NOTE / NOTE

Estimating forest conversion rates with annual forest inventory data

Paul C. Van Deusen and Francis A. Roesch

Abstract: The rate of land-use conversion from forest to nonforest or natural forest to forest plantation is of interest for forest certification purposes and also as part of the process of assessing forest sustainability. Conversion rates can be estimated from remeasured inventory plots in general, but the emphasis here is on annual inventory data. A new estimator is proposed based on analysis of plot-level variables that indicate when a change in forest condition occurs between inventory remeasurements. A weighted maximum likelihood estimator is derived that incorporates the binomial nature of the indicator variables, mapped plot conditions, and varying remeasurement periods. Example applications demonstrate the utility of the proposed methodology. This approach is broadly useful for estimating the annual rate of change from an initial condition to another condition from annual forest inventory data.

Résumé : Il est important de connaître le taux de conversion de l'utilisation du territoire, de la forêt vers une utilisation non forestière ou de la forêt naturelle vers une plantation forestière, à des fins de certification forestière et dans le cadre du processus d'estimation de la durabilité de la forêt. En général, des taux de conversion peuvent être estimés à partir des mesures répétées de placettes d'inventaire, mais dans cet article l'accent est mis sur les données d'inventaire annuel. Nous proposons un nouvel estimateur basé sur l'analyse de variables à l'échelle de la placette qui indiquent à quel moment un changement des conditions forestières survient entre les mesures répétées d'un inventaire. Nous avons dérivé un estimateur basé sur le maximum de vraisemblance pondéré qui tient compte de la nature binomiale des variables indicatrices, des conditions de la placette estimées à l'aide d'outils cartographiques et des intervalles de temps variables entre deux mesures. Des exemples d'application montrent l'utilité de la méthode proposée. Cette approche est généralement utile pour estimer le taux annuel de changement d'une condition initiale vers une autre condition à partir des données d'inventaire forestier annuel.

[Traduit par la Rédaction]

1. Introduction

Sustainable forestry practices are meant to ensure that forests are managed for an even flow of products and other forest amenities (Brand 1997; Hall 2001; Rametsteiner and Simula 2003; Cashore et al. 2004; Hansen et al. 2006). As part of a certification process, it may be necessary to estimate the rate at which natural forest is being converted to nonforest or forest plantation. In general, it may be important to estimate the rate of change from a particular preferred condition to a less desirable condition, or vice versa.

Remeasured forest inventory plots provide a data source for estimating conversion rates. A portion of remeasured plots will be on natural forest land that is converted to nonforest or plantation. These plots provide the data required to estimate conversion rates, but new statistical methodology is required. We develop a statistical procedure that encompasses the distributional properties of variables that indicate when plots have changed. We also take into consideration the fact that not all plots are remeasured at the same interval and that plot mapping requires differential weighting of the input data. Plot mapping involves partitioning a plot into different conditions. The conditions may be defined by forest type, stand density, dominant tree age, or a number of other characteristics. Mapping sometimes results in plots being partitioned into small slivers. Clearly, a sliver should have less influence on parameter estimates than a full plot. This issue is addressed by weighting the results according to mapped plot size proportions.

A weighted maximum likelihood solution algorithm is developed to provide conversion rate estimates that incorporate the special characteristics of an annual forest inventory. The method is demonstrated with an application to publicly available USDA Forest Service Forest Inventory and Analysis (FIA) data (Bechtold and Patterson 2005; US Department of Agriculture, Forest Service 2005).

2. A statistical model

Consider the following model for an indicator variable,

Received 8 December 2008. Accepted 20 May 2009. Published on the NRC Research Press Web site at cjfr.nrc.ca on 30 September 2009.

P.C. VanDeusen.[1] NCASI, 600 Suffolk Street, Fifth Floor, Lowell, MA 01854, USA.
F.A. Roesch. USDA Forest Service, Southern Research Station, 200 WT Weaver Boulevard, Asheville, NC 28804-3454, USA.

[1]Corresponding author (e-mail: PVanDeusen@ncasi.org).

