventories." Journal of Arboriculture **14**(2): 36-42.
STRATUM. (2012). "STRATUM Population Estimators and Standard Error Equations for Simple Random Street Segment Sampling." i-Tree Archives Retrieved August 2010, from http://www.itreetools.org/resources/archives.php.
Thompson, R. P. a. J. J. A. (2000). The state of urban and community forestry in California: Status in 1997 and trends since 1988. California Deapartment of Forestry and Fire Protection, Technical Report No. 9. San Luis Obispo, Urban Forest Ecosystems Institute, California Polytechnic State University**:** pp. 48.
TIGER/Line Shapefiles. (2011). "TIGER/Line Shapefiles." Retrieved August 2011, from http://www.census.gov/geo/www/tiger/.
US Census Bureau. (2011). "TECHNICAL DOCUMENTATION: 2011 TIGER/Line Shapefiles." Retrieved August 2011, from http://www.census.gov/geo/www/tiger/tgrshp2009/TGRSHP09.pdf.
US Census Bureau. (2012). "US Census Bureau: American Fact Finder." Retrieved August 2011, from http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml.
US Forest Service. (2010). "Publications and Data." Retrieved 7 February, 2012, from http://www.nrs.fs.fed.us/data/urban/.
Wiegert, R. G. (1962). "

The selection of an optimum quadrat size for sampling the standing crop of grasses and forbs." Ecology **43**(1): 125-129.
Wright, G. T. (2011). An assessment of the potential impacts of emerald ash borer (*Agrilus planipennis* Fairmaire) on Virginia's municipal street trees. Forest Resources and Environmental Conservation. Blacksburg, VA, Virginia Polytechnic Institute and State University. **Master of Science Thesis:** 78.
Xiao, X., G. Gertner, G. Wang, and A.B. Anderson (2004). "Optimal sampling scheme for estimation landscape mapping

of vegetation cover." Landscape Ecology **20**(4): 375-387.
Zahl, S. (1989). "Line transect sampling with unknown probability of detection along the transect." Biometrics **45**(2): 453-470.
Zeide, B. (1980). "Plot size optimization." Forest Science **26**(2): 251-257.

APPENDIX A

Sources of municipal boundary shapefiles used in ArcGIS to define the streets and trees in municipalities that submitted useable datasets.

Municipality	Source of Municipal Boundary Layer
Ann Arbor, MI	Municipal contact
Auburn Hill, MI	http://www.mcgi.state.mi.us/mgdl/?rel=ext&action=sext
Berkely, CA	Municipal contact
Charlotte, NC	http://www.ncdot.org/it/gis/DataDistribution/DOTData/default.html
Clarement, CA	http://planning.lacounty.gov/gis/download
Dillon, SC	http://www.dnr.sc.gov/GIS/gisdnrdata.html
Inianapolis	Municipal contact
Lakeland, FL	Municipal contact
Lodi, NJ	https://njgin.state.nj.us/NJ_NJGINExplorer/DataDownloads.jsp
Lower Merion, PA	http://www.pasda.psu.edu/uci/SearchResults.aspx?shortcutKeyword=base&searchType=shortcut&sessionID=51234364820111012124138
Minneapolis, MN	Municipal contact
Orlando, FL	http://www.cityoforlando.net/gis/dataform.htm
Queens, NY	Municipal boundary not necessary: TIGER/Lines available by NYC boroughs
Santa Monica, CA	http://www.smgov.net/Departments/ISD/content.aspx?id=17850
Southgate, KY	http://www.uky.edu/KGS/gis/bounds.html

APPENDIX B

Characteristics of the urban forest environment in each of the seven research municipalities used in this study.

Locality	STRATUM Climate Zone	Avg. Annual Precipitation (mm)	Impervious Cover (%)	Canopy Cover (%)	Potential Natural Vegetation
Ann Arbor, MI	Northeast (Queens)/ Midwest (Minneapolis)	81 - 91	33.6	21.74	39: Mosaic Bluestem/Oak - hickory, 45: Oak - hickory
Berkeley, CA	Northern California Coast (Berkeley)	51 - 76	48.72	11.13	18: CA Mix Evergreen (CA), 26: Chaparral,
Dillon, SC	South (Charlotte)	114 - 127	18.39	20.11	55: Oak-Hickory-Pine, 56: Southern Mixed Forest
Lakeland, FL	Central Florida (Orlando)	132 - 137	23.28	21.47	56: Southern Mixed Forest
Lodi, NJ	Northeast (Queens)	122 - 127	58.78	9.82	49: Appalachian Oak
Santa Monica, CA	Southern California Coast (Santa Monica)	25 - 50	63.73	0.44	26: Chaparral, 30: Annual grassland
Southgate, KY	Lower Midwest (Indianapolis)	107 - 112	16.48	21.81	33: Prairie, 45: Oak-Hickory

APPENDIX C

GIS procedure for standardizing street segment length prior to sampling and guidelines for handling intersections when assigning Linear_ID values (Presentation 1) and re-splicing (Presentation 2).

1. Splicing segment sampling frames. This process ensures that segments are large enough to be evenly divided in step 2 to values close to the target length and minimize the divisions that must take place.
   1.1. Acquire a street segments sampling frame. If a preferred sampling frame is not available, the use of TIGER/Line Shapefiles (TIGER/Line Shapefiles 2011) is recommended.
   1.2. Combine connected street segments with the same name using ArcGIS Dissolve by street name with Unsplit Lines.
   1.3. Randomly assign unique Linear_ID values to dissolved segments based on the guidelines depicted in Presentation 1 of this overview. Perform a second Dissolve by Linear_ID to produce a sampling frame of lengthened "spliced" segments.
2. Evenly divide each spliced segment greater than 1.5 times target length using a cheat sheet such as the one below. Use the AcrGIS Split (equal parts) function in the Editor tools.
3. Merge segments less than .5 the target length with the most suitable adjoining segment using Presentation 2.

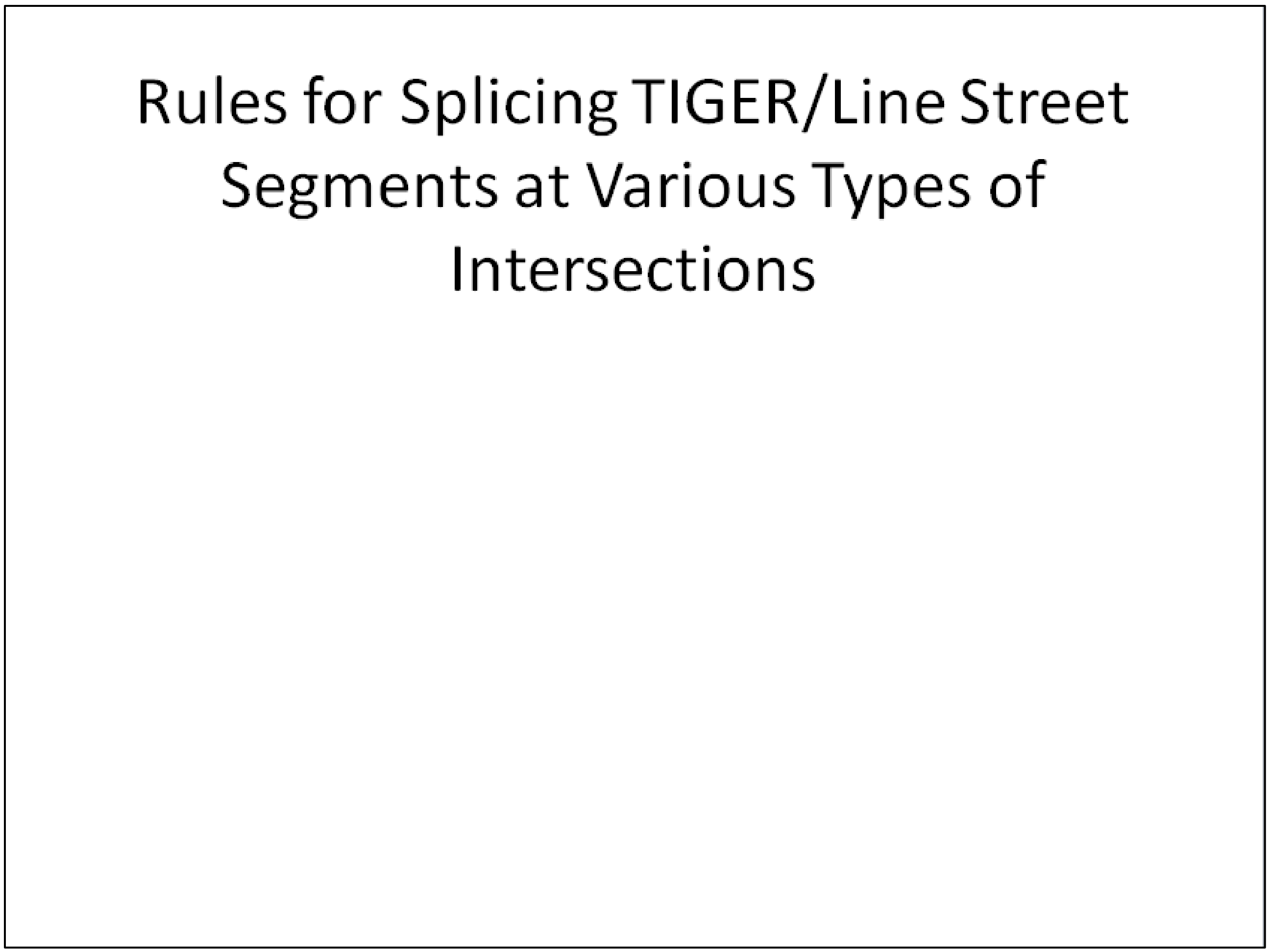
APPENDIX C, CONT.'D

Example of a cheat sheet for dividing spliced segments to a target length of 300 units:

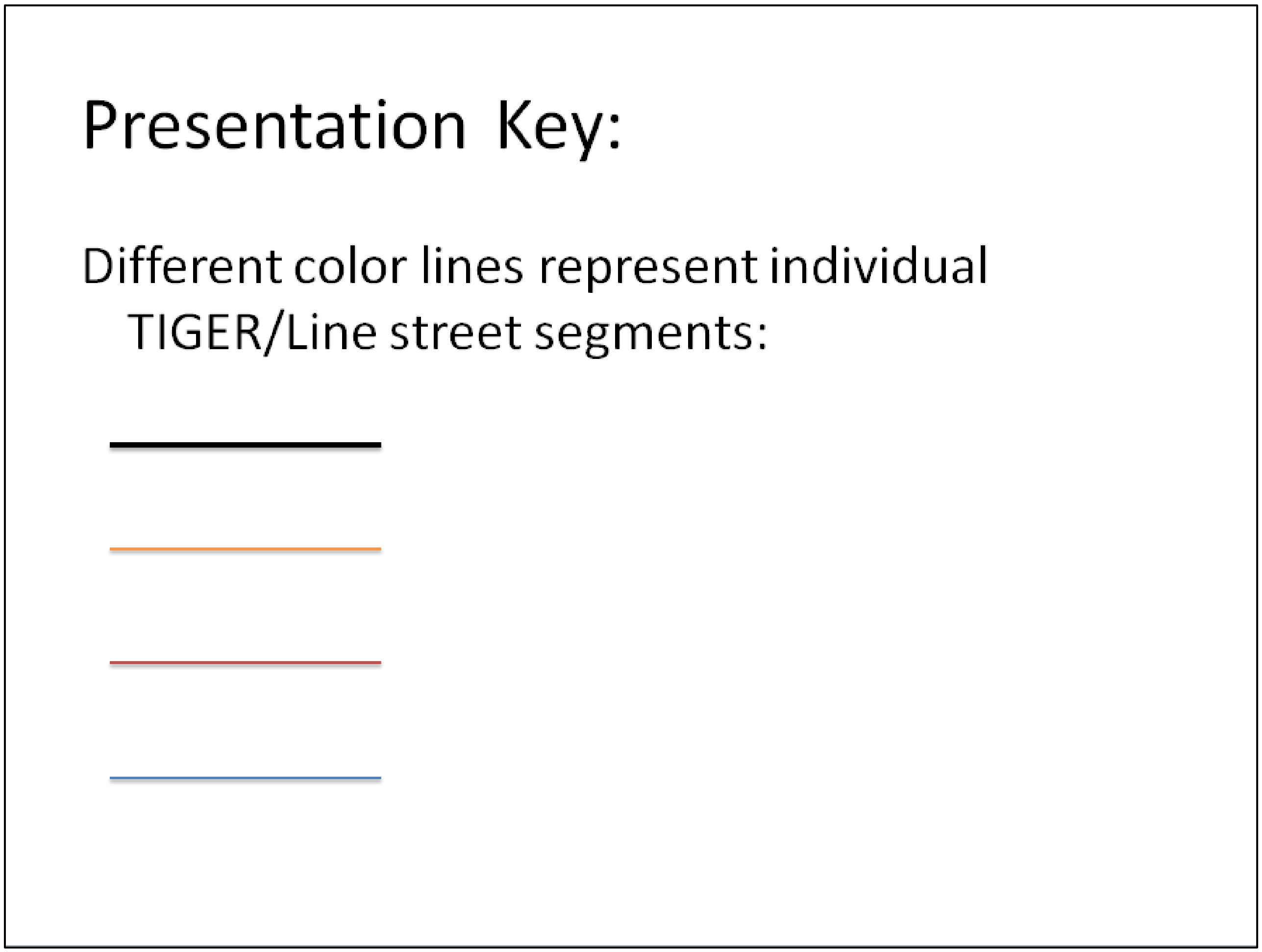
Lower Bound Spliced Length	Upper Bound Spliced Length	# of Equal Part to Split Spliced Segment	Lower Bound of Resulting Diced Length	Upper Bound of Resulting Diced Length
0	450	Do not split segment	0	450
450	750	2	225	375
750	1050	3	250	350
1050	1350	4	263	338
1350	1650	5	270	330
1650	1950	6	275	325
1950	2250	7	279	321
2250	2550	8	281	319
2550	2850	9	283	317
2850	3150	10	285	315
3150	3450	11	286	314
3450	3750	12	288	313
3750	4050	13	288	312
4050	4350	14	289	311
4350	4650	15	290	310
4650	4950	16	291	309
4950	5250	17	291	309
5250	5550	18	292	308
5550	5850	19	292	308

