# MONITORING LAND/FOREST COVER USING THE KALMAN FILTER: A PROPOSAL

Raymond L. Czaplewski, Ralph J. Alig and Noel D. Cost[1]

ABSTRACT. Although growth and yield models have been used to update forest inventories for large regions, such models poorly predict cover changes from land use conversions, regeneration, and harvest. These changes could be monitored directly for large areas using remote sensing, which can be expensive, or estimates made by agricultural agencies, which are not detailed for condition of timberlands. The Kalman filter, which is a flexible statistical estimator, might increase statistical efficiency and produce annual estimates of cover by combining such direct monitoring with past knowledge (i.e., previous forest inventory) and expected change (i.e., model for annual change in land cover). This paper describes a potentially useful application of this topical estimator and presents specific proposals for practical methods.

## INTRODUCTION

Inventories performed by the USDA Forest Service, Forest Inventory and Analysis (FIA) Projects (Cost and Knight, 1983; Frayer and Beltz, 1986) are adequate for their intended applications but are 1 to 10 years old. Models have been used to update (Smith and Hahn 1986) and reduce cost (MacLean 1981) of FIA inventories. However, most models do not predict changes in area caused by harvesting, regeneration, and conversions in land use. This paper describes the Kalman filter, and discusses how it might be used to monitor land cover-ownership (Table 1). The Kalman filter includes a deterministic model for change in a multivariate system over time, and a measurement model for a time series of monitoring data. Estimates from the most current FIA inventory could serve as initial conditions (i.e., time $t$=0). Subsequently ($t$>0), remotely sensed data and areal estimates of cover from other agencies would serve as monitoring data. Areal models of annual change in land cover, possibly based on econometric methods, would be used in the Kalman filter to combine this time series of data into a single, dynamic estimator. However, estimation errors would eventually build to unacceptable levels, and a new FIA inventory using traditional methods would be required. Until then, efficient annual estimates of cover would be produced for resource planning at the Federal and State levels.

## KALMAN FILTER

The Kalman filter (Kalman 1960) was originally developed for aerospace control problems. Its derivation is given by Lee (1964); Jazwinski (1970), Harrison and Stevens (1976), Maybeck (1979), and Sorenson (1985) provide contemporary descriptions. Dixon and Howitt (1979) proposed it for continuous forest inventory, and Visser (1986) and VanDeussen (1987) have used it in dendro-ecological studies. It is a vector analog of the familiar estimator used

---

[1]Mathematical Statistician, USDA Forest Service, Rocky Mountain Forest and Range Experiment Station, 240 West Prospect Street, Fort Collins, CO, 80526; Project Leader, USDA Forest Service, Forestry Sciences Laboratory, Box 12254, Research Triangle Park, NC, 27709; and Research Forester, USDA Forest Service, Southeastern Forest Experiment Station, 200 Weaver Boulevard, Asheville, NC, 28802, USA.

Presented at the IUFRO Forest Growth Modelling and Prediction Conference, Minneapolis, MN, August 24-28, 1987.

TABLE 1. Land classification system

| Land Ownership | Land Cover | Forest Condition | |
|---|---|---|---|
| | | FOREST TYPE | SIZE CLASS |
| PUBLIC (censused) | FOREST | | |
| National Forest | Timberland | Pine | Sawtimber |
| Other federal | Woodland | Oak-pine | Poletimber |
| BLM | | Upland hardwood | Seedling-sapling |
| Indian | | Bottomland hardwood | Nonstocked |
| Misc. Federal | | | |
| State | | | ORIGIN |
| County & Municipal | | | |
| | | | Natural |
| FOREST INDUSTRY (censused) | | | Planted |
| Owned | NONFOREST | | |
| Leased | | | |
| | Cropland | | |
| PRIVATE (not censused) | Grassland | | |
| | Barrenland | | |
| Farmer | Urban | | |
| Other private | Water | | |
| Other corporate | | | |

in sampling with partial replacement, in which components are weighted in inverse proportion to their variances (Gregoire, personal communications).

The Kalman filter is given by model (1), which represents the linear change ($\underline{\Phi}_t$) in a multivariate system ($\underline{x}$) between time $t-1$ and $t$, and model (2), which describes the linear relationship ($\underline{H}_t$) between measurements of the system ($\underline{y}_t$) and the true, but unknown state of the system ($\underline{x}_t$).

$$\underline{x}_t = \underline{\Phi}_t \, \underline{x}_{t-1} + \underline{w}_t \qquad (1)$$

$$\underline{y}_t = \underline{H}_t \, \underline{x}_t + \underline{v}_t \qquad (2)$$

$\underline{y}_t$ can represent a time series of vectors for independent sources of measurement data. The vector of prediction errors $\underline{w}_t$ has a zero mean vector and covariance matrix $\underline{Q}_t$. The vector of measurement errors $\underline{v}_t$ has a zero mean vector and covariance matrix $\underline{R}_t$. $\underline{w}_t$ and $\underline{v}_t$ are vectors of random variables; their exact values are unknown, but estimates can be made for $\underline{Q}_t$ and $\underline{R}_t$, and assumptions made regarding their true distributions.

The state vector ($\underline{x}_t$) might represent hectares in each of $n$ land/forest cover categories (Table 1) at time $t$. The true, but unknown, state vector ($\underline{x}_t$) equals an estimated state at time $t$ ($\hat{\underline{x}}_t$) plus a vector of estimation errors ($\underline{\epsilon}_t$).

$$\underline{x}_t = \hat{\underline{x}}_t + \underline{\epsilon}_t$$

$\underline{\epsilon}_t$ has a zero expected value and covariance matrix $\underline{P}_t$. If error vectors $\underline{w}_t$, $\underline{v}_t$ and $\underline{\epsilon}_0$ are multivariate normal for all values of $t$, then $\underline{\epsilon}_t$ will be multivariate normal, and confidence intervals for $\underline{x}_t$ are readily computed using $\underline{P}_t$.

The state vector at time $t$ is estimated using the state vector at a prior time ($t-1$) and expected change between $t-1$ and $t$

$$\hat{\underline{x}}_t = \underline{\Phi}_t \hat{\underline{x}}_{t-1} \quad (4)$$

The covariance matrix for propagated estimation errors ($\underline{P}_t$) is a function of the previous covariance matrix for estimation errors at time t-1 ($\underline{P}_{t-1}$), the expected change in the state of the system ($\underline{\Phi}_t$) from model 1, and the covariance matrix for prediction errors between t and t-1 ($\underline{Q}_t$).

$$\underline{P}_t = \underline{\Phi}_t \underline{P}_{t-1} \underline{\Phi}_t^T + \underline{Q}_t. \quad (5)$$

This requires estimates of initial conditions $\underline{x}_0$ and $\underline{P}_0$.

Measurements at time t can improve the updated estimate. Measurements might be estimates of hectares in m land/forest cover types from remote sensing or from other agencies. The composite estimator, which combines prior knowledge (model 1) with a new vector of measurements (model 2), is

$$\hat{\underline{x}}_t^* = [\underline{I} - \underline{G}_t \underline{H}_t] \hat{\underline{x}}_t + \underline{G}_t \underline{y}_t, \quad (6)$$

where $\underline{G}_t$ is computed to produce the minimum variance estimate (see Appendix). This estimator has less variance in estimation error ($\underline{P}_t^*$) than the prior estimate ($\underline{P}_t$) because

$$\underline{P}_t^* = \underline{P}_t (\underline{I} - \underline{G}_t \underline{H}_t). \quad (7)$$

where ($\underline{G}_t \underline{H}_t$) has values from zero to one.

## FIA DATA

FIA units provide the most detailed, reliable, statistical estimates of forest cover for the entire United States. Every 5 to 10 years, areal estimates are made for condition of forestland in multi-county physiographic regions (i.e., Inventory Units) in each State. The following describes FIA methods used in the southeastern United States, with specific examples from the 1984 inventory of North Carolina (NC) from Sheffield and Knight (1986).