$I_{i,t}$, for plot i at time t and its relationship to an underlying annual conversion probability, P,

$$[1] \qquad I_{i,t} = P + e_{i,t}$$

where $e_{i,t}$ is an error term that does not follow the normal distribution. The indicator, $I_{i,t}$, changes from 0 to 1 when the status of the condition of interest on the point changes. The status change might be from forest to nonforest or from natural forest to plantation forest. Therefore, P represents the annual conversion rate from one state to another state, conditional on the plot having been in the first state when the inventory started. The developments that follow are based on the assumption that the beginning and ending states of the forest inventory plots are known. No provision is made for plots that might have transitioned out of and then back into the initial state.

A similar equation is presented elsewhere (Van Deusen 2004) for continuous variables where the expected amount of the variable on plot i is proportional to a_i, the proportion of the plot in the condition of interest. For example, if $a_i = 0.5$, then we expect to have half as much volume present as would be present on a full plot in the same condition. However, the $E(I_{i,t})$ is assumed to be independent of the proportion of the plot in a particular condition. The effect of a_i is incorporated in the weighted likelihood function described below.

These methods can also be applied to inventory systems that do no follow the FIA paradigm. Plot mapping is not often used for non-FIA inventories. This discrenpancy can be accounted for by setting $a_i = 1$ in the formulas that follow.

2.1. Distribution of observed indicators

The complete indicator data set for plot i at time t is a sequence of zero and ones that is measured at the beginning and end of the plot remeasurement period, r_i. As such, there will be only two possible observed outcomes. The sequence begins and ends with 0's, or the sequence begins with a 0 and ends with a 1. Call the first sequence or history $h0$, and the other possibility $h1$. Note that $h1$ begins with a 0 and changes to a 1, but it cannot change back to 0. Call the year when it switches from 0 to 1 s_i. It is known that $1 \le s_i \le r_i$, but the actual year when the plot status changed is typically unknown. The probability of observing $h1$ if s_i is known is

$$[2] \qquad p(h1_i|s_i) = P(1-P)^{s_i - 1}$$

and the probability of observing sequence $h0$ is

$$[3] \qquad p(h0_i) = (1-P)^{r_i} = q^{r_i}$$

In practice, the complete indicator data are not observed for a plot. However, it is known whether the plot sequence is an $h0$ history or an $h1$ history. The unconditional probability of an $h1$ history is simply $1 - p(h0)$,

$$[4] \qquad p(h1_i) = 1 - q^{r_i}$$

In the remainder of this paper, the actual plot transition times, s_i, are assumed to be unavailable. However, they could be easily incorporated if they were available.

3. Estimation

The unknown annual change rate, P, can be estimated by finding the value that maximizes the observed data likelihood. The plot condition proportion, a_i, is incorporated as a weight in the likelihood function, thereby allowing plots where a_i is large to have the most influence on the estimate of P. Weighted likelihoods (Wang et al. 2004; Bowater 2004) can be viewed as a generalization of weighted regression and are the vehicle we use to incorporate variable plot sizes for this analysis.

3.1. Likelihood function

The weighted likelihood function of the observed set of histories is

$$[5] \qquad L_{\mathrm{w}} = \prod_{i \in 0} p(h0_i)^{a_i} \prod_{i \in 1} p(h1_i)^{a_i}$$

where the first product in eq. 5 is over the $h0$ histories, and the second product is over the $h1$ histories.

Maximum likelihood procedures are typically applied to the log of the weighted likelihood function

$$[6] \qquad L(P) = \sum_{i \in 0} a_i r_i \log(q) + \sum_{i \in 1} a_i \log(1 - q^{r_i})$$

The maximum likelihood estimate is the value where the Jacobian (first derivative) of the weighted log likelihood equals 0. The Jacobian of eq. 6 is

$$[7] \qquad \frac{\partial L}{\partial P} = g(P) = -\sum_{i \in 0} \frac{a_i r_i}{q} + \sum_{i \in 1} \frac{a_i r_i q^{r_i - 1}}{1 - q^{r_i}}$$