APPENDIX C, CONT.'D

Presentation 1: Rules for assigning Linear_ID values to dissolved segments at points of intersection.

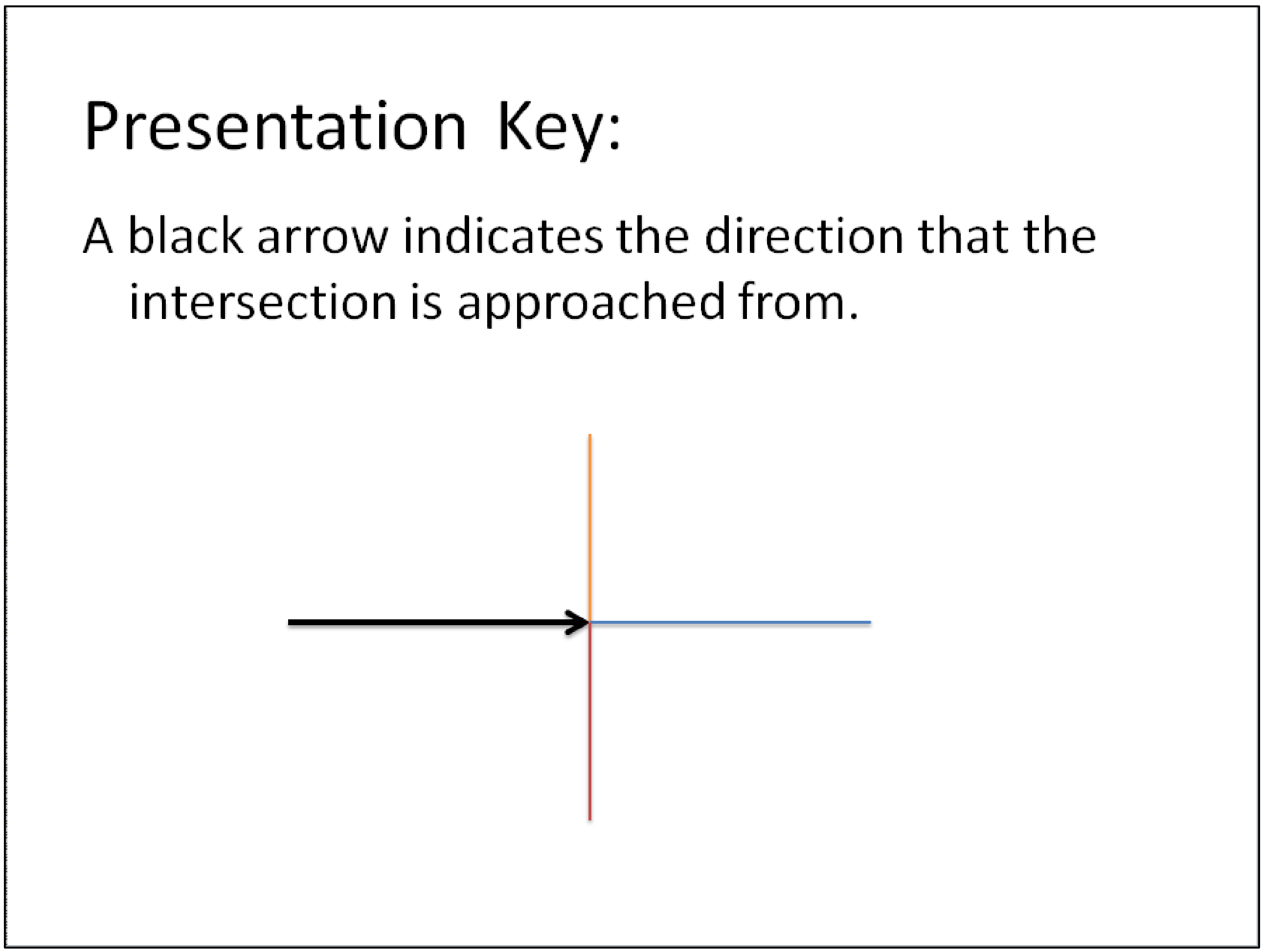


Presentation 1: Rules for assigning Linear_ID values to dissolved segments at points of intersection.



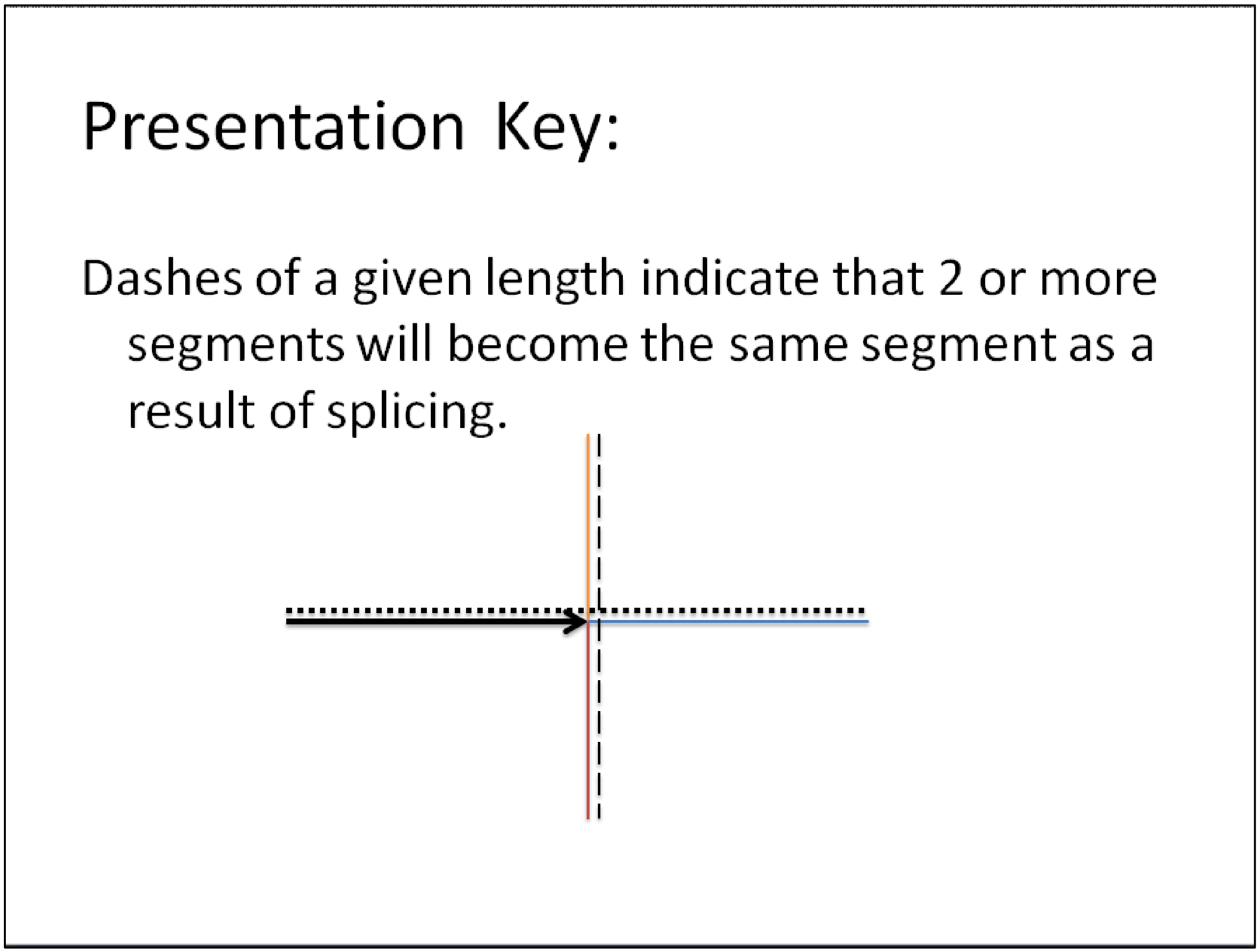
APPENDIX C, CONT.'D

Presentation 1: Rules for assigning Linear_ID values to dissolved segments at points of intersection.



