

University of New Hampshire

University of New Hampshire Scholars' Repository

Master's Theses and Capstones

Student Scholarship

Fall 2018

Invisible trees and meeting management needs: A comparison of distance sampling and traditional forest inventory methods.

Connor Breton

University of New Hampshire, Durham

Follow this and additional works at: <https://scholars.unh.edu/thesis>

Recommended Citation

Breton, Connor, "Invisible trees and meeting management needs: A comparison of distance sampling and traditional forest inventory methods." (2018). *Master's Theses and Capstones*. 1221.

<https://scholars.unh.edu/thesis/1221>

This Thesis is brought to you for free and open access by the Student Scholarship at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Master's Theses and Capstones by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact Scholarly.Communication@unh.edu.

INVISIBLE TREES AND MEETING MANAGEMENT NEEDS: A COMPARISON OF
DISTANCE SAMPLING AND TRADITIONAL FOREST INVENTORY METHODS

BY

CONNOR BRETON
BS, University of New Hampshire, 2016

THESIS

Submitted to the University of New Hampshire
in Partial Fulfillment of
the Requirements for the Degree of

Master of Science
in
Natural Resources - Forestry

September, 2018

This thesis has been examined and approved in partial fulfillment of the requirements for the degree of Master of Science in Natural Resources - Forestry by:

Thesis Director, Mark J. Ducey, Professor of Natural Resources and the Environment

Thomas D. Lee, Associate Professor of Natural Resources and the Environment

Jeffrey H. Gove, Research Forester, U.S. Forest Service

On August 8, 2018

Original approval signatures are on file with the University of New Hampshire Graduate School.

TABLE OF CONTENTS

DEDICATION	iv
ACKNOWLEDGEMENTS	v
LIST OF TABLES	vi
LIST OF FIGURES	vii
ABSTRACT	viii

CHAPTER	PAGE
INVISIBLE TREES AND MEETING MANAGEMENT NEEDS: A COMPARISON OF DISTANCE SAMPLING AND TRADITIONAL FOREST INVENTORY METHODS	1
Introduction.....	1
Methods.....	5
Fixed Area Sampling	7
Horizontal Point Sampling.....	9
Line Transect Sampling	10
Point Transect Sampling.....	12
Distance Sampling Estimators and Reference Populations	13
Efficiency Analysis.....	15
Results.....	17
Precision and Accuracy.....	17
Distance Sampling and HPS Permutations.....	22
Distance Sampling Models	24
Distance Sampling Efficiency.....	27
Discussion.....	30
Sampling Method Comparisons.....	30
Sampling Efficiency.....	32
Sampling and Management Implications.....	34
Conclusion	35
References.....	36
Appendix I	40
Appendix II.....	42

DEDICATION

I dedicate this work to my family, especially my parents, Pat and Kevin Breton.

*Without your unconditional love and support to fall back on, I never would have been able to
continue pushing forward. This one's for you.*

ACKNOWLEDGEMENTS

I cannot imagine a better committee to have relied on than Mark Ducey, Tom Lee, and Jeff Gove, all of whom I must thank for their encouragement, support, and guidance at each stage of this project. I would especially like to thank Mark Ducey for taking me on as his graduate student, encouraging me to work with the U.S. Fish and Wildlife Service, and for giving me the chance to prove myself. I would also like to thank those people, such as Heidi Asbjornsen and Inge Seaboyer, who have previously invested in me and given me the stepping-stone opportunities I have needed to reach my current position.

Thank you to Ethan Belair for sharing his knowledge of the Bartlett Experimental Forest, his work on the Bartlett cruise plots, and for permitting me to ask him as many questions as I could think of. Thank you to Kimberly Lavoie, Robert Keefe, Hannah (Rue) Teel, Julia Smith, Heidi Giguere, Madison Poe, and all others who have also managed and worked on the Bartlett cruise plots from which data for this project were collected.

I am also profoundly grateful for my friends and family. Thank you to Cody FitzGerald for helping ground me whenever I got overwhelmed. Thank you to Katherine Sinacore for originally acting as my academic and professional mentor, but also for becoming part of my family. Thank you to Abbey and Phil Ouellette for always being there for me and for always providing words of encouragement when I needed them the most. Thank you to Kira Kenny for her love and constant support, as well as her patience and understanding when I had to focus on work. Finally, I would like to express my deepest gratitude to my parents, Pat and Kevin Breton, whose unwavering support and countless sacrifices have enabled me to pursue my dream of working in the woods.

LIST OF TABLES

Table 1. Estimates of abundance (trees) and density (trees/ha) in the northern hardwood compartment.	20
Table 2. Estimates of abundance (trees) and density (trees/ha) in the hemlock-mixed hardwood compartment.	21
Table 3. Distance sampling models for the northern hardwood compartment	40
Table 4. Distance sampling models for the hemlock-mixed hardwood compartment.	41

LIST OF FIGURES

Figure 1. Inventory method sample designs	7
Figure 2. Slopover correction..	12
Figure 3. Permutation population estimates and associated variances	23
Figure 4. Best fit detection probability models for large trees of all species.....	25
Figure 5. Best fit detection probability models for large trees of select species..	27
Figure 6. Distance sampling time regressions	29

ABSTRACT

INVISIBLE TREES AND MEETING MANAGEMENT NEEDS: A COMPARISON OF DISTANCE SAMPLING AND TRADITIONAL FOREST INVENTORY METHODS

by

Connor Breton

University of New Hampshire, September, 2018

Forest inventory is an important part of forest planning and management. While land managers risk making misinformed management decisions when relying on low quality inventory data, they also must balance the data they collect with the time investment necessary for sampling. We compared abundance and density estimates of large trees 45.72 cm (18 in) or greater in diameter at breast height sampled with fixed area sampling, horizontal point sampling, and two forms of distance sampling (line transect sampling and point transect sampling), a sampling method primarily used in wildlife surveys. To compare the time investments of each form of distance sampling, we also recorded implementation time in the field and created linear regressions from which the required time to sample a specific number of trees with a given method could be predicted. Results suggest that in most cases, fixed area sampling and (or) horizontal point sampling out-perform distance sampling and produce more precise and accurate estimates of large trees, even when the sampled trees are present at various levels of density. However, it was also found that in some cases, distance sampling has the potential to out-perform traditional inventory methods and land managers are likely to prefer using point transect sampling over line transect sampling.

CHAPTER I

INVISIBLE TREES AND MEETING MANAGEMENT NEEDS: A COMPARISON OF DISTANCE SAMPLING AND TRADITIONAL FOREST INVENTORY METHODS

Introduction

Forest inventory is at the core of forest planning and management. The information provided by an inventory allows land managers to not only develop feasible management objectives, but to also make the well-informed decisions necessary for meeting those objectives. Forest inventories also provide information from which management success may be determined. By comparing information collected prior to and following any landscape alterations (e.g. harvesting), land managers are often able to determine whether they have met their objectives or need to engage in adaptive management. The ability to plan, manage, and adapt, however, depends not only on the type of inventory conducted, but also on how precisely and accurately the inventory accounts for forest components relevant to management.

Forest components in moderate to high abundance are often adequately accounted for with traditional inventory methods, yet traditional methods such as horizontal point sampling (HPS) (Bitterlich 1947, 1984; Grosenbaugh 1955, 1958) may yield biased estimates of scarce forest components that go undetected during sampling (Ritter et al. 2013). This is of management concern given that, despite their distribution, sparse and clustered forest components such as veneer quality trees or trees with wildlife cavities have significant economic (Cassens 2004) and ecological value (Tubbs et al. 1987). Without the ability to reliably account for these types of forest components, land managers are more likely to underestimate forest value and make misinformed management decisions, leading to lost profits or undesired ecological outcomes.

Many sampling methods have been proposed for use when trying to account for sparse and clustered forest components including modified horizontal line sampling (Ducey et al. 2002; Kenning et al. 2005), guided transect sampling (Ståhl et al. 2000), unrestricted guided transect sampling (Ringvall et al. 2007), and adaptive cluster sampling (Roesch 1993; Talvitie et al. 2006; Thompson 1990, 1991). While designed to increase the probability that sparse and clustered forest components will be sampled, sample design complexity may prevent some of these methods from also being used to account for common forest components. Therefore, when scarce and common forest components are to be sampled, a land manager may choose to use a non-traditional inventory method in conjunction with a traditional inventory method such as fixed area sampling (FAS), though some sampling combinations may be inefficient (Bäuerle et al. 2009). For example, despite recognizing that foresters commonly use plots when conducting timber inventories, Ståhl et al. (2000) proposed a method of inventory for sparsely distributed forest components that makes use of strips in one stage of sampling and lines in a second stage. Combining such a complex sample design with a plot-based inventory would likely require foresters to invest a significant amount of time and effort to account for separate groups of forest components.

Additionally, without a flexible, yet efficient, quantitative sampling method, those conducting forest inventories often adopt qualitative methods (Ståhl et al. 2000) such as site descriptions and subjective categorical labels (e.g. low, moderate, and dense) to describe forest conditions. While these data can be useful, it is difficult to compare qualitative data over different time periods (Ståhl et al 2000). Furthermore, consistency in using subjective categorical labels is difficult to achieve when separate groups and individuals are tasked with inventorying the same tract of land. Invasive species serve as a good example of when this inconsistency

challenges the strength of the data collected. For example, an individual may find an invasive species over their inventory tract and describe it as present at a “moderate” level of density. A decade later a new person may inventory the same tract of land and again find the invasive to be present. This time, however, the new person may have a differing perspective of what constitutes specific levels of density and may also say that the invasive is present at a “moderate” level, when in fact density has increased substantially. The new person in this case would be unable to compare their conclusion with that of the previous individual and would therefore have no way of determining how the invasive’s population has changed.

A promising method for inventorying sparse and clustered forest components is distance sampling, a quantitative sampling method which derives estimates of density and abundance from models of sample object detection probabilities (Buckland et al. 1993). Distance sampling was primarily designed to overcome the challenges of surveying wildlife with the ability to hide from or escape an observer’s detection (Ducey and Astrup 2013); however, the method’s potential for surveying non-wildlife objects, including vegetation, has also been recognized (Buckland et al. 1993; Burnham et al. 1980; Hayes and Buckland 1983; Drummer and McDonald 1987). Distance sampling is also considerably flexible with regards to implementation as it can be conducted with a line-based sample design known as line transect sampling (LTS) or a point-based sample design known as point transect sampling (PTS). Such flexibility in sample focus and design allows virtually any forest component to be sampled, regardless of its mobility or if it is commonly accounted for in traditional forest inventories.

Previous studies have used LTS to sample a variety of forest components including invasive plants (Kenny 2015), cowslips (*Primula veris* L.) (Buckland et al. 2007), Scottish primroses (*P. scotica* Hook.) (Shewry et al. 2002), Brazilnut trees (*Bertholletia excelsa* Humb. &

Bonpl.) (Peres and Baider 1997), iroko trees (*Milicia excelsa* Welw. C.C. Berg) (Ouinsavi and Sokpon 2010), *Myrianthus holstii* Engl. trees (Kissa and Sheil 2012), habitat trees (Bäuerle et al. 2009; Bäuerle and Nothdurft 2011), dead wood (Bäuerle et al. 2009), and tree holes (Didas 2009). Alternatively, PTS has been used to sample dead wood (Ritter et al. 2013; Ritter and Saborowski 2012, 2014) and an assortment of fruiting tree species (Marsden and Pilgrim 2003). Detection probability models rooted in PTS estimation processes have also been used to correct for the bias associated with a variety of forest parameter estimates derived from non-distance sampling inventory methods (Ritter et al. 2013; Ducey and Astrup 2013; Astrup et al. 2014). Additionally, given certain environmental conditions, distance sampling has been found to be capable of more efficiently sampling specific forest components than belt transect sampling (Kissa and Sheil 2012), HPS (Ritter and Saborowski 2012), and FAS (Ritter and Saborowski 2012, 2014).

Though these studies have demonstrated the flexibility and utility of distance sampling when used in the context of forestry, land managers interested in the method may still be uncertain about when to use it and how to decide between using LTS or PTS. To resolve such uncertainty, we sampled trees 45.72 cm (18 in) or greater in diameter at breast height (DBH) (hereafter referred to as “large trees”) over the same forest tracts with two traditional forest inventory methods and both forms of distance sampling. Large trees were selected as the focus of sampling efforts as they are frequently of interest to land managers due to their timber value (Cassens 2004), seed production potential (Healy and Houf 1989), benefits to wildlife (Tubbs et al. 1987), and role in providing forest structure (Bäuerle and Nothdurft 2011). Resulting abundance estimates associated with each inventory technique were then compared against a tract-specific reference population, and the sampling method which most precisely and

accurately accounted for large trees was determined. In addition, we recorded the sampling time associated with LTS and PTS so that linear regression models of sampling time could be produced, and the efficiency of each form of distance sampling could be assessed.

A comparison of forest inventory methods may provide land managers with a sense of how adequately their preferred method accounts for various forest components, as well as whether they should consider using distance sampling. Furthermore, a comparison of LTS and PTS efficiency when used in a forestry context will also provide land managers with insight into how they can best accommodate their needs given resource constraints and the field conditions they routinely encounter. Finally, another demonstration of distance sampling accounting for valuable, yet sparsely distributed forest components may cause land managers to use the method as a replacement for, or as B auerle et al. (2009) suggested, as a supplement to traditional forest inventory methods.

Methods

This study was conducted within three forest compartments of the Bartlett Experimental Forest (BEF) in Bartlett, NH USA. For this study, two of the three compartments were combined and considered a single 40.47-ha (100-ac) compartment (hereafter referred to as the “northern hardwood compartment”) because they were contiguous tracts of northern hardwood forest. The first tract of land forming the northern hardwood compartment was BEF Compartment 36 which had an area of 21.45 ha (53 ac) and a clearcut harvest in its southern portion in 1999. The second tract of land forming the northern hardwood compartment was BEF Compartment 37 which had an area of 19.02 ha (47 ac) and a diameter limit cut also conducted around 1999. The species defining the northern hardwood compartment were sugar maple (*Acer saccharum* Marshall), American beech (*Fagus grandifolia* Ehrh.), and yellow birch (*Betula alleghaniensis* Britton),

though other trees such as eastern hemlock (*Tsuga canadensis* (L.) Carrière) and red maple (*Acer rubrum* L.) were also frequently encountered throughout the area. The third study compartment was BEF Compartment 26 (hereafter referred to as the “hemlock-mixed hardwood compartment”) which was 31.97 ha (79 ac) of designated natural area with no recent harvesting. Many of the species found in the hemlock-mixed hardwood compartment were the same as those found in the northern hardwood compartment; however, eastern hemlock trees represented a greater proportion of all stems in the former.

