Fraver et al. Forest Ecosystems (2018) 5:39
https://doi.org/10.1186/s40663-018-0156-9

Forest Ecosystems

RESEARCH Open Access

Influence of transect length and downed woody debris abundance on precision of the line-intersect sampling method

Shawn Fraver[1*], Mark J. Ducey[2], Christopher W. Woodall[3], Anthony W. D'Amato[4], Amy M. Milo[5] and Brian J. Palik[6]

Abstract

Background: Accurate downed woody debris (DWD) volume or mass estimates are needed for numerous applications such as fuel loading, forest carbon, and biodiversity/habitat assessments. The line-intersect sampling (LIS) method of inventorying DWD is widely used in forest inventories and ecological studies because it is time-efficient and unbiased. Despite its widespread use, the appropriate transect length needed to achieve a desired precision at a particular location has received relatively little attention.

Methods: We conducted intensive LIS sampling at 33 locations representing eight mature or old-growth forest types in northeastern USA, providing a range of forest conditions and DWD volumes (from 17 to 323 $m^3 \cdot ha^{-1}$). We used these empirical field data to test, through simulations, the effect of increasing transect length (up to 340 m at each location) on precision of associated LIS volume estimates. Importantly, we used a novel application of copula models to account for within-transect spatial autocorrelation of DWD volumes during our simulations, thereby properly addressing variance estimates.

Results: As expected, precision consistently improved with increasing cumulative transect length, and locations with lower DWD volumes required longer transects to achieve a given level of precision. We developed models relating precision, transect length, and DWD volume that allows us to gauge a suitable LIS transect length for desired precision levels.

Conclusions: LIS provides an attractive method for estimating DWD volume for a given localized area of interest. For the forest types sampled here, and for the particular copula model framework employed, transect lengths of ca. 120 m provide a reasonable level of precision, ranging from 18% to 60% coefficients of variation.

Keywords: Coarse woody debris, Copula model, Dead wood, Fuel loads, Forest structure, Forest carbon

Background

Research over the past decades has clearly established the vital role of downed woody debris (DWD) in forest ecosystems, where it influences biodiversity, nutrient cycling, soil development, and wildfire behavior (Harmon et al. 1986; Schoennagel et al. 2004; Stokland et al. 2012). Recent attention has shifted to carbon stores and fluxes in DWD given the growing interest in carbon dynamics in the context of climate change (Russell et al. 2014; Woodall et al. 2015). Methods of reliably estimating DWD abundance in forest systems are thus critical for carbon estimates (e.g. Russell et al. 2015), as well as a broader range of applications.

As with live-tree sampling, a complete census could be conducted of all DWD pieces above a minimum diameter threshold within a location of interest; however, this may be impractical for inventories at stand or larger scales. As a consequence, sampling methods such as line-intersect sampling (LIS) are employed to inventory DWD abundance (Woodall et al. 2009). LIS produces an area-based estimate of DWD volume derived from a sample of the logs present. In the field, the intersect method proceeds by measuring diameters of pieces (above a given size threshold) at the point where an inventory transect crosses the pieces' central longitudinal

* Correspondence: shawn.fraver@maine.edu
[1]School of Forest Resources, University of Maine, Orono, ME, USA
Full list of author information is available at the end of the article

axes. The resulting data (tallied diameters and total transect length) are converted to area-based estimates of DWD volume, typically using equations provided by Van Wagner (1968) and Brown (1974). The theory underlying LIS is well developed (de Vries 1979; Kaiser 1983). No assumption of piece shape is required, as LIS leads naturally to an unbiased Monte Carlo estimate of volume based on the cross-section of pieces where they are intersected by the transect line. Moreover, despite some derivations that assume pieces are randomly distributed (e.g. Warren and Olsen 1964), from a design-based perspective, the positions of pieces are regarded as fixed even if they originated from a random process. Because it is time-efficient and design-based unbiased, LIS is widely used in ecological studies and fuel inventories (e.g. Brown 1974), as well as in many national-scale forest inventories (Woodall et al. 2009). Kershaw et al. (2016, ch. 12) provide a practical and theoretical review of LIS along with alternative DWD inventory approaches.

Despite its widespread use in large-scale forest inventories, the appropriate LIS transect lengths needed under local forest settings has received relatively little attention (but see Woldendorp et al. 2004; Sikkink and Keane 2008). Considering that many ecological studies intend to characterize DWD attributes at particular locations, where forest stands or small reserves are themselves the populations of interest (Jönsson et al. 2011; Fraver and Palik 2012), surprisingly few studies have evaluated DWD sampling methods at this local scale. Stokland et al. (2004) caution that a DWD sampling design that is efficient for assessing average ecological values over large scales may be poorly matched to the task of assessing those same values at local scales.

Previous empirical work as well as sampling theory demonstrates that for a given site the precision of the LIS estimate consistently improves (lower variance) with increasing cumulative transect length (Pickford and Hazard 1978; Woldendorp et al. 2004; Miehs et al. 2010) and that sites with lower DWD volume generally require longer transects to achieve a desired level of precision (Brown 1974; Pickford and Hazard 1978; Woldendorp et al. 2004). The objective of the present study was two-fold: 1) to better quantify the effect of increasing transect length and DWD volumes on the precision of LIS volume estimates, and 2) to provide practical guidance for selecting transect lengths that achieve a desired level of precision regarding LIS volume estimates. We addressed these objectives through simulations constrained by empirical LIS data from 33 locations representing eight diverse forest types in the northeastern USA. These locations provide a range of forest conditions and DWD volumes, making them ideal for addressing these objectives. Additionally, we present a novel application of copula models to account for within-transect spatial autocorrelation of DWD volumes; without accounting for such autocorrelation, simulations would underestimate the variance within transects, resulting in spurious assessments of appropriate transect lengths.

