Forest Ecology and Management 430 (2018) 10–20

Contents lists available at ScienceDirect

Forest Ecology and Management

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

Plot-level variability in biomass for tropical forest inventory designs

Nicolas Picard*, Javier G.P. Gamarra, Luca Birigazzi, Anne Branthomme

Food and Agriculture Organization of the United Nations, Forestry Department, Vialle delle Terme di Caracalla, 00153 Rome, Italy

ARTICLE INFO

Keywords:
Aboveground biomass
Cluster sampling
Forest inventory
Point process
Sampling design
Tropical rain forest

ABSTRACT

The spatial distribution of biomass is key to optimize forest inventory designs to estimate forest aboveground biomass. Point process theory sets an appropriate mathematical framework to model the spatial distribution of trees, then to derive analytical expressions for the relationship between the variance of biomass in plots and the characteristics (size and shape) of plots, possibly accounting also for plot autocorrelation in biomass. Models derived from point process theory provided a better fit to data from twenty spatially homogeneous sites in tropical rain forests than the commonly used Taylor power model for biomass variance. The model $CV = \sqrt{\omega + \kappa/|A|}$ with CV the coefficient of variation of biomass, $|A|$ the plot area, and ω and κ parameters to estimate, provided in particular a better fit than the power model when the range of autocorrelation in biomass was greater than the plot width. The twenty tropical forest sites greatly differed in the observed relationship between biomass variance and plot size, reflecting differences in the spatial pattern of biomass according to the fitted point process. Accordingly, optimized forest inventory designs also greatly differed between forest sites, with positive biomass autocorrelation favouring cluster sampling design with a distance between subplots in the order of the range of the biomass autocorrelation. In a spatially heterogeneous context consisting of different homogeneous forest strata, large-scale heterogeneity prevailed upon local biomass autocorrelation in determining the optimized plot size and shape. If uncontrolled through stratification, large-scale heterogeneity resulted in much smaller (approximately 0.1–0.2 ha) optimized plot sizes than the homogeneous case (approximately 1–2 ha).

1. Introduction

Forest inventory provides information necessary for forest management or for designing evidence-based forest policies at different scales. The important role played by forests in mitigating and adapting to climate change, as repositories of biodiversity and as ecosystems to combat desertification at national and global scale has put additional emphasis on the necessity to have reliable forest monitoring systems at national levels, with forest inventories being again at the core of these monitoring systems (FAO, 2017). Sampling designs when performing field inventories in forests are not equally efficient, and searching for optimized designs has long been addressed by foresters (Ranneby et al., 1987; Schreuder et al., 1993; Mandallaz, 2008; Fattorini, 2015). Optimization refers to different trade-offs to be solved, such as the trade-off between the number and size of plots, which depends on another nested trade-off between within-plot and between-plot variability (Hall et al., 1998). Solving trade-offs is further complicated in multipurpose inventories that involve simultaneous optimization of goals (Scott, 1993).

Optimization aims to maximize some relative efficiency that measures the information gained per unit effort expended in sampling, and generally defined as the product of the relative variability and relative cost (Wiegert, 1962):

$$e = \frac{\mathrm{CV}(A)^2}{\mathrm{CV}(1)^2} \times \frac{C_u(A)}{C_u(1)} \quad (1)$$

where $\mathrm{CV}(A)$ is the coefficient of variation (CV) of the target population attribute using sampling unit type A, $C_u(A)$ is the associated sampling unit-level cost, while $A = 1$ indicates some basis sampling unit of reference. The most appropriate sampling unit size is the one that minimizes e, i.e. when $e < 1$, sampling unit A is more efficient than the reference sampling unit 1 (Zeide, 1980; Husch et al., 2003; Köhl et al., 2006). Both CV and unit cost are dependent on the sampling unit size and shape, and on the target population attribute such as stock density, biomass, carbon or biodiversity, so that minimizing e determines the overall configuration of the sampling units.

A key element to know in order to optimize the efficiency is the relationship between the CV and the sampling unit. We hereafter focus on sampling designs whose units are fixed-area plots, possibly consisting of several subplots (cluster design, where a plot is a cluster of subplots). A classical formula for modelling the relationship between

* Corresponding author.
E-mail addresses: nicolas.picard@fao.org (N. Picard), javier.garciaperez@fao.org (J.G.P. Gamarra), luca.birigazzi@fao.org (L. Birigazzi), anne.branthomme@fao.org (A. Branthomme).

https://doi.org/10.1016/j.foreco.2018.07.052
Received 29 December 2017; Received in revised form 26 July 2018; Accepted 28 July 2018

the CV and A is the power model, also known as Taylor's law (Smith, 1938; Taylor, 1961; O'Regan and Arvanitis, 1966):

$$\mathrm{CV}(A) = \mathrm{CV}(1) \times |A|^{-\beta} \quad (2)$$

where $|A|$ is the area of the sampling plot expressed in units of the reference sampling plot and β a real-valued exponent. The value of β depends on the forest characteristics but also on the geometry of the reference sampling plot. When the spatial distribution of trees is completely random with no spatial correlation in the plot-level population attribute, $\beta = 0.5$. This value has also been used as a default value for β when little information was available on the spatial distribution of trees (Freese, 1961; Zeide, 1980). Then only the value of CV(1) needs to be estimated, which can be obtained from a pre-inventory consisting of a few reference sampling plots. The assumption of a constant β implies scale-invariance in biomass variation due to a lack of spatial structure (Muller-Landau et al., 2014). Yet, tropical forest biomass is known to present characteristic scales reflecting differential clustering patterns in tree size (Plotkin et al., 2002). Hence, the assumption may not necessarily hold true across all scales.