Within each forest compartment, large trees were sampled with two traditional forest inventory methods (HPS and FAS) and two forms of distance sampling (LTS and PTS) (Figure 1). To assess each method’s precision and accuracy, abundance estimates associated with each form of sampling were compared with a compartment-specific reference population. An analysis of sampling time was also conducted for LTS and PTS, allowing efficiency to be determined and for stronger conclusions to be drawn regarding the appropriate conditions for distance sampling to be employed as a forest inventory technique. Specifics regarding efficiency analysis and the four employed sampling methods follow.

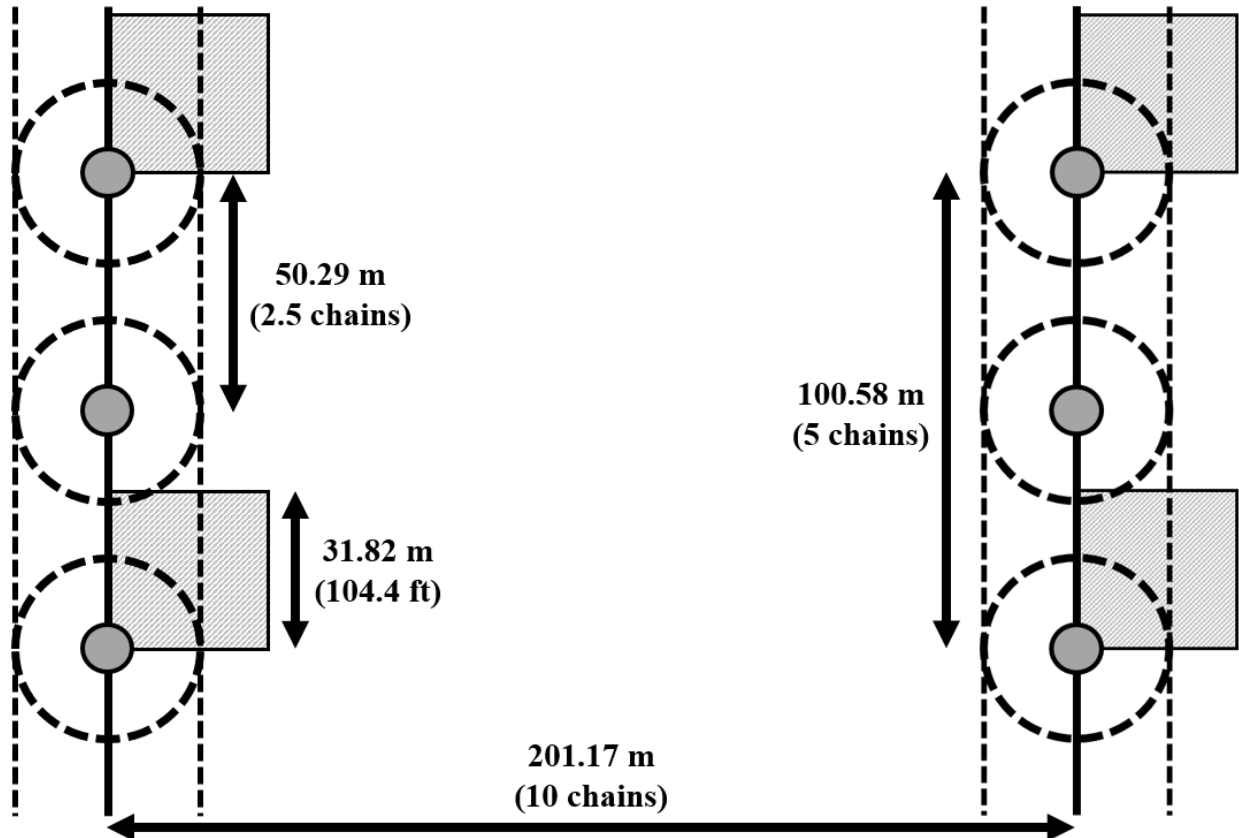


Figure 1. Inventory method sample designs. Designs include 0.1 ha (0.25 ac) fixed area plots (gray squares), horizontal point sampling points (gray circles), distance sampling line transects (black solid lines) with truncated search distance of 20.12 m (1 chain, black dashed lines), and distance sampling point transects (same location as horizontal point sampling points) with truncated search distance of 20.12 m (1 chain, black dashed circles).

Fixed Area Sampling

Square, 0.1 ha (0.25 ac) fixed area plots have been in place across the BEF landscape since the mid-1900s. These plots were laid out with corresponding corners located along a grid of 100.58 m (5 chains) by 201.17 m (10 chains), though the orientation of the grid's dimensions differs throughout the forest to capture elevation gradients and maximum landscape variability. In the northern hardwood compartment of this study, the grid of plots ran 100.58 m (5 chains) in the east-west direction and 201.17 m (10 chains) in the north-south direction, while in the

hemlock-mixed hardwood compartment the short and long dimensions of the grid were oppositely oriented.

Since their installment, the plots have been periodically maintained and measured, with the most recent measurements between 2015 and 2018 including a complete tally of all stems determined with a diameter tape to be 11.3 cm (4.45 in) or greater in DBH. Stems were tallied as belonging to 2.54 cm (1 in) wide diameter classes, and all stems with a partial DBH measurement of 1.14 cm (0.45 in) or greater were rounded up and tallied in the next highest diameter class. For example, if a tree was measured to have a DBH of 11.3 cm (4.45 in), it was tallied as a 12.7 cm (5 in) tree during the plot's complete tally.

To stay consistent with the minimum DBH of trees sampled with HPS (11.43 cm, or 4.5 in), counts of trees in the 12.7 cm (5 in) class were multiplied by 0.95 prior to calculating the total number of trees with a DBH of 11.43 cm (4.5 in) or greater within each compartment. This calculation accounts for the fact that only 95% of the trees tallied within the within the 12.7 cm (5 in) diameter class would be expected to have a DBH of 11.43cm (4.5 in) or greater. Counts of trees in classes higher than 12.7 cm (5 in) were left unadjusted. Likewise, prior to separately calculating the number of large trees within each compartment, counts of trees in the 45.72 cm (18 in) class were multiplied by 0.45 since only 45% of the trees within that diameter class would be expected to have a DBH of 45.72 cm (18 in) or greater. When calculating the number of large trees within each compartment, counts of trees in classes higher than 45.72 cm (18 in) were left unadjusted.

Only plots that were completely contained within each study compartment boundary were considered during fixed area plot calculations. For the northern hardwood compartment, two fixed area plots that were only partially in the study compartment were excluded from

analysis (in addition, HPS and distance sampling were not conducted where the normal sampling grid was interrupted by the compartment boundary). Moreover, though HPS and distance sampling were conducted close to the northern boundary of the northern hardwood compartment, the sample locations were close to a woods road which prevented four corresponding fixed area plots from being installed.

Horizontal Point Sampling

Starting at the corner of each compartment's first fixed area plot, HPS points were systematically established every 50.29 m (2.5 chains) along the short dimension of each compartment's fixed area plot grid. At each point, a basal area factor (BAF) 20 prism was used to sample the surrounding trees determined with a diameter tape to be 11.43 cm (4.5 in) or greater in DBH. Species, DBH, and live/dead status were recorded for each "in" tree, along with any notes indicating information such as deformities. Many of the trees sampled had physical damage or deformities; therefore, if DBH could not be measured at 1.37 m (4.5 ft) up the tree's stem, it was measured at the closest spot below 1.37 m (4.5 ft) free of deformity. Though rare, DBH measurements on deformed trees were also sometimes taken above 1.37 m (4.5 ft) if a deformity-free spot within reasonable reach of the observer was thought to better approximate the tree's true DBH (had it not had a deformity) than the closest deformity-free spot below 1.37 m (4.5 ft). Dead tree DBH was measured at 1.37 m (4.5 ft) regardless of missing chunks of wood and previously described DBH measurement adjustments were only made in cases of another type of deformity or fungal growth. In the case of a tree being considered "borderline," the distance between the sample point and the pith of the tree in question was measured with a tape, or a Vertex DME unit (Haglöf Sweden) which uses sound waves to measure distance. Boundary

slopoover of individual HPS points was corrected for using the walkthrough method of Ducey et al. (2004).

To further compare HPS with distance sampling (particularly LTS), and to test for sample design influence on abundance estimation, we also calculated estimates of large tree abundance for each compartment as if HPS data were collected along lines as opposed to points. While abundance estimates were derived from the conventional calculations associated with HPS, the selection of data to be used in the estimation process mimicked a line-based sample design as opposed to a point-based design. Within each compartment, the grid of fixed area plots (over which HPS points were also established) provided four lines with an associated series of HPS points. We considered each of these four lines as a sampling unit and calculated what our compartment-specific estimate of large tree abundance would have been if we had sampled those four lines with replacement. For example, if line one was selected twice for a given set of samples, followed by lines two and three, then the data of all HPS points associated with those lines would be used in abundance estimation. Furthermore, the data from the series of HPS points associated with line one would appear twice in the abundance calculation, while data from the series of HPS points associated with lines two and three would only appear once. We followed this process for each possible permutation of sample lines and produced 256 estimates of large tree abundance and their associated variances per compartment. Following estimation, the 256 large tree abundance estimates for each compartment were compared with the LTS and PTS confidence intervals believed to best represent the same large trees.

Line Transect Sampling

In each study compartment, a single observer conducted LTS as described by Buckland et al. (1993). Using a sighting compass, line transects were laid out along the short dimension of

the fixed area plot grid in each compartment; east-west in the northern hardwood compartment and north-south in the hemlock-mixed hardwood compartment. Unless the observer walking the transect encountered boundary line, transects were periodically broken up every 50.29 m (2.5 chains) so that a distance sampling point transect and HPS point could be established and sampled. Following PTS and HPS, the observer continued along the line transect from where it last ended (i.e. the point transect and HPS point). While walking along each line transect segment 50.29 m (2.5 chains) in length, the observer pulled an unwinding tape and searched for large trees located 20.12 perpendicular m (1 perpendicular chain) or less from the given line transect, effectively forming a sample strip with a width of 40.23 m (2 chains). In some cases, where the intersection of a line transect by boundary caused unintended truncation of search area, we attempted to account for the missing area by sampling trees in the compartment and within sight of the observer but located beyond the intersection point of the boundary and transect (Figure 2).

For each large tree detected, the observer recorded the distance between the tree and its associated line transect using a Vertex DME (or a tape if the sound of running water prevented the DME from taking an accurate distance measurement). For each detection, the observer also used a sighting compass to measure the angle between the tree and its associated line transect. While distance and angle are the only observation-specific measurements necessary to compute LTS abundance estimates (Buckland et al. 1993), the observer also recorded each observation's species, DBH (as measured with a diameter tape), and live/dead status, as well whether it was first detected due to the observer walking off the line to measure another tree, and any relevant notes. Measurements of DBH were taken as in HPS.

The observer occasionally made slight adjustments to the directional layout of the transects they were walking to ensure that the location of every other point transect and HPS point lined up with the appropriate corner of the next fixed area plot. Though rare, some line transects ran long due to imperfect spacing between fixed area plots, and as a result point transects and HPS points associated with the shifted fixed area plots were also shifted. The extra length of these transects was accounted for in all distance sampling analyses. The four most western fixed area plots of transect three in the northern hardwood compartment were also imperfectly spaced and positioned along the compartment's fixed area grid. Due to the shift in these plots, line transect segments 100.58 m (5 chains) in length did not lead to point transects and HPS points being located at their corners as had been the case for all other unshifted plots in the compartment.

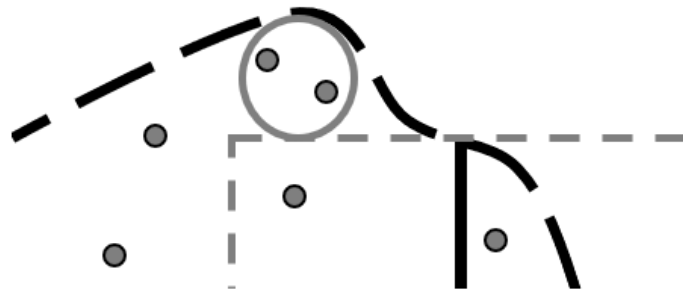


Figure 2. Slopover correction. In some cases, compartment boundary (black dashed line) would intersect a line transect (solid black line) and physically truncate search area (gray dashed line). Trees (gray circles with black outline) within the search area were eligible for sampling, however, trees in the compartment that were within sight of the observer but located beyond the intersection of boundary and transect (“trees” within the gray outlined circle) were also sampled if irregular, unintended truncation of the search area occurred.

Point Transect Sampling

Like LTS, a single observer conducted PTS in each study compartment according to Buckland et al. (1993) with a truncated search distance of 20.12 m (1 chain) from the point transect. Point transect locations coincided with HPS points, and the measurements taken for

PTS were the same as those taken for LTS, except for an angle to each observed tree since radial distance is used in abundance estimation as opposed to perpendicular distance. We acknowledge that because they exist at the same location, conducting HPS prior to PTS could allow the observer to find large trees that may otherwise go undetected during PTS; therefore, PTS was always conducted prior to HPS.

Additionally, though we recognize the possibility that trees within the overlapping search area of PTS and LTS may have been more easily detected during the observer's second pass through, we did not consider the joint likelihood for the distances to these trees as in Buckland et al. (2007) for several reasons. First, the purpose of our study was to consider LTS and PTS as separate sampling processes and therefore we were not interested in combining the data collected between them. Second, because of their size, large trees were likely to have a relatively high detection probability over a range of distances even without scanning an area twice. Finally, in the case of an observer conducting PTS without LTS, an initial coarse scan of the search area would still occur with the observer walking between PTS locations.

Distance Sampling Estimators and Reference Populations

All analyses for this study were conducted in R (R Core Team 2016). We relied on the R package *Distance* (Miller 2017) for modeling detection probabilities associated with large trees located 20.12 m (1 chain) or less from sampled line or point transects. Final population estimates of large trees and related summary statistics were also obtained from the *Distance* package. For each estimated population of large trees, we attempted to model detection probabilities with the same set of eight detection functions. Included in the set were those described as “generally useful” in Buckland et al. (Section 2.4, 1993): a uniform key function with either a cosine or simple polynomial adjustment, a half-normal key function with either a cosine or Hermite

polynomial adjustment, and a hazard-rate function with either a cosine or simple polynomial adjustment. The half-normal and hazard-rate key functions were also included without any adjustment terms. The uniform key function was excluded from the set of functions we used for modeling because, as Miller et al. (2016) noted, the function requires an adjustment. For each estimated population of large trees, candidate detection probability models were compared via Akaike's Information Criterion (AIC) (Akaike 1973) and the simplest model with a delta AIC (ΔAIC) of less than 2 was selected for use in density estimation. When two key-only functions (i.e. functions without adjustment terms) had ΔAIC values of less than 2, the model with the lowest ΔAIC value was selected.