First, complete aerial photo coverage is used to estimate area of timberland in each county; it is 1:40,000-scale black and white, or 1:58,000-scale color infrared (CIR) prints, and can be up to 10 years old. Temporary plots, nominally 12 ha, are systematically located on the photographs, and the proportion of timberland for each plot is photointerpreted. There were 91,765 photo samples (1 per 138 ha) in the 1984 inventory for NC. Photointerpretation is verified on the ground for a subsample (e.g., 8,123 out of 91,765 photo samples in NC). Estimates of total timberland are adjusted for misinterpretations and changes since the date of photography using double sampling with regression.

Second, exact area of timberland for certain ownerships (Table 1) is censused for each county using information from forest-product industries and government agencies. It is impractical to census area of other timberlands managed by non-industrial, private forest (NIPF) landowners (typically 75% of the timberland in the Southeast). Third, permanent ground plots, nominally 0.4-ha, are systematically installed in timberland by field crews (5,355 plots in 1984 for NC). Field crews remeasure forested permanent plots (4,878 of the 5,355 in NC) for estimates of growth, removals, and mortality. Ground data are used to classify each plot using the system in Table 1. In addition, a few temporary

0.4-ha ownership plots are measured by field crews to describe area in ownerships that do not contain a permanent ground plot in a county. This procedure produces reliable estimates of area for classes in Table 1 within each Inventory Unit.

## REMOTELY SENSED DATA

The following assessment of remote sensing strategies for the southern United States is based on Catts et al. (1987) and Czaplewski et al. (1987). It emphasizes statistical estimates for large regions rather than in-place vegetation maps required for land management or extension efforts. Three sources are considered: digital satellite data, high-altitude aerial photography, and medium- or large-scale aerial photography.

Forest type and some classes of land cover could be interpreted using digital satellite data. However, seedling stands would be confused with cropland and grassland. Accurate monitoring of regeneration is critical for forecasting future timber supplies. Ownership, size class (poletimber vs. sawtimber), stand origin, old clearcuts, and residential urbanland would be very difficult to interpret with digital satellite data. Small-scale, high-altitude aerial photography could also be used to reliably interpret some categories in Table 1. Seedling stands would be confused with cropland, grassland, and shrubland. Such photography is unsuitable for interpreting ownership and stand origin. These data are available every 5 to 10 years, but would include a wide range of dates for a single Inventory Unit. Alternatively, high resolution, medium- or large-scale aerial photography could be used to interpret all categories in Table 1 except for ownership. It is the only source of new monitoring data which could be acquired directly by FIA, and not subject to priorities and policies of other organizations. Regeneration and type conversions could be detected after 2 years. However, such photography is expensive, and would have to be used for sampling photoplots. Medium-scale aerial photography could be acquired annually, but a 3- to 5-year cycle would be more realistic.

## COVER ESTIMATES FROM OTHER AGENCIES

Government agencies routinely gather information about forestlands by mail, telephone canvass, and ground inventories. However, these rarely cover all ownerships, and vary greatly in detail on forest condition, frequency, timeliness, reliability, and definitions. The following lists periodic estimates which might be combined with FIA data to monitor land cover.

The National Resource Inventory (NRI) provides estimates by State of land use/land cover for nonfederal lands every 5 years (USDA Soil Conservation Service 1982). The Census of Agriculture also estimates area of farmland and other land uses by State and county every 5 years (USDC Bureau of the Census 1981). The National Agricultural Statistical Service (NASS) produces annual estimates of area of cropland and total forestland at the State level (USDA Statistical Reporting Service 1983). NASS has produced special estimates of land cover and forest type compatible with FIA definitions (Table 1) for Kansas, Arkansas, and Missouri (May et al. 1986), and a survey that examined harvest and reforestation decisions of NIPF landowners (Fecso et al. 1982). Estimates from these agencies vary from those made by FIA because of differences in definition, timing, sampling, and procedure. However, they might provide a measure of trends in forested area if differences are quantified.

Each year the Forest Service estimates area planted or seeded with trees by State and ownership (USDA Forest Service 1987). Data come from many sources,

and vary considerably in quality and accuracy. The area of successful artificial regeneration is overestimated because failures are not considered, and some planted areas were already stocked using FIA definitions. Trends might be monitored with these data if adjusted annually for seedling survival, perhaps using a climatological index.

Information on timber products output (TPO) describes trends and current levels of tree harvest. Most States in the Southeast canvass primary wood-using plants every 1 or 2 years to determine the volume produced and geographic source of all industrial roundwood products (Tansey 1984). Similar information is available for the Midsouth (McWilliams 1986). However, the type of material from which the products are derived is not specified. Some material is not considered growing stock by FIA. TPO does not include logging residue and growing stock removed during conversion in land use.

## PREDICTIVE MODELS FOR COVER

Econometric models have been used to predict change in area for land use/forest ownership using economic, social, and policy variables. White and Fleming (1980), Stoll et al. (1984), Brooks (1985), and Alig (1986) all rely on land rent theory. Brooks (1985) also uses models of land allocation to perennial crops as in Houck and Subotnick (1969), French and Matthews (1971), and Minami et al. (1979). Alig (1986) takes advantage of exact prior information to reflect zero sum changes for the total land base. At least three groups of input data are usually required: (1) cross-section and time series data, including FIA data on area by forest ownership (needed to estimate model parameters); (2) current area estimate of the existing forest land base, and; (3) projections or estimates of economic driving variables (necessary to predict changes in land use) such as population estimates and per capita income.

Econometric models have been linked to models of forest succession and cover type dynamics (Alig 1985). These require: (1) current data on the area in each forest type by ownership; (2) probability of a primary disturbance for each type of ownership and forest type; and (3) the conditional probabilities of forest type shifts in response to a primary disturbance (Alig and Wyant 1985). Regional models for the Nation are being developed to predict area changes on a 10-year time increment for long-range planning. Annual projections are possible if an annual time series of estimated economic driving variables is available.

## IMPLEMENTING THE KALMAN FILTER

Initial conditions for forestland monitoring using the Kalman filter can be determined from the most current FIA inventory for each Inventory Unit. Areal estimates for all categories of land in Table 1 at time $t=0$ are available for $\hat{\underline{x}}_0$ in equation (4). However, the covariance matrix for estimation error ($\underline{P}_0$ in equation 5) has never been computed because of the complex design used by FIA, which includes systematic sampling, census, and regression data. This problem is being investigated in a related study.

## STATE TRANSITION MODEL AND PREDICTION ERROR

Econometric models for an annual time increment, and annual estimates for socioeconomic driving variables, are presently unavailable. Therefore, a more empirical approach is proposed until such are available. Assuming annual transition probabilities among cover types are constant, an estimated state

transition matrix ($\hat{\underline{\Phi}}_k$) is readily computed from a previous FIA inventory ($t$=-$k$) to the most recent FIA inventory (t=0) using field classification of permanent plots at both times. For example, 4,878 0.4-ha FIA forested plots were classified in both 1974 ($t$=-10) and 1984 ($t$=0) in NC ($k$=10 years). The $ij$-th element of $\hat{\underline{\Phi}}_k$ is

$$(\hat{\Phi}_{ij})_k = \frac{N(A_0=i \,|\, A_{-k}=j)}{N(A_{-k}=j)}, \qquad (8)$$

where $N(A_0=i \,|\, A_{-k}=j)$ is the number of plots classified as category $i$ at $t$=0 (1984 in NC), given they were classified as category $j$ at $t$=-$k$ (1974 in NC), and $N(A_{-k}=j)$ is the total number of plots classified as category $j$ at $t$=-k.

Therefore, the columns in $\hat{\underline{\Phi}}_k$ are vectors of independent, estimated, conditional probabilities (see Appendix). Because these probabilities are for $k$ years, they can be converted to a time-invariant transition matrix for one year using

$$\hat{\underline{\Phi}}_1 = (\hat{\underline{\Phi}}_k)^{1/k}, \qquad (9)$$

which is derived in the Appendix from equation (1).