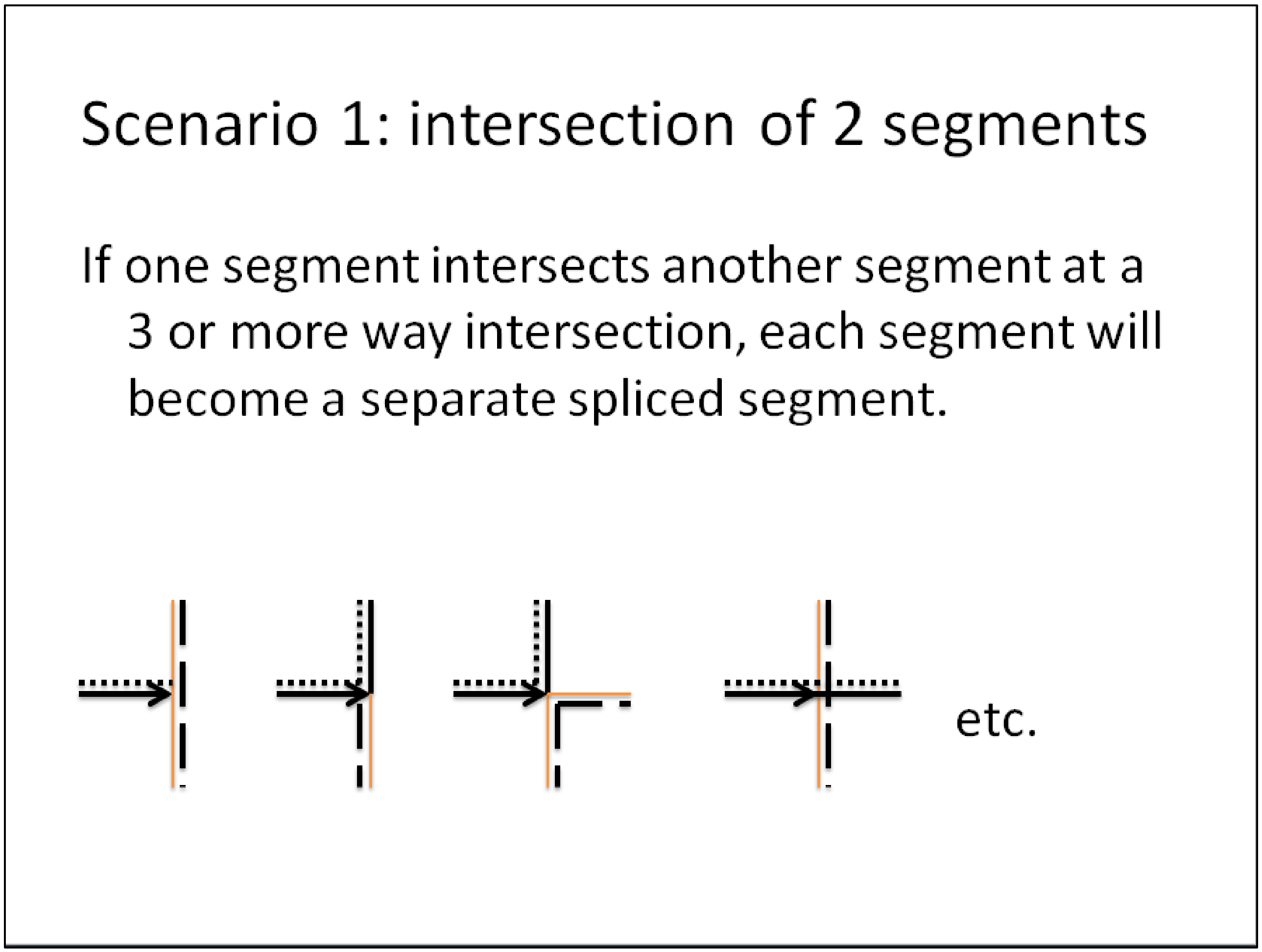
APPENDIX C, CONT.'D

Presentation 1: Rules for assigning Linear_ID values to dissolved segments at points of intersection.



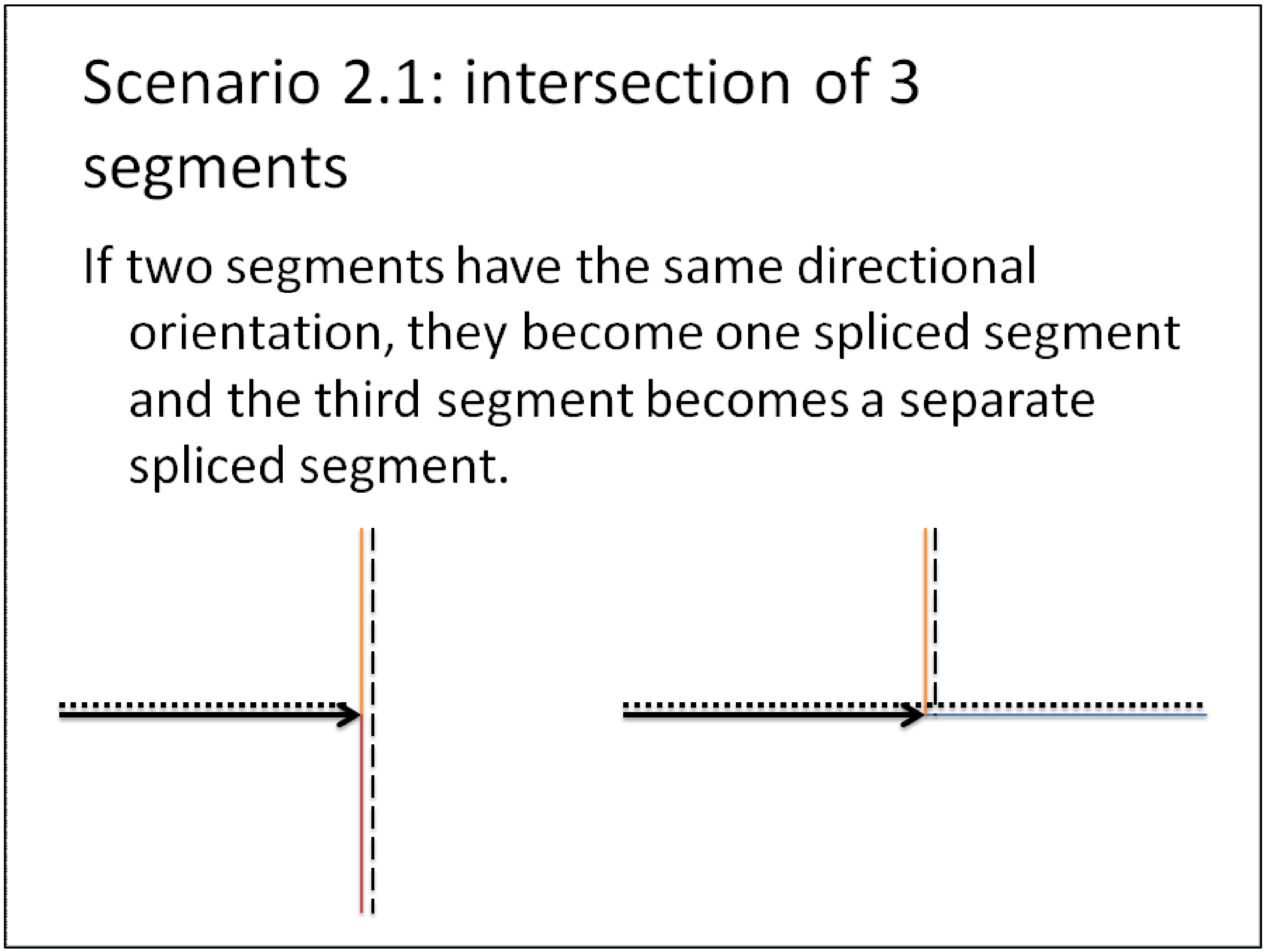
APPENDIX C, CONT.'D

Presentation 1: Rules for assigning Linear_ID values to dissolved segments at points of intersection.



APPENDIX C, CONT.'D

Presentation 1: Rules for assigning Linear_ID values to dissolved segments at points of intersection.

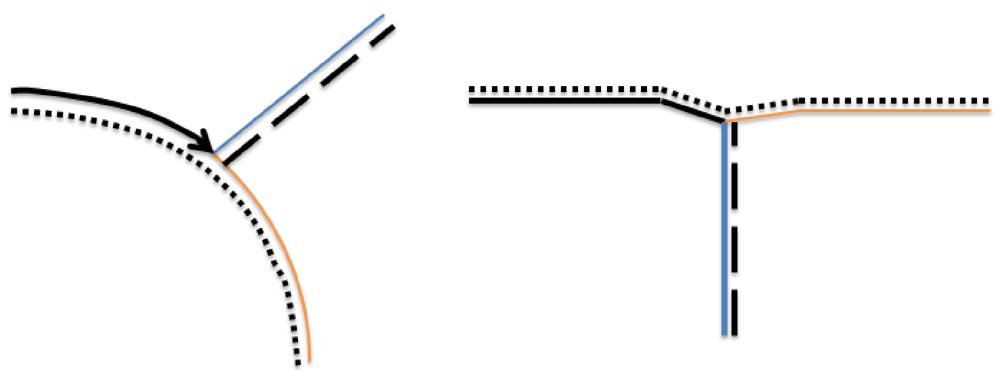


APPENDIX C, CONT.'D

Presentation 1: Rules for assigning Linear_ID values to dissolved segments at points of intersection.

Scenario 2.2: intersection of 3 segments

If none of the segments have the same directional orientation and it is apparent that 2 of the segments have similar directional orientations, they will be combined into one spliced segment and the remaining segment will become a separate spliced segment.

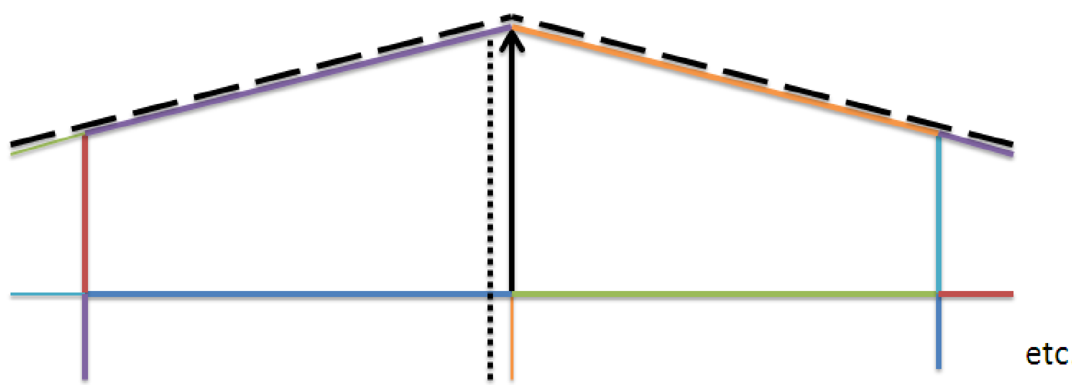


APPENDIX C, CONT.'D

Presentation 1: Rules for assigning Linear_ID values to dissolved segments at points of intersection.

Scenario 2.3: intersection of 3 segments

If none of the segments have the same directional orientation and it is apparent that 2 of the segments have similar directional orientations, they will be combined into one spliced segment and the remaining segment will become a separate spliced segment.

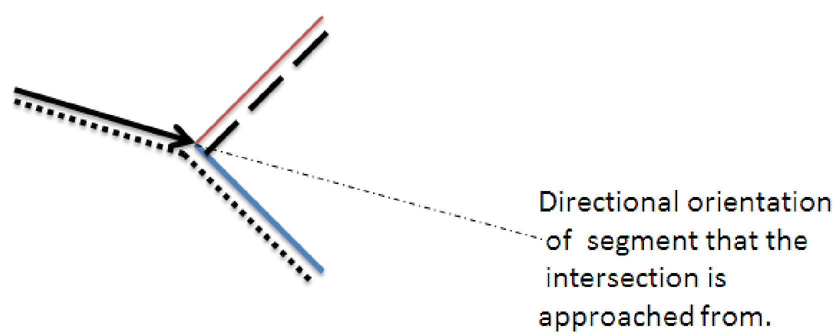


APPENDIX C, CONT.'D

Presentation 1: Rules for assigning Linear_ID values to dissolved segments at points of intersection.

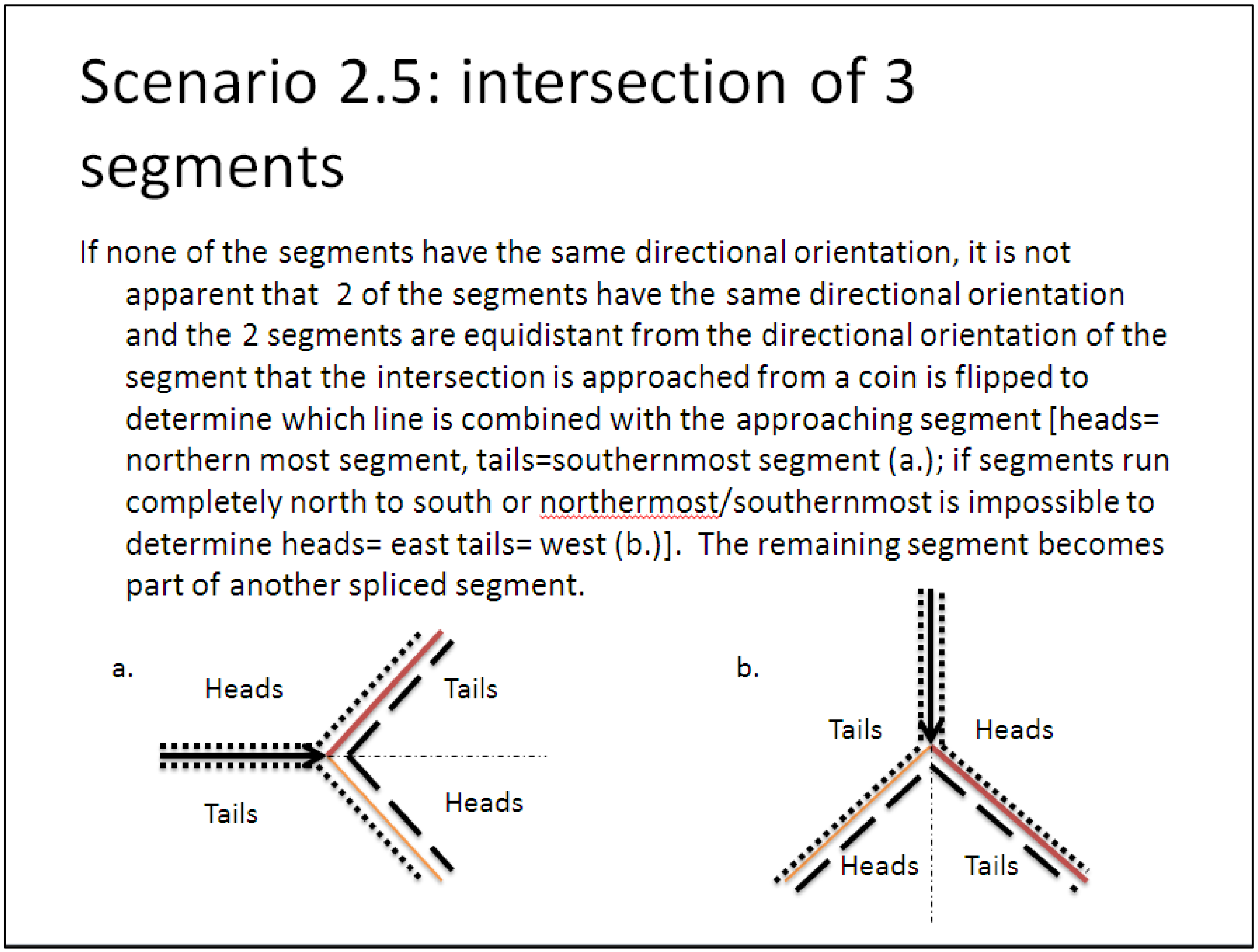
Scenario 2.4: intersection of 3 segments