In equations 1-3 below, we use the notation of Buckland et al. (1993); however, we indicate which sampling method the \hat{P}_a variable is associated with in equations 1 and 2 and that we are working with estimates of abundance and density in equation 3. Density estimates of large trees sampled with LTS were calculated as,

$$\hat{D} = \frac{n}{2L \int_0^w \hat{g}(x) dx} = \frac{n}{2wL\hat{P}_{aLTS}} \quad (\text{Eq. 1})$$

where n is the number of sample objects observed, L is the combined length of the sampled line transects, w is the truncated search distance, $\int_0^w \hat{g}(x) dx$ is the integral of the estimated detection function from 0 m out to the truncated search distance, and \hat{P}_{aLTS} is the probability of making an observation within a sample strip of area 2wL (Buckland et al. 1993). Similarly, density estimates of large trees sampled with PTS were calculated as,

$$\hat{D} = \frac{n}{2k\pi \int_0^w r \hat{g}(r) dr} = \frac{n}{k\pi w^2 \hat{P}_{aPTS}} \quad (\text{Eq. 2})$$

where k is the number of sampled point transects, $\int_0^w r \hat{g}(r) dr$ is the integral of radial distance r multiplied by estimated detection function $\hat{g}(r)$, \hat{P}_{aPTS} is the probability of making an

observation within a sample area of $k\pi w^2$, and all other variables are as previously defined (Buckland et al. 1993). Following density estimation, total abundance over the study tract could be calculated as,

$$\hat{N} = \hat{D}A \quad (\text{Eq. 3})$$

where \hat{D} is estimated density and A is total tract area.

Reference populations and associated variances were calculated using the following equations from Maybeck (Section 1.5, 1979). Presented notation, however, is adapted to suit our needs. Compartment-specific reference populations from which all inventory techniques were compared combined the large tree abundance estimates of HPS and FAS using the equation,

$$\hat{X}_{Composite} = \hat{X}_{HPS} + \left(\frac{\sigma_{HPS}^2}{\sigma_{HPS}^2 + \sigma_{FAS}^2} \right) (\hat{X}_{FAS} - \hat{X}_{HPS}) \quad (\text{Eq. 4})$$

where $\hat{X}_{Composite}$ is the combined HPS and FAS reference estimate of large trees, \hat{X}_{HPS} is the HPS estimate of large trees, \hat{X}_{FAS} is the FAS estimate of large trees, σ_{HPS}^2 is the variance associated with the HPS abundance estimate, and σ_{FAS}^2 is the variance associated with the FAS abundance estimate. Similarly, the variance of each composite reference estimate was calculated as,

$$\sigma_{Composite}^2 = \sigma_{HPS}^2 - \left(\frac{\sigma_{HPS}^2}{\sigma_{HPS}^2 + \sigma_{FAS}^2} \right) \sigma_{HPS}^2 \quad (\text{Eq. 5})$$

where, $\sigma_{Composite}^2$ is the combined HPS and FAS reference variance, and σ_{HPS}^2 and σ_{FAS}^2 are as previously defined.

Efficiency Analysis

The time to conduct each form of distance sampling was recorded in both study compartments. The recorded time for LTS included the time to search for and sample large trees, as well as the time to walk between point locations where PTS and HPS were conducted. The

time it took to reel in the tape that was laid out as the observer walked along each line transect was excluded. The recorded time for PTS included the time to search for and sample large trees, but excluded the time it took to walk between sample locations because LTS was conducted between points.

Since the search area associated with PTS overlapped a portion of the search area associated with LTS, trees located within the overlap should have had their measurements taken twice, once for each form of distance sampling. Instead, trees within the overlap were measured once during whichever form of sampling led the observer to first detect the tree. If the same tree was detected again during the second form of distance sampling, most of the previously recorded information for that tree was recorded again without re-taking measurements such as DBH. However, since the information required for each observation slightly differs between PTS and LTS, one new measurement almost always had to be taken for a tree detected twice. If a tree was first detected along a line transect, the observer usually had to take a new distance measurement to that tree during PTS. Conversely, if a tree was first detected on a point transect, the observer always had to measure the angle between that tree and its associated line transect during LTS.

Because some trees were not completely re-measured the second time they were detected, linear regressions of recorded sampling time and the number of trees first detected (i.e. first measured) on individual transects were produced for each form of distance sampling in each compartment. In some cases, trees that should have been ignored by the observer were sampled. These cases were most frequently due to the observer taking angled distance measurements to trees in LTS which were found post-data collection to have a perpendicular distance greater than 20.12 m (66 ft). Time regressions associated with LTS included these trees because the observer did take the field time to measure and record them. Additionally, because LTS was found to have

a higher fixed cost than PTS, only line transects with a total length within and including $\pm 5\%$ of 50.29 m (165 ft) were used to create LTS time regressions. Resulting regression equations were then used to estimate the total amount of time it would have taken the observer to fully measure all trees detected along each line transect or at each point transect, including those being observed a second time.

Regression equations for each form of distance sampling in each compartment were likely to over-estimate the amount of time it would take for an observer to sample a given number of trees. Over-estimation was likely due to the observer recording time which included re-writing previously detected tree information, taking new measurements associated with the second form of sampling a tree was detected with, and taking measurements on trees which were ultimately left unsampled. Trees left unsampled include those which the observer checked but were found in the field to have a DBH less than 45.72 cm (18 in) or be located outside the truncated search area for the form of sampling being used.

Results

Precision and Accuracy

Estimates of tree abundance varied within and between sampling compartments; however, the composite reference of HPS and FAS always resulted in the estimate with the lowest standard error. For a given series of estimates, the most precise sampling technique yielded the estimate with the standard error closest to that of the composite reference; however, the most precise sampling technique varied depending on the population of trees estimated.

Composite reference estimates for trees in the 40.47-ha (100-ac) northern hardwood compartment included 20,785 trees 11.43 cm (4.5 in) or greater in DBH, 1,196 trees 45.72 cm (18 in) or greater in DBH, and 736 northern hardwood (i.e. sugar maple, American beech, and

yellow birch) trees 45.72 cm (18 in) or greater in DBH (Table 1). Under these composite reference estimates, all large trees and large northern hardwood trees represented approximately 5.8% and 3.5%, respectively, of all trees in the compartment with a DBH of 11.43 cm (4.5 in) or greater. Composite reference abundance estimates were higher for the 31.97-ha (79-ac) hemlock-mixed hardwood compartment and included 22,872 trees 11.43 cm (4.5 in) or greater in DBH, 2,288 trees 45.72 cm (18 in) or greater in DBH, and 1,150 eastern hemlock trees 45.72 cm (18 in) or greater in DBH (Table 2). Under these composite reference estimates, all large trees and large eastern hemlock trees represented approximately 10% and 5%, respectively, of all trees in the compartment with a DBH of 11.43 cm (4.5 in) or greater.

Only HPS and FAS were compared when estimating the abundance of trees 11.43 cm (4.5 in) or greater in DBH within each compartment; however, all four sampling techniques conducted during the study were compared when estimating the abundance of large trees. When accounting for all trees 11.43 cm (4.5 in) or greater in DBH within each compartment and all trees 18 in or greater in DBH in the northern hardwood compartment, FAS was the most precise of the compared sampling techniques. When accounting for northern hardwood trees 45.72 cm (18 in) or greater in DBH in the northern hardwood compartment and all trees 45.72 cm (18 in) or greater in DBH in the hemlock-mixed hardwood compartment, HPS was the most precise of the compared sampling techniques.

When all four sampling techniques were compared, at least one of the two traditional forest inventory methods more precisely accounted for large trees than LTS and PTS, except in a solitary case. When accounting for eastern hemlock trees 45.72 cm (18 in) or greater in DBH within the hemlock-mixed hardwood compartment, PTS was the most precise sampling method while LTS was the least precise. No case resulted in which LTS most precisely accounted for a

given forest component; however, for a couple of forest components LTS was not the least precise of the conducted sampling techniques.

Except in two cases, the most precise sampling technique for each estimated forest component also yielded the estimates of abundance and density which most accurately reflected the corresponding estimates of the reference population. Considering all large trees in the hemlock-mixed hardwood compartment, HPS had the lowest standard error of the sampling techniques, but the abundance and density estimates of LTS most accurately reflected the estimates of the reference population. Similarly, when considering large northern hardwood trees in the northern hardwood compartment, HPS again had the lowest standard error of the sampling techniques, but the abundance and density estimates of FAS most accurately reflected the estimates of the reference population.

Table 1. Estimates of abundance (trees) and density (trees/ha) in the northern hardwood compartment. Northern hardwood tree species are sugar maple, yellow birch, and American beech.

Trees with a DBH \geq 11.43 cm (4.5 in)					
Estimate	HPS/FAS Composite	HPS	FAS	LTS	PTS
Abundance	20785	22246	20240	NA	NA
Confidence Interval	\pm 2368	\pm 4541	\pm 4397	NA	NA
Standard Error	1177.5	2258.3	2107.9	NA	NA
Density	514	550	500	NA	NA
Confidence Interval	\pm 59	\pm 112	\pm 109	NA	NA
Standard Error	29.1	55.8	52.1	NA	NA
Trees with a DBH \geq 45.72 cm (18 in)					
Estimate	HPS/FAS Composite	HPS	FAS	LTS	PTS
Abundance	1196	1384	1118	1386	1433
Confidence Interval	\pm 223	\pm 411	\pm 421	-629, +1151	-423, +599
Standard Error	110.9	204.3	201.6	290.0	255.2
Density	30	34	28	34	35
Confidence Interval	\pm 6	\pm 10	\pm 10	-16, +28	-10, +15
Standard Error	2.7	5.0	5.0	7.2	6.3
Northern Hardwood (i.e. Sugar Maple, American Beech, and Yellow Birch) Trees with a DBH \geq 45.72 cm (18 in)					
Estimate	HPS/FAS Composite	HPS	FAS	LTS	PTS
Abundance	736	665	779	618	632
Confidence Interval	\pm 183	\pm 299	\pm 367	-315, +645	-254, +425
Standard Error	91.2	148.9	176.2	160.5	166.9
Density	18	16	19	15	16
Confidence Interval	\pm 5	\pm 7	\pm 9	-8, +16	-6, +11
Standard Error	2.3	3.7	4.4	4.0	4.1

*Interval values are based on a 95% confidence interval, except for LTS and PTS which are based on information supplied by the R package *Distance*.

Table 2. Estimates of abundance (trees) and density (trees/ha) in the hemlock-mixed hardwood compartment.

Trees with a DBH \geq 11.43 cm (4.5 in)					
Estimate	HPS/FAS Composite	HPS	FAS	LTS	PTS
Abundance	22872	24197	22565	NA	NA
Confidence Interval	\pm 2101	\pm 4839	\pm 3634	NA	NA
Standard Error	1023.8	2358.3	1667.9	NA	NA
Density	715	757	706	NA	NA
Confidence Interval	\pm 66	\pm 151	\pm 114	NA	NA
Standard Error	32.0	73.8	52.2	NA	NA
Trees with a DBH \geq 45.72 cm (18 in)					
Estimate	HPS/FAS Composite	HPS	FAS	LTS	PTS
Abundance	2288	2088	2404	2293	2086
Confidence Interval	\pm 261	\pm 430	\pm 511	-542, +710	-497, +652
Standard Error	127.0	209.5	234.4	248.4	289.2
Density	72	65	75	72	65
Confidence Interval	\pm 8	\pm 13	\pm 16	-17, +22	-16, +20
Standard Error	4.0	6.6	7.3	7.8	9.0
Eastern Hemlock Trees with a DBH \geq 45.72 cm (18 in)					
Estimate	HPS/FAS Composite	HPS	FAS	LTS	PTS
Abundance	1150	1059	1205	1285	1152
Confidence Interval	\pm 244	\pm 399	\pm 480	-522, +880	-316, +435
Standard Error	118.8	194.5	220.1	246.4	183.6
Density	36	33	38	40	36
Confidence Interval	\pm 8	\pm 12	\pm 15	-16, +28	-10, +14
Standard Error	3.7	6.1	6.9	7.7	5.7

*Interval values are based on a 95% confidence interval, except for LTS and PTS which are based on information supplied by the R package *Distance*.

Distance Sampling and HPS Permutations

When HPS data were treated as if they were collected along lines as opposed to points, and permutations of line transects were used to estimate the abundance of all large trees in each compartment, the range of estimates and associated variances was greater in the northern hardwood compartment than the hemlock-mixed hardwood compartment. When the permutation estimates were compared with the confidence intervals for the best fit LTS and PTS models of all large trees (presented in tables 1 and 2), the intervals overlapped most of the population estimates associated with the northern hardwood compartment and all estimates associated with the hemlock-mixed hardwood compartment (Figure 3).

