The FIA Panel Design and Compatible Estimators for the Components of Change

Francis A. Roesch, Mathematical Statistician, Southern Research Station, USDA Forest Service, Asheville, NC

Abstract—*The FIA annual panel design and its relation to compatible estimation systems for the components of change are discussed. Estimation for the traditional components of growth, as presented by Meyer (1953, Forest Mensuration) is bypassed in favor of a focus on estimation for the discrete analogs to Eriksson's (1995, For. Sci. 41(4):796-822) time invariant redefinitions of the components of change, as evaluated in Roesch (2004), and Roesch (in review). A simple extension of the mixed estimation framework for forest inventory trends presented in Van Deusen (1996, Can. J. For. Res. 26: 1709-1713.) and Van Deusen (1999, Can. J. For. Res. 29(12): 1824-1828) is used to unify estimation of the individual components.*

Introduction

The USDA Forest Service's Forest Inventory and Analysis Units (FIA) report on the condition of forests within the United States and its territories. To this end, the Forest Service recently initiated an annualized forest inventory sampling design in order to improve estimation of both the current resource inventory and changes in the resource. Roesch (in review) gives the theory for a generalized three-dimensional explanation of the current sample frame. In two dimensions, the sample plots are located relative to a systematic triangular grid consisting of *k* mutually-exclusive interpenetrating panels. If the number of sample plots equals *n*, then each panel consists of approximately *n*/*k* plots. Time is the third dimension, incorporated by measuring one panel per year for *k* consecutive years, after which the panel measurement sequence reinitiates. That is, if panel 1 was measured in 1997, it will also be measured in 1997+k, 1997+2k, and so on. Panel 2 would then be measured in 1998, 1998+k, 1998+2k, etc. (Figure 1.) The panels are assigned to previously measured and new plots in a spatially systematic manner.

The sample is drawn from a three-dimensional cube, two dimensions constitute the land area and the third dimension is time. Roesch (in review) describes the two areal dimensions of the design as the joint selection of previously existing and new sample points by a randomly applied triangular grid. It is assumed that the sample points from the entire collection of previous periodic inventories constitute a random sample from the infinite set of points within the geographically-defined population described above. The sample unit is a series of line segments, linear in time. That is, when the time dimension is collapsed down onto the area dimensions, each series of line segments collectively appear as a single point on the area. Each line segment within a series is of an approximate length of 1 day. Individual segments occur every (k +/- 1) year, within each series. Within a sufficiently small segment of time, all points within the land area dimensions of the volume common to each area segment created by the overlapping inclusion areas of all possible subsets of trees occurring on the land area (sensei Roesch et. al. 1993) could be viewed as a temporally-specific sampling unit. However, because these segments change as time progresses, the sample unit appears as a point in the temporally-specific land area dimensions of the volume. That is, if the population is sliced into, say, annual volumes, and then the annual sub-population is viewed from the top, a set of N/k points on the land area base will be observed. The thinner the temporal slice, the smaller the sample per land area of interest and the wider the slice, the fuzzier the segment

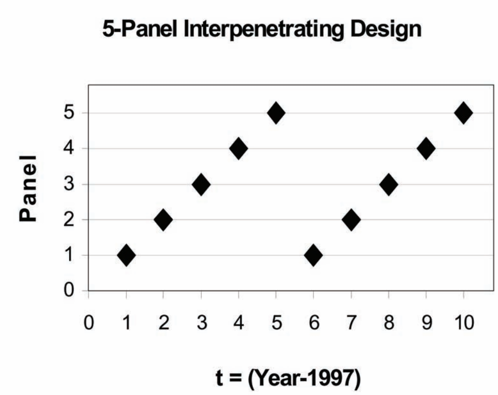

Figure 1. A rotating panel design with 5 panels and a five-year remeasurement period.

boundaries, due to changes in both the land area and the subpopulations of trees within the land area. The temporal slices should be thin enough that all daily line segments (sample units) associated with a point on the land area can be considered exchangeable. We usually (but not always) assume that the assignment of measurement day within panel (slice) is ignorable. When we can make this assumption, the plot measurements support the entire line segment within the panel volume (that is all days constituting the slice.) For most forestry purposes, annual slices will constitute the minimum height that forms a reasonable compromise between temporal specificity and land area generality.

Below, I briefly review the traditional components of growth as presented by Meyer (1953) and subsequently show the intuitively appealing redefinition of the components of change given by Eriksson (1995). I then discuss how discrete analogs to Erickson's components could be estimated from the annual inventory design.

Components of Growth

Meyer (1953) expressed the components of forest growth as:

$$V_2 - V_1 = S + I - M - C, \qquad (1)$$

where:

V_i = the total value at time i, i=1,2,
S = survivor growth,
I = ingrowth,
M = mortality, and
C = cut.