If none of the segments have the same directional orientation and it is not apparent that 2 of the segments have the same directional orientation, the line with the directional orientation closest to that of the segment that the intersection is approached from is combined into one spliced segment.



APPENDIX C, CONT.'D

Presentation 1: Rules for assigning Linear_ID values to dissolved segments at points of intersection.

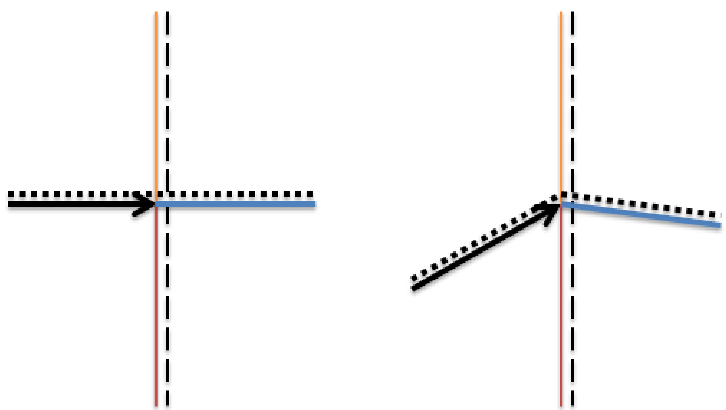


APPENDIX C, CONT.'D

Presentation 1: Rules for assigning Linear_ID values to dissolved segments at points of intersection.

Scenario 3.1: intersection of 4 segments

If a pair of segments have the same directional orientation that pair must become on spliced segment. The remaining pair will also become one spliced segment.



Presentation 1: Rules for assigning Linear_ID values to dissolved segments at points of intersection.

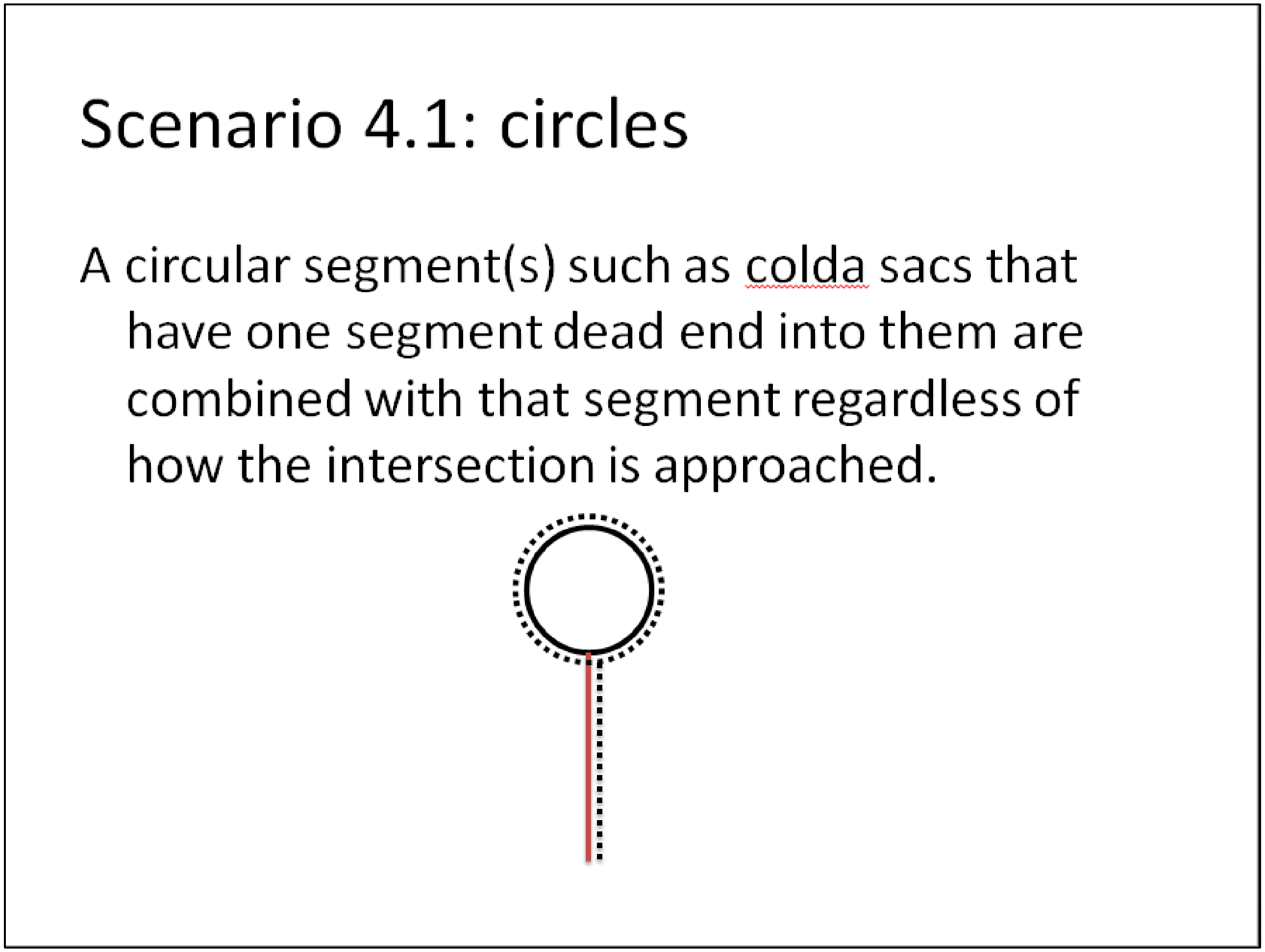
Scenario 3.2: intersection of 4 segments

If none of the segments have the same directional orientation and it is apparent that 2 of the segments have similar directional orientations, they will be combined into one spliced segment and the remaining segments will become a separate spliced segment.

If none of the segments have the same directional orientation and it is not apparent that 2 of the segments have the same directional orientation, the line with the directional orientation closest to that of the segment that the intersection is approached from is combined into one spliced segment.

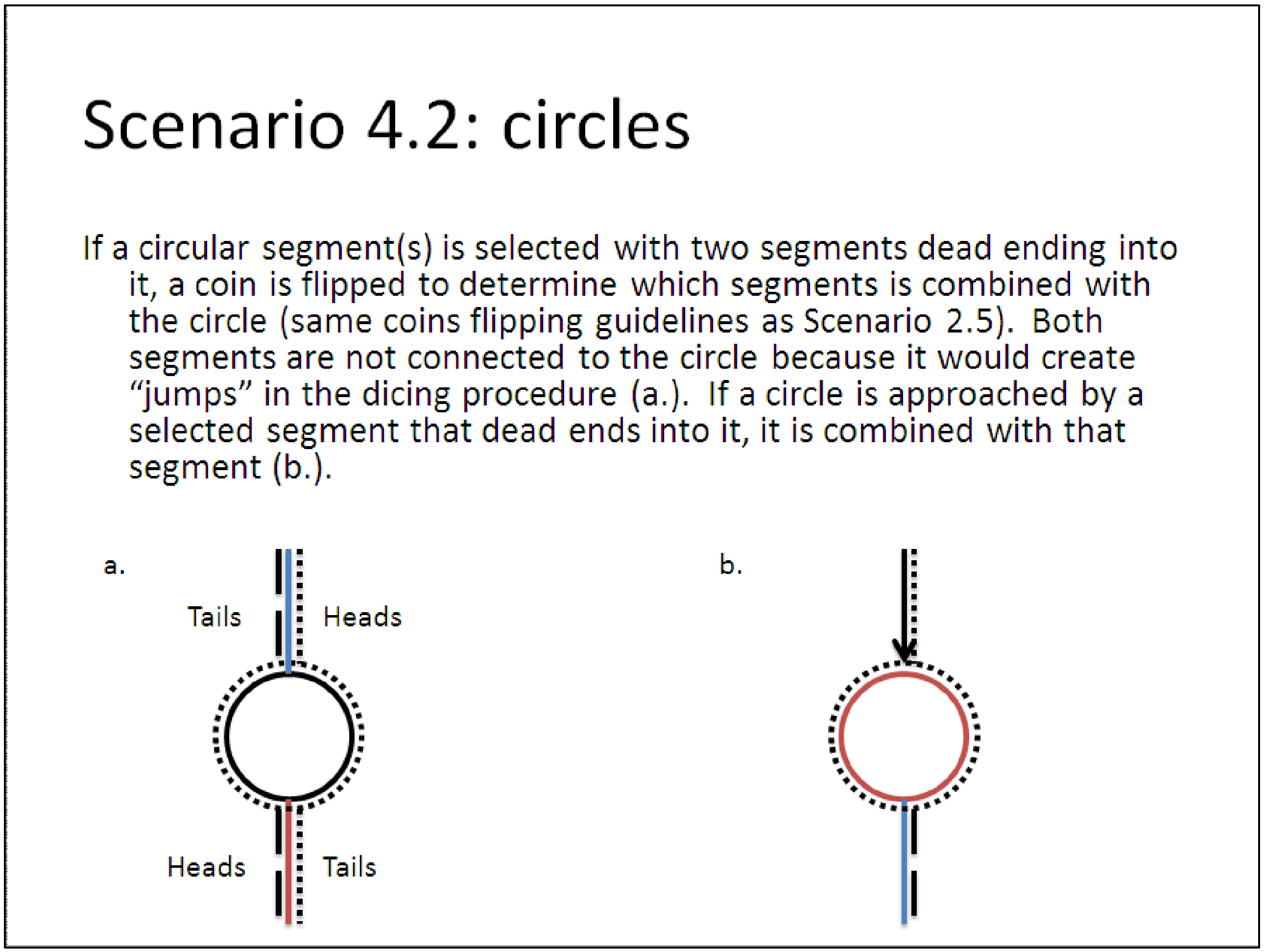
APPENDIX C, CONT.'D

Presentation 1: Rules for assigning Linear_ID values to dissolved segments at points of intersection.

