COMBINING INVENTORIES OF LAND COVER AND FOREST RESOURCES WITH PREDICTION MODELS AND REMOTELY SENSED DATA

Raymond L. Czaplewski, Ph.D.
USDA Forest Service Research
Rocky Mountain Forest and Range Experiment Station
240 W. Prospect Street
Fort Collins, CO 80526 USA

ABSTRACT

It is difficult to design systems for national and global resource inventory and analysis that efficiently satisfy changing, and increasingly complex objectives. It is proposed that individual inventory, monitoring, modeling, and remote sensing systems be specialized to achieve portions of the objectives. These separate systems can be statistically linked to accomplish all objectives, which might simplify statistical designs and improve efficiency and timeliness. This might be accomplished using an estimation technique known as the Kalman filter, which combines a theoretical or empirical prediction model with several multivariate time series of statistical estimates and remotely sensed data. An example is used to demonstrate one application that monitors land cover.

INTRODUCTION

A principal objective of this conference "is to improve the integration and coordination of national and international inventory and monitoring systems...to support environmental assessments in the 1990s and beyond," and specifically, "identify technical advances...to benefit inventory and monitoring activities on local or global scales." This paper describes one such technical advance: the Kalman filter. The Kalman filter is a statistical method that simultaneously combines different multivariate time series with a theoretical or empirical prediction model. It might provide flexibility, continuity, and efficiency, which are needed for timely input to national and global debates, detection of change, and support of environmental analyses. The Kalman filter might help integrate data collection systems from various organizations, within and among nations, to better address national and global issues.

Cunia (1983, 1985) discuss characteristics of inventories for national and global assessments of renewable natural resources. Global and national inventory systems should achieve diverse objectives, and rapidly respond to changes in multiresource data needs. The sampling design should be flexible and dynamic, which is necessary to accommodate changes in inventory objectives, sampling units, precision standards, and efficiencies of alternative measurement and sampling technologies. Such inventories should be efficient, nearly unbiased, periodically updated, and produce valid error estimates.

National inventory systems should estimate the current state and health of ecosystems (Barnard 1989), including: forest area; current wood volume and future supply; tree growth, removals, and mortality; condition of water, forage, wildlife habitat, and recreational resources; productivity; pest and disease prevalence; and biological diversity (USDA-FS 1987, 1989). These systems should monitor changes and trends to determine effects of environmental policies, expected natural trends, and anthropogenic perturbations.

National inventory systems should be able to statistically test hypotheses concerning observed trends and components of change (Birdsey and Schreuder 1989). They should be able to evaluate theoretical ecosystem models that have gained widespread political attention (Brunner et al. 1989), such as gap phase models that predict major changes in ecosystem distributions in response to global climate changes (Joyce et al. 1989). They should provide timely data for controversial short-term issues and unidentified future issues (Alston 1989, USDA-FS 1989). This might require gathering additional data on existing plots (Cunia 1985), or initiating special-purpose surveys that integrate with existing data collection systems (USDA-FS 1989).

To estimate change, Cunia (1985) recommends remeasurement of permanent plots. Temporary plots should be added to adequately estimate current renewable resources.

Cunia (1983) states that national and global inventories should be statistically independent of land management inventories, but it should be possible to integrate these inventories to efficiently exchange information. Inventory integration can be accomplished through common definitions, and statistical methods that combine several estimates into a single estimate. Although global and national inventories might not produce data required for local land management, they can help determine the type of management and policies that are necessary (Frayer and Mroz 1989).

Cunia (1983, 1985) advocate aerial photography or satellite imagery be used in multi-level sampling designs (e.g., Frayer 1979). This can increase efficiency and timeliness of data for national and global assessments, and provide more frequent data about smaller geographic areas for land management. Incorporation of remotely sensed data in an existing sampling design can synergistically achieve objectives for both national assessments and land management, yet maintain the continuity of existing surveys that have used the same permanent ground plots for years. Although remotely sensed measurements are associated with measurements taken on the ground, the nature and strength of these associations must be accurately quantified using ground procedures, standards, and definitions as truth.

Random measurement error at the first level of multi-level sampling designs should be corrected using calibration data from subsequent level(s). For example, Cunia (1985) shows that volume equations can be considered a second level, where the first level is diameter and height measurements of sample trees. However, error in estimating volume using volume equations is a type of measurement error that is usually ignored. Global and national inventory systems should correct for measurement error that occurs during integration of different levels of multi-level survey designs, different measurement technologies, or different data collection systems.