Methods

Study locations and field data collection

We tested the effect of increasing transect length (up to ca. 340 m) and DWD volume on LIS precision using 33 locations representing eight mature or old-growth forest types in northeastern USA (Table 1). These forest types and locations were utilized opportunistically, given that we needed DWD inventories for these locations as part of on-going research (e.g., Fraver and Palik 2012; Fraver et al. 2017). We thus opted to intensify sampling (i.e., long transects) to provide data for the modelling efforts presented here. These locations represent diverse forest compositions and structures, and they conveniently capture a large range of DWD volumes necessary for modelling. Forest types include mature red-spruce (*Picea rubens*), mature spruce–hemlock (*P. rubens*, *Tsuga canadensis*), mature hemlock–pine (*T. canadensis*, *Pinus strobus*), mature mixed conifer–hardwood (*T. canadensis*, *Acer saccharum*, *Fagus grandifolia*, *Betula alleghaniensis*), mature northern hardwoods (*A. saccharum*, *F. grandifolia*, *B. alleghaniensis*, with *P. rubens*), old-growth northern white-cedar (*Thuja occidentalis*), old-growth red pine (*Pinus resinosa*), and old-growth black ash (*Fraxinus nigra*). A minimum of three locations was established in each forest type. Additional site information is provided in Table 1.

At each location we conducted LIS sampling using a cluster of four transects that intersected at their midpoints as follows: two diagonal 100-m transects (arranged southwest to northeast and northwest to southeast) and two 70-m transects (arranged in cardinal directions). This configuration formed a spoke-like 'union jack' with equal angles between transects and a total transect length of 340 m. At four of the eight sites, the centers of the union jacks were placed on centers of existing long-term continuous forest inventory plots. At the remaining four sites, where previous plots did not exist, centers were placed randomly within sampled stands. This union jack configuration was initially employed at the red-pine locations (the first to be sampled; Fraver and Palik 2012), as an exploratory comparison of co-located LIS and fixed area inventories on 0.5-ha (70.7 m × 70.7 m) plots (unpublished). For sampling consistency, we then used this same configuration at all remaining locations. A multi-directional, spoke-like arrangement is often used to reduce potential sample variance when DWD pieces exhibit fairly uniform fall directions (Van Wagner 1968). The most common estimating equations used with LIS assume the angle of intersection between individual pieces and the sample line is random; when the sample line orientation itself is

Table 1 Characteristics of the eight forest types from which empirical data were collected for evaluation of LIS precision

Site	Forest type	No. locations	Mean (range) DWD vol. ($m^3 \cdot ha^{-1}$)
Howland Forest, Maine	Spruce-hemlock	3	25 (17–36)
BBEW, Maine	Northern hardwoods	4	46 (32–61)
PEF, Maine	Hemlock-white pine	3	52 (35–66)
Baxter SFMA, Maine	Red spruce	3	55 (42–70)
Various, Minnesota	Red pine	7	81 (51–137)
BEF, New Hampshire	Mixed conifer-hardwood	3	84 (58–111)
Various, Minnesota	Black ash	6	126 (70–243)
BRFR, Maine	Northern white-cedar	4	258 (205–323)

Note: *BBEW* Bear Brook Experimental Watershed, *PEF* Penobscot Experimental Forest, *SFMA* Scientific Forest Management Area of Baxter State Park, *BEF* Bartlett Experimental Forest, *BRFR* Big Reed Forest Reserve. See Methods for descriptions of forest types

random, these equations provide unbiased estimates, but strong correlation in the pieces' fall directions can inflate sample variance (Kaiser 1983; Kershaw et al. 2016, ch. 12). Spoke-like transect designs have been the subject of some debate in large-area inventories (Affleck et al. 2005; but see Woodall and Monleon 2008); however, their use represents common field practice (e.g., Dunn and Bailey 2015), and we feel the design is adequate to evaluate the effect of varying transect lengths, the focus of our study. For each DWD piece intersected by the sampling transect with a diameter ≥ 10 cm at the intersection, we recorded diameter at intersection, distance from center, species, and decay class (five-class system as per Sollins 1982). By convention, if a piece intersected two transects, or if it crossed the same transect twice, it was recorded twice (Van Wagner 1968, Brown 1974).

We calculated DWD volume for each site as follows,

$$V = \left(\pi^2 \sum \frac{d^2}{8L} \right) \times 10\,000$$

where V is the area-based volume ($m^3 \cdot ha^{-1}$), d is the DWD piece diameter (m) at point of intersection, and L is the total transect length at the site (m) (from Van Wagner 1968). Similar calculations were used to provide volume estimates for short 5-m segments of each transect, to be used in the sampling simulations described below.

Simulation of sampling variability

To assess how transect length and population characteristics influence sampling variability, we employed a model-based approach using simulation based on a zero-inflated, autoregressive copula. To motivate this approach, let us first consider a naive approach that would be simple to implement, and would have some desirable features, but could lead to inaccurate results. Given the line intersect data from a site, with the positions of all the intersections recorded, one could subdivide transects into shorter segments (e.g., 5 m) that would allow simulation of the sampling variance for different transect lengths, with a resolution on length that would be meaningful for operational inventory design. Then, to simulate transects of different total length, one would draw segments at random with replacement from the data for that site, concatenating them to form transects of the desired length. For example, to form a 20-m transect, one would choose four segments of 5 m each with replacement, and compute any desired values (such as DWD volume) for the new, simulated transect. By repeating this process many times, and varying the transect length, one would develop a Monte Carlo estimate of the sampling distribution of DWD volume estimates for LIS at that site and its dependence on transect length, with a richer potential inference beyond second-order statistics such as variance (which, under the assumptions implicit in this approach, could also be calculated using the usual formulae for simple random sampling; e.g., Thompson 2012). Essentially, this approach uses the same rationale as the bootstrap (Efron and Tibshirani 1993), substituting the empirical distribution of DWD volume on short segments for the unknown population distribution of possible segments, to arrive at the sample distribution through resampling.