The Hessian (second derivative) of the weighted log likelihood with respect to P is useful for the maximization process and to provide an asymptotic variance estimate. The Hessian of eq. 6 is

$$[8] \qquad \frac{\partial^2 L}{\partial P} = G(P) = -\sum_{i \in 0} \frac{a_i r_i}{q^2} - \sum_{i \in 1} \frac{a_i r_i (r_i - 1) q^{r_i - 2}}{1 - q^{r_i}} - \sum_{i \in 1} a_i \left[\frac{r_i q^{r_i - 1}}{1 - q^{r_i}} \right]^2$$

The estimated variance of $\widehat{P}$ is $-1/G(\widehat{P})$, which is the negative of the inverse Hessian (Agresti 1990) evaluated at $\widehat{P}$.

3.2. Newton–Raphson algorithm

The following Newton–Raphson algorithm provides an estimate of P using the Jacobian and Hessian from eq. 7 and eq. 8:

$$[9] \qquad P^{(1)} = P^{(0)} - \lambda \frac{g(\widehat{P})}{G(\widehat{P})}$$

where λ is a value between 0 and 1 that is used to control the convergence to the maximum likelihood value. The algorithm is applied iteratively. Typically, λ starts at 1.0 and is cut in half if $L(P^{(1)})$ is not greater than $L(P^{(0)})$. The algorithm is deemed to have converged when the change between $L(P^{(1)})$ and $L(P^{(0)})$ is small. The sign of λ can also be periodically changed to ensure that the algorithm does not overshoot the maximum likelihood value. In this case, $0 \le P \le 1$, and the maximum likelihood estimate cannot be allowed outside of the known range.

3.2.1. Starting value

A initial starting value for the Newton–Raphson algorithm is derived by assuming that there is a common remeasurement period, $\bar{r}$. The starting value is found by first replacing r_i in eq. 7 with $\bar{r}$. The fact that eq. 7 equals 0.0 when evaluated at the maximum likelihood solution is then used, $P(\bar{r})$. This results in

$$[10] \qquad P(\bar{r}) = 1 - \left(\frac{S_0}{S_0 + S_1}\right)^{\frac{1}{\bar{r}}}$$

where $S_0 = \sum_{i \in 0} a_i$ and $S_1 = \sum_{i \in 1} a_i$.

The example applications use the following estimate for $\bar{r}$,

$$[11] \qquad \bar{r} = \frac{\sum a_i r_i}{\sum a_i}$$

where the summations are over all plots with relevant conditions. $P(\bar{r})$ was very close to the final estimate for the example applications.

4. Application

The rate of conversion from forest to other uses can be estimated from remeasured FIA plots. In addition, conversion from natural forest to plantation forest may be of interest for forest certification purposes (Rametsteiner and Simula 2003). This interest is due to a concern about continued harvesting in areas where the annual conversion rate is too high. FIA data for Maine, Pennsylvania, and Tennessee are used to estimate conversion rates for each state separately and for the three states combined (Table 1).

Annual percent conversion estimates (Table 1) are derived from the remeasured FIA plots in each state. At this time, for many growth plots the current measurement cannot be matched with the previous measurement for data privacy reasons. This situation complicates making these estimates, because it is not possible to know whether a plot was previously a plantation. However, it is possible to know that a plot is currently in a plantation. The new plantation plot conditions are assumed to be those conditions that are currently plantation, had some nonzero amount of removals since the last measurement, and had most of those removals from species that are not the current plantation species. Therefore, these estimates are tentative and serve mainly to demonstrate the utility of the methodology.

Similar reasoning is used to estimate the proportion of forest area that was converted to nonforest. The recently converted nonforest plot conditions are those conditions that are currently nonforest, but are assumed to have been previously forested if they had a nonzero amount of removals since the previous measurement. As with the approach for estimating plantation conversion, a more accurate estimate could be obtained if the links to all previous plot measurements were available.

The plots used for this analysis are measured over a range of years. The minimum, maximum, median, mean, and first quartile and third quartile of the available inventory years (Table 2) can be used to assess the years over which the conversion estimates apply.