If prediction error is caused solely by sampling error in the estimated transition probabilities for $k$ years ($\hat{\underline{\Phi}}_k$), then the covariance matrix for prediction error ($\hat{\underline{Q}}_k$) can be determined using equation (1), the multinominal distribution for each column vector in $\hat{\underline{\Phi}}_k$, and the estimated area variables at time -$k$ and 0 (see Appendix). The covariance matrix for prediction errors in annual estimates ($\hat{\underline{Q}}_1$) is computed from the corresponding matrix for k years using the following (see Appendix)

$$\hat{\underline{Q}}_1 = [\hat{\underline{\Phi}}_1^{k-1} + \hat{\underline{\Phi}}_1^{k-2} + \ldots + \hat{\underline{\Phi}}_1 + \underline{I}]^{-1} \hat{\underline{Q}}_k [(\hat{\underline{\Phi}}_1^{k-1} + \hat{\underline{\Phi}}_1^{k-2} + \ldots + \hat{\underline{\Phi}}_1 + \underline{I})^T]^{-1}. \qquad (10)$$

This is an underestimate of prediction error because it ignores change in transition probabilities over time. Realistic estimates of $\underline{Q}_1$ are needed so that proper weight is placed on model predictions relative to measurements. An objective method for scalar inflating $\underline{Q}_1$ to better represent all prediction errors is in the Discussion and Appendix. In the future, this oversimplified model will be replaced with a better econometric model.

## COVER MEASUREMENTS USING REMOTE SENSING

Satellite imagery could be periodically classified by FIA for multi-scene areas to map land cover and produce area estimates. In other countries, photointerpreted stand maps might serve the same purpose (e.g., Bonner and Magnussen 1987). Some categories in Table 1 could not be accurately classified (e.g., ownership), and $\underline{H}_t$ would sum these into a less detailed classification system for satellite data (i.e., $m<n$). The measurement relationship matrix ($\underline{H}_t$) could be further determined from an analysis of classification accuracy as commonly done in remote sensing studies (Prisley and Smith 1987). There would

be no sampling error, but error in estimating $\underline{H}_t$ would be included in the measurement error ($\underline{R}_t$) using a multivariate regression estimator or conditional probabilities in a "classification error matrix" (see Appendix).

FIA could obtain aerial photography of photoplots, and photointerpret land cover. Assuming negligible interpretation error, $\underline{H}_t$ would contain ones and zeros, representing the logical association between Table 1 and a less detailed classification system for FIA photointerpretation (e.g., ownership could not be photointerpreted). Sampling error for the photoplots would be used for $\underline{R}_t$. However, classification error can be substantial, and $\underline{H}_t$ would have to be further estimated using ground reference data. The error in estimating $\underline{H}_t$ would be added to the sampling error to compute total measurement error.

COVER MEASUREMENTS FROM OTHER AGENCIES

Extensive estimates of land cover, made annually by NASS and at 5-year intervals by other Federal agencies, have an estimated covariance matrix for sampling errors. This matrix is needed for $\underline{R}_t$ to compute $\underline{G}_t$ in equations (6) and (7). However, these agencies do not produce a detailed classification of forestland. Therefore, the measurement matrix ($\underline{H}_t$) in model (2) must sum many state variables into a few categories. If definitions are compatible, $\underline{H}_t$ will contain zeros or ones, depending upon the logical relationship between land classification systems. However, effects of different definitions and procedures must be quantified; these could be estimated using multivariate regression and corresponding estimates from many geographic areas (e.g., counties or Inventory Units). Alternatively, FIA and another land classification systems might be applied to a randomized set of small plots. The resulting contingency table could be treated as a matrix of conditional probabilities, similar to $\hat{\underline{\Phi}}_k$ and $\hat{\underline{Q}}_k$ (see Appendix). Again, error in estimating $\underline{H}_t$ would have to be added to any sampling error to estimate total measurement error.

## VERIFICATION

Many assumptions, approximations, numerical calculations, and estimates are required to apply the Kalman filter. If any are inaccurate, then estimates of land cover or their confidence intervals can be biased, producing serious consequences if undetected. However, the actual distribution of residuals will likely deviate from their expected distribution in presence of significant inaccuracies. The entire time series of multivariate measurement residuals can be standardized and orthogonalized; goodness-of-fit tests could detect significant departures from their expected distribution, and tests of hypotheses used to investigate independence (Appendix). This would identify existence of a problem, but problem isolation and solution could be difficult. Comparison of estimates from the Kalman filter to estimates from a new FIA inventory is the most rigorous test of reliability. This would be possible every 5 to 10 years given present schedules. In a production system, detection of a failure in the filter for one State might suggest potential problems for other States. Serious scrutiny of residuals is necessary for responsible application of the Kalman filter.

## DISCUSSION

The Kalman filter has potential for frequent, efficient monitoring of land cover between FIA inventories; however, it also presents formidable challenges. An unexpected distribution of standardized residuals might indicate a poor assumption. The assumption most likely incorrect is time-invariant transition probabilities. This might be remedied using econometric models for cultural changes, and growth and yield models for vegetative succession. Also, the covariance matrix for prediction error, which considers only the sampling error in estimating time-invariant probabilities, could be inflated by a scalar value to represent additional prediction errors caused by changes in transition probabilities. The scalar value might be estimated to maximize the goodness-of-fit of standardized residuals to their expected distribution (Czaplewski 1986). Also, change in transition probabilities over time might be observed directly. Change in cover on permanent photoplots could be efficiently interpreted using a time-series of comparable aerial photography. Future NASS or Bureau of Census surveys could include estimates of timber harvest, forest regeneration, conversion of land use, and change in type of ownership, especially for NIPF lands. Annual estimates of seeding and planting rates, supplemented by survival estimates, might improve predictions of regeneration rates for planted stands. Annual estimates of TPO might improve predicted area of forest converted to other cover types because of harvest. These would require the Kalman filter to simultaneously estimate state variables and rate parameters.

Some existing sources of monitoring data might not be useful because they do not use FIA definitions (e.g., SCS, NASS, Bureau of Census, TPO, and Planting Reports). Definitions and criteria for minimum stand size vary. Measurements of aggregations of many variables might poorly represent changes which vary considerably among components (e.g., clearcut rate on industrial lands might be much different than on non-industrial lands, but this difference would be ignored if direct observation of clearcut area was not identified by ownership). The expense in quantifying these differences might not justify a marginal increase in efficiency. Rather, new sources of frequent monitoring data, such as remote sensing or surveys of forest landowners might be more cost effective.

The land classification system in Table 1 contains 418 categories. Each requires a separate state variable. More variables in the state vector would be required for simultaneous estimation of rate parameters. The Kalman filter is rarely successful for systems with more than 60 such variables, and the classification system must be simplified. This could be reduced to 48 classes if non-forested cover types are not differentiated by ownership, nor planted and nonstocked stands by forest type; and land ownership classes condensed into 3 categories: public, forest industry, other private.

This proposal is speculative, and there are major questions regarding implementation. Will insurmountable problems be encountered using remote sensing and ancillary data of varying quality as monitoring data, and FIA inventories as initial conditions and the basis for an empirical model of change in cover? Will monitoring data for land/forest cover and rates of change and inflation of prediction error compensate for a predictive model that assumes constant transition probabilities among cover types. Can econometric models be formulated to better predict annual changes in cover. These questions are being addressed in an ongoing study for the State of North Carolina.

## LITERATURE CITED

## APPENDIX

Available from authors upon request.

LITERATURE CITED

Alig, Ralph J. 1985. Modeling acreage changes in forest ownerships and cover types in the Southeast. USDA Forest Service Rocky Mtn. Forest and Range Experiment Station Research Paper RM-260. 14 pp.

Alig, Ralph J. 1986. Econometric analysis of the factors influencing forest area trends in the Southeast. Forest Science 32(1):119-134.

Alig, Ralph J. and James G. Wyant. 1985. Projecting regional area changes in forestland cover in the U.S.A. Ecological Modeling 29:27-34.

Bonner, G. M. and S. Magnussen. 1987. Forest inventories in Canada: a framework for change. The Forestry Chronicle. June 1987.

Brooks, D. J. 1985. Public policy and long-term timber supply in the South. Forest Science 31(2):342-357.