Unfortunately, the bootstrap approach makes a strong independence assumption that is inappropriate for DWD data under either a design- or model-based approach (for details on the distinction, see Gregoire 1998). Within a design-based framework (Kaiser 1983; Gregoire and Valentine 2008), DWD pieces are fixed but sample selection is random. The entire transect is a single randomly selected sample unit, and treating segments within a sample unit as if they were drawn independently is anathema. Without substantial additional information (such as the joint inclusion probabilities of the individual pieces of DWD) it is difficult to make progress from this perspective, beyond inference based on the design that was actually employed (except for changes in the number, rather than the length, of sample transects). While a design-based approach has substantial benefits, both in terms of formulating a design for a specific inventory, and computing defensible estimates once the data are in hand, to understand how different designs (such as varying transect lengths) would impact estimates and their uncertainty, a model is invaluable. However, from a model-based perspective, the independence of short, contiguous segments is also suspect. One might expect positive autocorrelation of the values from adjacent segments, not only because of the potential for spatially clumped inputs from tree mortality, but also because local terrain might cause DWD pieces that are close to each other also to share similar orientation. In the presence of positive autocorrelation, the variance of estimates from transects of a given length would be higher than that implied by independent segments, and the transect length needed to reach a given level of accuracy would be longer. To summarize: from a design-based perspective, successive segments within a randomly selected transect are likely to display autocorrelation because of the fixed characteristics of the downed woody debris population. From a model-based perspective, the independence of successive short segments is also suspect, because of likely clumping in the random process that generates the population. Whichever perspective is taken in inference, it would appear important to incorporate autocorrelation within any simulations.

To address the challenge of autocorrelation among successive segments, while remaining faithful to the observed marginal distribution of estimates from those segments, we employ a copula modeling approach. Within the forestry literature, we believe this is the first application to employ an autoregressive model (to account for correlation between successive segments), and also to address a zero-inflated distribution (many segments are empty, with no DWD intersections). Copula models are relatively new in the forestry literature; they have been used to describe relationships between continuous variables such as tree diameter and height (Wang et al. 2008, 2010; MacPhee et al. 2017), to inform spatially-explicit stand simulations (Kershaw et al. 2010), and for imputation-based growth and yield models (Kershaw et al. 2017). Eskelson et al. (2011) and Fortin et al. (2013) extended copula models in forestry to account for spatial autocorrelation. More recently, copula methods have been used to simulate spatial populations for remote sensing-assisted inventory (Ene et al. 2012, 2013a, 2013b; Grafström et al. 2014). For a somewhat theoretical overview of copulas, see Genest and Mackay (1986) and Nelsen (2006).

Strictly defined, a bivariate copula describes the relationship between two random variables x and y in terms of two random deviates that are uniformly distributed on [0, 1]. Specifically, if U_1 and $U_2 \sim$ Uniform (0, 1), then the copula C is defined as

$$C(u_1, u_2) = \Pr\{U_1 \leq u_1, U_2 \leq u_2\}$$

If $F_X(x)$ and $F_Y(y)$ are the cumulative distribution functions of x and y, then

$$F_{XY}(x, y) = C(F_X(x), F_Y(y))$$

In our application, x and y are the volume estimates ($m^3 \cdot ha^{-1}$) associated with successive pairs of short LIS segments within the same dead wood population; thus, they share the same cumulative distribution function $F(x) = F_X(x) = F_Y(y)$. A variety of formulations have been proposed to model the copula function C. For simulation purposes, one of the simplest is the Gaussian or normal copula (Wang 1998). Let Z_1 and Z_2 have a standard normal marginal distribution, i.e., Z_1 and $Z_2 \sim N$ (0, 1), with correlation coefficient ρ. Let $\Phi(\cdot)$ be the cumulative distribution function of the standard normal distribution. Then $\Phi(Z_1)$ and $\Phi(Z_2)$ will be bivariate uniform, with Kendall's $\tau = (2/\pi)\text{acrsin}(\rho)$. More generally, if we simulate a sequence of Z_i, $i = 1, \ldots, m$, the sequence will be a first-order autoregressive [AR(1)] time series, and each pair $\Phi(Z_1)$, $\Phi(Z_{i+1})$ will be described by the same copula function C. To complete a simulation of a single transect composed of m segments, each having marginal distribution $F(x)$, we employ the following algorithm:

1) Generate a standard normal AR(1) time series of length m, with correlation coefficient ρ. This can easily be done through the sequential generation of independent random deviates, with appropriate summation to generate successive entries in the time series.
2) Convert the resulting series to an autocorrelated series of uniform deviates $u_i = \Phi(Z_1)$.
3) Apply the inverse of the empirical cumulative distribution function to translate the uniform series to the marginal distribution of downed wood volume estimates, $x_i = F^{-1}(u_i)$.

The only slight complication in the simulation phase of our application arises due to the zero-inflated nature of DWD volume data. Specifically, let p_0 be the probability of obtaining $x_i = 0$ in a given population. Then, in step 3 above, $x_i = 0$ whenever $u_i < p_0$.

The zero-inflated nature of the data does cause a more important problem in the estimation stage that must precede simulation, however. Specifically, it complicates the estimation of the autocorrelation parameter ρ. In the absence of zero-inflation, the Z_i, u_i, and x_i would share a common value of Kendall's τ (or Spearman's rank-correlation coefficient), both of which share a one-to-one relationship with ρ. Thus, it would be possible to calculate a reasonable value of ρ based on the value of τ calculated from the sample pairs x_i, x_{i+1} in the transects for a given population. Alternatively, one could reverse the process used in simulation: calculate $u_i = F(x_i)$ to translate the sample series to an autocorrelated uniform series, then $Z_i = \Phi^{-1}(u_i)$ to obtain a standard normal AR(1) series, and finally estimate ρ from the serial correlation of the Z_i. However, this approach fails for zero-inflated data, because it is unclear what value to assign to u_i (and hence Z_i) when $x_i = 0$.

