Accuracy and efficiency of area classifications based on tree tally

Michael S. Williams, Hans T. Schreuder, and Raymond L. Czaplewski

Abstract: Inventory data are often used to estimate the area of the land base that is classified as a specific condition class. Examples include areas classified as old-growth forest, private ownership, or suitable habitat for a given species. Many inventory programs rely on classification algorithms of varying complexity to determine condition class. These algorithms can be simple decision trees applied in the field or computer calculations applied on a field data recorder or after the data are collected. The advantages to using these algorithms are consistent classification of the condition class, reduced crew training, and the ability to define new condition classes after the data are collected, which will be referred to as postclassification. We discuss three types of the errors that can occur when these types of algorithms are employed and quantify the potential for error with examples. The examples are substantial oversimplifications of the true problem, but they show how difficult it is to determine anything but the most general condition classes using plot data alone. A discussion of how condition class is scale dependent and some general guidelines and recommendations are given.

Résumé : Les données d'inventaire sont souvent utilisées pour estimer la superficie du territoire caractérisée par une classe de situations spécifiques. On retrouve par exemple des superficies classées comme vieille forêt, forêt privée ou habitat propice à une espèce donnée. Plusieurs programmes d'inventaire dépendent d'algorithmes de classification, de complexité variable, pour déterminer la classe de situations. Ces algorithmes peuvent être de simples arbres de décision utilisés sur le terrain ou des calculs programmés intégrés aux appareils d'enregistrement de données sur le terrain ou effectués après que les données ont été collectées. L'utilisation de ces algorithmes présente plusieurs avantages dont une classification consistante des classes de situations, un entraînement moindre des équipes de terrain et la possibilité de définir de nouvelles classes de situations après que les données ont été collectées, autrement appelée classification à posteriori. Nous discutons de trois types d'erreurs qui peuvent survenir lorsque ce type d'algorithmes est utilisé et nous quantifions ces erreurs potentielles à l'aide d'exemples. Ces exemples sont très simplifiés comparativement au vrai problème mais ils illustrent à quel point il est difficile de déterminer autre chose que les classes de situations les plus générales en utilisant seulement les données des parcelles. La façon dont les classes de situations dépendent de l'échelle à laquelle on se situe ainsi que des suggestions et recommandations générales font l'objet de la discussion.

[Traduit par la Rédaction]

Introduction

Traditionally, forest sampling has dealt with the estimation of forest attributes that fall into two categories. The first is those derived directly from tree tally. Examples include number of trees, basal area, and volume. The second category is composed of area or length attributes. Examples are land-use class, cover type, stand structure, productivity class, ownership, habitat, and length of boundaries. The general term condition class refers to this class of attributes. Estimators for condition class have been discussed under different sampling paradigms in articles by Scott and Bechtold (1995), Williams and Schreuder (1995), and Eriksson (2001). When samples are collected using fixed-area plot sampling, with m plots uniformly spread over the area sampled, one estimator of the area in condition class c is

$$[1] \qquad \hat{A}_c = A\hat{P}(c) = A\frac{\sum_{i=1}^{m} a_i(c)}{\sum_{i=1}^{m} a_i}$$

where A is the total area over which sampling is performed, $\hat{P}(c)$ is the estimated proportion of the area in condition class c, $a_i(c)$ is the area of plot i covering condition class c, and a_i is the area of the plot within the area being sampled. While the properties of condition-class estimators are well understood, they rely on the assumption that the condition class can be determined perfectly, which we will show is not always a reasonable assumption because of sampling error and issues of scale.