Catts, G. P., N. D. Cost, R. L. Czaplewski, and P. W. Snook. 1987. Preliminary results from a method to update timber resource statistics in North Carolina. In: 11th Biennial Workshop on Color Aerial Photography in the Plant Sciences. April 27-May 1, 1987; Weslaco, Texas. In press.

Cost, N. D. and H. A. Knight. 1983. Forest classification considerations for monitoring changes and trends. In: Proceedings of international conference on Renewable Resources Inventories for monitoring changes and trends: August 15-19, 1983; Corvallis, Oregon.

Czaplewski, R. L. 1986. Acceptability of the Kalman filter to monitor pronghorn population size. Ph.D. dissertation; Colorado State University; Fort Collins, Co. 135 p.

Czaplewski, R. L., G. P. Catts, and P. W. Snook. 1987. National land cover monitoring using large, permanent photo plots. In: Proceedings of international conference on Land and Resource Evaluation for National Planning in the Tropics: January 25-31, 1987; Chetumal, Quintana Roo, Mexico. pp 197-202. USDA Forest Service GTR WO-39, Washington, DC. 524 p.

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

Fesco, R. S., H. F. Kaiser, J. P. Royer, and M. Weidenhamer. 1982. Management practices and reforestation decisions for harvested southern pinelands. Statistical Research Division, Statistical Reporting Service; U.S. Department of Agriculture. Staff Report No. AGES821230. 74 p.

Frayer, W. E. and Beltz, R. C. 1986. Loss of bottomland hardwoods in the Mississippi Alluvial Plain. In: Inventorying and Monitoring Endangered Forests. Proceedings, IUFRO conference, 1985, Zurich, Switzerland. Birmensdorf, Eidg: Anstalt fur das forstliche Versuchswesen; p. 307-309.

French, B. C. and J. L. Matthews. 1971. A supply response model for perennial crops. American Journal of Agricultural Economics 62:910-916.

Harrison, P. J. and C. F. Stevens. 1976. Bayesian forecasting. Journal of Royal Statistical Society Series B 38:205-247.

Jazwinski, A. H. 1970. Stochastic Processes and Filtering Theory. Academic Press; New York. 376 p.

Kalman, R. E. 1960. A new approach to linear filtering and prediction problems. Transactions of the ASME, Journal of Basic Engineering, March 1960.

Lee, R. C. K. 1964. Optimal estimation, Identification, and Control. Research Monograph 28. The MIT Press; Cambridge, Massachusetts. 152 p.

MacLean, C. D. 1981. Walk-through inventory: a short-cut substitute for remeasuring slow-growing inventory plots. In: Proceedings of international conference on Arid Land Inventories; November 30 - December 6, 1980; La Paz, Mexico. 620 p.

May, G. A., M. L. Holko, and N. Jones, Jr. 1986. Landsat large-area estimates for land cover. IEEE Transactions on Geoscience and Remote Sensing. GE-24:175-184.

Maybeck, P. S. 1979. Stochastic Models, Estimation, and Control. Volume 1. Academic Press; New York. 423 p.

McWilliams, W. H. 1986. Primary sources of timber removals information for the midsouth. In proceedings of symposium on Use of Auxiliary Information in Natural Resources Inventories. Blacksburg, Virginia; Oct. 1 and 2, 1985. Society of American Foresters Publication No. SAF 86-01. pp 38-44.

Minami, D. D., B. B. French, and G. A. King. 1979. An econometric analysis of market control in the California cling peach industry. Calif. Ag. Exp. Stn., Giannini Found. Monogr. 39. 139 p.

Parks, P. J. 1986. A multinominal probit model for land use decisions: Recognizing interdependence among uses and neighboring landowners. Paper presented at 1986 Western Forest Economists Meeting, Wemme Oregon. May 5-8, 1986. 11 pp.

Prisley, Stephen P. and James L. Smith. 1987. Using classification error matrices to improve the accuracy of weighted land-cover models. Photogrammetric Engineering and Remote Sensing. In press.

Sheffield, R. M. and H. A. Knight. 1986. North Carolina's forests. U.S. Department of Agriculture, Forest Service, Resource Bulletin SE-88. Asheville, NC, 97 p.

Smith, W. B. and J. T. Hahn. 1986. Michigan's forest statistics. U.S. Department of Agriculture, Forest Service, General Technical Report NC-112. St. Paul, MN. 44 p.

Sorenson, H. W. (ed.) 1985. Kalman filtering: Theory and Applications. The Institute of Electrical and Electronics Engineers Press, Inc.; New York, 457 p.

Stoll, J. R., R. F. Ziemer, F. C. White, J. C. Bergstrom, and R. S. Kao. 1984. Differential assessment of agricultural lands in the South. Southern Rural Development Center SRDC Series No. 69. Mississippi State University, Mississippi State, MS. 98 pp.

Tansey, J. B. 1984. South Carolina's industrial timber products output, 1977-1981. U.S. Department of Agriculture, Forest Service, Resource Bulletin SE-71. Asheville, NC, 21 p.

USDA Forest Service. 1987. 1986 U.S. Forest Planting Report. U.S. Department of Agriculture, Forest Service, Washington, DC, 13 p.

USDA Soil Conservation Service. 1982. Appraisal 1980, Parts I and II, soil, water and related resources in the United States: status, condition, and trends. U.S. Department of Agriculture, Soil Conservation Service, Washington, DC., Part I 328 p., Part II 296 p.

USDA Statistical Reporting Service. 1983. Scope and methods of the Statistical Reporting Service. Misc. Pub. No. 1308. Washington D.C. 140 p.

USDC Bureau of the Census. 1978 Census of Agriculture, State and County Data, North Carolina. AC78-A-33. Washington, DC., 1981. 705 p.

Visser, Hans. 1986. Analysis of tree ring data using the Kalman filter technique. IAWA Bulletins 71 289-297.

TABLE 1. Land cover-ownership categories. These include combinations of the following classifications.

| Land ownership | Land cover | Forest condition |
|---|---|---|
| Censused categories | | Forest type |
| National forest | Forest | Pine |
| Other Federal | Cropland | Oak-pine |
| Bureau of Land Management | Grassland | Upland hardwood |
| Indian | Urban | Bottomland hardwood |
| Miscellaneous Federal | Water | |
| State | Woodland | Size class |
| County and municipal | Barrenland | Sawtimber |
| Forest industry | | Poletimber |
| Forest industry-leased | | Sapling and seedling |
| Other private (not censused) | | Nonstocked |
| Farmer | | |
| Other individual | | Origin |
| Other corporate | | Natural |
| | | Planted |

## APPENDIX

For review purposes only, will not be published because of page limit. Available from authors upon request.

## KALMAN FILTER

The Kalman filter is given by.

$$\underline{x}_t = \underline{\Phi}_s \, \underline{x}_{t-k} + \underline{w}_{t-k}, \qquad \text{(A1)}$$

$$\underline{y}_t = \underline{H}_t \, \underline{x}_t + \underline{v}_t. \qquad \text{(A2)}$$

In equation (A1), $\underline{x}_t$ is the *nx1* state vector at time *t*. The *n* state variables can represent area for each combination of land ownership, land cover, and forest condition in Table 1. $\underline{\Phi}_s$ is the *nxn* state transition matrix, which is a multivariate linear model of change between time *t-k* and *t*. $\underline{\Phi}_s$ may be time-invariant (s=constant), or may change with time. For example, $\underline{\Phi}_s$ might include proportion of poletimber area that annually changes to sawtimber because of succession. $\underline{w}_{t-k}$ is the *nx1* vector containing prediction errors from *t-k* to *t*. The distribution of $\underline{w}_{t-k}$ can be estimated, but its exact values are unknown.

Equation (A2) is a model for *m* measurements at time *t*. $\underline{H}_t$ is the *mxn* measurement matrix, which describes the linear relationship between $\underline{x}_t$ and $\underline{y}_t$ at time *t*. $\underline{H}_t$ can have off-diagonal elements, which represent linear combinations of state variables (e.g., an estimated area of timberland includes the sum of many different types of timberland). $\underline{v}_t$ is an *mx1* vector of measurement errors, which is the difference between the measurements ($\underline{y}_t$) and the best prediction of the measurements before they are actually made ($\underline{H}_t \, \underline{x}_t$). The exact values of $\underline{v}_t$ are unknown; however, parameters describing their joint distributions can be estimated.