To overcome this challenge, we adopted a maximum-likelihood approach based on the estimation of correlation coefficients in the presence of censoring, as originally developed by Lyles et al. (2001). In fact, our situation is somewhat simpler than that described by Lyles et al. (2001), because the Z_i are defined to have standard normal distribution; we do not need to concern ourselves with the mean and variance as nuisance parameters. We reverse-engineer the Z_i values as described above, but record the values as censored whenever $x_i = 0$; that is, we only know that $Z_i \le \Phi^{-1}(p_0)$. Then, we may recognize four distinct categories of outcomes for Z_i, Z_{i+1} pairs:

Category 1: In this category, neither Z_i nor Z_{i+1} are censored. Pairs of this type contribute $\ln f(Z_i, Z_{i+1}) = \ln f(Z_{i+1}|Z_i) + \ln f(Z_i)$ to the log-likelihood. Following Lyles et al. (2001), we denote this contribution as $\ln(t_{i1})$:

$$\ln(t_{i1}) = -\frac{1}{2(1-\rho^2)}\left(Z_i^2 + Z_{i+1}^2 - 2\rho Z_i Z_{i+1}\right) - \ln\left[2\pi\left(1-\rho^2\right)\right]$$

Category 2: In this category, Z_i is known but Z_{i+1} is censored. Let $Z' = \Phi^{-1}(p_0)$. The contribution of these pairs to the log-likelihood is

$$\ln(t_{i2}) = \ln\left[\Pr\left(Z_{i+1} \le Z' | Z_i\right)\right] + \ln f(Z_i)$$

which equals

$$\ln(t_{i2}) = \frac{Z_i^2}{2} - \ln(2\pi) + \ln\Phi\left[\frac{Z' - Z_i}{\sqrt{1-\rho^2}}\right]$$

Category 3: In this category, Z_{i+1} is known but Z_i is censored; by the same rationale as Category 2, the contribution to the log-likelihood is

$$\ln(t_{i3}) = \frac{Z_{i+1}^2}{2} - \ln(2\pi) + \ln\Phi\left[\frac{Z' - Z_{i+1}}{\sqrt{1-\rho^2}}\right]$$

Category 4: In this category, both Z_i and Z_{i+1} are censored. Then the contribution to the log-likelihood is $\ln\{\Pr[(Z_i \le Z') \cap (Z_{i+1} \le Z')]\}$, i.e.

$$\ln(t_{i4}) = \ln\left\{\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{Z'} \Phi\left[\frac{Z' - Z}{\sqrt{1-\rho^2}}\right] \exp\left(\frac{Z^2}{2}\right) dZ\right\}$$

Unfortunately, the integral has no convenient, closed-form expression, but it yields readily to numerical integration. By summing the contributions to the log-likelihood across all unique adjacent pairs of segments within a site, one obtains the log-likelihood function; maximization with respect to ρ yields the maximum likelihood estimate, and the inverse of the negative of its second derivative provides the asymptotic variance of the estimate, from which the standard error and confidence limits may be computed.

To summarize: For each location, we used all unique pairs of adjacent 5-m segments, and estimated a site-specific ρ and its standard error by maximum likelihood. Then, we simulated artificial transects of varying length using a zero-inflated autogressive copula, based on the estimated value of ρ and the empirical distribution of volume estimates for 5-m segments from that location. We simulated transect lengths from 20 m to 340 m by 20-m intervals, with 1000 simulated replications at each length for each location. For comparison purposes, we performed identical simulations with $\rho = 0$, i.e., no autocorrelation, for each location. In this case, the simulation is mathematically identical to the independent bootstrap approach. We summarized the distributional characteristics of the 1000 replicates and used the results as response variables for further modeling to characterize the dependence of sampling variability on transect length and DWD volume.

Modeling of simulation results

Using output from our simulation set, we tallied mean DWD volume estimates and standard deviations for use in the modeling described below. That is, each transect length from each of the 33 locations yielded a mean and a standard deviation, calculated from the 1000 runs unique to each transect length–location combination. We used linear and non-linear mixed-effects models to evaluate the influence of transect length and total DWD volume (predictor variables, unique to each location) on the coefficient of variation (CV) of these estimates. Total DWD volume refers to that calculated from the full 340-m transect length for each location. We first tested the predictors in individual models and then combined them to determine the best-fit two-term model. Forest type was included as a random factor in all models, given that we are not making inferences about particular forest types, for which our sample size (3–7 locations per type) is inadequate. Logarithmic and square-root transformations of response and predictor variables were explored for stabilizing variance and improving residual-versus-predicted diagnostics. The fact that our candidate models included transformed and non-transformed variables limited our reliance on Akaike's information criterion (Burnham and Anderson 2002) for evaluating model performance. Instead, we relied on pseudo R^2 (see below) and root mean square error (RMSE, back transformed as needed) values, as well as graphs of observed-versus-predicted and residual-versus-predicted values to determine which models were best supported by the data. For the one-term models, we tested linear, negative exponential, and power function model forms; for the two-term model, we tested linear, negative exponential, power, quadratic, and partial quadratic forms. Analyses were conducted in the *nlme* package (Pinheiro et al., 2016) in R (version 3.0.3; Core Team, 2016), using the weighting option to compensate for non-homogenous variance. Goodness-of-fit (pseudo R^2) for the models was expressed as the correlation between observed and predicted CVs (cf. Canham et al. 2004).

Results

Simulations of various transect lengths revealed that, as expected, the precision (here, CV) of the LIS estimate consistently improved with increasing cumulative transect length. Model comparisons revealed that the relationship between CV and transect length was best described by a power function ($P < 0.0001$, $R^2 = 0.801$, RMSE = 0.127) as shown in Fig. 1. Importantly, our analysis of DWD values along transects indicated the presence of spatial autocorrelation, with correlation coefficients (one for each of 33 locations) forming a bimodal distribution; that is, some locations showed positive and others negative autocorrelation (Fig. 2).

Simulations also revealed that precision of the LIS estimate improved with increasing DWD volume present, meaning that for a given transect length, greater precision could be achieved where higher DWD volumes are encountered. The relationship between CV and DWD volume was best described by a linear function, as shown in Fig. 3. Although this model was statistically significant ($P < 0.01$), the relationship is quite weak ($R^2 = 0.228$, RMSE = 0.206).