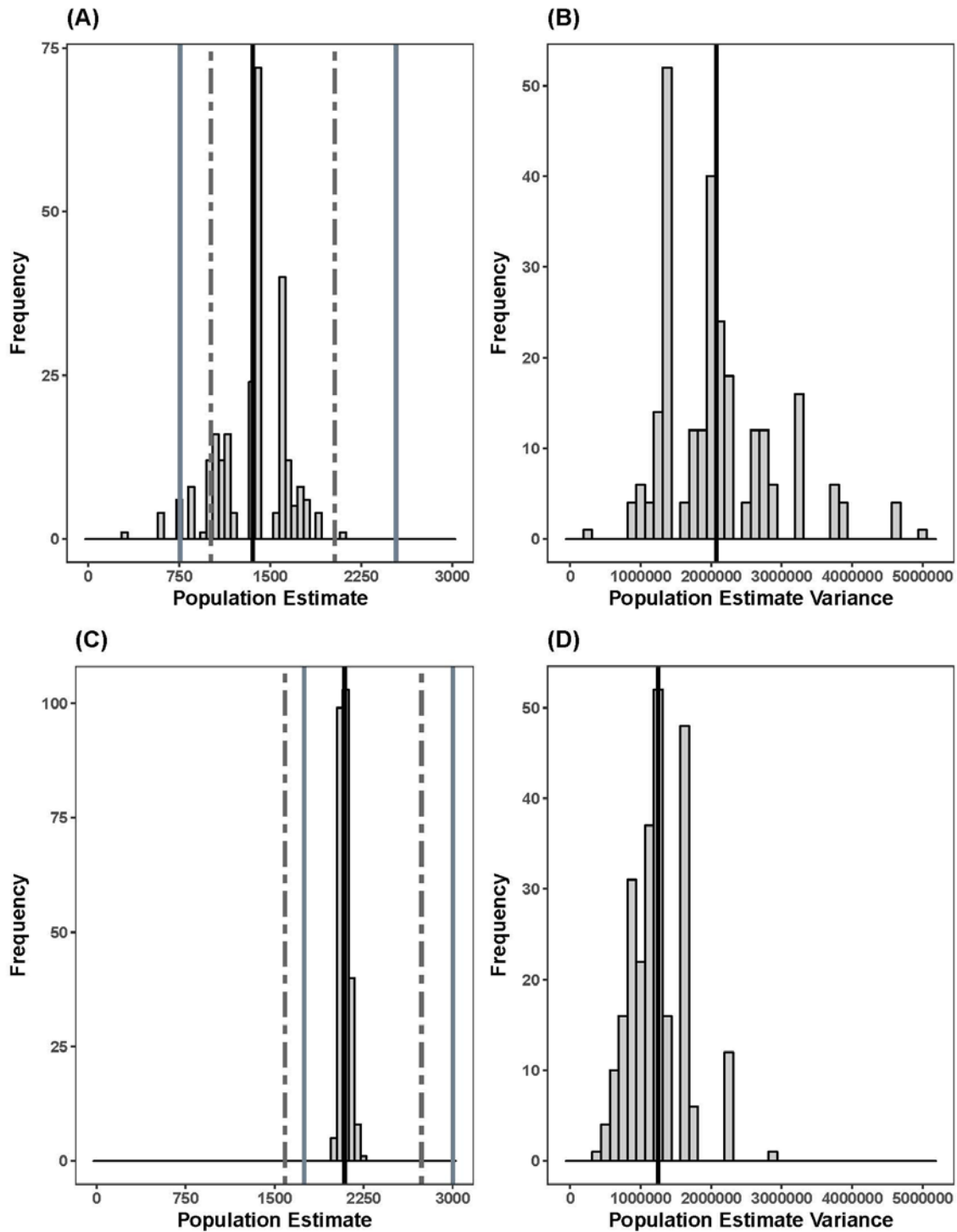


Figure 3. Permutation population estimates and associated variances. Panels show (A) permutation population estimates and (B) associated variances for the northern hardwood compartment, as well as (C) permutation population estimates and (D) associated variances for the hemlock-mixed hardwood compartment. Figures show permutation means (solid black lines), as well as the upper and lower confidence limits associated with the best fit LTS (solid light gray lines) and PTS (dashed gray lines) models of all large trees in each compartment.

Distance Sampling Models

For both LTS and PTS, the same key detection functions and their corresponding adjustment terms were used to model large tree detection probability over distance, though not all functions could always be fitted to the observed distance data for a given forest component. As a result, some of the forest components accounted for with distance sampling had an unequal number of models to compare; however, each had a minimum of three detection models compared prior to one being selected as the best fit (all successfully fitted models are presented in Appendix I). Each of the best fit detection models for the forest components accounted for with distance sampling were based on a half-normal detection function with no adjustment terms or a hazard-rate detection function with no adjustment terms.

Despite best fit models arising from only two types of detection functions, the number of observations associated with these models was dependent upon the population of trees sampled and the form of distance sampling used in the process. When considering all large trees in the northern hardwood compartment, detection functions were fitted to 292 observations for LTS and 177 observations for PTS. In the hemlock-mixed hardwood compartment the total number of large trees observed was higher; therefore, detection functions were fitted to 392 observations for LTS and 229 observations for PTS. All detection functions leading to the best fit models of trees 45.72 cm (18 in) or greater in DBH regardless of species in both the northern hardwood and hemlock-mixed hardwood compartment were half-normal detection functions with no adjustment terms (Figure 4).

In the northern hardwood compartment, large trees were modeled to have a minimum detection probability of approximately 60% when located at the perpendicular truncation distance of a given line transect or the corresponding radial distance of a given point transect.

Minimum large tree detection probabilities were higher for each form of distance sampling in the hemlock-mixed compartment than the northern hardwood compartment. When located at the perpendicular truncation distance of a given line transect, large trees in the hemlock-mixed hardwood compartment were modeled to have a minimum detection probability of approximately 80%. In the same compartment, large trees located at the radial truncation distance of a given point transect were modeled to have a minimum detection probability greater than 90%.

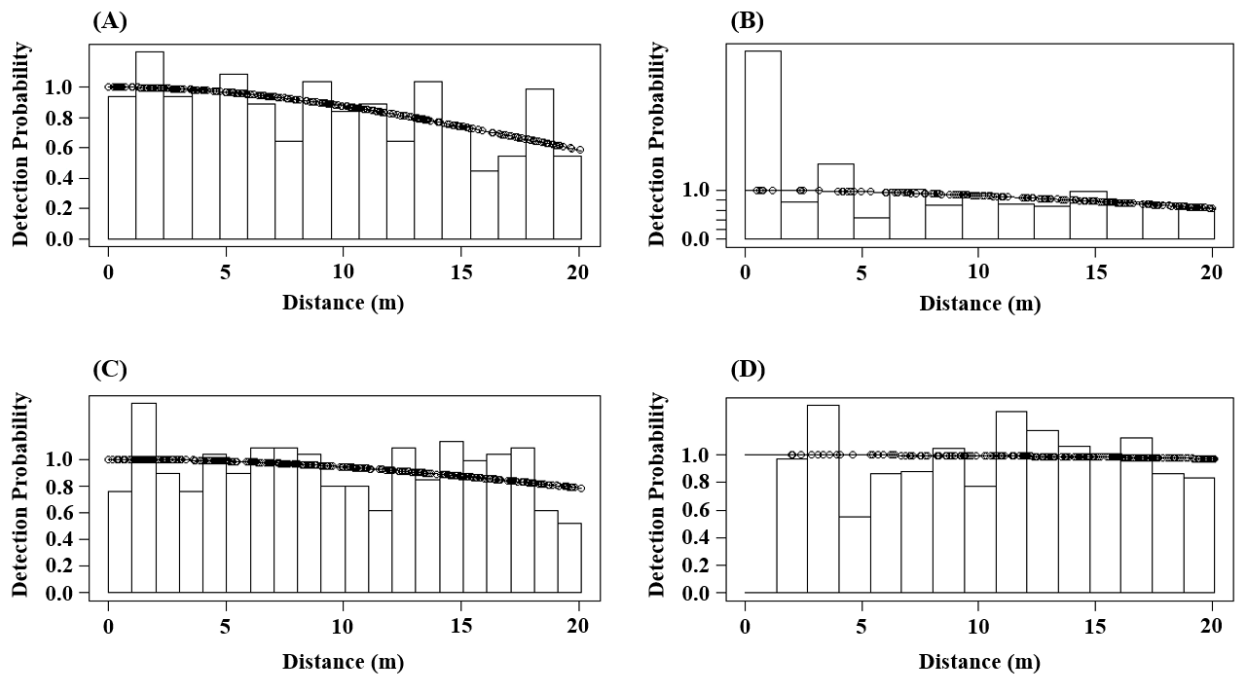


Figure 4. Best fit detection probability models for large trees of all species. Models are associated with all large trees sampled in the northern hardwood compartment with (A) LTS and (B) PTS, and all corresponding trees in the hemlock-mixed hardwood compartment sampled with (C) LTS and (D) PTS. All detection functions are half-normal key functions with no adjustment terms.

When only considering large northern hardwood trees in the northern hardwood compartment, detection functions were fitted to 131 observations for LTS and 81 observations for PTS. In the hemlock-mixed hardwood compartment, when only considering large eastern hemlock trees, detection functions were fitted to 204 observations for LTS and 118 observations

for PTS. Detection functions associated with the northern hardwood compartment leading to the best fit LTS and PTS models of northern hardwood trees 45.72 cm (18 in) or greater in DBH were half-normal detection functions with no adjustment terms. A half-normal detection function with no adjustment terms also lead to the best fit LTS model of eastern hemlock trees 45.72 cm (18 in) or greater in DBH in the hemlock-mixed hardwood compartment. For PTS, however, the detection function leading to the best fit model of large eastern hemlock trees in the hemlock-mixed hardwood compartment was a hazard-rate detection function with no adjustment terms (Figure 5).

In the northern hardwood compartment, large northern hardwood trees were modeled to have a minimum detection probability of approximately 60% when located at the perpendicular truncation distance of a given line transect. Large northern hardwood trees in the northern hardwood compartment were also modeled to have a minimum detection probability between 60% and 70% when located at the radial truncation distance of a given point transect. In the hemlock-mixed hardwood compartment, large eastern hemlock trees were modeled to have a minimum detection probability of approximately 60% when located at the perpendicular truncation distance of a given line transect or the corresponding radial distance of a given point transect.

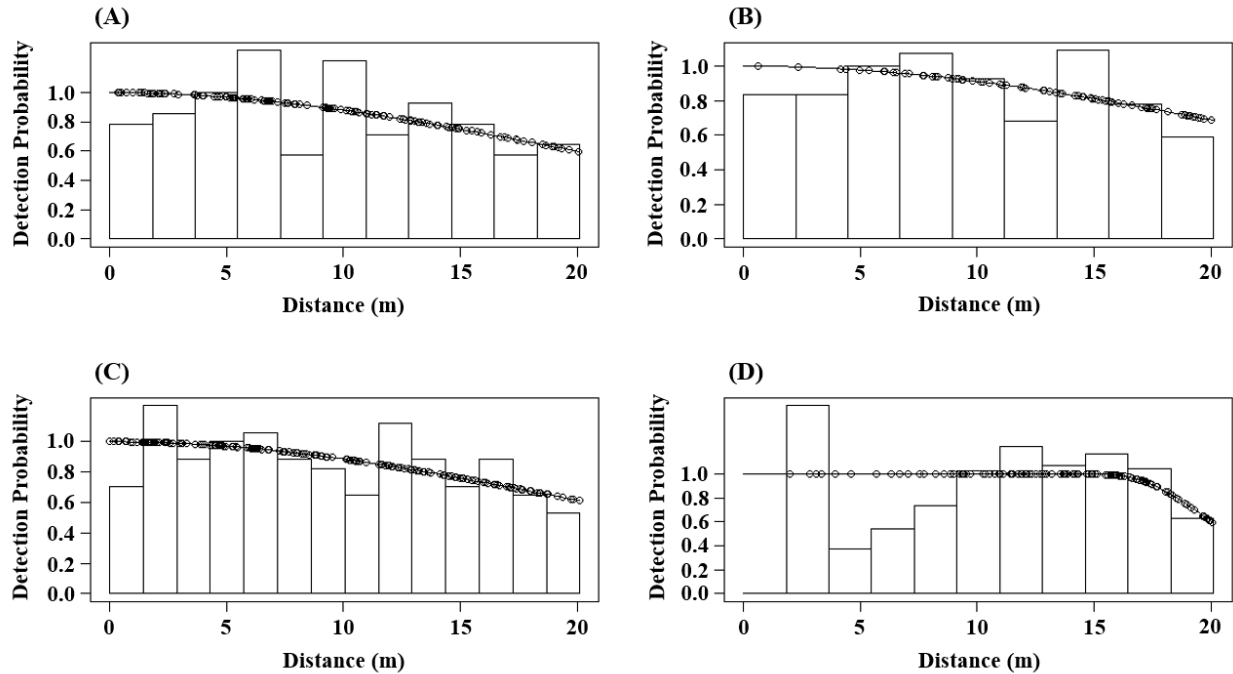


Figure 5. Best fit detection probability models for large trees of select species. Models are associated with large northern hardwood trees sampled in the northern hardwood compartment with (A) LTS and (B) PTS, and large eastern hemlock trees sampled in the hemlock-mixed hardwood compartment with (C) LTS and (D) PTS. All detection functions are half-normal functions with no adjustment terms, except the hazard-rate function with no adjustment terms modeling the detectability of large eastern hemlock trees sampled in the hemlock-mixed hardwood compartment with PTS.

Distance Sampling Efficiency

The time required to sample trees 45.72 cm (18 in) or greater in DBH regardless of species within each compartment not only relied on the compartment being sampled, but also the implemented form of distance sampling. Fixed and variable time costs were associated with both LTS and PTS; however, while variable costs were similar regardless of the form of sampling and the compartment sampled, the fixed cost of sampling an individual line transect was always estimated to be higher than the fixed cost of sampling an individual point transect. When comparing all time regressions, the highest fixed cost associated with PTS was 8.65 min/transect; however, this cost was still less than the lowest fixed cost associated with LTS which was 16.37 min/transect. While LTS was associated with the highest fixed cost of all time regressions, it was

also associated with the lowest variable cost of all time regressions which was 4.58 min/observation. The highest variable cost of all time regressions was 5.54 min/observation and was associated with PTS (Figure 6).

Based on the presented time regressions and the assumption that all trees observed from a given transect were fully measured, it was predicted that it would have taken the observer longer to sample large trees regardless of species in each compartment using LTS as opposed to PTS. In the 100-ac northern hardwood compartment, the total predicted time to complete LTS was 42.8 h, while the total predicted time to complete PTS was 18.6 h; however the predicted sampling time associated with PTS did not include the time required to walk between sample points. Using predicted sampling time for each transect and including transects with lengths other than 50.29 m (165 ft) or search areas truncated by compartment boundary rather than by design, mean sampling time per transect in the northern hardwood compartment was 45.9 min and 22.8 min for line and point transects, respectively. In the more densely populated 31.97-ha (79-ac) hemlock-mixed hardwood compartment, the total predicted time to complete LTS was 45.0 h, while the total predicted time to complete PTS was 22.1 h, though again, the predicted sampling time associated with PTS did not include the time required to walk between sample points. Using predicted sampling time for each transect and including transects of corresponding condition to those used for the northern hardwood compartment, mean sampling time per transect in the hemlock-mixed hardwood compartment was 84.3 min and 47.4 min for line and point transects, respectively.