Estimators for the components of growth are considered compatible if they can replace the population parameters in equation (1) without destroying the equality.

Roesch (2004) and Roesch (in review) claim that the weakness of the traditional definitions of the components of growth lies in their inherent dependence on the length of the measurement interval because the definition of the components of growth for the population is dependent upon the time that the relative sample stages are executed. By these definitions, the components are not strictly population parameters to be estimated. Rather, they were a convenient marriage of a population and its sample. Unfortunately, the marriage ceased to remain convenient when the sample design changed in a significant way.

Components of Change

Eriksson (1995) recommended a new set of definitions, labeled the components of change, that were applicable over a temporal continuum, as opposed to the traditional growth component definitions that were not time-additive over multiple period lengths. For example, over a ten-year period in which all plots are measured in years 0, 5, and 10, the sums of the expected values of the estimators of each component over the two intervals (years 0-5 and 5-10) would not equal the expected value of the estimators sans the year 5 measurement. The components of change given by Eriksson (1995) are defined by population attributes and are therefore not sample dependent. Non-additivity is a valid concern due to a fundamental flaw in the original definition of the components of growth. Additionally, the redefinitions become extremely compelling in the realm of annual inventories. The discussion above at least suggests that the original definitions are inadequate for time-interpenetrating sample designs, such as this rotating panel design.

Eriksson (1995) noted that the traditional component of ingrowth consists of both the value of the trees that attain the minimum merchantability limit and the growth in value subsequent to attaining the minimum merchantability limit. Obviously, the later should actually be attributed to survivor growth. Paraphrasing the definitions of Eriksson (1995), live tree growth is the growth in value that occurs on trees after the minimum merchantability limit has been achieved. Entry is the value of trees as they attain the minimum merchantability limit. Mortality is the value of trees as they die, and Cut is the value of trees as they are harvested. Roesch (2004, in review) used discrete analogs to Eriksson's (1995) definitions with a small (1 year) interval length for two reasons, (1) an assumption that 1 year is about the minimum interval length required for the growth signal to overpower measurement error, and (2) to facilitate a tractable temporal partitioning of the observations.

Implicit in the use of the discrete definitions is the assumption that no growth occurs on mortality and cut trees during year of death or harvest. During the time interval of interest, a tree can contribute to multiple components of change. For example, an individual may enter the population, live for two years and then die in between observation instances.

The discrete intervals allow the definition of a set of indicator matrices, one for each component, having one row for each tree in the population during the forest inventory. For example, the indicator matrix for the entry component:

time = P P-1…1 tree

$$\mathbf{I}_{E=} \begin{array}{c} \begin{bmatrix} 0 & 0 \cdots 1 \\ 0 & 0 \ddots .0 \\ 0 & 1 \ddots .0 \\ 1 & 0 \cdots 0 \end{bmatrix} \end{array} \begin{array}{l} 1 \\ \vdots \\ \vdots \\ N \end{array} .$$

In $\mathbf{I_E}$, the first column is for the most recent year of the inventory, and each successive column is one year prior to the previous column. For each tree, all columns are zero except for the year of entry, which would contain a 1. Analogous indicator matrices for tree mortality, $\mathbf{I_M}$, and for tree harvest, $\mathbf{I_C}$, are defined. The indicator matrix for the live category contains a 1 for each year that a tree is alive subsequent to its entry year and prior to its year of harvest or death, and a 0 otherwise:

$$\mathbf{I_L}= \begin{array}{c} \begin{bmatrix} 0 & 0\cdots 0 \\ 1 & 1\ddots .1 \\ 0 & 1\ddots .0 \\ 1 & 1\cdots 1 \end{bmatrix} \end{array} \begin{array}{l} 1 \\ \vdots \\ \vdots \\ N \end{array}$$

The four indicator matrices are of equal dimension and sum to the population indicator matrix, $\mathbf{I_P}=\mathbf{I_E}+\mathbf{I_L}+\mathbf{I_M}+\mathbf{I_C}$. Likewise, we form a value matrix of the same dimension:

$$\mathbf{v}=\begin{bmatrix} v_{1,P} & v_{1,P-1} \cdots & v_{1,1} \\ v_{2,P} & v_{2,P-1} & \\ \vdots & \ddots & \vdots \\ v_{N,P} & \vdots & v_{N,1} \end{bmatrix},$$

and an (Nx1) entry value vector: $\begin{bmatrix} v_1^E \\ v_2^E \\ \vdots \\ \vdots \\ v_N^E \end{bmatrix}$

The period of interest of length t beginning in year h ($h\geq 1, h+t\leq P$) is selected by defining a column vector $\mathbf{Y_{h+t,h}}$ in which rows represent time in reverse annual order. A row contains a 1 for a year of interest and a 0 otherwise.