When the CV was modeled as a function of both transect length and DWD volume, model comparisons revealed that the exponential relationship was best supported by the data, expressed as $CV = a \times e^{b \times \text{volume}} \times e^{c \times \text{sqrt\_length}}$,

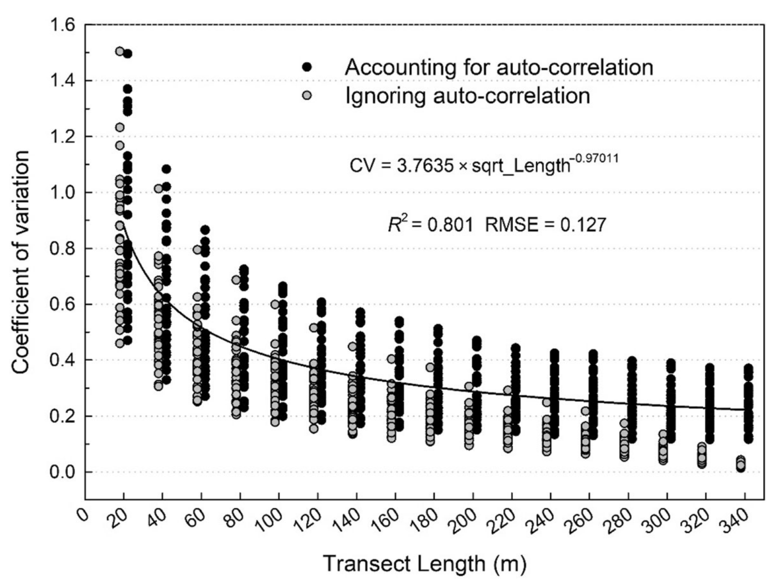

Fig. 1 Modelled relationship (black line, accounting for spatial auto-correlation) between line intersect sampling precision (expressed as coefficient of variation, CV) and transect length, showing that precision increases (lower CV) with increasing transect length. Ignoring spatial autocorrelation of DWD values (grey dots) falsely suggests greater precision can be obtained for a given transect length

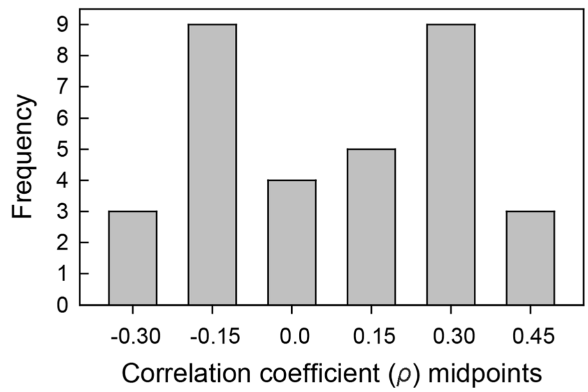

Fig. 2 Frequency distribution of ρ coefficients for the 33 locations, suggesting tendencies toward both negative and positive spatial autocorrelation of DWD values within sampling transects

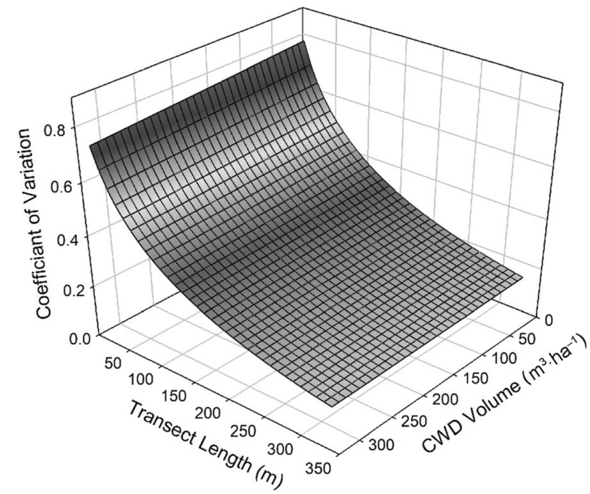

Fig. 4 Modelled relationship between line intersect sampling precision (expressed as coefficient of variation), transect length, and downed woody debris (DWD) volume, using simulations constrained by empirical field data from 33 locations within eight forest types of northeastern USA. Top model form: $CV = 1.30268 \times e^{-0.0002277 \times volume} \times e^{-0.1059549 \times sqrt\_length}$, where sqrt_length is the square root of transect length. Diagnostics for this model ($R^2 = 0.795$, RMSE = 0.128) were poorer than those for the length-only model (Fig. 1), and the volume term was non-significant

where sqrt_length is the square root of transect length (Fig. 4). However, diagnostics for this model were poorer than those for the length-only model (here, $R^2 = 0.795$, RMSE = 0.128), and the volume term became non-significant. For this reason, the length-only model was used for inferences regarding appropriate transect lengths.

Discussion

Our study was motivated by the need to provide guidance in selecting a LIS transect length appropriate for estimating DWD volumes for particular locations of interest. We note that a sampling design efficient for assessing average ecological values of DWD over large scales (as in national forest inventories) may not be well suited for assessing those same values at local scales, such as small forest reserves (Stokland et al. 2004), where the reserve itself may be the population of interest.

Our finding that the precision of LIS estimates consistently improved with increasing transect length corroborates a number of previous studies from quite different forest types (e.g., Van Wagner 1982; Pickford and Hazard 1978;

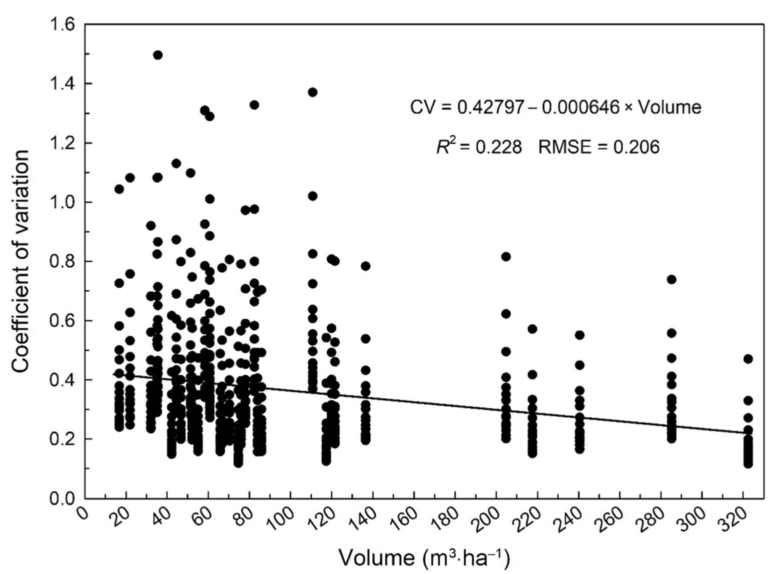

Fig. 3 Relationship between line intersect sampling precision (expressed as coefficient of variation) and downed woody debris (DWD) volume. For a given transect length, sites with greater volumes generally produce more precise estimates