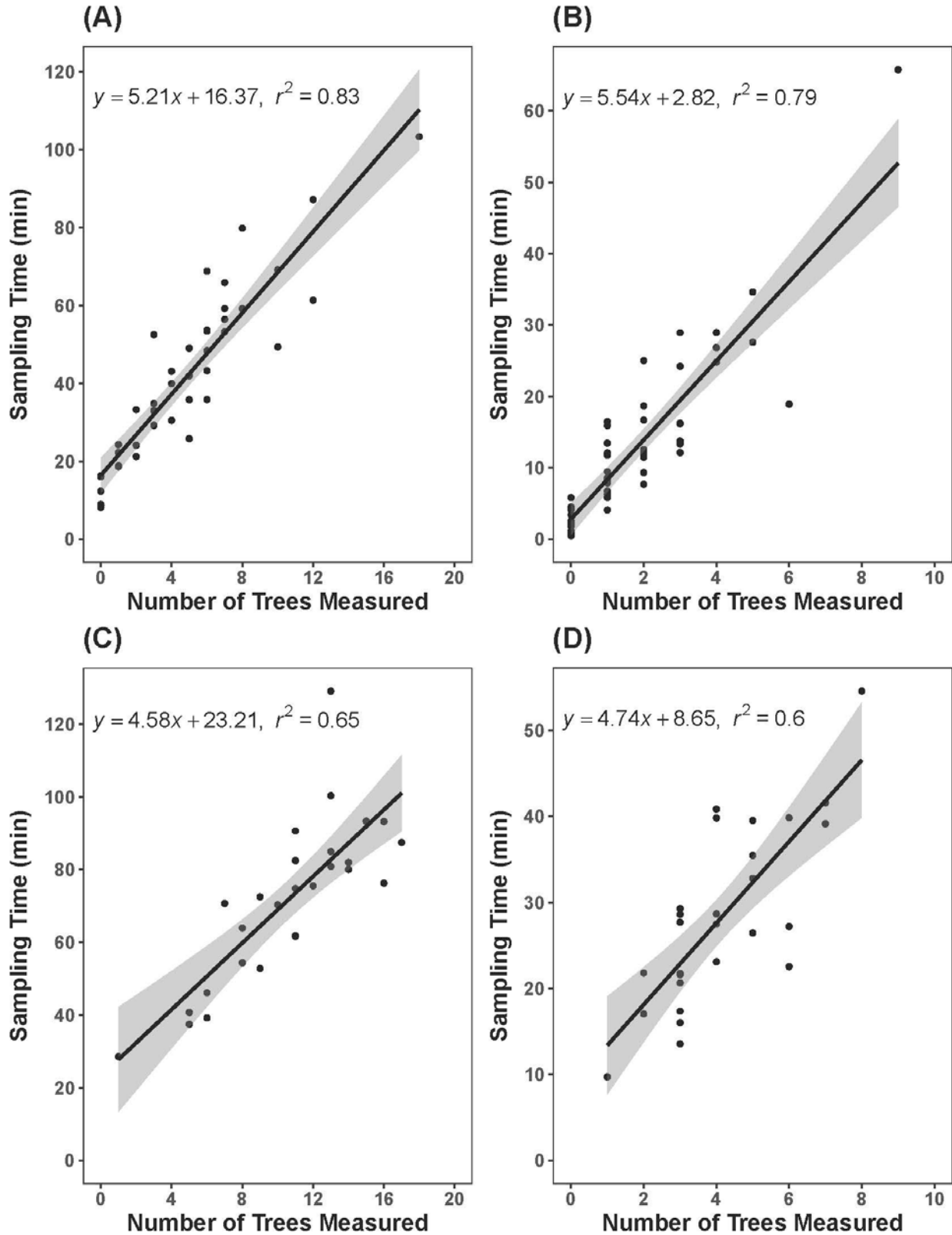


Figure 6. Distance sampling time regressions. Regressions are based on the sampling of trees 18 in. or greater in DBH regardless of species with (A) LTS and (B) PTS in the northern hardwood compartment, and (C) LTS and (D) PTS in the hemlock-mixed hardwood compartment.

Discussion

Sampling Method Comparisons

Our results suggest that in most cases, traditional forest inventory methods lead to more precise and accurate estimates of large tree abundance than distance sampling, even when the trees exist at low density over the sampled landscape. However, our results also demonstrate that distance sampling can be implemented as a legitimate method of forest inventory. For example, confidence intervals for the best fit LTS and PTS models of all large trees in each compartment overlapped with most, or all compartment-specific HPS permutation abundance estimates. This overlap indicates that both LTS and PTS can provide population estimates that are comparable to HPS. Additionally, the similarity of abundance estimates across sampling techniques when estimating a given forest component, and the number of HPS permutation estimates overlapped by distance sampling confidence intervals, suggests that sample design did not have a strong influence on abundance estimation. We also found that in some cases, distance sampling can lead to more precise and accurate estimates of large tree abundance than traditional inventory methods; however, our results suggest that these cases may be limited.

Regardless of the outcome, the ability to compare the accuracy and precision of competing inventory techniques (i.e. HPS, FAS, LTS, and PTS) relied on estimates of abundance and the standard errors associated with those estimates, respectively. While estimates of abundance and standard error are based on more than just the number of objects sampled and their detection probabilities, we acknowledge several factors which may have influenced these variables, and thus the precision and accuracy of LTS and PTS in this study.

The first factor we acknowledge is that our estimates of abundance may have been improved by integrating covariates for variables such as DBH or live/dead status into our large

tree detection probability models. Though we collected potential covariate data such as DBH for each observation, we kept our analyses simple and did not use these covariates as an exploratory measure or a final modeling component. This decision was primarily due to the number of forest components we tried to estimate, and the number of models we tried to fit for each of these components. Despite excluding covariate data, models selected as the best fit for each sampled forest component predicted relatively high detection probabilities over the span of the truncated search distance for each transect. These high detection probabilities suggest that even if covariates were used in the modeling process, the result may have only been marginal model improvements, especially for the LTS and PTS models of all large trees sampled in the hemlock-mixed hardwood compartment.

The second factor we acknowledge is that we included trees first detected away from their associated transect in our distance sampling analyses. These trees were generally detected due to the observer leaving the transect to measure another tree; however, they were counted and measured as if they were detected while the observer was on the transect. The rationale for including these trees was that a land manager implementing distance sampling would be unlikely to ignore them simply because they were off their sample transect. Furthermore, many of these trees were likely to have been detected had the observer walked the entire line transect or scanned the entire point transect prior to measuring any detections. As a result, it was determined that the number of observations within a given data set should not be reduced by intentionally ignoring trees which met the sampling criteria.

The third factor we acknowledge is that the search area associated with a transect was sometimes physically truncated by compartment boundary. This unintended truncation is likely to have reduced the number of detections the observer would have otherwise made, and as a

result, distance sampling analyses should have been corrected for the realized search area. Due to the frequency in which this situation arose, we believe an efficient and easily implemented field correction for boundary slopover would make distance sampling more suitable for use in a forestry context.

When a boundary truncated the length of any 20 m (65.62 ft) line transect, Didas (2009) accounted for its full length by implementing a boundary slopover correction based on the work of Ducey et al. (2002). In practice, Didas (2009) sampled the missing length in the opposite direction of original travel, beginning from the center of the 40 m (131.23 ft) transect from which the cutoff transect was a part. While this approach was able to account for truncated line length, it is not likely to be as useful when accounting for irregularly truncated search areas. Buckland et al. (Chapter 3, 1993) proposed a variable (“c”) which can be used in density estimation to account for the fraction of line transect or point transect search areas actually sampled. This approach, however, is not likely to solve the problem of irregular search area truncation as it is a single variable representing a fraction of all transect search areas. With the correction procedure implemented in this study, we were unable to quantify whether we partially, fully, or over-accounted for slopover, therefore the development of a more adequate solution to this problem would further support the use of distance sampling as a method of forest inventory.

Sampling Efficiency

Though LTS sometimes led to more precise and (or) accurate estimates of large trees than PTS, our time regressions and total predicted sampling times suggest that land managers working in conditions like those we encountered would likely prefer PTS over LTS because of the time in which it can be conducted. While such results provide insight into how a distance sampling inventory can function, we caution land managers from using our time data and corresponding

conclusions as a direct indication of how their inventory will go. To this point, our results also suggest that sampling time may be influenced by many factors including the form of distance sampling implemented, sample design, selected truncation distance, how thoroughly the observer searches for sample objects, forest type, terrain ruggedness, and the density and abundance of the sample object. Previous studies have also acknowledged some of these factors as potential controls of sampling time. For example, it has been recognized that if search distance is left high or untruncated, distance sampling may become inefficient due to the number of objects eligible for sampling (Kissa and Sheil 2012, Ritter and Saborowski 2012). Similarly, it has been suggested that when sample object density is high, the pace at which LTS can be conducted may be reduced (Kenny 2015). Moreover, while it was not something we encountered, seasonality has also been shown to influence the sampling time of PTS in areas that experience snow in winter and more foliated vegetation in the summer (Ritter and Saborowski 2012, 2014). Land managers interested in conducting distance sampling should consider these factors when selecting which form of the method to use and establishing their sample design.

There are ways, however, in which sampling time can be controlled such as limiting the number or length of transects sampled, changing the auxiliary information collected for each observation, and selecting an appropriate truncation distance. While trying to limit sampling time, however, land managers should also be aware that the forest component they are interest in can also be too sparsely distributed for distance sampling to be an appropriate sampling method. Buckland et al. (1993) suggest that to fit a proper model of detection probabilities and acquire reliable population estimates, a minimum of 60 to 80 observations and 75 to 100 observations should be made for LTS and PTS campaigns, respectively. Such a requirement presents the

challenge that land managers must have an idea of how sparsely distributed the forest component they wish to sample is, or risk investing time which will not lead to meaningful estimates.

Sampling and Management Implications

Best fit detection probability models suggested that large trees were relatively easy to detect during LTS and PTS, even when located at the truncated search distance associated with each sample design. Though, as congruent with the theory of distance sampling, trees located further from a given transect (i.e. closer to the transect's truncated search distance) were modeled as being more difficult to detect than those nearest the sample line or point. This decrease in detection probability over distance leads to concerns regarding forest sampling with methods that do not account for non-detection bias, especially since large trees may be easier to detect at a given distance than many other forest components because of their size (Drummer and McDonald 1987). This decrease in large tree detection probability also justifies the concerns of Chen et al. (2009) that non-detection is an under-estimated and critical issue in vegetation sampling.

However, given the observed precision, accuracy, and efficiency of both LTS and PTS in this study, the main advantage of using distance sampling over traditional inventory methods is not likely to be the estimates it provides or the pace in which it can be conducted, but the flexibility it permits in sample focus and design. For example, distance sampling provides a reliable point-based alternative to HPS when the forest component of interest cannot be sampled with a prism. Moreover, LTS and PTS both have a simpler sample design than many inventory methods proposed for sampling sparse and clustered forest components, and neither LTS nor PTS requires the observer to spend time establishing plots as in FAS. It should be noted, however, that regardless of any advantage in estimation or sampling efficiency, distance

sampling should not be preferred over FAS or HPS for continuous forest inventory (CFI) as CFI is used for repeatedly sampling and monitoring the same vegetation over time. Additionally, as with all forest sampling and management, a single approach will not work for all forest components in all ecosystems; therefore, land managers conducting an inventory should decide whether the purpose of their sampling justifies the cost.

Conclusion

Traditional forest inventory methods can account for large trees more precisely and more accurately than either form of distance sampling in a variety of conditions. Moreover, given the observed time requirements to conduct a basic distance sampling survey, land managers interested in acquiring abundance and density estimates of these trees are likely to remain using HPS or FAS for data collection. Our results demonstrate, however, that under certain conditions, distance sampling can more precisely and accurately account for sparse and clustered forest components than HPS and FAS. In the hemlock-mixed hardwood compartment, LTS most accurately accounted for all large trees, while PTS most precisely and accurately accounted for large eastern hemlock trees.

Furthermore, distance sampling is likely to have merit over traditional forest inventory methods when the object of interest cannot be sampled with a prism, or when the sampler believes that non-detection is an issue. Our results demonstrate that non-detection can be an issue even when sampling more obvious forest components such as large trees; therefore, land managers should re-evaluate their preferred inventory methods and ensure that bias arising from non-detection is limited. In deciding whether to use distance sampling, land managers should consider not only their time and resource restrictions, but also the conditions in which they will be working and the value of the forest component they wish to sample.

References

- Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle, in *International Symposium on Information Theory* (second edition) (eds B.N. Petran and F. Csàaki), Akadèemiai Kiadi, Budapest, Hungary, pp. 267-281.
- Astrup, R., Ducey, M.J., Granhus, A., Ritter, T., and Lüpke, N. von. 2014. Approaches for estimating stand-level volume using terrestrial laser scanning in a single-scan mode. *Canadian Journal of Forest Research*, 44(6): 666-676.
- Bäuerle, H., and Nothdurft, A. 2011. Spatial modeling of habitat trees based on line transect sampling and point pattern reconstruction. *Canadian Journal of Forest Research*, 41(4): 715-727.
- Bäuerle, H., Nothdurft, A., Kändler, G., and Bauhus, J. 2009. Monitoring von Habitatbäumen und Totholz auf Basis von Stichproben. *Allgemeine Forst- und Jagdzeitung*, 180(11/12): 249-260.
- Bitterlich, W. 1947. Die winkelmessung. *Allgemeine Forst- Holzwirtschaft. Ztg.*, 58: 94-96.
- Bitterlich, W. 1984. The relascope idea: Relative measurements in forestry. CAB International, Slough, England. 242 p.
- Buckland, S.T., Anderson, D.R., Burnham, K.P., and Laake, J.L. 1993. *Distance sampling: estimating abundance of biological populations (first edition)*. Chapman & Hall, London.
- Buckland, S.T., Borchers, D.L., Johnston, A., Henrys, P.A., and Marques, T.A. 2007. Line transect methods for plant surveys. *Biometrics*, 63(4): 989-998.
- Burnham, K.P., Anderson, D.R., and Laake, J.L. 1980. Estimation of density from line transect sampling of biological populations. *Wildlife Monographs*, (72): 3-202.
- Cassens, D.L. 2004. Factors determining the suitability of trees and logs for the face veneer industry. In: Yaussy, D.A., Hix, D.M., Long, R.P., and Goebel, P.C., eds. Proceedings, 14th Central Hardwood Forest Conference; 2004 March 16-19; Wooster, OH. Gen. Tech. Rep. NE-316. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station: 130-139.
- Chen, G., Kéry, M., Zhang, J., and Ma, K. 2009. Factors affecting detection probability in plant distribution studies. *Journal of Ecology*, 97(6): 1383-1389.
- Didas, C. M. 2009. Sampling and classification of tree holes within a Northeast temperate forest system. M.S. Thesis, University of New Hampshire, Durham.
- Drummer, T.D., and McDonald. 1987. Size bias in line transect sampling. *Biometrics*, 43(1): 13-21.