Additionally, we define the first difference matrix with P-1 columns and P rows, such that the ordered pair $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ appears once and only once in each column and all other entries are zero. The ordered pair occupies the first two positions in the first column, and moves one position down in each subsequent column:

$$\mathbf{d}^1= \begin{array}{c} \begin{array}{cccc} \text{P-1} & \text{P-2}\cdots & \cdots & 1 \end{array} \\ \begin{bmatrix} 1 & 0\cdots 0 & 0 \\ -1 & 1\cdots 0 & 0 \\ 0 & -1\ddots \vdots & \vdots \\ 0 & 0\cdots -1 & 1 \\ 0 & 0\cdots 0 & -1 \end{bmatrix} \end{array} \begin{array}{c} \\ P \\ P-1 \\ \vdots \\ \vdots \\ 1 \end{array}$$

The change components are represented as:

Entry: $\mathbf{E_{h,h+t}}=(\mathbf{I_E}'\mathbf{v^E})'\mathbf{Y_{h+t,h+1}}$,

Live growth: $\mathbf{L_{h,h+t}}=(\mathbf{I_L}'(\mathbf{vd}^1))'\mathbf{Y_{h+t,h+1}}+[\mathbf{I_E}'(\mathbf{v}-\mathbf{v^E})]'\mathbf{Y_{h+t,h+1}}$,

Merchantable Mortality: $\mathbf{M_{h,h+t}}=(\mathbf{I_M}'\mathbf{v})'\mathbf{Y_{h+t-1,h}}$, and

Merchantable Cut: $\mathbf{C_{h,h+t}}=(\mathbf{I_C}'\mathbf{v})'\mathbf{Y_{h+t-1,h}}$.

We also estimate Merchantable Volume at times h+t to h: $\mathbf{V_{h+t,h}}=(\mathbf{I_P}'\mathbf{v})'\mathbf{Y_{h+t,h}}$.

It is important to clearly distinguish the attributes that truly belong to the population from those that are artifacts of the sample design. My suggestions for estimation recognize (1) a measurement interval length (say five years) that is longer than the minimum growth interval of one year, and (2) annually overlapping measurement intervals that result from the rotating panel design.

Mixed Estimator

The components of change definitions are superior to the traditional components of growth for a number of reasons, the time additivity advantage pointed out by Eriksson (1995) being the most obvious. However, that does not necessarily translate into a stronger argument for forcing compatibility of the estimates of the components. Suppose that our sample design consists of five overlapping continuously remeasured panels, one panel measured each year and then remeasured five years hence. The strongest signal for value during any particular year will come from the panel actually measured in that year, while the strongest signal for the live growth component will come from the panel with a remeasurement interval centered on that year. A mixed estimator would seem to be a good way to balance the desire for the "best" estimate for each component with a desire for compatible estimates. The mixed estimator draws strength from overlapping panels. It's a generalized least squares estimator in which model constraints are appended to the data matrices. Van Deusen (1996, 1999, 2000) showed mixed estimators for successive annual estimates. Roesch (2001) tested mixed estimators using both real and simulated data, finding the mixed estimators to perform quite well relative alternative techniques. Roesch (in review) argues for an approach that involves building compatibility constraints for the components of change into a mixed estimator after noting that compatibility requires, for any i and any t, the strict equality: $\hat{V}_{i+t}-\hat{V}_t=\hat{L}_{i,i+t}+\hat{E}_{i,i+t}-\hat{M}_{i,i+t}-\hat{C}_{i,i+t}$.

Consider a model in which the observed midpoint values are used to constrain the component estimates. That is, for k, an integer, and $t\geq k+1$ let

$$\delta_t^k=\begin{cases} (V_{t-(k-1/2)}-V_{t-(k+1/2)}), \text{ if k is odd:} \\ (V_{t-(k-2/2)}-V_{t-(k+2/2)}), \text{ if k is even;} \end{cases}$$

and form a four-column row vector of components such that:

$$\chi_t^k = \begin{cases} \frac{L_{t-k,t}}{k} \mid \frac{E_{t-k,t}}{k} \mid \frac{-M_{t-k,t}}{k} \mid \frac{-C_{t-k,t}}{k}, \text{ if k is odd;} \\ \frac{2L_{t-k,t}}{k} \mid \frac{2E_{t-k,t}}{k} \mid \frac{-2M_{t-k,t}}{k} \mid \frac{-2C_{t-k,t}}{k}, \text{if k is even;} \end{cases}$$

Assume an observation model at each time t: $\delta_t^k = \chi_t^k \beta_t + e_t$,
where β_t is a vector of coefficients with a row for each component and e_t is iid (0, σ^2_t/m_t), and combine it with a reasonably constrained transition model.