Woldendorp et al. 2004; Miehs et al. 2010; Keane and Gray 2013). Similarly, our finding that precision of LIS estimates improved with increasing DWD volume present has been reported from a range of distinct forest systems (Brown 1974; Pickford and Hazard 1978; Woldendorp et al. 2004). However, the best-fit model that included both transect length and DWD volume as predictors showed poorer diagnostics than that of the length-only model described above, suggesting that the length-only model (shown in Fig. 1) could be used for inferences regarding appropriate transect lengths.

If DWD pieces were distributed at random, basic sampling theory (e.g. Thompson 2012) predicts that the variance of the total DWD intersected by a transect would scale linearly with transect length; thus, the CV would scale with the − 0.5 power of transect length for a given DWD density. However, clumping or regularity in DWD pieces would produce a departure from this scaling relationship. Indeed, our analysis of DWD values along transects revealed the presence of spatial autocorrelation, with some locations tending toward positive and others toward negative autocorrelation (Fig. 2). We found no discernable pattern among locations regarding forest type or DWD volume that might explain the bimodal distribution. We minimized the confounding effects of autocorrelation through simulations based on a zero-inflated autoregressive copula model, which accounted for correlations between adjacent 5-m transect segments unique to each location. Without accounting for such autocorrelation, our simulations would have underestimated variance (Fig. 1), leading to spurious assessments of appropriate transect lengths. To the best of our knowledge, this is the first attempt to employ an autoregressive model to account for spatial autocorrelation, while also addressing a zero-inflated distribution (segments may contain no DWD).

A frequent question in the design of research or monitoring protocols is how long transects should be to provide a reliable DWD volume estimate for a given location. A perusal of literature where LIS was used to estimate DWD volume for particular locations reveals substantial variation in the transect length employed: total transect lengths can range from < 40 m (Bradford and Kastendick 2010; Buma et al. 2014) to 200 m or more (Ekbom et al. 2006; Lõhmus and Kraut 2010). For the forest types and conditions sampled here, and considering the particular copula model framework employed, transect lengths of ca. 120 m provide what we feel is a reasonable level of precision, ranging from 18% to 60% CV from the observed simulations, and 37% from the fitted model results, across the range of DWD volumes encountered (Fig. 1). Practitioners working in these or similar forest types and requiring greater precision could chose a transect length based on the curve or equation presented in Fig. 1. For our own inventories, we find it convenient to arrange three 40-m transects radiating equi-angularly from a center point. This total length is somewhat comparable to the 100 m recommended by Woldendorp et al. (2004) for eucalypt forests of Australia to remain below a 100% CV threshold. Regarding the national DWD inventory of the US (Woodall and Monleon 2008), the total transect length sampled in a fully forested plot is 88 m, which would place the estimated CV below 43% for the locations sampled here. This level of uncertainty may be acceptable, as the strategic-scale inventory is designed to produce reliable DWD estimates (e.g., CV's below 10%) at the scale of entire forest types or regions (Woodall et al. 2013), not for particular locations of interest, the focus of our study.

LIS provides an attractive method for estimating DWD volume for a given localized area of interest, particularly when transects of reasonable length are employed. LIS sampling has been repeatedly shown to be much more time efficient than the other common sampling alternative, namely censuses of fixed area plots (Bailey 1970; Jordan et al. 2004). Nevertheless, a complete census of large plots, including mapping of DWD pieces, may remain the method of choice in studies where spatial information is needed, or where a particular plot size is needed for comparability to preexisting benchmarks (see Rouvinen and Kouki 2002; Jönsson et al. 2008). It should not be assumed, however, that a complete census of a large plot yields a precise estimate of DWD volume, in part because of the chance of overlooking hard-to-detect DWD pieces (Jordan et al. 2004; Kershaw et al. 2016), and more importantly because of the bias inherent in an assumed piece shape (Fraver et al. 2007; Ducey and Fraver 2018), both problems against which LIS provides natural safeguards. That is, LIS requires tallying only those pieces that cross a specific, well-defined transect, thus avoiding the need to search a large area for hard-to-detect pieces, and as above, no assumption of piece shape is required for LIS sampling.

Conclusions

As interest in forest carbon accounting, woody fuels assessments, and forest structure inventories increases, so does the need for practical DWD sampling guidelines based upon metrics of uncertainty, as we have provided in this study. Our simulations, based on empirical field data, allowed us to test the influence of transect length and DWD volume on the precision of LIS volume estimates, using a range of forest conditions. Our study is among the first to provide practical guidelines for appropriate LIS transect lengths for particular locations, where forest stands or small reserves may themselves represent the entire populations of interest. To the best of our knowledge, our study is the first to apply copula models to account for within-transect spatial autocorrelation of DWD volumes during our simulations, while also addressing a zero-inflated distribution (transect segments may contain no DWD), thereby properly addressing variance estimates. As expected, precision of DWD volume estimates improved dramatically with increasing cumulative transect length; precision improved only slightly with increasing DWD volumes. For the forest types sampled here, as well as the particular copula model framework employed, transect lengths of ca. 120 m provide a reasonable level of precision, ranging from 18% to 60% coefficients of variation. Finally, we note that although we have focused our analyses on DWD volume, these same results would apply to estimates of DWD biomass or carbon density, as these are typically calculated directly from volume data.