For condition classes such as ownership, estimators such as eq. 1 will perform well, because the condition class can usually be determined without error. However, many condition classes are defined by the tree tally. These condition classes must also have some minimum area associated with the count. The minimum area requirement is needed because

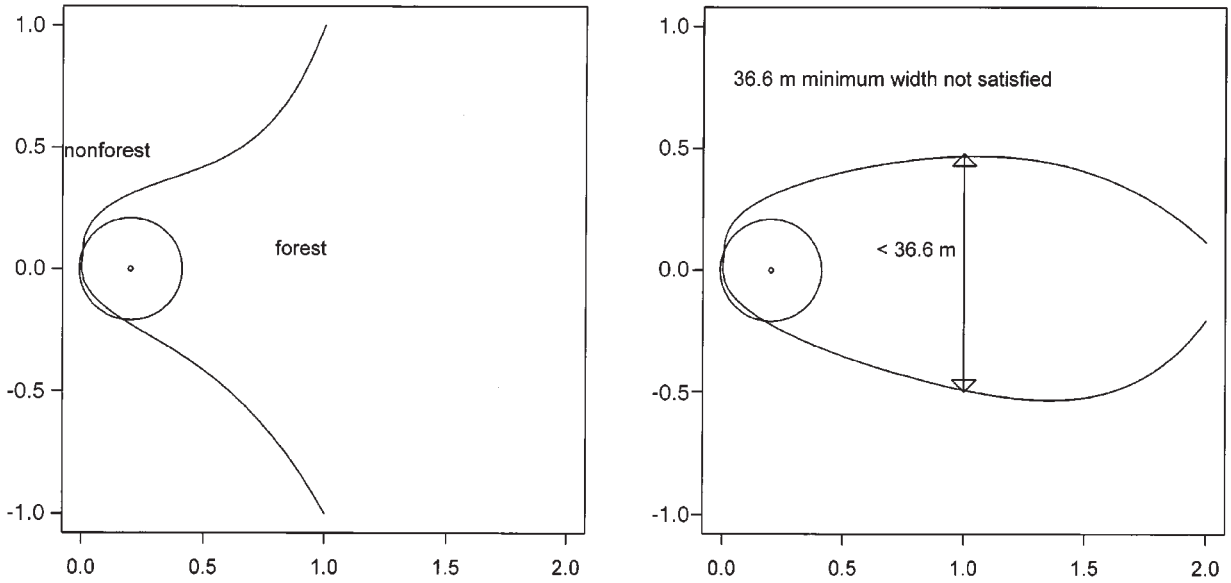
Received April 17, 2000. Accepted December 1, 2000. Published on the NRC Research Press Web site on March 13, 2001.

M.S. Williams,[1] H.T. Schreuder, and R.L. Czaplewski. Rocky Mountain Research Station, USDA Forest Service, 2150 A Center Drive, Suite 350, Fort Collins, CO 80526, U.S.A.

[1]Corresponding author (e-mail:usdafs@lamar.colostate.edu).

 DOI: 10.1139/cjfr-31-3-556

Fig. 1. Two examples of the classification of identical plots. Both plots cover identical areas and the conditions on the plot, but in the second case the minimum width requirement is not met.



the count per unit area varies wildly with the size of the area considered. For example, consider the problem of determining a condition class based on basal area per hectare for a population consisting of a single tree with basal area equal to 0.5 m^2. If the tree is measured on a 1-m^2 area (plot) the resulting basal area per hectare is 1250 m^2/ha. If the same tree is used to classify a plot covering 50 m^2, the resulting basal area per hectare is 100 m^2/ha. This problem has been addressed by Lappi (1991), where methods are studied for determining the distribution of stand densities from plot data. The shortcoming of these techniques is that two distinctly different condition classes can have an identical stand densities. For example, Lappi's results for estimating the number of tree per hectare cannot be used to differentiate between idle agricultural land that has 120 saplings per hectare and an old-growth stand with the same number of trees per unit area. To further complicate matters, additional restrictions are sometimes placed on condition classes. An example is forest area, where in the United States, a stand must be at least 36.6 m wide in addition to the tally and area requirements (Helms 1998). Thus, given the conditions for area and tally, it is possible to determine the correct classification of the plot only by examining an area as large or larger than the area limit. An example is given in Fig. 1, where it is assumed that the area sampled and tree tally is identical for two different plots, but in the second case the width requirement for forest is not met. In practice, a smaller area can usually be measured, because the minimum area and tally requirements are obviously met, e.g., the entire plot is encompassed by a large forested stand that easily meets the tally, area, and width requirements. However, in the case of condition classes that depend on rare occurrences, such as the presence of large trees, or large areas, such as habitat for foraging animals, it may be impossible to accurately determine the correct condition class. We discuss three types of errors that can occur and their effect on the estimation of A_c. Some simple examples based on Poisson process models are used to show how classification of area based on plot tally can be a difficult endeavor.