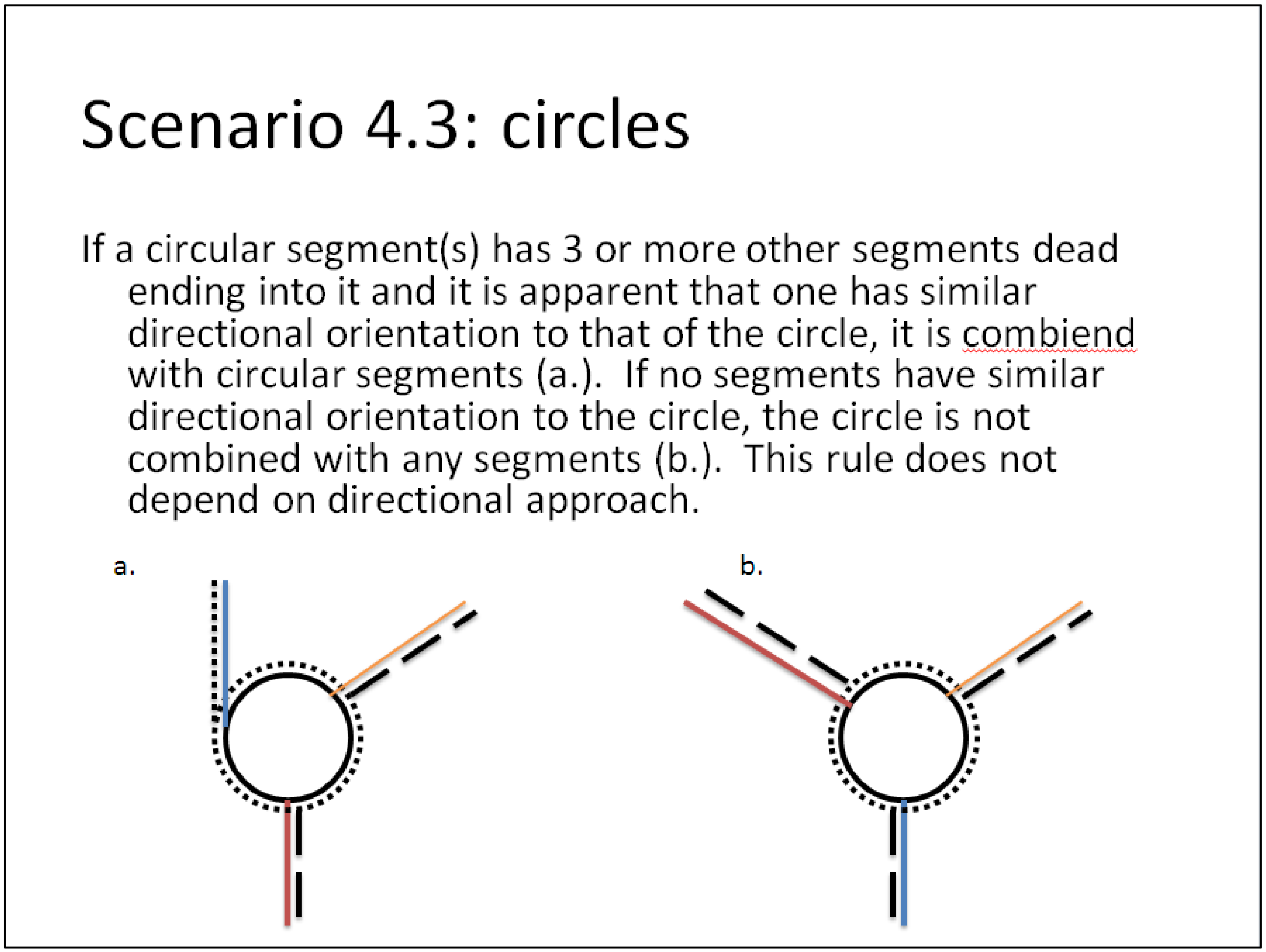


APPENDIX C, CONT.'D

Presentation 1: Rules for assigning Linear_ID values to dissolved segments at points of intersection.

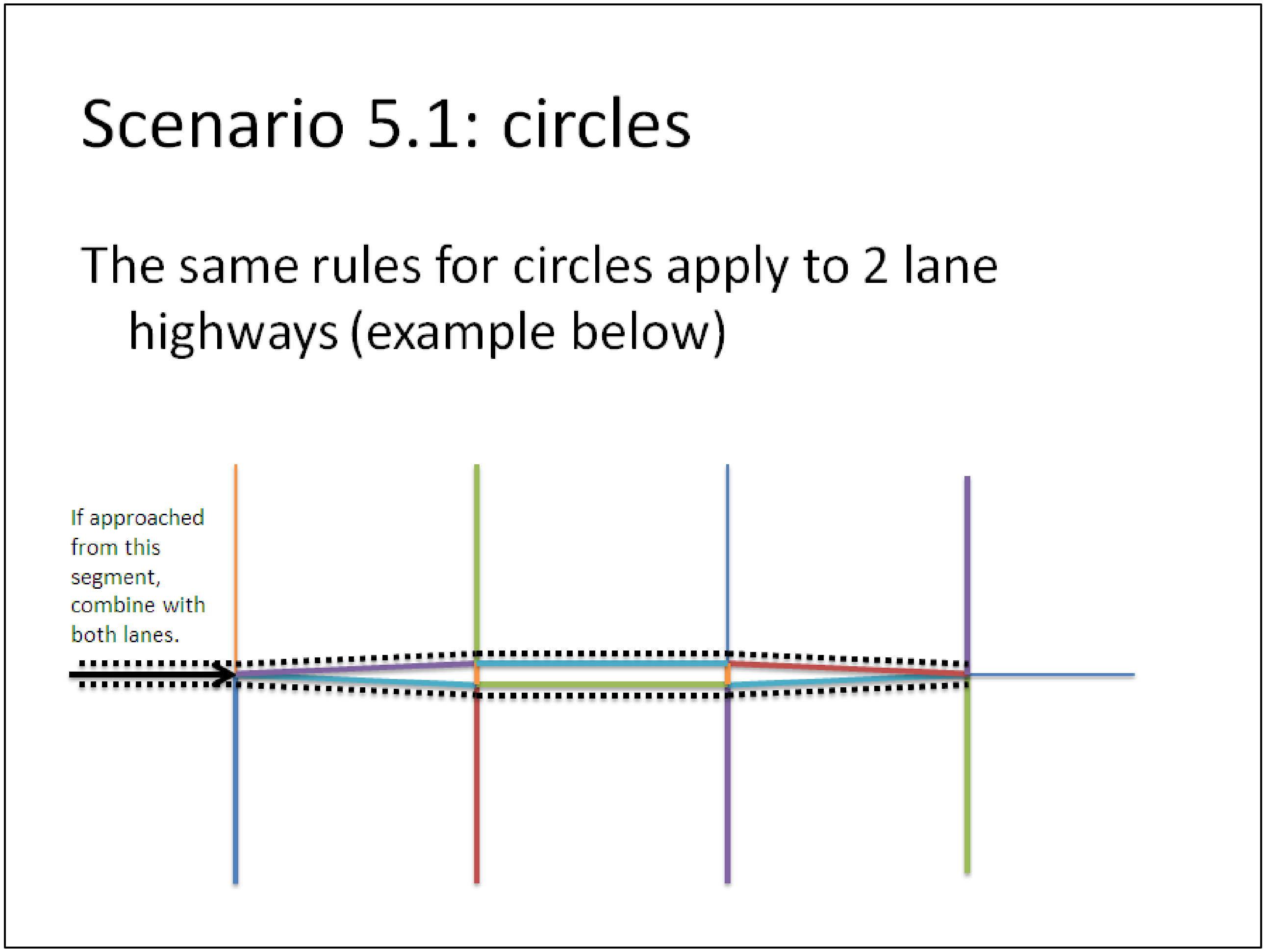


Presentation 1: Rules for assigning Linear_ID values to dissolved segments at points of intersection.

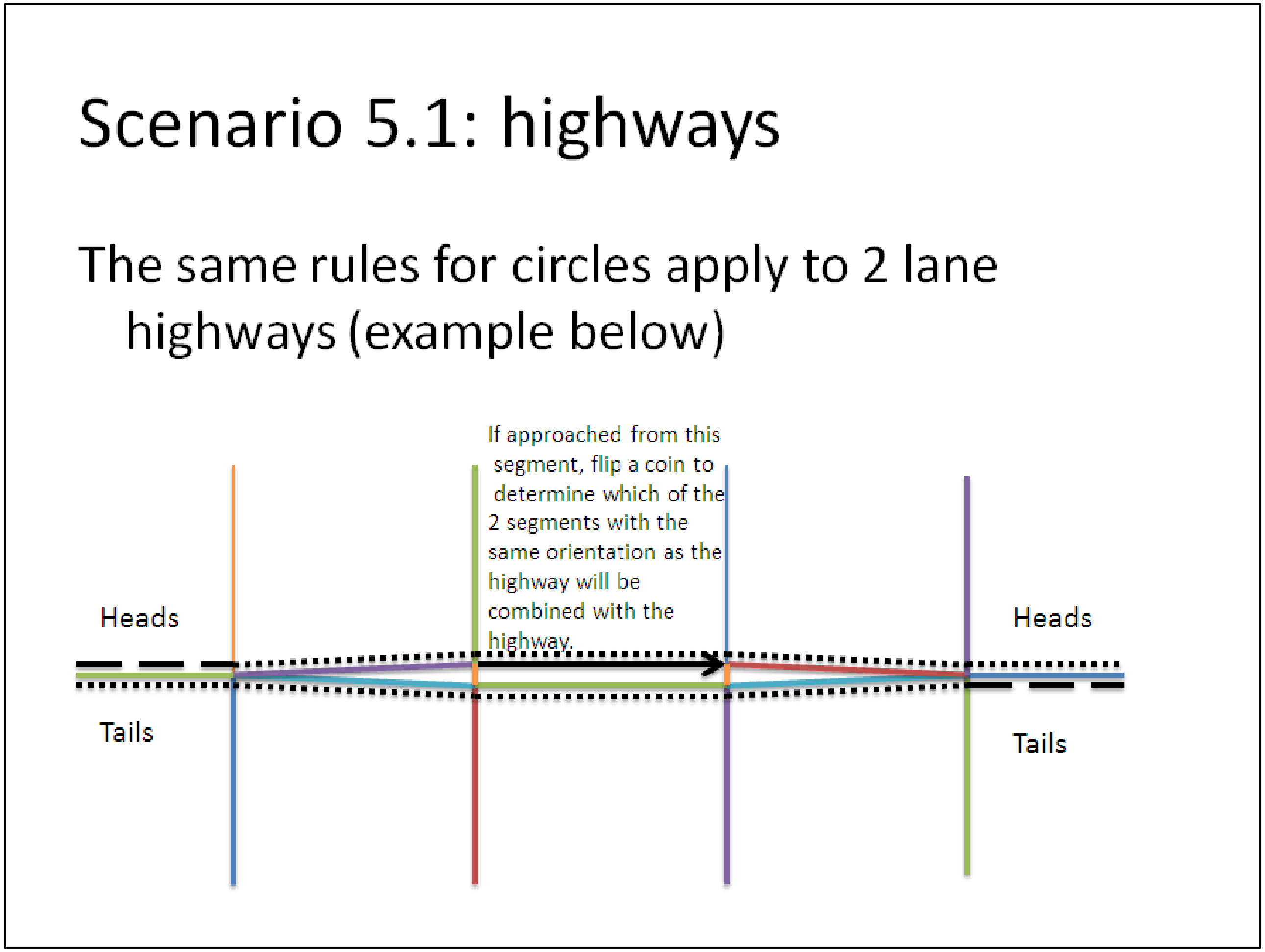


APPENDIX C, CONT.'D

Presentation 1: Rules for assigning Linear_ID values to dissolved segments at points of intersection.