Abbreviations
AR: Auto-regressive; CV: Coefficient of variation; DWD: Downed woody debris; LIS: Line-intersect sampling

Acknowledgements
We thank Aakala T, Clark J, Curzon M, Doud E, Fien E, Foster J, Gill K, Kilbride J, Klockow P, Kuehne C, Manley H, Reinikainen M, Souder C, and Teets A for assistance in the field or laboratory. We thank Russell M, Sikkink P, and Stanovick J for thoughtful comments on earlier drafts of the manuscript. We thank the many staff and scientists at the Penobscot Experimental Forest, Bartlett Experimental Forest, Howland Research Forest, Bear Brook Experimental Watershed, The Northeast Wilderness Trust, The Nature Conservancy, and Baxter State Park Scientific Forest Management Area for maintaining these sites and facilitating our field research.

Funding
Support was provided by the Maine Agricultural and Forest Experiment Station, the Joint Fire Science Program (Project 08–1–5-04), and the USDA Forest Service Northern Research Station.

Availability of data and materials
The materials described in the manuscript including all relevant raw data will be freely available upon request from the corresponding author.

Authors' contributions
SF conceived and designed the study, conducted/supervised the field inventory, and led the analyses of simulation results. MD conceived, conducted, and wrote text addressing the copula models. SF and MD drafted the first version of the manuscript. CWW and AWD assisted with interpretation of results and contributed to writing. All authors contributed to editing and revising the final manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]School of Forest Resources, University of Maine, Orono, ME, USA. [2]Department of Natural Resources and the Environment, University of New Hampshire, Durham, NH, USA. [3]US Forest Service, Northern Research Station, Durham, NH, USA. [4]Rubenstein School of Environment and Natural Resources, University of Vermont, Burlington, VT, USA. [5]Department of Biological Sciences, George Washington University, Washington, DC, USA. [6]US Forest Service, Northern Research Station, Grand Rapids, MN, USA.

Received: 30 April 2018 Accepted: 5 November 2018
Published online: 05 December 2018

References
Affleck DLR, Gregoire TG, Valentine HT (2005) Design unbiased estimation in line intersect sampling using segmented transects. Environ Ecol Stat 12(2):139–154
Bailey GR (1970) A simplified method of sampling logging residue. Forest Chron 46(4):288–303
Bradford JB, Kastendick DN (2010) Age-related patterns of forest complexity and carbon storage in pine and aspen–birch ecosystems of northern Minnesota, USA. Can J For Res 40:401–409
Brown JK (1974) Handbook for inventorying downed woody material. USDA For Serv Gen Tech Rep INT-16, Ogden
Buma B, Poore RE, Wessman CA (2014) Disturbances, their interactions, and cumulative effects on carbon and charcoal stocks in a forested ecosystem. Ecosystems 17(6):947–959
Burnham KP, Anderson DA (2002) Model selection and multimodal inference: a practical information-theoretic approach. Springer-Verlag, Inc., New York
Canham CD, LePage PT, Coates KD (2004) A neighborhood analysis of canopy tree competition: effects of shading versus crowding. Can J For Res 34:778–787
Core Team R (2016) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
de Vries PG (1979) Line intersect sampling – statistical theory, applications, and suggestions for extended use in ecological inventory. In: Cormack RM, Patil GP, Robson DS (eds) Sampling biological populations. International Cooperative Publishing House, Fairland
Ducey MJ, Fraver S (2018) The conic-paraboloid formulae for coarse woody material volume and taper and their approximation. Can J For Res 48:966–975
Dunn CJ, Bailey JD (2015) Modeling the direct effects of salvage logging on long-term temporal fuel dynamics in dry-mixed conifer forests. For Ecol Manag 341:93–109
Efron B, Tibshirani RJ (1993) An introduction to the bootstrap. Chapman & Hall, New York
Ekbom B, Schroeder LM, Larsson S (2006) Stand specific occurrence of coarse woody debris in a managed boreal forest landscape in Central Sweden. For Ecol Manag 221:2–12
Ene LT, Næsset E, Gobakken T (2013a) Model-based inference for k-nearest neighbours predictions using a canonical vine copula. Scand J For Res 28: 266–281
Ene LT, Næsset E, Gobakken T, Gregoire TG, Ståhl G, Holm S (2013b) A simulation approach for accuracy assessment of two-phase post-stratified estimation in large-area LiDAR biomass surveys. Remote Sens Environ 133:210–224
Ene LT, Næsset E, Gobakken T, Gregoire TG, Ståhl G, Nelson R (2012) Assessing the accuracy of regional LiDAR-based biomass estimation using a simulation approach. Remote Sens Environ 123:579–592
Eskelson BNI, Madsen L, Hagar JC, Temesgen H (2011) Estimating riparian understory vegetation cover with beta regression and copula models. For Sci 57:212–221
Fortin M, Delisle-Boulianne S, Pothier D (2013) Considering spatial correlations between binary response variables in forestry: an example applied to tree harvest modeling. For Sci 59(3):253–260
Fraver S, Dodds K, Kenefic LS, Morrill R, Seymour RS, Sypitkowski E (2017) Forest structure following tornado damage and salvage logging in northern Maine, USA. Can J For Res 47:560–564
Fraver S, Palik BJ (2012) Stand and cohort structures of old-growth *Pinus resinosa*-dominated forests of northern Minnesota, USA. J Veg Sci 23:249–259
Fraver S, Ringvall A, Jonsson BG (2007) Refining volume estimates of down woody debris. Can J For Res 37:627–633
Genest C, MacKay J (1986) The joy of copulas: bivariate distributions with uniform marginals. Am Stat 40:280–283
Grafström A, Saarela S, Ene LT (2014) Efficient sampling strategies for forest inventories by spreading the sample in auxiliary space. Can J For Res 44: 1156–1164
Gregoire TG (1998) Design-based and model-based inference in survey sampling: appreciating the difference. Can J For Res 28:1429–1447
Gregoire TG, Valentine HT (2008) Sampling strategies for natural resources and the environment. Chapman & Hall/CRC, New York
Harmon ME, Franklin JF, Swanson FJ, Sollins P, Gregory SV, Lattin JD, Anderson N, Cline SP, Aumen NG, Sedell JR, Lienkaemper GW, Cromack K Jr, Cummins KW (1986) Ecology of coarse woody debris in temperate ecosystems. Adv Ecol Res 15:133–302