One-way classification errors

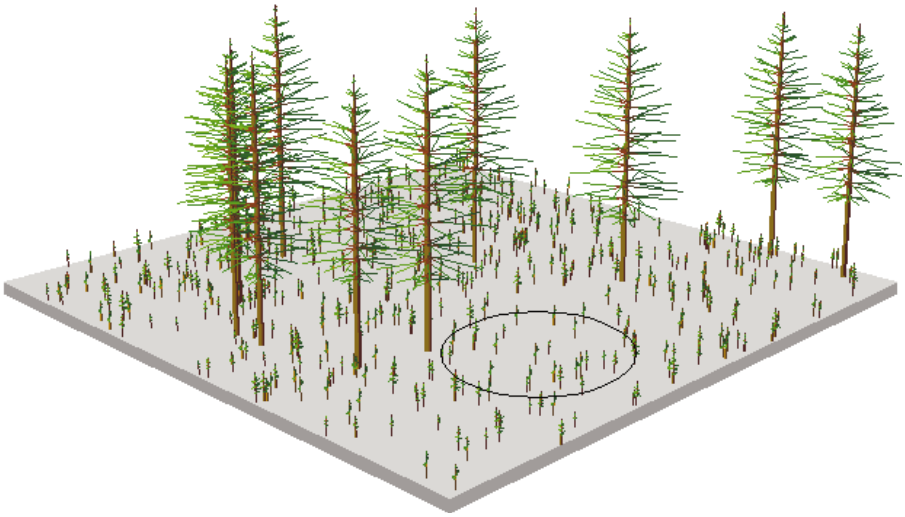
We define a one-way error as the case where an error in classification will always bias the estimator in one direction. This situation occurs when the classification is based on the presence or absence of one or more trees. An example could be the classification of an old-growth stand or nesting habitat that relies on the presence of at least one large tree on the plot. Figure 2 illustrates the case where the plot falls in an old-growth stand, but no old-growth trees are sampled. In this situation only an objective decision based on observing the surrounding area or a census of the hectare can correctly determine the condition class.

Now that the problem has been identified, how can we quantify the probability of an error in condition classification? A simple point process model can be assumed to gain some insight into the probability of misclassification. Suppose old-growth trees are located within a 1-ha area in accordance with a Poisson process (Taylor and Karlin 1994) with intensity λ = number of old-growth trees/hectare and suppose the plot is classified as old-growth provided at least one old-growth tree is in the sample. Under this model the probability of not observing any old-growth trees in the sample is

$$P(n = 0) = e^{-\lambda a}$$

where a is the area of the plot and n is the number of old-growth trees. Figure 3 depicts the probability of misclassification for the plot design used by the Forest Inventory and Analysis (FIA) program in the United States, whose plot encompasses 0.07 ha. Note that the probability of misclassification exceeds 70% when the intensity is only five old-growth trees per hectare ($\lambda = 5$) and still exceeds 5% until $\lambda > 43$. Doubling and tripling the plot size reduces the

Fig. 2. Sampling scenario wherein a plot falls in a hectare that is classified as old-growth, but no old-growth trees are tallied.



probability of misclassification, but the error rates still exceed 20% for low densities of trees.

In the presence of one-way classification error, $\hat{A}_c$ would tend to underestimate the true area of condition class c, with the bias being roughly equal to the probability of misclassification. This bias persists for populations of any size. The effects of similar errors on forest inventories has been studied by Gertner (1990) and references therein. It should be noted that one-way classification errors can occur anytime the classification is based on tree tally. However, the probability of misclassification decreases as the tree tally increases.