A SPECIFIC ESTIMATION APPROACH

The Kalman filter is a statistical composite estimator, which might help meet the complex technical demands discussed above. The Kalman filter combines multivariate time series of inventory and monitoring data using a deterministic theoretical prediction model. Therefore, individual sampling or remote sensing systems can efficiently satisfy subsets of objectives, and these specialized systems can be integrated to meet overall objectives.

Composite Estimators

Composite estimators are often used in forestry (Gregoire and Walters 1988), including sampling with partial replacement (e.g., Ware and Cunia 1962). A composite estimator (Fig. 1) combines two or more estimates, each of which is weighted inversely proportional to its variance (or in the multivariate case, its covariance matrix). The weight can be derived using maximum likelihood, minimum variance, or Bayesian theory. If all assumptions are reasonable, error in a composite estimate is less than error in either prior estimate.

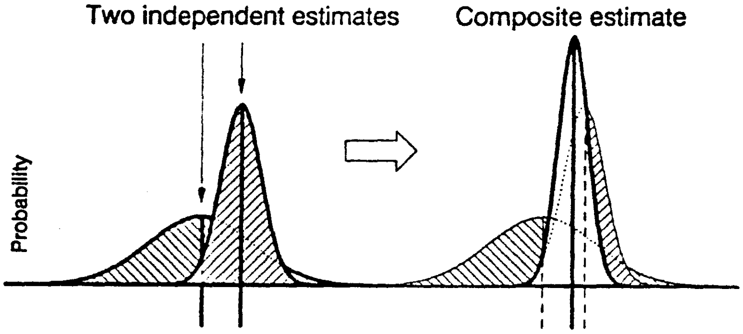

Fig. 1 Probability densities for 2 independent estimates. These are weighted inversely proportional to their variances, and combined into a single, more precise, composite estimate.

Kalman Filter

The Kalman filter can be considered a composite of two independent estimates (Gregoire and Walters 1988), as portrayed in Fig. 2. One estimate is made at time t (e.g., remotely sensed data). The other estimate is made at a previous time t-1, but is updated for expected changes using a theoretical deterministic prediction model, or an empirical historical prediction model. Variance for the updated estimate includes effects of (1) errors in the previous estimate that are propagated over time, and (2) prediction errors from the model that links the previous and current estimates. Dixon and Howitt (1979) describe how the Kalman filter can be applied to sampling with partial replacement in continuous forest inventories.

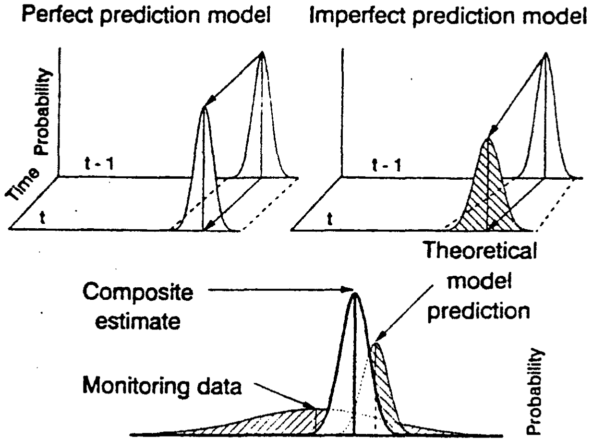

Fig. 2 Probability density for a Kalman estimate that is a composite of measurement data at time *t*, and a prior estimate at time *t*-1, which is updated using a prediction model. Given a perfect prediction model, only the estimation error at *t*-1 is propagated to time *t*. More realistically, the prediction model is imperfect, and a prediction error also occurs. The Kalman filter combines measurements and model predictions into a composite estimate, weighted inversely proportional to their variances.

The Kalman filter is usually applied to a time series of measurements (Fig. 3). With each new measurement, a composite estimate is made, which serves as new initial conditions for the next deterministic prediction from the model (e.g., Fig. 3, year 4).