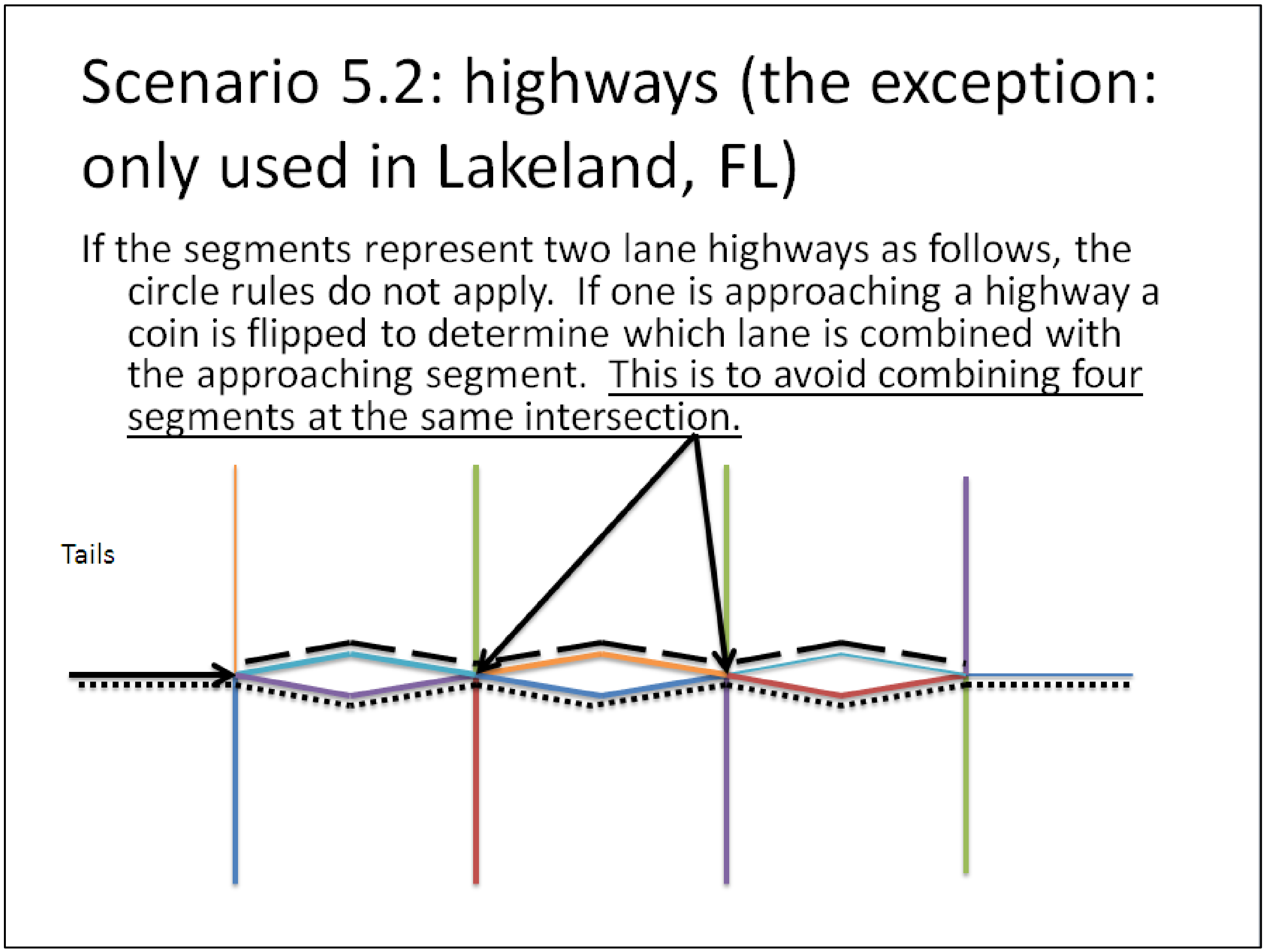


Presentation 1: Rules for assigning Linear_ID values to dissolved segments at points of intersection.



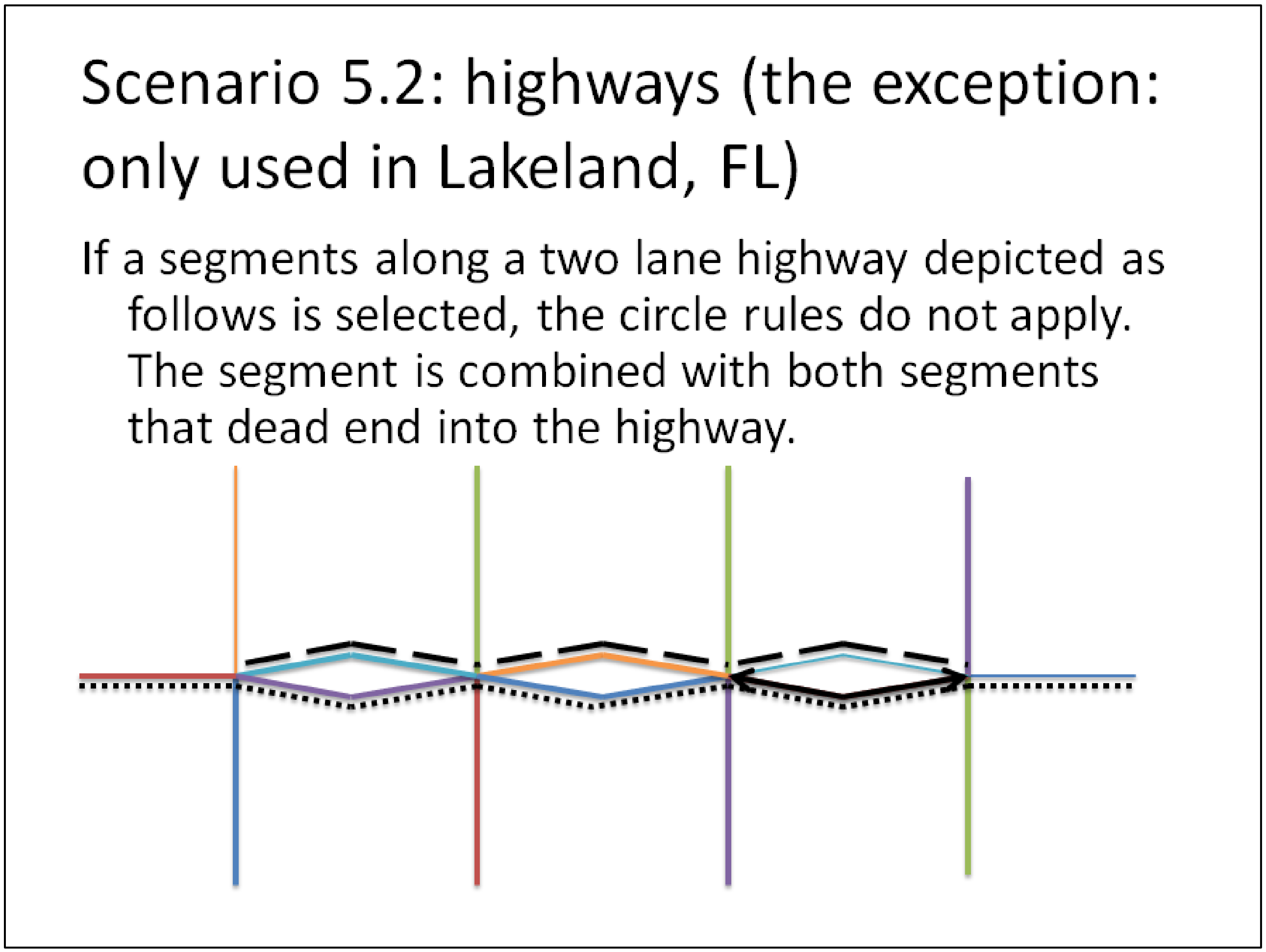
APPENDIX C, CONT.'D

Presentation 1: Rules for assigning Linear_ID values to dissolved segments at points of intersection.



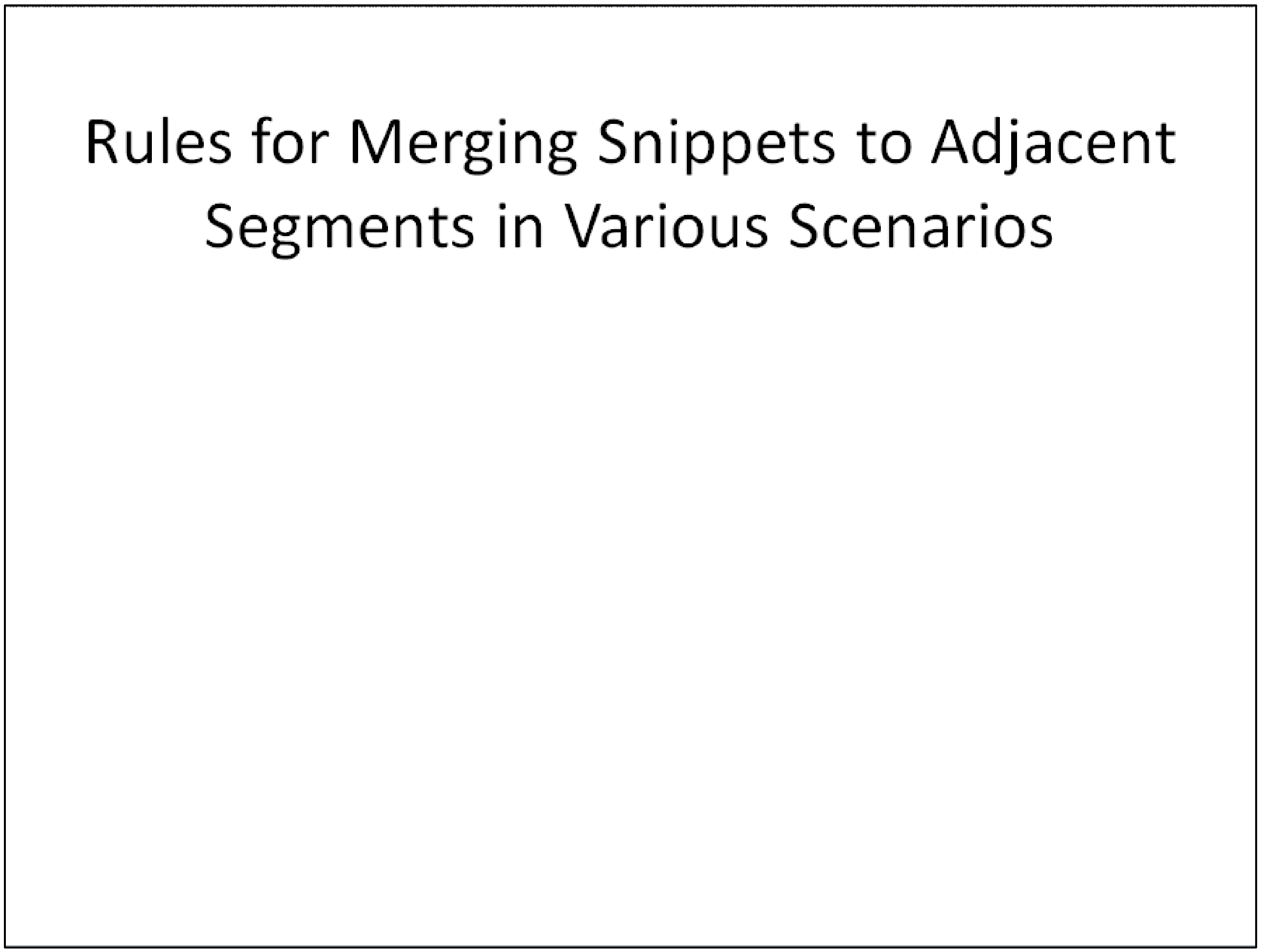
APPENDIX C, CONT.'D

Presentation 1: Rules for assigning Linear_ID values to dissolved segments at points of intersection.