- Ducey, M.J., and Astrup, R. 2013. Adjusting for nondetection in forest inventories derived from terrestrial laser scanning. *Canadian Journal of Remote Sensing*, 39(5): 410-425.
- Ducey, M.J., Jordan, G.J., Gove, J.H., and Valentine, H.T. 2002. A practical modification of horizontal line sampling for snag and cavity tree inventory. *Canadian Journal of Forest Research*, 32: 1217-1224.
- Ducey, M.J., Jordan, G.J., Gove, J.H., and Valentine, H.T. 2004. A walkthrough solution to the boundary overlap problem. *Forest Science*, 50(4): 427-435.
- Grosenbaugh, L.R. 1955. Better diagnosis and prescription in southern forest management. USDA Forest Service, Southern Forest Experiment Station, Occasional Paper 145: 1-27.
- Grosenbaugh, L.R. 1958. Point-sampling and line-sampling: probability theory, geometric implications, synthesis. USDA Forest Service, Southern Forest Experiment Station, Occasional Paper 160: 1-34.
- Hayes, R.J., and Buckland, S.T. 1983. Radial-distance models for the line-transect method. *Biometrics*, 39(1): 29-42.
- Healy, W.H., and Houf, G.F. 1989. Treating mature stands for wildlife. In: Hutchinson, J.G., ed. Central hardwood notes. St. Paul, MN.: U.S. Dept. of Agriculture, Forest Service, North Central Forest Experiment Station. 9.05.
- Kenning, R.S., Ducey, M.J., Brissette, J.C., and Gove, J.H. 2005. Field efficiency and bias of snag inventory methods. *Canadian Journal of Forest Research*, 35: 2900-2910.
- Kenny, C. 2015. The effects of recreation and disturbance on the invasibility of forest interiors. M.S. Thesis, University of New Hampshire, Durham.
- Kissa, D.O., and Sheil, D. 2012. Visual detection based distance sampling offers efficient density estimation for distinctive low abundance tropical forest tree species in complex terrain. *Forest Ecology and Management*, 263: 114-121.
- Marsden, S.J., and Pilgrim, J.D. 2003. Diversity and abundance of fruiting trees in primary forest, selectively logged forest, and gardens on New Britain, Papua New Guinea. *Tropical Biodiversity*, 8(1), 15-29.
- Maybeck, P.S. 1979. Introduction. In: *Stochastic models, estimation, and control (volume 1)*. Academic Press, New York.
- Miller, D.L. 2017. Distance: Distance sampling detection function and abundance estimation. R package Version 0.9.7. <https://CRAN.R-project.org/package=Distance>

- Miller, D.L., Rexstad, E., Thomas, L., Marshall, L., and Laake, J.L. 2016. Distance sampling in R. bioRxiv, <http://dx.doi.org/10.1101/063891>.
- Ouinsavi, C., and Sokpon, N. 2010. Morphological variation and ecological structure of iroko (*Milicia excelsa* Welw. C.C. Berg) populations across different biogeographical zones in Benin. *International Journal of Forestry Research*, Article ID 658396: 1-10. <https://doi.org/10.1155/2010/658396>
- Peres, C.A., and Baider, C. 1997. Seed dispersal, spatial distribution and population structure of Brazilnut trees (*Bertholletia excelsa*) in southeastern Amazonia. *Journal of Tropical Ecology*, 13: 595-616.
- R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Ringvall, A., Snäll, T., Ekström, M., and Ståhl, G. 2007. Unrestricted guided transect sampling for surveying sparse species. *Canadian Journal of Forest Research*, 37(12): 2575-2586.
- Ritter, T., Nothdurft, A., and Saborowski, J. 2013. Correcting the nondetection bias of angle count sampling. *Canadian Journal of Forest Research*, 43(999): 344-354.
- Ritter, T., and Saborowski, J. 2012. Point transect sampling of deadwood: A comparison with well-established sampling techniques for the estimation of volume and carbon storage in managed forests. *European Journal of Forest Research*, 131(6): 1845-1856.
- Ritter, T., and Saborowski, J. 2014. Efficient integration of a deadwood inventory into an existing forest inventory carried out as a two-phase sampling for stratification. *Forestry*, 87(4): 571-581.
- Roesch, Jr., F.A. 1993. Adaptive cluster sampling for forest inventories. *Forest Science*, 39(4): 655-669.
- Shewry, M.C., Buckland, S.T., and Shaw, P. 2002. Distance Sampling and Its Application to Monitoring Selected Species. Research Survey and Monitoring Report No. 177. Scottish Natural Heritage, Perth, UK.
- Snäll, T., Ribeiro, P.J.Jr., and Rydin, H. 2003. Spatial occurrence and colonisations in patch-tracking metapopulations: local conditions versus dispersal. *Oikos*, 103(3): 566-578.
- Ståhl, G., Ringvall, A., and Lämås, T. 2000. Guided transect sampling for assessing sparse populations. *Forest Science*, 46(1): 108-115.
- Talvitie, M., Leino, O., and Holopainen, M. 2006. Inventory of sparse forest populations using adaptive cluster sampling. *Silva Fennica*, 40(1): 101-108.

- Thompson, S.K. 1990. Adaptive cluster sampling. *Journal of the American Statistical Association*, 85(412): 1050-1059.
- Thompson, S.K. 1991. Stratified adaptive cluster sampling. *Biometrika*, 78(2): 389-397.
- Tubbs, C.H., DeGraaf, R.M., Yamasaki, M., Healy, W.M. 1987. Guide to wildlife tree management in New England northern hardwoods. Gen. Tech. Rep. NE-118. Broomall, PA: U. S. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station. 30 p.

Appendix I

Table 3. Distance sampling models for the northern hardwood compartment.

LTS Models, Trees with a DBH \geq 45.72 cm (18 in)				
Detection Function	Abundance	Confidence	Standard Error	Delta AIC
Uniform w/ 1st Order Cosine	1429	-644, +1173	302.1	0.0122
Uniform w/ 2nd Order Simple Polynomial	1368	-624, +1149	283.3	0.1212
Half-Normal	1386	-629, +1151	290.0	0.0000
Hazard-Rate	1396	-616, +1102	311.1	2.0239
PTS Models, Trees with a DBH \geq 45.72 cm (18 in)				
Detection Function	Abundance	Confidence	Standard Error	Delta AIC
Uniform w/ 1st Order Cosine	1505	-475, +694	291.7	0.2459
Uniform w/ 2nd Order Simple Polynomial	1414	-397, +552	237.1	0.0000
Half-Normal	1433	-423, +599	255.2	0.0171
Hazard-Rate	1468	-595, +1002	394.3	2.2069
Hazard-Rate w/ 2nd Order Cosine	1899	-799, +1380	536.5	1.8402
LTS Models, Northern Hardwood Trees with a DBH \geq 45.72 cm (18 in)				
Detection Function	Abundance	Confidence	Standard Error	Delta AIC
Uniform w/ 1st Order Cosine	628	-318, +645	166.4	0.7419
Uniform w/ 2nd Order Simple Polynomial	615	-317, +653	157.1	0.0000
Half-Normal	618	-315, +645	160.5	0.1285
Hazard-Rate	588	-303, +623	150.3	1.6150
PTS Models, Northern Hardwood (i.e. Sugar Maple, American Beech, and Yellow Birch) Trees with a DBH \geq 45.72 cm (18 in)				
Detection Function	Abundance	Confidence	Standard Error	Delta AIC
Uniform w/ 1st Order Cosine	648	-284, +506	193.2	0.3097
Uniform w/ 2nd Order Simple Polynomial	629	-242, +394	156.7	0.0000
Half-Normal	632	-254, +425	166.9	0.0400
Hazard-Rate	571	-196, +299	122.2	1.7746

* Models selected as best fit in bold. Interval values are based on information supplied by the R package *Distance*.

Table 4. Distance sampling models for the hemlock-mixed hardwood compartment.

LTS Models, Trees with a DBH \geq 45.72 cm (18 in)				
Detection Function	Abundance	Confidence	Standard Error	Delta AIC
Uniform w/ 1st Order Cosine	2286	-542, +710	261.3	0.9406
Uniform w/ 2nd Order Simple Polynomial	2298	-544, +713	245.0	0.0000
Half-Normal	2293	-542, +710	248.4	0.1212
PTS Models, Trees with a DBH \geq 45.72 cm (18 in)				
Detection Function	Abundance	Confidence	Standard Error	Delta AIC
Uniform w/ 1st Order Cosine	2057	-615, +878	374.3	0.0172
Uniform w/ 2nd Order Simple Polynomial	2089	-494, +647	287.2	0.0000
Half-Normal	2086	-497, +652	289.2	0.0014
Hazard-Rate	2126	-350, +418	191.7	1.3361
LTS Models, Eastern Hemlock Trees with a DBH \geq 45.72 cm (18 in)				
Detection Function	Abundance	Confidence	Standard Error	Delta AIC
Uniform w/ 1st Order Cosine	1308	-528, +885	256.0	0.7201
Uniform w/ 2nd Order Simple Polynomial	1279	-524, +888	241.0	0.0000
Half-Normal	1285	-522, +880	246.4	0.1229
Hazard-Rate	1181	-494, +849	215.3	2.1621
PTS Models, Eastern Hemlock Trees with a DBH \geq 45.72 cm (18 in)				
Detection Function	Abundance	Confidence	Standard Error	Delta AIC
Uniform w/ 1st Order Cosine	1078	-456, +790	305.5	4.4750*10 ⁻⁶
Uniform w/ 2nd Order Simple Polynomial	1078	-387, +604	245.2	0.0000
Hazard-Rate	1152	-316, +435	183.6	2.1604*10⁻¹

* Models selected as best fit in bold. Interval values are based on information supplied by the R package *Distance*.

Appendix II

This R markdown document serves as an example and will give the reader an intro to the "Distance" package in R (Miller 2017, see references section of this thesis). For access to all of the data and R code used in this thesis, please visit the following repository.

https://github.com/crbreton/Breton_UNHMS_Thesis_Code_2018.git

The code below gives examples of data prep, detection function modeling, and abundance/density estimation for both line transect sampling (LTS) and point transect sampling (PTS). Examples are based on all trees 45.72 cm (18 in) or greater in diameter at breast height (DBH) in the northern hardwood compartment studied for this thesis. Code comments are in black, chunks of code are in blue, and R output values are in brown. Similar code was used for other forest compartments and other population subsets. Original data are classified under the following columns:

*# **Transect** - Compartment-specific transect number.*
*# **cover.type** - Compartment forest type (northern hardwood or hemlock-mixed hardwood).*
*# **Measured** - Number of trees measured for the first time from the associated transect.*
*# **Total** - Number of trees observed from the transect, regardless of whether any were previously observed and measured on another transect.*
*# **Rec.Time** - Field recorded sample time for the transect. Data format is hh:mm:ss.*
*# **Method** - Whether transect is associated with LTS or PTS.*
*# **Direction** - Direction (azimuth) the observer walked during sampling. Units are deg.*
*# **Ob. Angle** - Azimuth from the observer to the observed tree. For some trees, the angle between Ob. Angle and Direction was calculated and recorded in this column, yet this is inconsequential since the values were transferred to the Angle column used in the following LTS example. Units are deg.*
*# **Angle** - Angle between Ob. Angle and Direction. Units are deg.*
*# **Ob. Distance** - Measured distance between the observer and each observed tree. Units are ft.*
*# **DBH** - Diameter at breast height of each observed tree. Units are in.*
*# **Species** - FIA numerical code for observed tree species (except for "unknown").*
*# **Group** - Yes/no indication of whether the tree was observed as the result of the observer leaving the associated transect to measure another observed tree.*
*# **Live.Dead** - Live/dead status of the observed tree.*

```

# Line.Length - Length of line transect. Units are ft.
# Delete - Yes/no indication of whether the observation needs to be
deleted prior to any analyses.

# Note that for the "Distance" package to function properly, some of
the preceding column names must be used exactly as presented. We
refer the reader to R documentation for the "Distance" package, and to
Miller et al. (2016) in the references section of this thesis for
further information.

#####
# Distance Sampling -- The Basics
#####

# Set working directory.
setwd("C:/Users/Connor/Documents/word_files/graduate_courses/thesis/bef_data")

# The "Distance" package is a flexible, yet easily implemented means
of conducting distance sampling analyses in R. Prior to using it for
the first time, install the package with the following code (do not
reinstall each R session).
install.packages("Distance")

# Following installation, load the package into each new R session.
library(Distance)

## Loading required package: mrds

## This is mrds 2.1.17
## Built: R 3.3.3; ; 2017-06-06 02:04:31 UTC; windows

# The "ds" function fits detection probability models to distance
sampling data. The required arguments should be checked prior to any
model fitting.
?ds

# Read in and check the csv file with all collected distance sampling
data.
all_distance <- read.table("bef_data_ordered_modified_final_ds.csv", s
ep = ",", header = TRUE, stringsAsFactors = FALSE)

str(all_distance)

## 'data.frame': 1178 obs. of 17 variables:
## $ Transect : int 1 1 1 1 1 1 1 1 1 1 ...
## $ cover.type : chr "NH" "NH" "NH" "NH" ...