Equations (A1) and (A2) are conceptual models. They are not used for estimation because $\underline{x}_{t-k}$, $\underline{w}_{t-k}$, and $\underline{v}_t$ are vectors of unknown random variables. Rather, estimates or assumptions are made for the first two moments of these variables. For example, prediction errors ($\underline{w}_{t-k}$) are assumed to have a zero mean vector, estimated covariance matrix ($\underline{Q}_s$), and are independent of prediction errors at time intervals other than *t-k* to *t*. Measurement errors ($\underline{v}_t$) are assumed to have a zero mean, estimated covariance matrix ($\underline{R}_t$), and are independent of measurement errors and prediction errors at times other than *t*.

The true, but unknown, state of the system is modeled as:

$$\underline{x}_t = \hat{\underline{x}}_{t|t-k} + \underline{\epsilon}_{t|t-k},$$

where $\hat{\underline{x}}_{t|t-k}$ is the known estimate of the state vector at time *t* given knowledge at time *t-k*, where *k* may equal zero. $\underline{\epsilon}_{t|t-k}$ represents the unknown errors in estimation. It is a vector of random variables, assumed to have a zero mean vector with covariance matrix of $\underline{P}_{t|t-k}$. The estimated state of the system at

time $t$ given knowledge at time $t-k$ is

$$\hat{\underline{x}}_{t|t-k} = \underline{\Phi}_s \hat{\underline{x}}_{t-k|t-k}. \quad \text{(A3)}$$

The covariance matrix for this estimate of the state vector is

$$\underline{P}_{t|t-1} = \underline{\Phi}_s \underline{P}_{t-1|t-1} \underline{\Phi}_s^T + \underline{Q}_s. \quad \text{(A4)}$$

$\underline{G}_t$ is a matrix which combines two independent sources of multivariate information on the system: the predicted estimate ($\hat{\underline{x}}_{t|t-k}$ from equation A3) and measurements ($\underline{y}_t$):

$$\hat{\underline{x}}_{t|t} = [\underline{I} - \underline{G}_t \underline{H}_t] \hat{\underline{x}}_{t|t-k} + \underline{G}_t \underline{y}_t. \quad \text{(A5)}$$

The minimum variance estimator for $\underline{x}_t$ is produced for

$$\underline{G}_t = \underline{P}_{t|t-1} \underline{H}_t^T [\underline{H}_t \underline{P}_{t|t-k} \underline{H}_t^T + \underline{R}_t]^{-1}. \quad \text{(A6)}$$

The error covariance matrix for the new estimate ($\hat{\underline{x}}_{t|t}$) is

$$\underline{P}_{t|t} = [\underline{I} - \underline{G}_t \underline{H}_t] \underline{P}_{t|t-k}. \quad \text{(A7)}$$

If no measurements are taken at time $t$, $\hat{\underline{x}}_{t|t} = \hat{\underline{x}}_{t|t-k}$ and $\underline{P}_{t|t} = \underline{P}_{t|t-k}$. If $\underline{\epsilon}_{t|t}$ has a multivariate normal distribution, then confidence intervals are readily computed using $\hat{\underline{x}}_{t|t}$ and $\underline{P}_{t|t}$.

Each time measurements are taken, a vector of known residuals ($\underline{r}_t$), also called innovations, is available:

$$\underline{r}_t = \underline{y}_t - \underline{H}_t \hat{\underline{x}}_{t|t-k}, \quad \text{(A8)}$$

which is independent of residuals at all other times, has an expected value vector of zero, and covariance matrix of

$$E[\underline{r}_t \underline{r}_t^T] = \underline{H}_t \underline{P}_{t|t-k} \underline{H}_t^T + \underline{R}_t. \quad \text{(A9)}$$

If the distribution of residuals deviates significantly from their expected distribution, then there is a likely problem in model identification.

## ANNUAL TRANSITION MATRIX

From equation (A1), a prediction for $k$ years is

$$\underline{x}_t = \underline{\Phi}_k \underline{x}_{t-k}, \quad \text{(A10)}$$

which can be represented by annual transition matrices ($\underline{\Phi}_1$) using

$$\underline{x}_t = (\underline{\Phi}_1 \underline{\Phi}_1 \cdots \underline{\Phi}_1) \underline{x}_{t-k}. \quad \text{(A11)}$$

Therefore,

$$\underline{\Phi}_k = (\underline{\Phi}_1)^k, \quad \text{(A12)}$$

and

$$\underline{\Phi}_1 = (\underline{\Phi}_k)^{1/k}, \quad \text{(A13)}$$

which yields equation (9). Eigenvalues are used to compute the root of a matrix that has large dimensions (Bodewig, 1959; Wiberg 1971).

## COVARIANCE MATRIX PREDICTION ERROR

The covariance matrix for predictions for $k$-years ($\Phi_k$) is derived by studying the vector of prediction errors for $k$-years ($\underline{w}_{t-k}$) in equation (A1):

$$\underline{w}_{t-k} = \underline{x}_t - \Phi_k \underline{x}_{t-k}. \tag{A14}$$

Their crossproducts are

$$\underline{w}_{t-k} \underline{w}_{t-k}^T = \underline{x}_t \underline{x}_t^T - \Phi_k \underline{x}_{t-k}^T - \underline{x}_t \underline{x}_t^T \Phi_k^T + \Phi_k \underline{x}_{t-k} \underline{x}_{t-k}^T \Phi_{t-k}^T. \tag{A15}$$

Assuming the state vectors are constants rather than random variables, the expected value of the crossproducts are

$$E [\underline{w}_{t-k} \underline{w}_{t-k}^T] = \underline{x}_t \underline{x}_t^T - E [\Phi_k] \underline{x}_{t-k} \underline{x}_t^T - \underline{x}_{t-k} \underline{x}_{t-k}^T E [\Phi_k^T] + E [\Phi_k \underline{x}_{t-k} \underline{x}_{t-k}^T \Phi_k^T] \tag{A16}$$

Taking for $E [\Phi_k] = \hat{\Phi}_k$, which is computed from FIA inventories from times $t$ and $t-k$ using equation (8), the $i\ell$-th term of the last portion of equation (A16) is

$$E [(\Phi_k x_{t-k} x_{t-k}^T \Phi_k^T)_{i\ell}] = \sum_{j=1}^{n} \sum_{p=1}^{n} (x_j)_{t-k} (x_p)_{t-k} \{Cov [(\Phi_{ij})_k (\Phi_{\ell p})_k] + (\hat{\Phi}_{ij})_k (\hat{\Phi}_{\ell p})_k\}, \tag{A17}$$

where $(x_j)_{t-k}$ is the $j$-th element of $\underline{x}_{t-k}$, and $(\Phi_{ij})_k$ is the $ij$-th element of $\Phi_k$. The covariance term in equation (A17) is estimated using the multinominal distribution:

$$\begin{aligned} Cov [(\Phi_{ij})_k (\Phi_{\ell p})_k] &= \frac{(\hat{\Phi}_{ij})_k [1 - (\hat{\Phi}_{ij})_k]}{N(A_{t-k}=j)} && \text{for } j = p,\ i = \ell \\ &= - \frac{(\hat{\Phi}_{ij})_k (\hat{\Phi}_{\ell j})_k}{N(A_{t-k}=j)} && \text{for } j = p,\ i \neq \ell \\ &= 0 && \text{for } j \neq p \end{aligned} \tag{A18}$$

where $N(A_{t-k}=j)$ is the number of plots in category $j$ at time $t-k$. The covariance of probabilities in different column vectors of $\Phi_k$ are independent because

$$P(A_t=i \,|A_{t-k}=j) \cap P(A_t=i \,|A_{t-k}=p) = P(A_t=i \,|A_{t-k}=j)\ P(A_t=i \,|A_{t-k}=p),$$

where $j \neq p$ and $P(A_t=i \,|A_{t-k}=j)$ is the probability of a plot being in category $i$ at time $t$ given it was in category $j$ at time $t-k$.