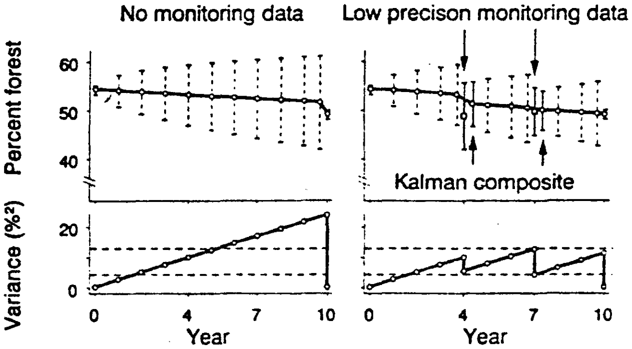

Fig. 3 Kalman estimates for percent forest and approximate 95% confidence intervals. Forest inventories were conducted in years 0 and 10; monitoring data were gathered in years 4 and 7. A time series of imprecise monitoring data can prolong utility of a previous forest inventory.

The Kalman filter is a multivariate vector estimator (Maybeck 1979). It can simultaneously estimate many state variables, such as proportions of vegetation cover types in a nation and average wood volume for each type. Measured rates of change can be statistically combined with rates of change predicted from theoretical econometric models or empirical prediction models, using a multivariate Taylor series approximation. The Kalman filter can model correlated errors among state variables and rate coefficients, correlated errors in predictions from the model, and random errors in the measurement data.

It is possible that two independent estimates disagree or "diverge" in that neither estimate is likely given the other (Fig. 4). Contradictory estimates can be combined, but the resulting composite estimate can be severely biased (Fig. 4). Discrepancies are probably caused by biased estimates of the error distribution (either location or spread) of (1) the measurement at time *t*, or (2) the estimate at time *t*-1 that is updated to time *t* using the prediction model.

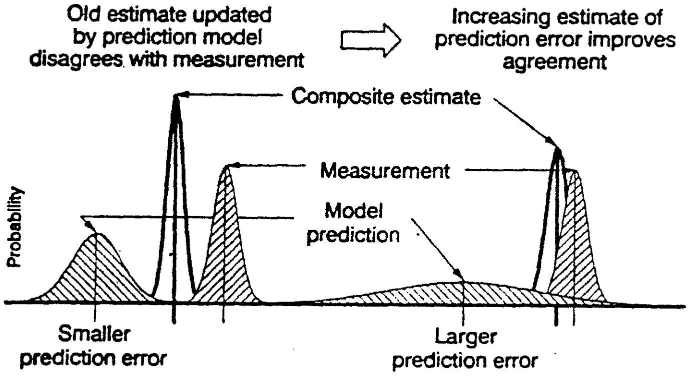

Fig. 4 Expected probability densities for two estimates that disagree. This is caused by bias in the estimated error distribution for the measurement or model prediction. Adaptive filters assume the estimated variance of model prediction error is inaccurate, and change this estimate until the disagreement is within acceptable bounds.

Concerning (1), it is possible that bias exists in the current measurement. For example, calibration experiments are needed to realistically model photointerpretation errors using misclassification probability models, such as "confusion" matrices (Hay 1988). If problems with the measurement model are suspected but cannot be empirically quantified, Meinhold and Singpurwalla (1989) provide an alternative weighting rule that emphasizes the prediction model. Concerning (2), it is very difficult to estimate the variance of errors from the prediction model; accurate estimates would require known differences between model predictions and the true, but unknown, state of the system. As an alternative, adaptive filters modify initial, but inaccurate, variance estimates until disagreements are within acceptable bounds (Fig. 4), often using a time series of residuals (Sorenson 1985). Some Bayesian techniques assume one of the estimates is biased, and choose a weight or shrinking coefficient that minimizes a risk function (e.g., Fienberg and Holland 1973). Both methods change the weight placed on model predictions relative to unbiased measurement data; as accuracy of model predictions increases, the weight placed on model predictions will increase, and as will accuracy of the Kalman filter.

Weight placed on the prediction model in the Kalman filter is a measure of model accuracy and fit to available data. By separately applying the Kalman filter using different prediction models, the weights can be used to rank alternative models. This can be used to test alternative sets of hypotheses, as expressed in theoretical prediction models.

Czaplewski (1989) presents a tutorial example of estimating forest cover using the Kalman filter. The purpose is to offer an intuitive understanding of the Kalman filter; the common sense nature of the Kalman filter is rarely recognized because the Kalman filter is usually introduced using abstract mathematical models and estimation equations.