```

```
## $ Measured : int 0 0 0 0 0 NA NA NA NA NA ...
## $ Total : int 0 0 0 0 0 NA NA NA NA NA ...
## $ Rec.Time : chr "0:09:07" "0:01:11" "0:16:01" "0:05:50" ...
## $ Method : chr "Line" "Point" "Line" "Point" ...
## $ Point.Line : chr "17H - B" "17H" "17H - 16H.5" "16H.5" ...
## $ Direction : int 106 NA 286 NA 286 NA NA NA NA NA ...
## $ Ob.Angle : int NA NA NA NA NA NA NA NA NA NA ...
## $ Angle : int NA NA NA NA NA NA NA NA NA NA ...
## $ Ob.Distance: num NA NA NA NA NA NA NA NA NA NA ...
## $ DBH : num NA NA NA NA NA NA NA NA NA NA ...
## $ Species : int NA NA NA NA NA NA NA NA NA NA ...
## $ Group : chr NA NA NA NA ...
## $ Live.Dead : chr NA NA NA NA ...
## $ Line.Length: num 90.4 NA 165 NA 129.8 ...
## $ Delete : chr "No" "No" "No" "No" ...
```

```
#####
# Data Prep for LTS
#####
```

```
# Subset to a new data set containing only LTS data from the northern
# hardwood compartment.
nh_lts <- all_distance[all_distance$cover.type == "NH" & all_distance$
Method == "Line", ]
```

```
# Check the subset.
str(nh_lts)
```

```
## 'data.frame': 330 obs. of 17 variables:
## $ Transect : int 1 1 1 1 1 1 1 1 1 1 ...
## $ cover.type : chr "NH" "NH" "NH" "NH" ...
## $ Measured : int 0 0 0 NA NA NA 0 1 NA NA ...
## $ Total : int 0 0 0 NA NA NA 0 3 NA NA ...
## $ Rec.Time : chr "0:09:07" "0:16:01" "0:19:13" NA ...
## $ Method : chr "Line" "Line" "Line" "Line" ...
## $ Point.Line : chr "17H - B" "17H - 16H.5" "16H.5 - 16H" "16H - 1
5H.5" ...
## $ Direction : int 106 286 286 NA NA NA 106 286 286 286 ...
## $ Ob.Angle : int NA NA NA NA NA NA NA 331 324 275 ...
## $ Angle : int NA NA NA NA NA NA NA 45 38 11 ...
## $ Ob.Distance: num NA NA NA NA NA NA NA 26.5 50.8 32.8 ...
## $ DBH : num NA NA NA NA NA NA NA 26.8 21.5 25.7 ...
## $ Species : int NA NA NA NA NA NA NA 318 318 318 ...
## $ Group : chr NA NA NA NA ...
## $ Live.Dead : chr NA NA NA NA ...
```

```

## $ Line.Length: num  90.4 165 129.8 NA NA ...
## $ Delete      : chr  "No" "No" "No" "Yes" ...

# Create new columns as required for the "ds" function.

# "Area" which is compartment area in sq.m.
nh_lts$Area <- (43560 * 100) * 0.092903
# 43560 is the number of sq.ft/ac, 100 is the number of acres in the
northern hardwood compartment, and 0.092903 is the number of
sq.m/sq.ft.

# "Sample.Label" which is the same as Transect.
nh_lts$Sample.Label <- nh_lts$Transect

# "Effort" which is the length of each line transect running through
the sample area. But, first create function to sum only non-NA values.
sum_na_rm_fxn <- function(x){

  sum_na_rm <- sum(na.omit(x))
  return(sum_na_rm)
}

# Aggregate effort by transect.
nh_lts_aggregate_effort <- aggregate.data.frame(nh_lts$Line.Length, li
st(Sample.Label = nh_lts$Sample.Label), sum_na_rm_fxn)

# Convert effort by Sample.Label from ft to m.
nh_lts_aggregate_effort$metric_effort <- nh_lts_aggregate_effort$x * 0
.3048
# 0.3048 is the number of m/ft.

# Finalize "Effort" column.
nh_lts$Effort <-
  ifelse(nh_lts$Sample.Label == 1, nh_lts_aggregate_effort[1,3],
  ifelse(nh_lts$Sample.Label == 2, nh_lts_aggregate_effort[2,3],
  ifelse(nh_lts$Sample.Label == 3, nh_lts_aggregate_effort[3,3],
  ifelse(nh_lts$Sample.Label == 4, nh_lts_aggregate_effort[4,3], nh_lt
s$Sample.Label))))

# "Region.Label" indicates stratum (compartments were not broken up
into multiple strata).
nh_lts$Region.Label <- 1

# Install and/or Load "plyr" package.
install.packages("plyr")

library(plyr)

```

Check how many rows should be deleted based on the Delete column. Delete rows are placeholders for transects which were not sampled due to boundary interruption, and rows that do not contain data which will be used in any analyses. Transects that were sampled but did not lead to any observations should not be deleted. Instead, transects lacking observations should be included in any analyses so that values such as total sample effort are accounted for.

```
count(nh_lts$Delete)
```

Subset out rows which should be deleted (i.e. Delete value is "Yes").

```
nh_lts_sub <- nh_lts[nh_lts$Delete == "No", ]
```

Check data type of values in each column.

```
str(nh_lts_sub)
```

```
## 'data.frame': 326 obs. of 21 variables:
## $ Transect : int 1 1 1 1 1 1 1 1 1 1 ...
## $ cover.type : chr "NH" "NH" "NH" "NH" ...
## $ Measured : int 0 0 0 0 1 NA NA 0 0 3 ...
## $ Total : int 0 0 0 0 3 NA NA 0 1 3 ...
## $ Rec.Time : chr "0:09:07" "0:16:01" "0:19:13" "0:05:10" ...
## $ Method : chr "Line" "Line" "Line" "Line" ...
## $ Point.Line : chr "17H - B" "17H - 16H.5" "16H.5 - 16H" "14H - 14H.5" ...
## $ Direction : int 106 286 286 106 286 286 286 286 284 286 ...
## $ Ob.Angle : int NA NA NA NA 331 324 275 NA 277 358 ...
## $ Angle : int NA NA NA NA 45 38 11 NA 7 72 ...
## $ Ob.Distance : num NA NA NA NA 26.5 50.8 32.8 NA 36.9 15.9 ...
## $ DBH : num NA NA NA NA 26.8 21.5 25.7 NA 22.7 22.2 ...
## $ Species : int NA NA NA NA 318 318 318 NA 318 318 ...
## $ Group : chr NA NA NA NA ...
## $ Live.Dead : chr NA NA NA NA ...
## $ Line.Length : num 90.4 165 129.8 10 165 ...
## $ Delete : chr "No" "No" "No" "No" ...
## $ Area : num 404685 404685 404685 404685 404685 ...
## $ Sample.Label: int 1 1 1 1 1 1 1 1 1 1 ...
## $ Effort : num 317 317 317 317 317 ...
## $ Region.Label: num 1 1 1 1 1 1 1 1 1 1 ...
```

Covert data type to numeric in specific columns.

```
nh_lts_sub[c(1,3:4,8:10,13,19)] <- lapply(nh_lts_sub[c(1,3:4,8:10,13,19)], as.numeric)
```

For LTS, perpendicular distances from the sample line to each observed tree need to be calculated. A new column called "distance" needs to be created prior to any analyses. Data in the Ob.Distance and Angle columns will be used to calculate perpendicular distances, but first it needs to be confirmed that no angle measurements are greater than 90 degrees.

Confirm that Angle measurements above 90 degrees do not exist in the data set.

```
max(nh_lts_sub$Angle, na.rm = TRUE)
```

If any observations have a maximum angle greater than 90 degrees, delete them. NA values within the data set should not be deleted since the "ds" function used later requires them for cells missing data. If the deletion of an observation also leads to the deletion of a sampled transect, replace the observed data with NA values and retain the transect in the data set.

Confirm that DBH measurements below 18 in do not exist in the data set.

```
min(nh_lts_sub$DBH, na.rm = TRUE)
```

If any observations have a DBH less than 18 in, delete them. NA values within the data set should not be deleted since the "ds" function used later requires them for cells missing data. Again, if the deletion of an observation also leads to the deletion of a sampled transect, replace the observed data with NA values and retain the transect in the data set.

The "sin" function required for calculating perpendicular distance uses radians instead of degrees. Values in the Angle column must be converted.

```
calc_radians <- function(x){
```

```
  rads <- (x * pi) / (180)
```

```
  return(rads)
```

```
}
```

Create new column "Radians" with converted Angle values.

```
nh_lts_sub$Radians <- calc_radians(nh_lts_sub$Angle)
```

Write function calculating perpendicular distance.

```
perp_dist <- function(x, y){
```

```
  dist <- x*sin(y)
```

```
  return(dist)
```

```
}
```



```

# Create new column "distance_ft" with calculated perpendicular
distances in ft.
nh_lts_sub$distance_ft <- perp_dist(nh_lts_sub$Ob.Distance, nh_lts_sub
$Radians)

# Confirm that perpendicular distance measurements above 66 ft do not
exist in the data set.
max(nh_lts_sub$distance_ft, na.rm = TRUE)
# If any observations have a perpendicular distance greater than
66 ft, delete them. NA values within the data set should not be
deleted since the "ds" function used later requires them for cells
missing data.

# Many rows can quickly be deleted by subsetting the data and only
keeping rows with NA values or values less than or equal to 66 ft in
the distance_ft column.
nh_lts_sub2 <- nh_lts_sub[nh_lts_sub$distance_ft <= 66 | is.na(nh_lts_
sub$distance_ft), ].
# Check the subset and see if any transects were deleted as the result
of deleted observations. If so, replace the observed data for those
transects with NA values and retain them in the data set (see below).

# Transects which would be lost due to the deletion of individual
observations can have specific data replaced with NA values. Replacing
the observed data will retain the transects while treating them as if
they had no associated field observations. This process will ensure
that the correct number of transects will be recognized during any
distance sampling analyses, and that detection functions will not be
influenced by observations that should not have been sampled.
nh_lts_sub[46,c(9:15,22:23)] <- NA
nh_lts_sub[46,3:4] <- 0

# Re-subset the data, keeping rows with NA values or values less than
or equal to 66 ft the in distance_ft column.
nh_lts_sub2 <- nh_lts_sub[nh_lts_sub$distance_ft <= 66 | is.na(nh_lts_
sub$distance_ft), ]

# Convert DBH from in to cm.
nh_lts_sub2$DBH_Metric <- nh_lts_sub2$DBH * 2.54
# 2.54 is the number of cm/in.

# Again, check that minimum DBH is not less than 45.72 cm.
min(nh_lts_sub2$DBH_Metric, na.rm = TRUE)

# Create new column distance, where distance measurements are in m ins
tead of ft.
nh_lts_sub2$distance <- nh_lts_sub2$distance_ft * 0.3048
# 0.3048 is the number of m/ft.

```

```

# Again, check that maximum perpendicular distance is not greater than
20.1168 m.
max(nh_lts_sub2$distance, na.rm = TRUE)

#####
# Fitting Detection Functions to LTS Data
#####

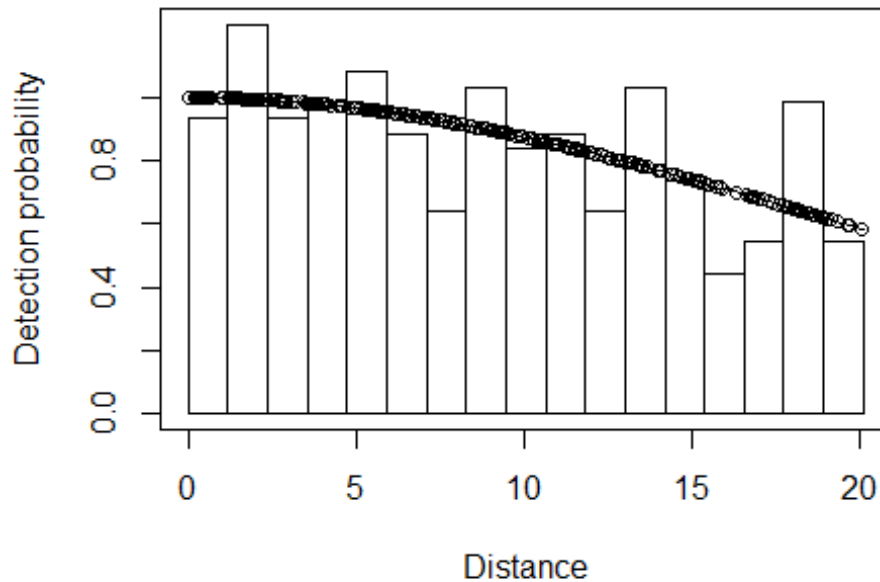
# The code below can be modified to fit a variety of key functions and
adjustment terms. Functions available for fitting with the "ds"
function include a uniform key function (key = "unif") with either a
cosine or simple polynomial adjustment (adjustment = "cos" or "poly"),
a half-normal key function (key = "hn") with either a cosine or
Hermite polynomial adjustment (adjustment = "cos" or "herm"), and a
hazard-rate function (key = "hr") with either a cosine or simple
polynomial adjustment (adjustment = "cos" or "poly"). The half-normal
and hazard-rate key functions are also available without any
adjustment term (adjustment = NULL).

# Fit half-normal detection function without adjustment term.
nh_line_halfnorm <- ds(nh_lts_sub2, transect = "line", key = "hn", adj
ustment = NULL, truncation = 20.1168, region.table = NULL, sample.tabl
e = NULL, obs.table = NULL, formula = ~1)

## Fitting half-normal key function
## Key only model: not constraining for monotonicity.
## AIC= 1747.913

# View fitted detection function model.
plot(nh_line_halfnorm)

```



```
# Summarize model and obtain abundance/density estimates.
```

```
summary(nh_line_halfnorm)
```

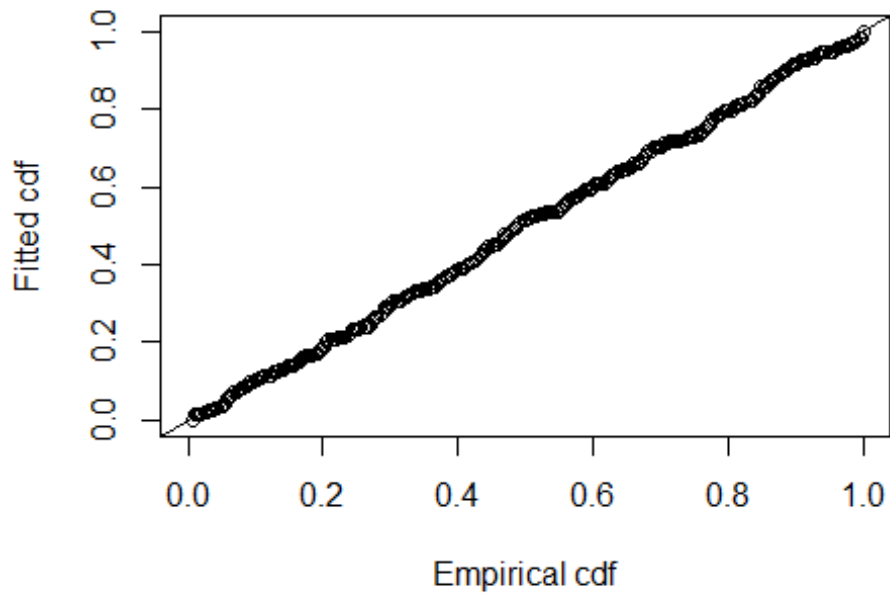
```
##
## Summary for distance analysis
## Number of observations : 292
## Distance range       : 0 - 20.1168
##
## Model : Half-normal key function
## AIC   : 1747.913
##
## Detection function parameters
## Scale coefficient(s):
##           estimate      se
## (Intercept) 2.963795 0.1916348
##
##           Estimate      SE      CV
## Average p      0.8459616 0.05035429 0.05952314
## N in covered region 345.1693446 22.02205707 0.06380073
##
## Summary statistics:
## Region Area CoveredArea Effort n k ER se.ER
## 1      1 404685.5 100764.5 2504.486 292 4 0.1165908 0.02337826
##           cv.ER
## 1 0.2005154
```

```

## Abundance:
##   Label Estimate      se      cv      lcl      ucl      df
## 1 Total 1386.253 289.9537 0.2091637 757.5388 2536.763 3.551733
##
## Density:
##   Label      Estimate      se      cv      lcl      ucl
## 1 Total 0.003425506 0.0007164915 0.2091637 0.00187192 0.00626848 3.551733

# Check goodness of fit and Cramer-von Mises P-value (acceptable model
will have non-significant P-value).
gof_ds(nh_line_halfnorm)