This assumes FIA plots, used to compute $\hat{\Phi}_k$ in equation (8), are independent. This is reasonable because they are systematically and widely dispersed (one 0.4-ha plot per 25 $km^2$ in NC). However, all land is being modeled in equation (A1), not FIA plots alone. Each site in the landscape is dependent upon surrounding sites, and this covariance is ignored in equation (A18). Therefore, covariance in cover type transition probabilities among contiguous sites is assumed negligible for an entire multi-county physiographic region or State. This is reasonable for the southern United States, where tracts are small and

different land uses are very intermingled. This assumption could be relaxed, but covariance in transition probabilities for contiguous sites could not by quantified using 0.4-ha FIA plots; the covariance might be quantified with much larger plots.

Assuming prediction errors are unbiased (i.e., $E[\underline{w}_{t-k}]=\underline{0}$) and $E[\underline{\Phi}_k]=\hat{\underline{\Phi}}_k$, the covariance matrix for prediction error from $t_{-k}$ to $t$ is

$$\underline{Q}_k = \underline{x}_t \underline{x}_t^T - \hat{\underline{\Phi}}_k \underline{x}_{t-k} \underline{x}_{t-k}^T - \underline{x}_t \underline{x}_{t-k}^T \hat{\underline{\Phi}}_k + E[\hat{\underline{\Phi}}_k \underline{x}_{t-k} \underline{x}_{t-k}^T \hat{\underline{\Phi}}_k^T] \qquad \text{(A19)}$$

using equations (A16) to (A18). However, the state vectors $\underline{x}_{t-k}$ and $\underline{x}_t$ are unknown; instead, their estimates from the last two FIA inventories ($\hat{\underline{x}}_{t-k}$ and $\hat{\underline{x}}_t$) are used in equation (A19) to estimate $\underline{Q}_k$. This ignores the estimation error in state vectors; however, such error is relatively small for FIA inventories.

## ERROR FOR ANNUAL PREDICTIONS

Assuming time-invariant transition probabilities, the prediction error covariance matrix for $k$-years is transformed to an annual time increment using equation (A1):

$$\underline{x}_t = \underline{\Phi}_1 (\underline{\Phi}_1 (\ldots (\underline{\Phi}_1 \underline{x}_{t-k} + \underline{w}_1) + \ldots) + \underline{w}_1) + \underline{w}_1, \qquad \text{(A20)}$$

where $\underline{w}_1$ is the prediction error vector for one year rather than $k$-years. Collecting terms,

$$\underline{x}_t = \underline{\Phi}_1^k \underline{x}_{t-k} + \underline{\Phi}^{k-1} \underline{w}_1 + \underline{\Phi}^{k-2} \underline{w}_1 + \ldots + \underline{\Phi} \underline{w}_1 + \underline{w}_1. \qquad \text{(A21)}$$

From equations (A1) and (A21) and for $\underline{\Phi}_k = \underline{\Phi}_1^k$,

$$\underline{w}_{t-k} = (\underline{\Phi}_1^{k-1} + \underline{\Phi}_1^{k-2} + \ldots + \underline{\Phi}_1 + \underline{I}) \underline{w}_1. \qquad \text{(A22)}$$

Using the expectation operator and assuming $\underline{\Phi}_1$ is a constant rather than a random variable,

$$E[\underline{w}_{t-k} \underline{w}_{t-k}^T] = (\underline{\Phi}_1^{k-1} + \underline{\Phi}_1^{k-2} + \ldots + \underline{\Phi}_1 + \underline{I}) E[\underline{w}_1 \underline{w}_1^T] (\underline{\Phi}_1^{k-1} + \underline{\Phi}_1^{k-2} + \ldots + \underline{\Phi}_1 + \underline{I})^T \qquad \text{(A23)}$$

For unbiased predictions for one and $k$-years, $\underline{Q}_k=E[\underline{w}_{t-k} \underline{w}_{t-k}^T]$ and $\underline{Q}_1=E[\underline{w}_1 \underline{w}_1^T]$. Solving equation (A23) for $\underline{Q}_1$ gives

$$\underline{Q}_1 = (\underline{\Phi}_1^{k-1} + \underline{\Phi}_1^{k-2} + \ldots + \underline{\Phi}_1 + \underline{I})^{-1} \underline{Q}_k [(\underline{\Phi}_1^{k-1} + \underline{\Phi}_1^{k-2} + \ldots + \underline{\Phi}_1 + \underline{I})^T]^{-1} \qquad \text{(A24)}$$

However, $\underline{\Phi}_1$ and $\underline{Q}_k$ can only be estimated as in equations (A13) and (A20), and $\hat{\underline{\Phi}}_1$ and $\hat{\underline{Q}}_k$ are used in equation (A24). This again ignores estimation error of $\hat{\underline{\Phi}}_1$ and $\hat{\underline{Q}}_k$, which must be assumed to have negligable effects on $\hat{\underline{Q}}_1$.

Estimation of $\underline{Q}$ is typically the most difficult task when using the Kalman filter. If all cells in the estimated covariance matrix $\hat{\underline{Q}}$ are approximately proportional to their true values, then $\hat{\underline{Q}}$ can be inflated by a scalar to compensate for the above simplifications and approximations (see Validation below).

## ESTIMATING MEASUREMENT MATRIX

The covariance matrix for estimation errors in certain measurement matrices ($\underline{H}_t$) can be estimated in a fashion similar to $\underline{Q}_k$. Consider the case when small plots are randomly selected and categorically classified by digital analysis of satellite data, manual interpretation of aerial photography, or with techniques and definitions from other agencies. Reference data using FIA definitions (Table 1) are available for each small plot, and are considered truth. Let $N(A_y=i \,|\, A_r=j)$ be the number of plots classified into category $i$ given their FIA classification is category $j$. $N(A_r=j)$ is the number of plots in FIA category $j$. The estimated measurement $\hat{\underline{H}}_t$ matrix would be computed similar to $\hat{\underline{\Phi}}_k$ in equation (8):

$$(\hat{H}_{ij})_t = \frac{N(A_y=i \,|\, A_r=j)}{N(A_r=j)} \qquad \text{(A25)}$$

where $(\hat{H}_{ij})_t$ is the $ij$-th element of the estimated measurement matrix $(\hat{\underline{H}}_t)$. Therefore, each column vector in $\hat{\underline{H}}$ is an independent, estimated vector of conditional probabilities, namely the probability that a plot will be in category $i$ of the classification system used for measurements (remote sensing or estimates from another agency) given it is in category $j$ of the FIA classification system (Table 1). However, errors in estimating $\underline{H}_t$ contribute to the total measurement error.

Assume the only measurement error is from estimating the measurement matrix. From equation (A2),

$$\underline{v}_t = \underline{y}_t - \underline{H}_t \, \underline{x}_t. \qquad \text{(A26)}$$

For $\underline{y}_t$ and $\underline{x}_t$ constants rather than random variables, unbiased errors in estimating $\underline{H}_t$, and $E[\underline{H}_t]=\hat{\underline{H}}_t$, then

$$\underline{R}_t = \underline{y}_t \, \underline{y}_t^T - \hat{\underline{H}}_t \, \underline{x}_t \, \underline{y}_t^T - \underline{y}_t \, \underline{x}_t^T \, \hat{\underline{H}}_t + E[\underline{H}_t \, \underline{x}_t \, \underline{x}_t \, \underline{H}^T]. \qquad \text{(A27)}$$

This is derived similar to $\underline{Q}_k$ in equations (A15), (A16) and (A19). The expected value in equation (A27) is estimated using the multinominal distribution, similar to equations (A17) and (A18):

$$E[(\underline{H}_t\ \underline{x}_t\ \underline{x}_t^T\ \underline{H}_t^T)_{i\ell}] = \sum_{j=1}^{n}\sum_{p=1}^{n} (x_j)_t\ (x_p)_t$$
$$\{\text{Cov}\ [(\hat{H}_{ij})_t\ (\hat{H}_{\ell p})_t] + (\hat{H}_{ij})_t\ (\hat{H}_{\ell p})_t\} \qquad \text{(A28)}$$

where

$$\text{Cov}\ [(\hat{H}_{ij})_t\ (\hat{H}_{\ell p})_t] = \frac{(\hat{H}_{ij})_t[1-(\hat{H}_{ij})_t]}{N(A_r=j)} \quad \text{for } j=p,\ i=\ell$$
$$= -\frac{(\hat{H}_{ij})_t(\hat{H}_{\ell j})_t}{N(A_r=j)} \quad \text{for } j=p,\ i\neq\ell$$
$$= 0 \quad \text{for } j\neq p \qquad \text{(A29)}$$

This assumes classification error for a plot is independent of classification of all other plots. However, this also assumes that estimation error for $\underline{H}_t$ is the only source of measurement error. If plots are used for measurements ($\underline{y}_t$), then the covariance matrix for sampling error would have to be added to equation (A27) to compute a covariance matrix for total measurement error.