AN EXAMPLE OF THE KALMAN FILTER FOR MONITORING

Czaplewski et al. (1988) propose use of the Kalman filter to update an existing national inventory (USDA-FS 1987), which is an aggregation of regional inventories, for harvesting, regeneration, land use change, catastrophic events, and natural succession. State variables include number of hectares in 25 classes of forest cover. Estimated initial conditions would be defined by the most recent regional inventories. A theoretical econometric model would predict change in forest area. In addition, rates of cover change since the most recent inventory would be measured directly using photointerpretation of permanent, large photoplots (400 ha), which is similar to methods used in wetland monitoring in the United States (Frayer and Mroz 1989) and ecosystem monitoring in the USSR (Vinogradov 1989). The Kalman filter would integrate past inventory data, a prediction model, and current photo-interpreted data for the large sample plots. This proposal is being evaluated using the entire State of North Carolina, U.S.A. (12×10^6 ha) as a study area. This study is discussed by Catts et al. 1987 and Czaplewski et al. 1987. Cost of 1:12,000 scale color infrared aerial photography of the large sample photoplots and visual interpretation was $0.005 per hectare of the entire study area. This included 411 large photoplots that sample 1.2% of North Carolina; aerial photography was acquired in 7 days.

It is not possible to determine forest type and size class of all stands in a large photoplot (e.g., 10 to 400 ha) using established field procedures and standards; only photointerpretation can measure such a large plot. However, a statistical calibration model can be built using photointerpretation of an independent sample of existing ground plots (0.4 ha) that are accurately registered to comparable aerial photography. This model can correct for bias caused by errors in photointerpreting a large photoplot. Calibration data are analogous to a final level in a multi-level sampling design, similar to Cunia's (1985) characterization of volume equations as the second phase in double sampling.

In the heterogeneous and fine-grained landscapes of southeastern United States, manual interpretation of high resolution aerial photography appears necessary to identify partial harvesting, early regeneration after harvest, early reversion of agricultural lands to timberlands, conversion of natural stands to planted stands, and conversion of timberland to forested urban lands (Czaplewski, et al. 1988). Frayer and Mroz (1989) also emphasize photointerpretation of large sample photoplots because they do not feel current satellite data have acceptable precision for wetland monitoring. However, multispectral satellite data would be useful for detecting large landclearings that are associated with timber harvest and changes in land use. Also, satellite data are more appropriate in global monitoring, or national inventories of developing countries, if adequate aerial photography of sample plots can not be feasibly acquired.

The system proposed by Czaplewski et al. (1988) might closely integrate with existing inventories: existing field standards and procedures would be used to define forest attributes, existing inventory estimates would serve as initial conditions, and calibration models would make remotely sensed estimates unbiased relative to the existing field inventory system. The Kalman filter might reduce inventory cycle time by updating existing regional inventories for area change, and provide national estimates for a single baseline year to access forest resources.

More frequent data might be provided to monitor changes and trends in land use. This might measure effectiveness of government policies, such as economic incentives that influence land management and land use decisions by private citizens. Monitoring data could more frequently test hypotheses contained in the econometric model used in the Kalman filter, which is currently used to nationally assess forest resources in the United States. If the assumptions used for the Kalman filter are reasonable, parameters that describe the distribution of estimation errors for the state vector can be used to perform tests of hypothesis, such as the trend in area of various cover types.

This system might efficiently provide data relevant to certain controversial short-term issues or unidentified future issues. High resolution aerial photographs contain some information on the state and health of forest ecosystems (e.g., canopy and gap structure, visible crown condition, pest and disease prevalence, wildlife habitat quality, biological diversity, etc.). A time series of photography for permanent plots preserves sample data on conditions and changes at the stand and landscape levels. Expense of measuring new data from the aerial photography could be deferred until data needs for new issues are clearly identified.

Satellite data can provide thematic maps that are valuable for land management and extension programs; this is not possible using sample plots alone. However, large photoplots could be integrated with satellite data; they are readily registered to satellite data, and photointerpreted cover type maps for photoplots can be used to label or train digital classifications of multispectral satellite data. Large photoplots could also be used to fit statistical calibration models, which correct misclassification bias in satellite area estimates. Calibrated area estimates from satellite data would be directly compatible with the definitions, procedures, and standards used for photointerpretation of high-resolution aerial photography. This could integrate national inventories and management inventories, while maintaining the mutual independence needed to accomplish their different objectives. Also, synergistic benefits might be provided: thematic maps for management, and calibrated area estimates that increase efficiency and timeliness of national inventory estimates.