Two-way classification errors

Two-way classification errors occur when the condition class is defined by a count per unit area, such as basal area or stems per hectare. Examples include classification of forest versus nonforest and size classes, such as pole- or sawtimber. The error is due to the fact that the count per unit area is estimated from a sample of the area rather than a census. Thus, the misclassification is due to within-plot sampling error.

Some insight can be gained through the Poisson process model, where we assume that condition class is determined by a count of stems per hectare, which is λ. Assume the condition class of interest comprises all areas where the number of stems per hectare fall in the range $200 < \lambda \leq 300$ and assume the locations of trees for the forest population in question are distributed according to a Poisson process with $\lambda = 250$. If the 0.07-ha FIA plot is used, the plot will be classified in the correct condition class whenever the number of stems tallied, n, falls in the range $14 \leq n \leq 21$. Under the Poisson process model the probability of correctly classifying the plot is

$$P(14 \leq n \leq 21) = \sum_{n=14}^{21} \frac{e^{-250/0.07}(250/0.07)^n}{n!} = 0.66$$

Thus, the probability of correct classification is approximately 66%. The probability of the calculated condition class being $\lambda \leq 200$ is

$$P(n < 14) = \sum_{n=1}^{13} \frac{e^{-250/0.07}(250/0.07)^n}{n!} = 0.17$$

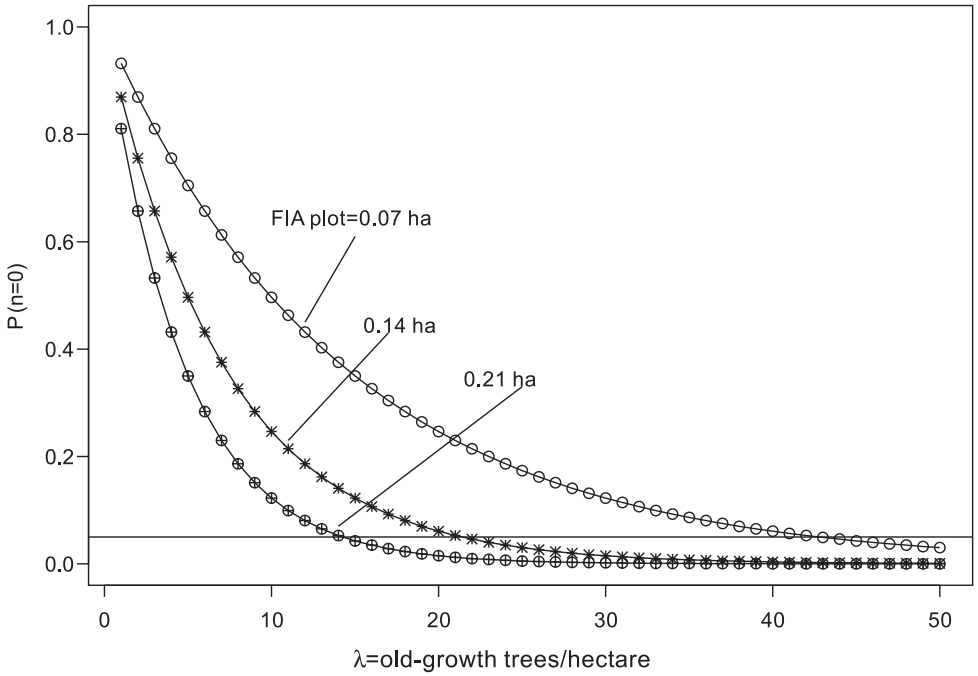
while the probability of the plot being classified in the $\lambda > 300$ class is $P(n > 21) = 0.17$. While the range of stem density may be fairly narrow in this example, it should be noted that λ falls exactly in the middle of the range. When $\lambda = 215$, the probability of correct classification drops to 0.59. The probability of the plot being classified in the $\lambda < 200$ class is 0.36, which shows that the probability of assigning the incorrect condition class is not symmetrical. In the presence of two-way classification error, the magnitude of the bias in $\hat{A}_c$ is difficult to assess, because it depends on both the underlying distribution of stand densities and the specific condition class chosen. It is possible for offsetting classification errors to reduce the bias. However, it seems unlikely that the bias will be negligible or that a confidence interval for $\hat{A}_c$ will achieve its nominal coverage rate.