It is important to precisely estimate the measurement matrix because errors in estimating $\underline{H}_t$ increase $\underline{R}_t$ in equation (A6), which increases estimation error for state variables in equation (A7). For example, measurements with small sampling errors might not substantially improve estimates of state variables if the relationship between measurements and state variables $\underline{H}_t$ is imprecisely known. Therefore, error matrices in remote sensing studies, and relationships among land classification systems, should be determined with adequate sample sizes.

## INDEPENDENCE OF ERRORS

The vectors of measurement errors ($\underline{v}_t$), prediction errors ($\underline{w}_{t-k}$), and estimation errors for intial conditions ($\underline{\epsilon}_{0|0}$) are assumed independent of each other in equations (6) and (7). This is reasonable for $\underline{v}_t$ because measurements come from independent sources. However, FIA data are used for both initial conditions and estimating transition probabilities in $\hat{\Phi}_1$. Covariance between $\underline{\epsilon}_{0|0}$ and $\underline{w}_t$ (t>0) might be assumed negligable because the former involves estimating the area for cover types, while the latter involves estimating probability of a change in cover type. Prediction errors for time *t* are assumed independent of prediction errors at all other times. This is reasonable assuming an accurate prediction model. Measurement errors are also assumed uncorrelated in time. This might be a poor assumption if permanent plots are used, as in the NASS, SCS, and some remote sensing systems. Techniques presented by Dixon and Howitt (1979) might be required for systems using some permanent plots.

## SQUARE ROOT FILTER

Computing problems are frequently encountered when using the Kalman filter because covariance matrices are used. The square root filter (Bierman 1977, Maybeck 1979) is equivalent to the Kalman filter, but is computationally more stable. It is frequently used instead of equations (A4) to (A7). There are many variations of the square root filter. These formulations differ in

computational speed and storage requirements, which are important for real-time, aerospace applications. However, the square root filter will be applied off-line to monitoring changes in land cover, and computational speed and array storage are lesser considerations. Therefore, the simplest algorithm is given.

The state vector $\underline{x}_{t|t-1}$ and its covariance matrix for estimation errors ($\underline{P}_{t|t-1}$) at time $t$ are predicted from their values at $t$-1 using the standard methods in equations (A3) and (A4). The latter step can be made more numerically accurate using the modified Gram-Schmidt orthogonalization or the Householder transformation; however, it is the measurement update, not the state space predictions in equations (A3) and (A4), which causes critical numerical problems in the filter.

Next, the measurement vector at time $t$ ($\underline{y}_t$) is orthogonalized ($\underline{y}^*$) using its covariance matrix for measurement errors ($\underline{R}_t$):

$$\underline{y}_t^* = (\underline{R}_t)^{-0.5} \underline{y}_t. \qquad \text{(A30)}$$

$\underline{R}_t$ can contain error covariances for estimating the measurement matrix ($\hat{\underline{H}}_t$) in addition to sampling errors.

From equations (A2) and (A30),

$$\underline{y} = \underline{H}_t \underline{x}_t + \underline{v}_t$$

$$(\underline{R}_t^{-0.5}) \underline{y} = \underline{R}_t^{-0.5} (\underline{H}_t \underline{x}_t + \underline{v}_t)$$

$$\underline{y}_t^* = \underline{H}_t^* \underline{x}_t + \underline{v}_t^*$$

where

$$\underline{H}_t^* = \underline{R}_t^{-0.5} \underline{H}_t$$

$$\underline{v}_t^* = \underline{R}_t^{-0.5} \underline{v}_t$$

The covariance matrix for the transformed measurement error ($\underline{v}_t^*$) is the identity matrix.

$$\begin{aligned} E[\underline{v}_t^* (\underline{v}_t^*)^T] &= \underline{R}_t^{-0.5} E[\underline{v}_t \underline{v}_t^T] (\underline{R}_t^{-0.5})^T \\ &= \underline{R}_t^{-0.5} \underline{R}_t (\underline{R}_t^{-0.5})^T \\ &= (\underline{R}_t^{-0.5} \underline{R}_t^{0.5}) (\underline{R}_t^{-0.5} \underline{R}_t^{0.5})^T \\ &= \underline{I} \underline{I}^T \\ &= \underline{I} \end{aligned}$$

However, $\underline{R}_t$ is unknown, and the estimated measurement error covariance matrix must be used instead.

Now that the transformed measurements at time $t$ from equation (A30) are mutually independent, the $i$-th standardized measurement $(y_i^*)_t$, $i$=1, 2, ..., $m$, can be processed sequentially as scalers rather than simultaneously as a vector (Potter 1964).

Let $\underline{S}_{t|t-1}$ be the matrix square root of the propagated covariance matrix for estimation error $\underline{P}_{t|t-1}$, where

$$\underline{P}_{t|t-1} = (\underline{S}_{t|t-1})\ (\underline{S}_{t|t-1})^T, \qquad \text{(A31)}$$

$\underline{S}_{t|t-1}$ is computed using the Cholesky decomposition or eigenvalues. Define the $n$x1 vector $\underline{a}_i$, corresponding to the $i$th orthogonalized measurement at time $t$, as

$$\underline{a}_i = [(\underline{H}_i^*)_t\ \underline{S}_{t|t-1}]^T, \qquad \text{(A32)}$$

where $(\underline{H}_i^*)_t$ is a 1x$n$ vector containing the $i$th row of the transformed measurement matrix

$$\underline{H}_t^* = \underline{R}_t^{-0.5}\ \underline{H}_t. \qquad \text{(A33)}$$

Let the scaler $b_i$, corresponding to the $i$th measurement at time $t$, be defined as

$$b_i = (\underline{a}_i^T\ \underline{a}_i + 1)^{-1}. \qquad \text{(A34)}$$

Then the $n$x1 gain matrix corresponding to the $i$th measurement is.

$$(\underline{G}_i^*)_t = b_i\ \underline{S}_{t|t-1}\ \underline{a}_i, \qquad \text{(A35)}$$

which is used as in equation (A5) to combine measurement data with the prior estimates:

$$(\hat{\underline{x}}_i)_{t|t} = (\hat{\underline{x}}_{i-1})_{t|t} + (\underline{G}_i^*)\ [(y_i^*)^t - (\underline{H}_i^*)_t\ (\hat{\underline{x}}_{i-1})_{t|t}] \qquad \text{(A36)}$$

$(\hat{\underline{x}}_j)_{t|t}$ is the composite estimator for the state vector at time $t$ that combines the previous estimate (equation A3) with measurements 1 to $j$ at time $t$, $(\hat{\underline{x}}_0)_{t|t} = \hat{\underline{x}}_{t|t-1}$, and $(\hat{\underline{x}}_m)_{t|t} = \hat{\underline{x}}_{t|t}$ in equation (A5).

The square root of the covariance matrix for estimation error in this composite estimator is

$$(\underline{S}_i)_{t|t} = (\underline{S}_{i-1})_{t|t} - [(\underline{G}_i^*)_t\ \underline{a}_i^T]\ (1 + b_i^{0.5})^{-1} \qquad \text{(A37)}$$

$(\underline{S}_j)_{t|t}$ includes measurements 1 to $j$, $(\underline{S}_0)_{t|t} = S_{t|t-1}$, and $(S_m)_{t|t} = \underline{S}_{t|t}$.