Calibration models could predict true classifications of field plots using imperfect classifications of satellite pixels. However, misregistration error can be confounded with misclassification error, and substantially affect calibration models. To minimize effects of misregistration, a calibration model could be developed between clusters of satellite pixels and interpreted cover type maps of large photoplots; a second calibration model could be developed between photointerpretation of permanent 0.4 ha plots and field classifications of the same plots; and these two linear calibration matrix models could be multiplied together. This would statistically correct estimates from satellite data, and make them directly compatible with existing definitions, standards, and field procedures, using plots that better match in scale and reduce registration error.

The Kalman filter might statistically integrate forest inventories with other data collection systems such as annual pest surveys; forest health surveys; midcycle updates; environmental surveys, censuses, and inventories conducted by other agencies within a nation; data systems from other nations; and global monitoring systems (e.g., Singh 1989, Vinogradov 1989). Integration would require careful calibration experiments using accurately registered plots or sites, much like the calibration experiments using photointerpreted and field classifications of small forest inventory ground plots, or satellite and photointerpreted measurements of large photoplots. Calibration models would permit unbiased statistical compatibility among data collection systems that use different definitions, standards, and procedures. However, sample size of calibration plots must be adequate to build reliable empirical calibration models. Integration of several data collection systems could provide synergistic benefits. For example, better growth and mortality estimates from annual forest health surveys would improve current volume estimates from the Kalman filter, and increase efficiency and timeliness of national inventory estimates.

CONCLUSIONS

The Kalman filter might reliably combine various time series of independent forest inventories, updates, or monitoring estimates, using a theoretical prediction model of expected change in forest condition. This might achieve many of the diverse and changing objectives of national data collection systems, while reducing complexity of individual subsystems. Integration of data collection systems, within and among nations, could improve global assessments in the 1990s and beyond. The synergism provided by statistically combining reliable data from different sources might improve efficiency and precision of individual data collection systems.

ACKNOWLEDGEMENTS

The monitoring system for updating existing inventories for changes in land use, harvesting, and regeneration, was originally conceived through a team effort; the team included Dr. Hans T. Schreuder, Mr. Glenn P. Catts, Mr. Paul W. Snook, and the author. Dr. Charles C. Van Sickle originally proposed that a national system be developed to update existing estimates to a common baseline year for national assessments. Mr. Noel D. Cost, and especially Mr. Catts, have expended considerable energy in solving the many practical problems that arise when a serious commitment is made to integrate photointerpreted data with an existing forest inventory system. Mr. Kurt Teuber, and especially Mr. Snook, have spent much effort in integrating satellite data with photointerpreted polygons on 400-ha plots. Dr. Ralph J. Alig has provided helpful advice in integrating econometric models for area change into a statistical estimator.

REFERENCES

Alston, R.M., 1989, Removing Institutional and Political Barriers to Improve Global Monitoring and Assessments: Proceedings of international symposium Global Natural Resource Monitoring and Assessments: Preparing for the 21st Century, Venice, Italy, September 24-30, 1989. In press.

Barnard, J.E., 1989, Environmental health concerns: a role for forest inventory and monitoring: Proceedings of international symposium State of the Art Methodology of Forest Inventory, Syracuse, N.Y., July 30 to August 5, 1989. In press.

Birdsey, R. and Schreuder, H. 1989, Forest Change - An Assessment of Nationwide Forest Inventory Methods: Proceedings of international symposium State of the Art Methodology of Forest Inventory, Syracuse, N.Y., July 30 to August 5, 1989. In press.

Brunner, B., Essmann, H., Krott, M., Lunnan, A., and Vehkamaki, S., 1989, Policy Aspects of Natural Resource Monitoring, a Comparative Analysis of Monitoring Forest Die-back in Austria, Germany, Finland, Norway and Switzerland: Proceedings of international symposium Global Natural Resource Monitoring and Assessments: Preparing for the 21st Century, Venice, Italy, September 24-30, 1989. In press.