Table 1. Annualized percentage of forest acres (1 acre = 0.404 ha) being converted by state.

	Forest to nonforest	Forest to plantation	All
Maine	0.128 (0.003)	0.044 (0.001)	0.172 (0.003)
Pennsylvania	0.370 (0.012)	0.019 (0.001)	0.389 (0.012)
Tennessee	0.385 (0.011)	0.106 (0.002)	0.491 (0.012)
All	0.310 (0.005)	0.068 (0.001)	0.378 (0.005)

Note: Standard errors are given in parentheses.

Table 2. Most recent inventory year descriptive statistics for the data used in Table 1.

Statistic	Year
Min.	2000
1st quartile	2004
Median	2005
Mean	2005
3rd quartile	2006
Max.	2007

5. Conclusions

A method is developed for estimating the annual rate of change from an initial condition to another condition with annual forest inventory data. For example, the method could be used to estimate the annual rate of conversion from forest to nonforest, or from natural forest to plantation forest.

There is no general closed-form solution for the estimator, but a relatively simple maximum likelihood algorithm is described. There is a closed-form solution for the special case where all plots have the same remeasurement period. This result also provides a good starting value for the general maximum likelihood algorithm. Standard errors can also be estimated, but no closed-form solution is provided.

The new method is applied to FIA data from Maine, Pennsylvania, and Tennessee to estimate annual conversion rates from natural forest to plantation forest and to nonforest. These estimates suggest that the annual conversion rate away from natural forest has been nominal for those states over the past decade. However, some links to previous plot measurements were missing from the data. This shortcoming adds some uncertainty to the conversion rate estimates, and therefore the results (Table 1) should be viewed as rough approximations that mainly demonstrate the value of the method.

Acknowledgements

The comments of the Associate Editor and reviewers were very helpful.

References

Agresti, A. 1990. Categorical data analysis. John Wiley & Sons, New York.

Bechtold, W.A., and Patterson, P.L. Editors. 2005. The enhanced forest inventory and analysis program — national sampling design and estimation procedures. USDA For. Serv. Gen. Tech. Rep. SRS-80.

Bowater, R.J. 2004. A foundational justification for a weighted likelihood approach to inference. Int. Stat. Rev. **72**(3): 307–330.

Brand, D.G. 1997. Criteria and indicators for the conservation and sustainable management of forests: progress to date and future directions. Biomass Bioenergy, **13**(4–5): 247–253. doi:10.1016/S0961-9534(97)10012-5.

Cashore, B., Cashore B.W., Auld, G., and Newsom, D. 2004. Governing through markets: forest certification and the emergence of non-state authority. Yale University Press, New Haven, Conn.

Hall, J.P. 2001. Criteria and indicators of sustainable forest management. Environ. Monit. Assess. **67**(1–2): 109–119. doi:10.1023/A:1006433132539. PMID:11339693.

Hansen, E., Fletcher, R., Cashore, B., and McDermott, C. 2006. Forest certification in North America. Oregon State University, Extension Service Pubication, Corvallis, Ore.

Rametsteiner, E., and Simula, M. 2003. Forest certification — an instrument to promote sustainable forest management? J. Environ. Manage. **67**(1): 87–98. doi:10.1016/S0301-4797(02)00191-3.

US Department of Agriculture, Forest Service. 2005. Forest inventory and analysis national core field guide, volume 1: field data collection procedures for phase 2 plots, version 3.0. U.S. Department of Agriculture, Forest Service, Washington Office. Internal report. On file with: U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis, Rosslyn Plaza, 1620 North Kent Street, Arlington, VA 22209, USA.

Van Deusen, P.C. 2004. Forest inventory estimation with mapped plots. Can. J. For. Res. **34**(2): 493–497. doi:10.1139/x03-209.

Wang, X., van Eeden, C., and Zidek, J.V. 2004. Asymptotic properties of maximum weighted likelihood estimators. J. Statist. Plann. Inference, **119**(1): 37–54. doi:10.1016/S0378-3758(02)00410-X.