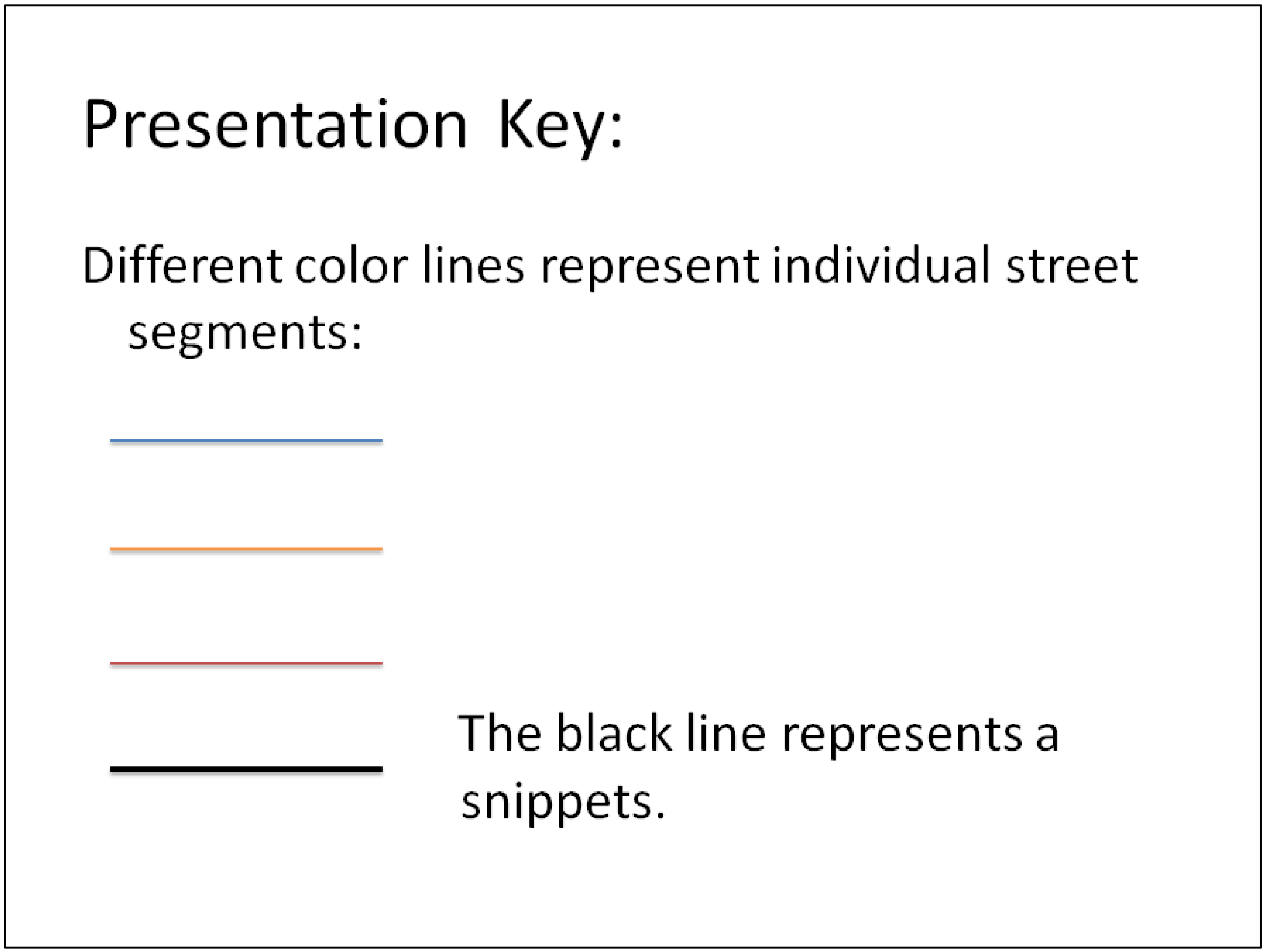
APPENDIX C, CONT.'D

Presentation 2: Rules for merging snippets with the appropriate adjoining segment.



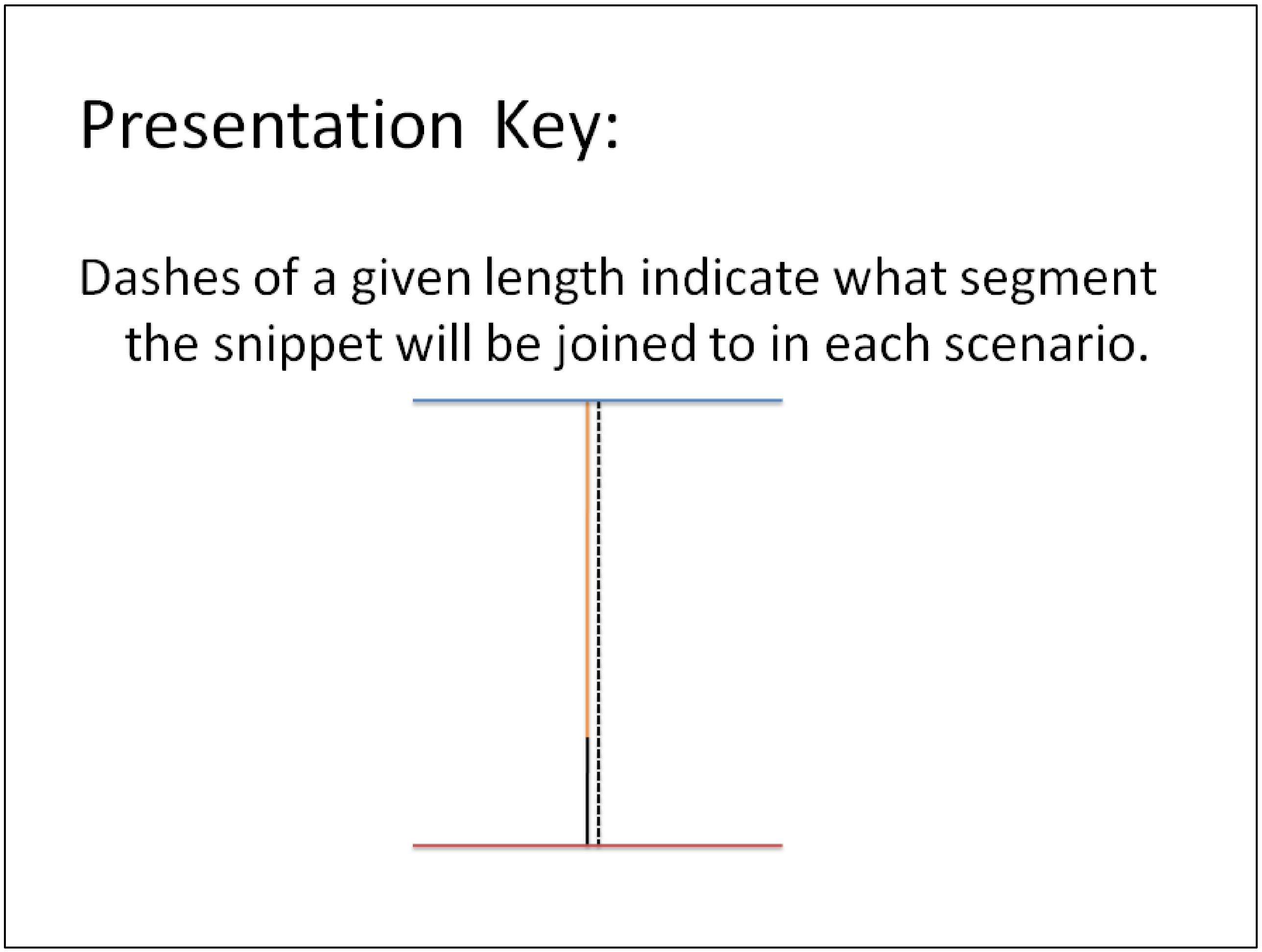
APPENDIX C, CONT.'D

Presentation 2: Rules for merging snippets with the appropriate adjoining segment.



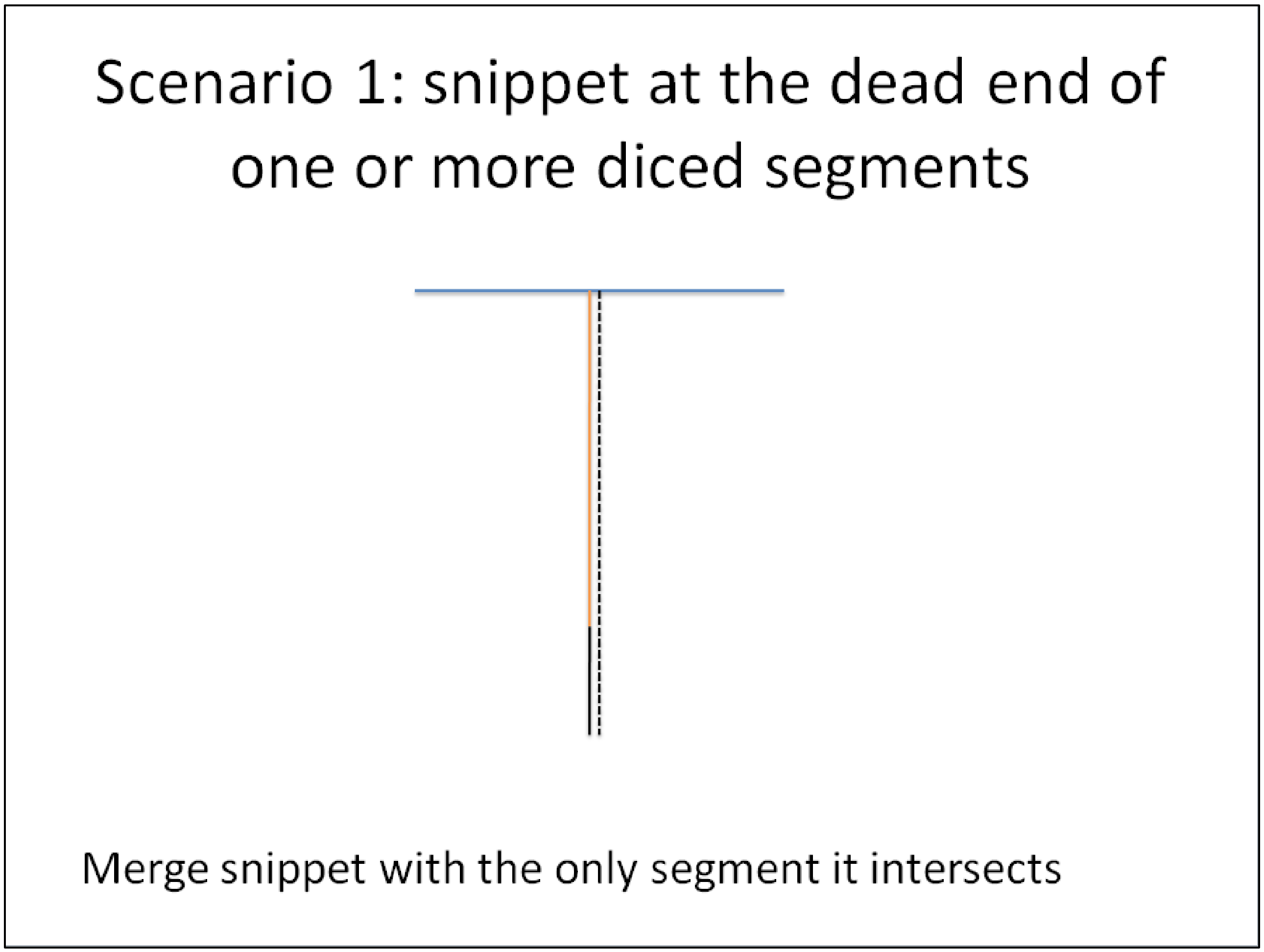
APPENDIX C, CONT.'D

Presentation 2: Rules for merging snippets with the appropriate adjoining segment.