Catts, G.P., Cost, N.D., Czaplewski, R.L., and Snook, P.W., 1987, Preliminary results from a method to update timber resource statistics in North Carolina: In proceedings 11th Biennial Workshop on Color Aerial Photography in the Plant Sciences, Weslaco, Texas, U.S.A. pp.39-52.

Cunia, T., 1983, Forest inventories as the basis for a continuous monitoring of forest biomass resources: In proceedings of NATO seminar Biomass Utilization, Alcabidecke, Portugal, 1982. W. A. Cote', Editor. Plenum Press, NY, NY.

Cunia, T., 1985, Main Objectives and Desirable Characteristics of National Forest Inventory Systems: Proceedings of international symposium on European National Forest Inventory Systems, Frieburg, Germany, June 10-14, 1985.

Czaplewski, R.L., Catts, G.P., and Snook, P.W., 1987, National land cover monitoring using large, permanent photoplots: In proceedings Land and Resource Evaluation for National Planning in the Tropics, Chetumal, Mexico, pp. 197-202. USDA Forest Service GTR WO-39, Washington, D.C.

Czaplewski, R.L., Alig, R.J., and Cost, N.D., 1988, Monitoring land/forest cover using the Kalman filter, a proposal: In proceedings Forest Growth Modelling and Prediction, Minneapolis, MN, U.S.A. pp. 1089-1096. USDA Forest Service GTR NC-120, Saint Paul, MN. 1153p.

Czaplewski, R.L., 1989, Kalman filter to update forest cover estimates: Proceedings of international symposium State of the Art Methodology of Forest Inventory, Syracuse, N.Y., July 30 to August 5, 1989. In press.

Dixon, B.L., and Howitt, R.E., 1979, Continuous forest inventory using a linear filter: Forest Science, vol. 25, pp. 675-689.

Fienberg, S.E., and Holland, P.W., 1973, Simultaneous estimation of multinomial cell probabilities: Journal of the American Statistical Association, vol. 68, pp. 683-691.

Frayer, W.E., 1979, Multi-level Sampling Designs for Resource Inventories. Department of Forest and Wood Sciences, Colorado State University, Fort Collins, Colorado, U.S.A. 113pp.

Frayer, W.E. and Mroz, G., 1989, Importance of global data in monitoring wetlands: Proceedings of international symposium Global Natural Resource Monitoring and Assessments: Preparing for the 21st Century, Venice, Italy, September 24-30, 1989. In press.

Gregoire, T.G. and Walters, D.K., 1988, Composite vector estimators derived by weighting inversely proportional to variance: Canadian Journal of Forest Research, vol. 18, pp. 282-284.

Hay, A.M., 1988, The derivation of global estimates from a confusion matrix: International Journal of Remote Sensing, vol. 9, pp. 1395-1398.

Joyce, L.A., Fosberg, M.T., and Comanor, J.M., 1989, Climate Change and America's Forests, USDA Forest Service, GTR RM-00, Fort Collins, Colorado, U.S.A. 32pp.

Maybeck, P.S., 1979, Stochastic Models, Estimation and Control, Vol. 1. Academic Press, New York. 423pp.

Meinhold, R.J. and Singpurwalla, N.D., 1989, Robustification of Kalman filter models: Journal of the American Statistical Association, vol. 84, pp. 479-486.

Singh, K.D., 1989, The conceptual framework of FAO Forest Resources Assessment 1990 (within the framework of Tropical Forestry Action Plan): Proceedings of international symposium Global Natural Resource Monitoring and Assessments: Preparing for the 21st Century, Venice, Italy, September 24-30, 1989. In press.

Sorenson, H.W., 1985, Kalman filtering: Theory and Applications: IEEE Press Inc., New York. 457pp.

USDA-FS, 1987, Forest Service Resource Inventory: An Overview, USDA Forest Service, Washington, D.C. 29pp.

USDA-FS, 1989, Draft 1990 RPA Program, USDA Forest Service, Washington, D.C. 335pp.

Vinogradov, B.V., 1989, Aerospace monitoring of the biosphere state, a global approach: Proceedings of international symposium Global Natural Resource Monitoring and Assessments: Preparing for the 21st Century, Venice, Italy, September 24-30, 1989. In press.

Ware, K.D. and Cunia, T., 1962, Continuous forest inventory with partial replacement: Forest Science Monograph No. 3.