These examples probably underestimate the probability of misclassification, because it is assumed that the populations are homogeneous and plots do not straddle a forest–nonforest or stand boundary. It should also be noted and that one-way classification error is just a special case of this situation.

Errors in classification based on plurality

Consider the problem of determining whether a stand should be classified as predominantly hardwood, softwood, or mixed hardwood–softwood. In practice, forest type is determined by the percent cover of the dominant species group (Helms 1998), where percent cover is closely related to the number of trees and size distribution. Simple point process models cannot account for the size distribution, but some idea of the accuracy of the classification into these categories can be derived by studying a simpler problem. For the purpose of illustration, suppose forest type is determined strictly from the

Fig. 3. Probability of one-way misclassification for three plot sizes. The number of old-growth trees per hectare is modeled by a Poisson process with intensity λ.



percentage of trees in each species group, with a stand being classified as pure hardwood or softwood when more than two-thirds of all trees fall into one of these categories. Thus, the stand will be classified as mixed hardwood–softwood whenever fewer than two-thirds of the trees are either hardwoods or softwoods. Poisson process models can be used to determine rough estimates of the likelihood of correctly classifying the stand based on the plot data.

Assume a hardwood stand contains $\lambda = 450$ trees/ha (182 trees/acre). Of these trees, let p be the proportion of hardwoods, with $0.67 \le p < 1$. Suppose the hardwood and softwood trees are located in accordance with two independent Poisson processes with intensities $\lambda_{HW} = p\lambda$ and $\lambda_{SW} = (1 - p)\lambda$, respectively. Define n, n_{HW}, and n_{SW} as the total number of trees in the sample, and the sample size for the hardwood and softwood trees, respectively. One requirement of this calculation is that the achieved sample size must be specified. If the FIA plot were used to draw the sample, the average sample size would be approximately $n = 32$ trees and in order for the hectare to be properly classified the number of hardwoods observed would have to be at least $n_{HW} \ge 22$. The result that will be employed is that the probability of observing n_{HW} given a sample of n trees, $P(n_{HW}|n)$, is distributed according to a binomial $(n, \lambda_{HW}/(\lambda_{HW} + \lambda_{SW}))$ (Taylor and Karlin 1994). Thus:

$$P(n_{HW} \ge 22 | n = 32) = \sum_{i=22}^{n} \binom{n}{i} \left(\frac{\lambda_{HW}}{\lambda_{HW} + \lambda_{SW}} \right)^{i} \left(1 - \frac{\lambda_{HW}}{\lambda_{HW} + \lambda_{SW}} \right)^{n-i}$$

The accuracy of the classification is a function of $p = \lambda_{HW}/(\lambda_{HW} + \lambda_{SW})$. As expected the accuracy is poor when p is close to 0.67 and improves as the percentage of hardwoods increases (Table 1). However, the probability of correct classification does not exceed 95% until $p \ge 0.8$. Once again, this method probably overestimates the true accuracy, because it assumes that the plot always falls completely within a single forest condition. Also, the assumption of independent Poisson processes is not likely to be correct, because species composition is more likely to be clustered within the hectare and composition tends to vary along the edge of stands (Chen et al. 1992).

Table 1. Probabilities of correctly classifying a hardwood stand based on the binomial distribution.

p ($n_{HW} \ge 22$, $n = 32$)*	p
0.524	0.675
0.644	0.700
0.754	0.725
0.846	0.750
0.914	0.775
0.959	0.800
0.983	0.825
0.995	0.850
0.998	0.875
0.999	0.900

*p, proportion of trees that are hardwoods; n and n_{HW}, the number of trees and number of hardwood trees tallied, respectively.

Discussion and conclusions

The problem of classification is directly related to the spatial scale (Cressie 1993) of the sampling design: that is, the observations of the survey are collected at a certain aggregation (e.g., plot size), and the aggregations are separated by a

certain distance (e.g., separation between plot). Thus, to accurately classify a condition class, the scale of the plot must be well matched to the scale of the condition class. Examples of correctly matched scales would be using plot data to estimate diameter growth increment or using remotely sensed data to determine average stand size. Problems are very likely to occur when the scale of the plot (or other data source) no longer matches the scale of the condition or attribute to be estimated. For example, it would be absurd to assume that diameter growth increment could be estimated from a satellite image.