```



```

##
## Goodness of fit results for ddf object
##
## Distance sampling Kolmogorov-Smirnov test
## Test statistic = 0.028526 P = 0.97141
##
## Distance sampling Cramer-von Mises test (unweighted)
## Test statistic = 0.030158 P = 0.97565

```

```
# Summarize all fitted models and select the best fit. Multiple models
can be called at once within the "summarize_ds_models" function.
summarize_ds_models(nh_line_halfnorm)
```

```
##                               Model Key function Formula
## 1 \\texttt{nh\\char`_line\\char`_halfnorm} Half-normal      ~1
## C-vM p-value  $\hat{P}_a$  se( $\hat{P}_a$ )  $\Delta AIC$ 
## 1      0.9756462      0.8459616      0.05035429      0
```

```
# Once best model is selected, density in trees/ha can be calculated
by multiplying the supplied density estimate (which is in trees/sq.m)
by 10000 (the number of sq.m/ha).
```

```
#####
```

```
# Data Prep for PTS
```

```
#####
```

```
# Subset to a new data set containing only PTS data from the northern
hardwood compartment.
```

```
nh_pts <- all_distance[all_distance$cover.type == "NH" & all_distance$
Method == "Point", ]
```

```
# Check the subset.
```

```
str(nh_pts)
```

```
## 'data.frame':      190 obs. of  17 variables:
## $ Transect      : int  1 1 1 1 1 1 1 1 1 1 ...
## $ cover.type    : chr  "NH" "NH" "NH" "NH" ...
## $ Measured      : int  0 0 NA NA NA NA 2 NA 0 1 ...
## $ Total         : int  0 0 NA NA NA NA 2 NA 0 1 ...
## $ Rec.Time      : chr  "0:01:11" "0:05:50" NA NA ...
## $ Method        : chr  "Point" "Point" "Point" "Point" ...
## $ Point.Line    : chr  "17H" "16H.5" "16H" "15H.5" ...
## $ Direction     : int  NA NA NA NA NA NA NA NA NA NA ...
## $ Ob.Angle      : int  NA NA NA NA NA NA NA NA NA NA ...
## $ Angle         : int  NA NA NA NA NA NA NA NA NA NA ...
## $ Ob.Distance   : num  NA NA NA NA NA NA 26.5 50.8 NA 36.9 ...
## $ DBH           : num  NA NA NA NA NA NA 26.8 21.5 NA 22.7 ...
## $ Species       : int  NA NA NA NA NA NA 318 318 NA 318 ...
## $ Group         : chr  NA NA NA NA ...
## $ Live.Dead     : chr  NA NA NA NA ...
## $ Line.Length   : num  NA NA NA NA NA NA NA NA NA NA ...
## $ Delete        : chr  "No" "No" "Yes" "Yes" ...
```

```

# Create new columns as required for the "ds" function.

# "Area" which is compartment area in square meters.
nh_pts$Area <- (43560 * 100) * 0.092903
# 43560 is the number of sq.ft/acre, 100 is the number of acres in the
northern hardwood compartment, and 0.092903 is the number of
sq.m/sq.ft.

# "Effort" which is the number of times each point transect was
sampled.
nh_pts$Effort <- 1

# "Region.Label" indicates stratum (compartments were not broken up
into multiple strata).
nh_pts$Region.Label <- 1

# "distance_ft" which is the distance in feet between each observation
and its associated point transect.
nh_pts$distance_ft <- nh_pts$Ob.Distance

# Check data type of values in each column.
str(nh_pts)

## 'data.frame':  190 obs. of  21 variables:
## $ Transect      : int  1 1 1 1 1 1 1 1 1 1 ...
## $ cover.type    : chr  "NH" "NH" "NH" "NH" ...
## $ Measured      : int  0 0 NA NA NA NA 2 NA 0 1 ...
## $ Total         : int  0 0 NA NA NA NA 2 NA 0 1 ...
## $ Rec.Time      : chr  "0:01:11" "0:05:50" NA NA ...
## $ Method        : chr  "Point" "Point" "Point" "Point" ...
## $ Point.Line    : chr  "17H" "16H.5" "16H" "15H.5" ...
## $ Direction     : int  NA NA NA NA NA NA NA NA NA NA ...
## $ Ob.Angle      : int  NA NA NA NA NA NA NA NA NA NA ...
## $ Angle         : int  NA NA NA NA NA NA NA NA NA NA ...
## $ Ob.Distance   : num  NA NA NA NA NA NA 26.5 50.8 NA 36.9 ...
## $ DBH           : num  NA NA NA NA NA NA 26.8 21.5 NA 22.7 ...
## $ Species       : int  NA NA NA NA NA NA 318 318 NA 318 ...
## $ Group         : chr  NA NA NA NA ...
## $ Live.Dead     : chr  NA NA NA NA ...
## $ Line.Length   : num  NA NA NA NA NA NA NA NA NA NA ...
## $ Delete        : chr  "No" "No" "Yes" "Yes" ...
## $ Area          : num  404685 404685 404685 404685 404685 ...
## $ Effort        : num  1 1 1 1 1 1 1 1 1 1 ...
## $ Region.Label : num  1 1 1 1 1 1 1 1 1 1 ...
## $ distance_ft   : num  NA NA NA NA NA NA 26.5 50.8 NA 36.9 ...

```

```

# Covert data type to numeric.
nh_pts[c(1,3:4,8:10,13)] <- lapply(nh_pts[c(1,3:4,8:10,13)], as.numeric)

# Check how many rows should be deleted based on the Delete column.
Delete rows are placeholders for transects which were not sampled due
to boundary interruption, and rows that contain data which will not be
used in any analyses. Transects that were sampled but did not lead to
any observations should not be deleted and should still be included in
distance sampling analyses.
count(nh_pts$Delete)

# Subset out rows which should be deleted (i.e. Delete value is
"Yes").
nh_pts_sub <- nh_pts[nh_pts$Delete == "No", ]

# Create column "Sample.Label" which is a unique ID for each point
transect.
nh_pts_sub <- transform(nh_pts_sub, Sample.Label = match(Point.Line, unique(Point.Line)))

# Covert data type to numeric.
nh_pts_sub[22] <- lapply(nh_pts_sub[22], as.numeric)

# Confirm that tree DBH measurements below 18 in do not exist in the
data set.
min(nh_pts_sub$DBH, na.rm = TRUE)

# If any observations have a DBH less than 18 in, delete them. NA
values within the data set should not be deleted since the "ds"
function used later requires them for cells missing data. If the
deletion of an observation also leads to the deletion of a sampled
transect, replace the observed data with NA values and retain the
transect in the data set.

# Confirm that distance measurements above 66 ft do not exist in the
data set.
max(nh_pts_sub$distance_ft, na.rm = TRUE)

# If any observations have an associated distance measurement greater
than 66 ft, delete them. NA values within the data set should not be
deleted since the "ds" function used later requires them for cells
missing data. If the deletion of an observation also leads to the
deletion of a sampled transect, replace the observed data with NA
values and retain the transect in the data set.

```

```

# Convert DBH from in to cm.
nh_pts_sub$DBH_Metric <- nh_pts_sub$DBH * 2.54
# 2.54 is the number of cm/in.

# Again, check that minimum DBH is not less than 45.72 cm.
min(nh_pts_sub$DBH_Metric, na.rm = TRUE)

# Create new column "distance", where distance measurements are in m
instead of ft.
nh_pts_sub$distance <- nh_pts_sub$distance_ft * 0.3048
# 0.3048 is the number of m/ft.

# Again, check that maximum distance is not greater than 20.1168 m.
max(nh_lts_sub2$distance, na.rm = TRUE)

#####
# Fitting Detection Functions to PTS Data
#####

# Code can be modified as in LTS. Note that transect must equal
"point" when fitting detection functions to PTS data.

# Fit half-normal detection function without adjustment term.
nh_point_halfnorm <- ds(nh_pts_sub, transect = "point", key = "hn", ad
justment = NULL, truncation = 20.1168, region.table = NULL, sample.tab
le = NULL, obs.table = NULL, formula = ~1)

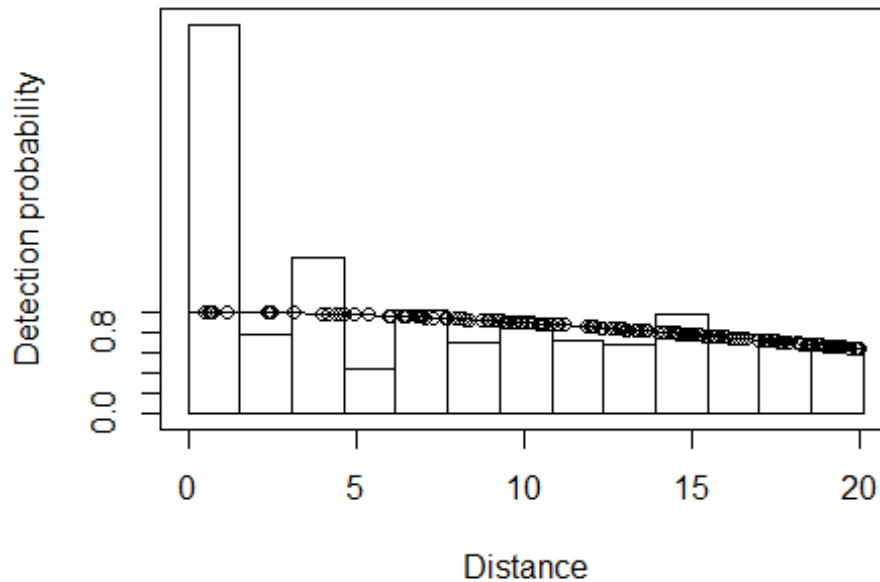
## Fitting half-normal key function

## Key only model: not constraining for monotonicity.

## AIC= 1029.125

# View fitted detection function model.
plot(nh_point_halfnorm)

```

```
# Summarize model and obtain abundance/density estimates.
```

```
summary(nh_point_halfnorm)
```

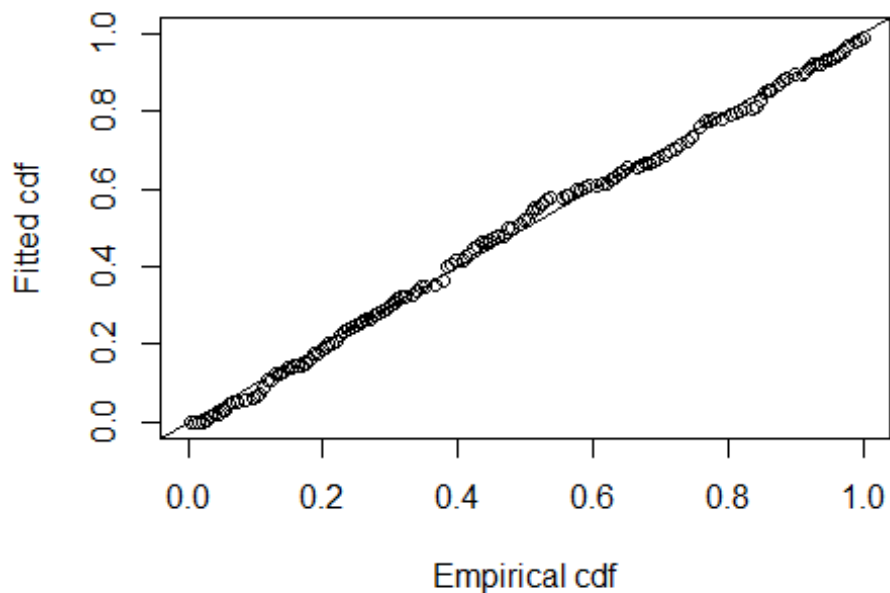
```
##
## Summary for distance analysis
## Number of observations : 177
## Distance range       : 0 - 20.1168
##
## Model : Half-normal key function
## AIC   : 1029.125
##
## Detection function parameters
## Scale coefficient(s):
##      estimate      se
## (Intercept) 3.044906 0.2855862
##
##              Estimate      SE      CV
## Average p      0.8021218 0.09702883 0.1209652
## N in covered region 220.6647329 27.69367575 0.1255011
##
## Summary statistics:
## Region Area CoveredArea Effort n k ER se.ER
cv.ER
## 1      1 404685.5 62296.51 49 177 49 3.612245 0.4718399 0.13
06224
```

```

## Abundance:
##   Label Estimate      se      cv      lcl      ucl      df
## 1 Total 1433.464 255.2 0.1780303 1010.874 2032.715 137.9601
##
## Density:
##   Label      Estimate      se      cv      lcl      ucl
## 1 Total 0.003542168 0.0006306132 0.1780303 0.002497926 0.00502295 1
## 37.9601

# Check goodness of fit and Cramer-von Mises P-value (acceptable model
# will have non-significant P-value).
gof_ds(nh_point_halfnorm)

```



```

##
## Goodness of fit results for ddf object
##
## Distance sampling Kolmogorov-Smirnov test
## Test statistic = 0.04937 P = 0.78136
##
## Distance sampling Cramer-von Mises test (unweighted)
## Test statistic = 0.048295 P = 0.88655

```

Summarize all fitted models and select the best fit. Multiple models can be called at once within the "summarize_ds_models" function.
`summarize_ds_models(nh_point_halfnorm)`

```
##                               Model Key function Formula
## 1 \\texttt{nh\\char`_point\\char`_halfnorm} Half-normal      ~1
##   C-vM p-value  $\hat{P}_a$  se( $\hat{P}_a$ )  $\Delta$ AIC
## 1   0.8865527   0.8021218   0.09702883         0
```

Once best model is selected, density in trees/ha can be calculated by multiplying the supplied density estimate (which is in trees/sq.m) by 10000 (the number of sq.m/ha).