Optimizing the sampling efficiency with respect to both plot size and shape requires a more comprehensive description of the way CV varies with size and shape. For cluster designs where a sampling plot consists of several subplots separated by a given distance, testing different subplots configuration in the field can be a way to directly compare the efficiency of different cluster designs (Yim et al., 2015), and a key descriptor in this case is the covariance function that describes the autocorrelation between two subplots at a certain distance from each other (Cochran, 1977; Kleinn, 1994, p.219). Previous studies have shown that the most efficient cluster designs are obtained when the distance between subplots is of the order of the range of the covariance function (Yim et al., 2015).

Implementing pre-inventories in the field to establish the relationship between the CV and the sampling plot as a prerequisite for design optimization can be costly, and an alternative consists in mapping the target attribute at the forest scale and then virtually implementing the forest inventory. This approach has great flexibility as any type of inventory based on sampling units of any size and shape can be simulated, but it requires the forest map to be realistic. Satellite images combined with field measurements have been used to produce such maps of forest attributes for design optimization (Tokola and Shrestha, 1999). Forest simulators that produce virtual forest stands have been used for that purpose too (Mackisack and Wood, 1990; Brink and Schreuder, 1992). Large research plots have also been used to derive the relationship between CV and plot size using nested grids of plots that partition the large plot (Wagner et al., 2010; Réjou-Méchain et al., 2014). One difficulty in this case is to predict the CV for plot shapes that are different from the grid, and another is to account for biomass correlation among plots.

The current study aims to derive the CV for any plot size and shape based on the population characteristics, and to apply this expression of the CV to the optimization of the plot design of a forest inventory. We mainly considered the case of an homogeneous forest, thus for an inventory of limited spatial extent, even if the heterogeneous case consisting of different homogeneous strata is also addressed. The theoretical framework to derive the CV is the point process theory that models the spatial distribution of trees as the outcome of a stochastic process (Cressie, 1991). We will show that the covariance function as defined by Kleinn (1994) can also be analytically derived from the point process, thus providing a consistent analytical framework to compute the influence of the distance between subplots in a cluster design on the design efficiency. To our knowledge, this analytical framework based on point process theory is novel for design optimization of forest inventories. We will focus on biomass as the target population attribute given its current importance in climate change issues, and will focus on tropical forests given the lack of design optimization studies for those forests. We used previously published data at twenty fully studied tropical forest sites (Réjou-Méchain et al., 2014).

2. Materials and methods

2.1. Forest types and data

We relied on the data published by Réjou-Méchain et al. (2014) giving the CV of biomass for different square plot sizes in different forests (see Tables S1 and S2 in Réjou-Méchain et al., 2014). We used data only from tropical forests, with 20 different sites in Africa, America and Asia. Réjou-Méchain et al. (2014) fitted the power model (2). Based on the value of the β exponent, forests can be classified into three categories that reflect the type of spatial pattern of trees: $\beta = 0.5$ corresponds to complete spatial randomness, $\beta > 0.5$ to regular patterns, and $\beta < 0.5$ to clustered patterns. The latter situation is the most common in natural forests. Accordingly, the 20 forests studied by Réjou-Méchain et al. (2014) could be ranked from the most regular (Ituri Edoro 1, Paracou) to the most clustered (Yasuni, Xishuangbanna, Palanan, Sinharaja, etc.) with intermediate forest types close to random (Ituri Lenda 2, Ituri Edoto 2, Manaus, etc.).

All the plots used by Réjou-Méchain et al. (2014) to get the CV data were square subplots that partitioned a single large rectangular or square plot (with a size ranging from 9 ha to 50 ha depending on the forest), as shown in Fig. 1. Because positive spatial autocorrelation in biomass is most often observed (Réjou-Méchain et al., 2014), subplots that partition a large plot are not independent and the CV of biomass computed from these subplots is less than what would be obtained with independent plots. This issue, as well as the issue of generalizing the CV-size relationship to other plot shapes than the square, were simultaneously solved by considering a latent model for the spatial distribution of trees, enabling us to explicitly account for spatial correlation between subplots and to adapt to any plot shape.

2.2. Coefficient of variation of biomass: homogeneous case

We first consider the case of an homogeneous forest. The latent model that we used to derive the expression of the CV of biomass is a marked point process where the position of a point is the location of a tree and the mark is the tree biomass (Cressie, 1991; see also Appendix A). To reflect the homogeneity of the forest, the point process considered in this section is homogeneous and isotropic. The biomass M of a plot A is the sum of the biomasses B of the trees found in this plot. In the jargon of point process theory, the plot biomass is a mark sum measure (Stoyan, 1984; Cressie, 1991, p.713–714). The mean plot biomass then is $\mathrm{E}(M) = \lambda |A| \mathrm{E}(B)$, while the covariance of plot biomass between two (possibly overlapping) plots A_1 and A_2 is (Cressie, 1991, p.713; see also Appendix A):

$$\mathrm{Cov}(M_1, M_2) = 2\pi\lambda^2 \mathrm{E}(B)^2 \int_0^\infty \gamma_{A_1 A_2}(r)\,[k(r)\,g(r) - 1]\,r\,dr + \lambda\,|A_1 \cap A_2|\,[\mathrm{Var}(B) + \mathrm{E}(B)^2] \quad (3)$$

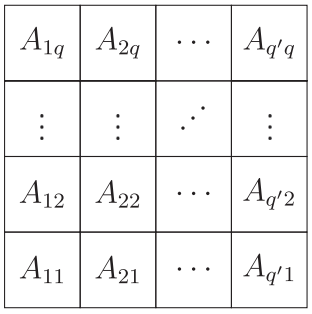

Fig. 1. Partition of a large square ($q = q'$) or rectangular ($q \neq q'$) plot into non-overlapping square subplots A_{ij}. Subplot indexes are based on row and column numbers.