The estimation of area by condition class is a difficult but necessary endeavor. However, there is a delicate balance between providing users with as much information as can be reasonably derived from the sampling design and producing estimates that are so flawed as to degrade the credibility of the program producing the estimates.

The problem we see with most of the current inventories is that data are collected at the ends of the scale spectrum. These inventories use high-altitude aerial photography or satellite imagery to accurately estimate area of forest and nonforest. Ground plots are used to estimate total basal area and growth. What is lacking is a data source with an intermediate scale, such as low-altitude aerial photography (Czaplewski 1999). These data, in conjunction with the other two data sources, could be used to estimate attributes at intermediate scales such as habitat, forest type, and area of partial cuts. Without an intermediate scale it is unlikely that reasonable estimates of these attributes can be provided.

There are a number of points that are worth mentioning.

(1) Users should be reasonable when specifying the number of possible condition classes. Practical experience and the results of this note suggest that current forest inventories can not accurately divide forest area into more than a handful of distinctly different condition classes.
(2) Estimates for condition classes that are defined on areas that are substantially larger than the plot size or rely on rare attributes should be avoided, particularly when postclassification is used.
(3) Condition classes computed from sample data are often used to determine the accuracy of remotely sensed data and correct area estimates (Card 1982). However, it is incorrect to assume field data always represent the "ground truth."
(4) Increasing the plot size will always reduce the classification error; however, the reduction may not justify the increase in cost, because the error rate may still be unreasonably high. While there are no widely accepted guidelines regarding what constitutes an unreasonably high error rate, we feel that most users would find an error rate greater than 5–15% unsatisfactory.
(5) The field crew should always be allowed to challenge a computed condition class. While the field crew is not infallible, natural populations can be far too complicated to be accurately classified by a model. Both the field crew and computed condition classes should be recorded for diagnostic purposes.
(6) Practitioners should always realize the limitations of the sample design. When the design is inappropriate for the estimation of a particular condition class, it would be wise to derive estimates from an alternative source, high-resolution photography being a logical choice for many of these attributes.

Acknowledgements

The authors would like to thank the Associate Editor and two anonymous reviewers for their insightful comments and suggestions.

References

Card, D.H. 1982. Using known map category marginal frequencies to improve estimates of thematic map accuracy. Photogramm. Eng. Remote Sens. **48**: 431–439.

Chen, J., Franklin, J.F., and Spies, T.A. 1992. Vegetation responses to edge environment in old-growth Douglas-fir forests. Ecol. Appl. **2**: 387–396.

Cressie, N.A.C. 1993. Statistics for spatial data. J. Wiley & Sons, New York.

Czaplewski, R.L. 1999. Multistage remote sensing: toward an annual national inventory. J. For. **97**: 44–49.

Eriksson, M. 2001. An examination of the mapped-plot design using a randomization-based associative paradigm. For. Sci. In press.

Gertner, G.Z. 1990. The sensitivity of measurement error in stand volume estimation. Can. J. For. Res. **20**: 800–804.

Helms, J.A. 1998. Dictionary of forestry. Society of American Foresters, Bethesda, Md.

Lappi, J. 1991. Estimating the distribution of a variable measured with error: stand densities in a forest inventory. Can. J. For. Res. **21**: 469–473.

Scott, C.T., and Bechtold, W.A. 1995. Techniques and computations for mapping plot clusters that straddle stand boundaries. For. Sci. Monogr. **31**: 46–61.

Taylor, H.T., and Karlin, S. 1994 An introduction to stochastic modeling. Academic Press, New York.

William, M.S., and Schreuder, H.T. 1995. Documentation and evaluation of growth and other estimators for the fully mapped design used by FIA: a simulation study. For. Sci. Monogr. **31**: 26–45.