Form a vector from the, δ_t^k's, $\Delta = \left[\delta_{k+1}^k, \ldots, \delta_T^k\right]'$, a matrix **X**, from the χ^k_t's, having ((T-k)*4) columns. The vector χ^k_t is placed in row i beginning in column ((i-1)*4)+1). The rest of the elements in the row are zeros. Concatenate successive elements of the column vectors β_t into the column vector $\beta = \begin{bmatrix} \beta_{k+1} \\ \vdots \\ \beta_T \end{bmatrix}$, having ((T-k)*4) rows. Form vectors from the error terms **e**=[e_{k+t},...,e_T]' and ε=[ε_{k+t},...,ε_T]'. Represent equation (3) with:

$$\Delta = \chi\beta + e \tag{4}$$

Represent the covariance matrix of Δ with Σ. The temporal constraints can be re-expressed as:

$$\mathbf{R}\beta = \varepsilon \tag{5}$$

where **R** is the appropriately sized matrix of constraints for the transition model.

Combining the observation model with the transition model:

$$\begin{bmatrix} \Delta \\ 0 \end{bmatrix} = \begin{bmatrix} X \\ R \end{bmatrix} \beta + \begin{bmatrix} e \\ \varepsilon \end{bmatrix} \tag{6}$$

Following Van Deusen (1999), the error vectors **e** and ε are assumed to have independent multivariate normal distributions. ε represents random deviations applied to the beta coefficients, which should be independent of the sampling errors represented by **e**. Theil's (1963, 1971) mixed estimator for β, is:

$$\hat{\beta} = \left[X'\Sigma^{-1}X + \frac{1}{p} R'\Omega^{-1}R \right]^{-1} X'\Sigma^{-1}\Delta$$

The transition covariance matrix Ω is assumed to be a scaled submatrix of Σ, allowing us to adapt a maximum likelihood estimator of the parameter p given by Van Deusen (1999), which determines the strictness of the constraints.

Conclusion

In this paper and the previous ones cited, I attempted to clearly distinguish the effect of scale in the definition of the components of change from the effect on our ability to estimate the components of change at different scales caused by the sample design. Once this distinction has been made, it is clear that the annual estimates of the components of change desired by users of FIA data can only be obtained through the use of models applied to the sample design. That is, there are no strictly design-unbiased estimators for the annual components of change available for this rotating panel design. The mixed estimation technique allows us to use simple models to make well-supported estimates at varying scales by drawing strength from measurements made on temporal "neighbors".

Literature Cited

Eriksson, M. 1995. Compatible and time-additive change component estimators for horizontal-point-sampled data. *Forest Sci*. 41(4):796-822.

Meyer, H.A. 1953. *Forest Mensuration*. Penn Valley Publishers, Inc. State College, Pa. 357 pp.

Roesch F. A. 2001. Mixed Estimation for a Forest Survey Sample Design. In *1999 Proceedings of the Section on Statistics and the Environment: 1-6*, American Statistical Association, Presented at the Joint Statistical Meetings Baltimore, MD August 8-12, 1999.

Roesch, F. A. 2004. Continuous Inventories and the Components of Change. Paper presented at the 2nd International Conference on Forest Measurements and Quantitative Methods and Management & the 2004 Southern Mensurationists Meeting. June 15-18, 2004. Hot Springs, AR.

Roesch, F. A. (in review). Compatible Estimators of the Components of Growth for an Annual Forest Inventory Design.

Roesch, F.A. Jr., E. J. Green, and C. T. Scott. 1993. A test of alternative estimators for volume at time 1 from remeasured horizontal point samples. *Can. J. For. Res*. 23(4): 598-607.

Theil, H. 1963. On the use of incomplete prior information in regression analysis. *JASA* 58:401-414.

Theil, H. 1971. *Principles of Econometrics*. John Wiley & Sons. New York.

Van Deusen, P. C. 1996. Incorporating predictions into an annual forest inventory. *Can. J. For. Res*. 26: 1709-1713.

Van Deusen, P. C. 1999. Modeling trends with annual survey data. *Can. J. For. Res*. 29(12): 1824-1828.

Van Deusen, P.C. 2000. Alternative sampling designs and estimators for annual surveys. In: *Integrated Tools For Natural Resources Inventories In The 21st Century,* USDA Forest Service, North Central Research Station GTR NC-212, Mark Hansen & Tom Burk editors.