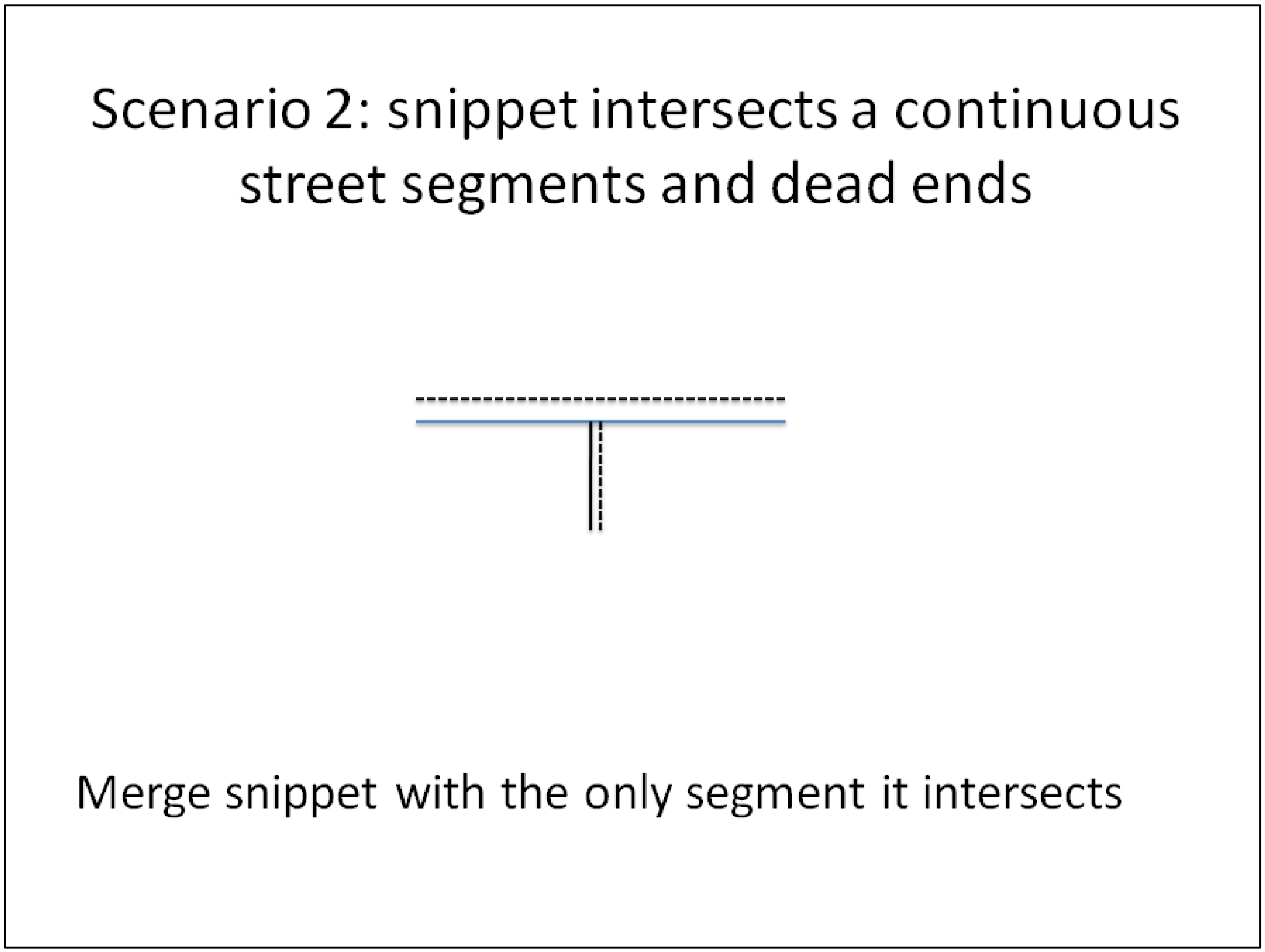
APPENDIX C, CONT.’D

Presentation 2: Rules for merging snippets with the appropriate adjoining segment.



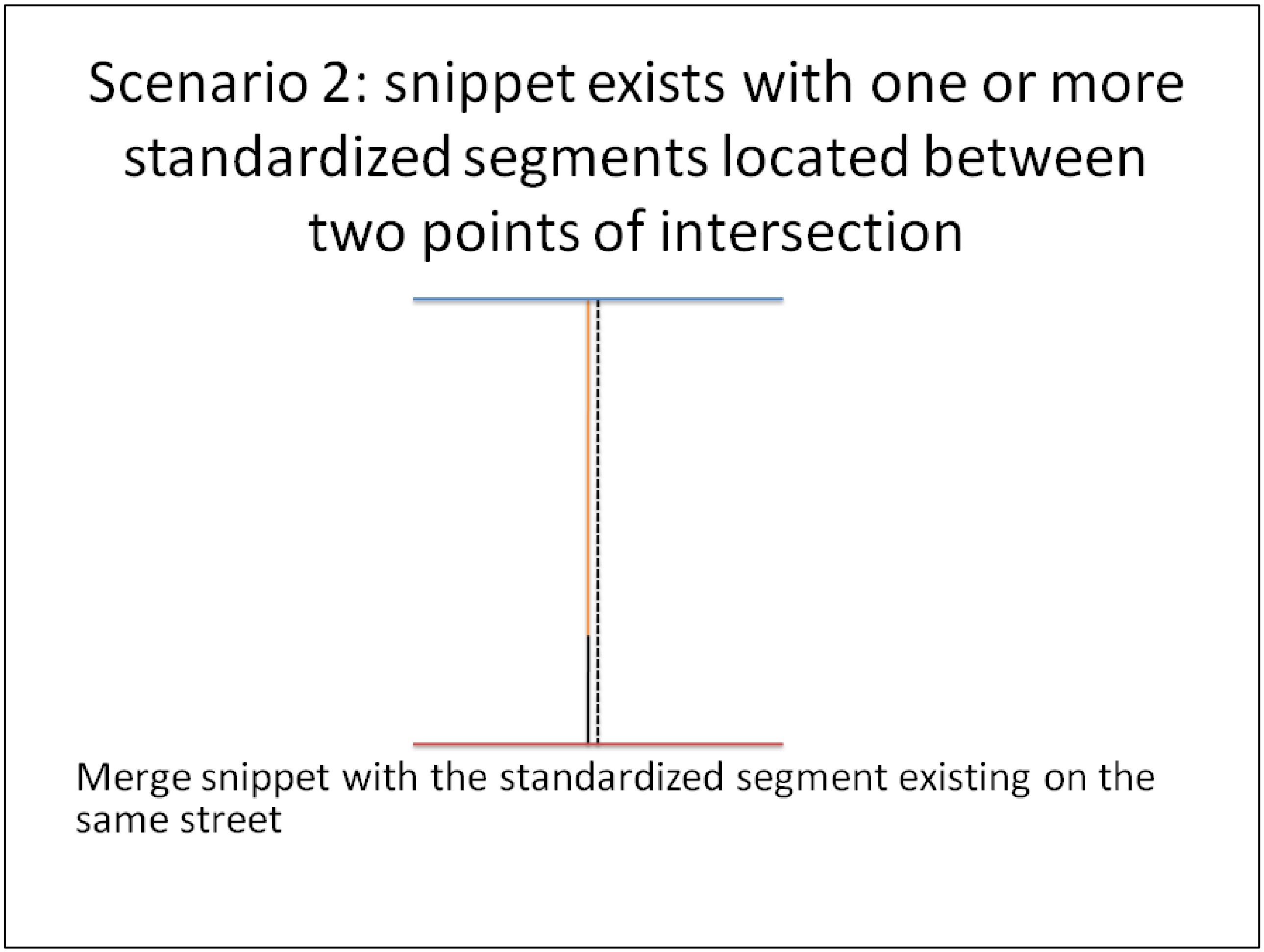
APPENDIX C, CONT.'D

Presentation 2: Rules for merging snippets with the appropriate adjoining segment.



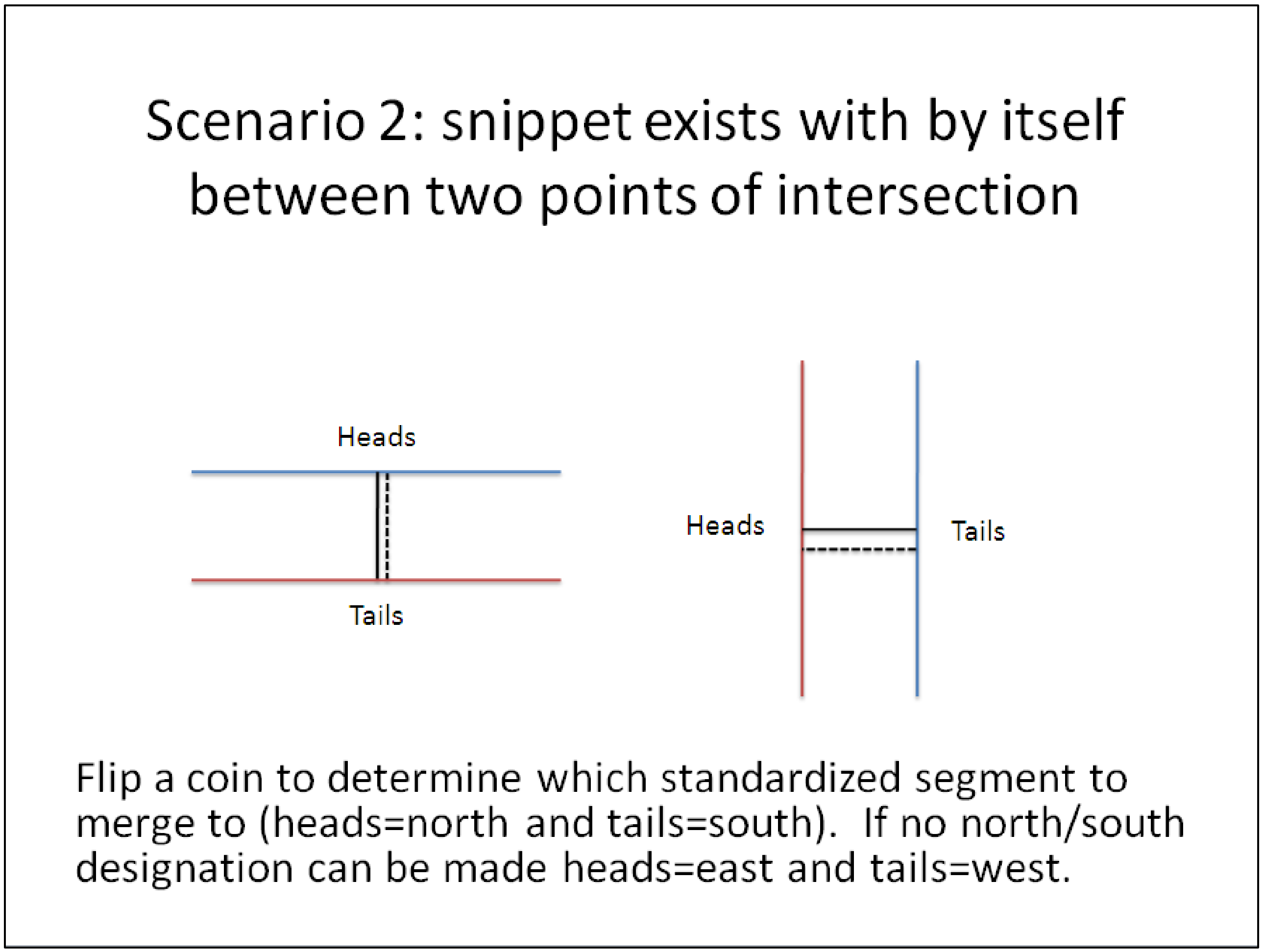
APPENDIX C, CONT.'D

Presentation 2: Rules for merging snippets with the appropriate adjoining segment.



APPENDIX C, CONT.'D

Presentation 2: Rules for merging snippets with the appropriate adjoining segment.



APPENDIX D

Screen capture of the computational script that was run using R Statistical Computing to simulate street tree inventories.

C:\Users\Mason\Desktop\Thesis\Simulation\SimScript.R - R Editor

```
fn <- "csv._file_location"

Locality_SamplingFrame <- read.csv(fn,header=T)

pop_data <- Locality_SamplingFrame

sample_intensity <- 5
sample_size <- round(nrow(pop_data)*sample_intensity/100)
pop_size <- nrow(pop_data)

srs <- function(pop_data,sample_intensity)
{
	pop_data[sample(1:nrow(pop_data),round(nrow(pop_data)*sample_intensity/100),replace=F),]
}

nreps <- 30

for(i in 1:nreps)
{
	if(i == 1)
	{
		s <- srs(pop_data,sample_intensity)
	}else
	{
		s <- cbind(s,srs(pop_data,sample_intensity))
	}
}

means <- colMeans(s[,seq(2,nreps*2,2)])
totals <- means*nrow(pop_data)
se_tot <- sapply(s[,seq(2,nreps*2,2)],sd)*nrow(pop_data)*sqrt((pop_size-sample_size)/(pop_size))/sqrt(sample_size)


write.table(totals,file="C:/Users/Mason/Desktop/Thesis/Simulation/Locality_SamplingFrame_Totals.csv",sep=",")
write.table(se_tot,file="C:/Users/Mason/Desktop/Thesis/Simulation/Locality_SamplingFrame_SE.csv",sep=",")
write.table(length,file="C:/Users/Mason/Desktop/Thesis/Simulation/Locality_SamplingFrame_Lngth.csv",sep=",")
```