Jönsson M, Fraver S, Jonsson BG (2011) Spatio-temporal variation of coarse woody debris input in woodland key habitats in Central Sweden. Silva Fenn 45(5):957–967
Jönsson MT, Edman M, Jonsson BG (2008) Colonization and extinction patterns of wood-decaying fungi in a boreal old-growth *Picea abies* Forest. J Ecol 96: 1065–1075
Jordan GJ, Ducey MJ, Gove JH (2004) Comparing line-intersect, fixed-area, and point relascope sampling for dead and downed coarse woody material in a managed northern hardwood forest. Can J For Res 34:1766–1775
Kaiser L (1983) Unbiased estimation in line-intercept sampling. Biometrics 39: 965–976
Keane RE, Gray K (2013) Comparing three sampling techniques for estimating fine woody down dead biomass. Int J Wildland Fire 22(8):1093–1107
Kershaw JA, Weiskittel AR, Lavigne MB, McGarrigle E (2017) An imputation/ copula-based stochastic individual tree growth model for mixed species Acadian forests: a case study using the Nova Scotia permanent sample plot network. Forest Ecosystems 4(1):15. https://doi.org/10.1186/s40663-017-0102
Kershaw JA Jr, Ducey MJ, Beers TW, Husch B (2016) Forest Mensuration, 5th edn. Wiley, New York
Kershaw JA Jr, Richards EW, McCarter JB, Oborn S (2010) Spatially correlated forest stand structures: a simulation approach using copulas. Comput Electron Agric 74:120–128
Lõhmus A, Kraut A (2010) Stand structure of hemiboreal old-growth forests: characteristic features, variation among site types, and a comparison with FSC-certified mature stands in Estonia. For Ecol Manag 260(1):155–165
Lyles RH, Williams JK, Chuachoowong R (2001) Correlating two viral load assays with known detection limits. Biometrics 57:1238–1244
MacPhee C, Kershaw JA, Weiskittel AR, Golding J, Lavigne MB (2017) Comparison of approaches for estimating individual tree height–diameter relationships in the Acadian forest region. Forestry 91:132–146
Miehs A, York A, Tolhurst K, Di Stefano J, Bell T (2010) Sampling downed coarse woody debris in fire-prone eucalypt woodlands. For Ecol Manag 259:440–445
Nelsen RB (2006) An introduction to copulas, 2nd edn. Springer, Inc. New York, p 272
Pickford SG, Hazard JW (1978) Simulation studies on line intersect sampling of forest residue. For Sci 24(4):469–483
Pinheiro J, Bates D, DebRoy S, Sarkar D (2016) Nlme: linear and nonlinear mixed effects models. CRAN. R-package. R Foundation for Statistical Computing, Vienna
Rouvinen S, Kouki J (2002) Spatiotemporal availability of dead wood in protected old-growth forests: a case study from boreal forests in eastern Finland. Scand J Forest Res 17:317–329
Russell M, Fraver S, Aakala T, Gove J, Woodall CW, D'Amato AW, Ducey MJ (2015) Quantifying carbon stores and decomposition in dead wood: a review. For Ecol Manag 350:107–128
Russell MB, Woodall CW, D'Amato AW, Fraver S, Bradford JB (2014) Technical note: linking climate change and downed woody debris decomposition across forests of the eastern United States. Biogeosciences 11:6417–6425
Schoennagel T, Veblen TT, Romme WH (2004) The interaction of fire, fuels, and climate across Rocky Mountain forests. Bioscience 54:661–676
Sikkink PG, Keane RE (2008) A comparison of five sampling techniques to estimate surface fuel loading in montane forests. Int J Wildland Fire 17:363–379
Sollins P (1982) Input and decay of coarse woody debris in coniferous stands in western Oregon and Washington. Can J For Res 12:18–28
Stokland JN, Siitonen J, Jonsson BG (2012) Biodiversity in dead wood. Cambridge University Press, Cambridge
Stokland JN, Tomter SM, Söderberg U (2004) Development of dead wood indicators for biodiversity monitoring: experiences from Scandinavia. In: Marchetti M (ed) Monitoring and indicators of forest biodiversity in Europe – from ideas to operationality. European Forest Institute proceedings 51, Joensuu, Finland, pp 207–226
Thompson SK (2012) Sampling, 3rd edn. John Wiley & Sons, New York
Van Wagner CE (1968) The line intersect method in forest fuel sampling. For Sci 14(1):20–26
Van Wagner CE (1982) Practical aspects of the line intersect method, vol 12. Petawawa National Forestry Institute, Chalk River
Wang M, Rennolls K, Tang S (2008) Bivariate distribution modeling of tree diameters and heights: dependency modeling using copulas. For Sci 54:284–293
Wang M, Updhyay A, Zhang L (2010) Trivariate distribution modeling of tree diameter, height, and volume. For Sci 56:290–300
Wang SS (1998) Aggregation of correlated risk portfolios: models and algorithms. Proc Casualty Actuarial Soc 85:848–937
Warren WG, Olsen PF (1964) A line intersect technique for assessing logging waste. For Sci 10:267–276
Woldendorp G, Keenan RJ, Barry S, Spencer RD (2004) Analysis of sampling methods for coarse woody debris. For Ecol Manag 198:133–148
Woodall CW, Monleon VJ (2008) Sampling protocols, estimation procedures, and analytical guidelines for down woody materials indicator of the Forest inventory and analysis program. USDA For Serv Gen Tech Rep NRS-22:68
Woodall CW, Rondeux J, Verkerk P, Ståhl G (2009) Estimating dead wood during national inventories: a review of inventory methodologies and suggestions for harmonization. Environ Manag 44:624–631
Woodall CW, Russell MB, Walters BJ, D'Amato AW, Fraver S, Domke G (2015) Downed dead wood net carbon flux in forests of the eastern United States. Oecologia 177:861–874
Woodall CW, Walters BF, Oswalt SN, Domke GM, Toney C, Gray AN (2013) Biomass and carbon attributes of downed woody materials in forests of the United States. For Ecol Manag 305:48–59