Equations (A32) to (A37) are repeated $m$ times, once for each of the $m$ measurements at time $t$. The covariance matrix for the error in the final estimate is

$$\underline{P}_{t|t} = \underline{S}_{t|t}\ \underline{S}_{t|t}^T,$$

which is the equivalent of equation (A7) on the original Kalman filter.

## VALIDATION

Many assumptions, approximations, and estimates are necessary to apply the Kalman filter for monitoring land cover. Also, numerical problems and computer programming errors during implementation are possible. Therefore, it is crucial to closely inspect the solution for inconsistencies. This can be accomplished using the time series of residuals, which use known measurements and their known expected values given the structure of and parameter estimates for the Kalman filter.

The multivariate residuals at time t are defined in equation (A8), and their covariance matrix in equation (A9). If the assumptions and estimates used in the Kalman filter are accurate, then vectors of residuals from different time periods are unbiased and independent of those at time $t$. Also, residuals from independent sources of measurement data at the same time should also be independent. The task is to combine all residuals for t>0 into a single set of independent transformed residuals, each with the same expected distribution so that their actual distribution and correlations can be inspected and compared to their expected statistics.

Each vector of residuals ($\underline{r}_t$) from the same multivariate measurement can be standardized and orthogonalized ($\underline{r}^*_t$) using the square root of their expected covariance matrix (equation A9):

$$\underline{r}^*_t = (\underline{H}_t \, \underline{P}_{t|t-1} \, \underline{H}^T_t + \underline{R}_t)^{-0.5} \, \underline{r}_t \qquad \text{(A37)}$$

This could be done using the Cholesky decomposition or eigenvalues, similar to procedures used an the square root filter. If $\underline{r}_t$ is unbiased and multivariate normal, and the Kalman filter is implemented accurately, then $\underline{r}^*_t$ will be unbiased and multivariate normal with a covariance matrix of unity (i.e., each element of $\underline{r}^*_t$ will be independent of all other elements of $\underline{r}^*_t$, each will have a variance of one and each will be normally distributed). These standardized measurement residuals may be pooled with similarly standardized residual vectors from all other time periods and measurement systems because these vectors are mutually independent and identically distributed assuming an accurate implementation of the filter.

These pooled, independent, standarized measurement residuals may be inspected in many ways. Goodness-of-fit tests, such as the Kolmogorov-Smirnov, Cramer-Von Mises, and Anderson-Darling tests (Reynolds 1984), could be used to test for a normal distribution with a known mean of zero and an known variance of one. Many different goodness-of-fit tests should be used because each performs best for certain types of deviations from expected distributions.

If these tests suggest rejection of the assumption of normally distributed standardized residuals with zero mean and unit variance, then there are 3 potential causes. The mean might not be zero, which can be tested using a t-test or a distribution-free equivalent (Mielke 1986). The variance might not equal one. This could be tested using an F-test (Sokal and Rohlf 1969) or a similar distribution-free test (Mielke 1986). If the assumptions of zero mean and unit variance are not rejected, then the distribution of standardized measurement residuals might not be normal. In this case, the Kalman filter still produces minimum variance estimates; however, confidence intervals for these estimates should not be computed under the assumption of multivariate normal estimation errors. This greatly complicates estimation of confidence intervals. In this case, transformations will be considered to normalize distribution of prediction and measurement errors.

If the assumptions, estimates, and approximations used to apply the Kalman filter are accurate, then the standardized measurement residuals should be mutually independent. Tests for correlations in the standardized residuals should be made. For example, NASS estimates for cropland might be consistently greater than those predicted by the filter, suggesting an inaccurate estimate of the measurement matrix for such data or a prediction model which consistently

underestimates area of cropland. Analysis of variance or similar distribution-free techniques (Mielke 1986) might be used to detect simple patterns. If there are enough residuals, more complicated pattern recognition techniques might be applied (e.g., Breiman et al. 1984). Significant patterns in the standardized residuals suggest an inefficient and possibly biased implementation of the filter, and opportunities likely exist to reduce variance of the estimation error.

The approach described in the first part of this paper assumes time-invariant transition probabilities (i.e., $\underline{\Phi}_1$ from equation A13). The time-invariant covariance matrix for prediction error ($\underline{Q}_1$) from equation (A24) considers only the sampling error in estimating these time-invariant transition probabilities. However, this assumption is very likely an oversimplification. Economic forces probably change the rates of harvest, regeneration, land use conversion; shifts among different ownership types; and differences in timber stand improvement over time, especially since the last FIA inventory. If these changes occur at rates different than those observed in the past, then measurements would tend to reduce the bias in the estimates, and would partially compensate for an inaccurate state transition model. However, the variance-covariance of prediction errors would be underestimated because prediction errors are caused by an oversimplified model in addition to sampling errors during estimation of transition probabilties. In terms of an optimal composite estimator, too much weight would be placed on model predictions and not enough on measurements. If the assumption of time-invariant transition probabilities is a seriously inaccurate then the standardized residuals (equation A37) will likely fail the validation tests already described.

The following addition to the estimation model is a simple attempt to recognize the likelihood of time-variant transition probabilities. Let the state transition model (equation A1) be modified as follows:

$$\underline{x}_t = \underline{\Phi}_1 \underline{x}_{t-1} + q \underline{v}_1 \qquad \text{(A38)}$$

where q is a time-invariant scalar, representing the additional prediction error caused by assuming $\underline{\Phi}_t$ does not change over time. The covariance matrix for prediction error ($\underline{Q}_1'$) in equation (A38) would be

$$\underline{Q}_1' = q^2 \underline{Q}_1,$$

where $Q_1$ is estimated using equation (A24).

The value of q could be estimated by iteratively exploring a wide range of possible scalar values (presumeably q>1), and inspecting the effect on the goodness-of-fit statistics (see Validation section in this Appendix). The best estimate of q would be that which maximizes the goodness-of-fit of the standardized residuals to their expected distribution. This tactic has been proposed by Czaplewski (1986) for an application of the Kalman filter to big game management, and is suggested as a useful model for estimating prediction error even when the null hypothesis is not rejected during the goodness-of-fit tests.

Another approach to improving the assumption of time-invariant transition probabilities is to use a more realistic, econometric model that is driven by annual economic indicators for the State or sub-State physiographic region. When such models are available, it will be necessary to describe their prediction errors, including a new estimate of $\underline{Q}_t$. Such models are also based

on historical trends, which might have changed since the econometric model was estimated. Therefore, a tactic similar to equation (A38) might be useful, even after better transition models are built.

## LITERATURE CITED IN APPENDIX
(Also see Literature Cited for main paper)

Bierman, G. J. 1977. Factorization Methods for Discrete Sequential Estimation. Academic Press; New York. 241 p.

Brieman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. 1984. Classification and Regression Trees. Wadsworth Internation Groupl Belmont, California. 358 p.

Bodewig, E. 1959. Matrix Calculus. North-Holland Publishing Co; New York, 452 p.

Mielke, P. W. 1986. Multiresponse permutation procedures. In Encyclopedia of Statistical Sciences, Volume 6. Wiley-Interscience New York. pp 724-727.

Potter, J. E. 1964. W. Matrix Augmentation. M.I.T. Instrumentation Laboratory Memo SGA 5-64. Cambridge, Massachusetts 1/64.

Reynolds, M. R. 1984. Estimating the error in model predictions. Forest Science 30:454-469.

Sokal, R. R. and F. J. Rohlf. 1969. Biometry. W. H. Freeman and Co. San Francisco. 775 p.

Van Deusen, P. C. 1987. Southern red qpruce tree-ring analysis. In: Tree rings and forest mensuration: how can they document trends in forest health and productivity. NCASI Technical Bulletin No. 523.

Wiberg, D. M. 1971. State space and Linear Systems. McGraw-Hill Book Company; New York. 237 p.