Kalman filter for statistical monitoring of forest cover across sub-continental regions

Raymond L. Czaplewski

USDA Forest Service

Rocky Mountain Forest and Range Experiment Station

240 West Prospect Street

Fort Collins, CO 80526-2098

U.S.A.

FAX: 303-498-1010

Invited paper presented at the 4th Symposium on Biometrical Problems in Agricultural Forestry and Animal Investigations, August 19-21, 1991, Jokioinen, Finland.

ABSTRACT

The Kalman filter is a generalization of the composite estimator. The univariate composite estimate ($\hat{\underline{z}}$) combines 2 prior estimates ($\hat{\underline{x}}$ and $\hat{\underline{y}}$) of population parameter ($\underline{z}$) with a weighted average, $\hat{\underline{z}}=(1-\underline{k})\hat{\underline{x}}+\underline{k}\hat{\underline{y}}$, where the scalar weight $\underline{k}$ is inversely proportional

to the variances: $\underline{k}=\hat{\underline{var}}(\hat{\underline{x}})/[\hat{\underline{var}}(\hat{\underline{x}})+\hat{\underline{var}}(\hat{\underline{y}})]$. The composite estimator is a minimum variance estimator that requires no distributional assumptions other than estimates of the first 2 moments. The Kalman filter recursively combines 2 estimates: a past estimate that is updated for expected change over time, and a current, direct estimate.

The multivariate Kalman filter is analogous to the univariate composite estimator. $\hat{\mathbf{x}}_{t|t}$ is an unbiased 1x$\underline{k}$ vector estimate of $\underline{k}$ population parameters (i.e. a state vector) at time $\underline{t}$, with $\underline{k}$x$\underline{k}$ covariance matrix $\hat{\mathbf{P}}_{t|t}$. For example, there might be $\underline{k}$=6 proportions of the population in various states of forest condition: undisturbed, degraded, cleared, plantation, fallow, and non-forest. $\hat{\mathbf{\Phi}}_{t+1|t}$ is a $\underline{k}$x$\underline{k}$ transition matrix that predicts change in the state vector between times $\underline{t}$ and $\underline{t}$+1; the vector of random prediction errors (i.e., difference between the true and estimated states at time $\underline{t}$+1) is assumed unbiased and independent of errors at other times, with $\underline{k}$x$\underline{k}$ covariance matrix $\hat{\mathbf{Q}}_{t+1|t}$. The Kalman filter updates the past estimate for changes expected between times $\underline{t}$ and $\underline{t}$+1 using $\hat{\mathbf{x}}_{t+1|t}=\hat{\mathbf{\Phi}}_{t+1|t}\hat{\mathbf{x}}_{t|t}$, with covariance matrix $\hat{\mathbf{P}}_{t+1|t}=\hat{\mathbf{\Phi}}_{t+1|t}\hat{\mathbf{P}}_{t|t}\hat{\mathbf{\Phi}}'_{t+1|t}+\hat{\mathbf{Q}}_{t+1|t}$. At time $\underline{t}$+1, a direct estimate ($\hat{\mathbf{y}}_{t+1}$) of the population might be available; $\hat{\mathbf{y}}_{t+1}$ is a 1x$\underline{m}$ vector, with the $\underline{m}x\underline{m}$ covariance matrix $\hat{\mathbf{R}}_{t+1}$, and $\underline{m}$ need not equal $\underline{k}$. $\hat{\mathbf{y}}_{t+1}$ might be a linear combination of the state vector $\mathbf{y}_{t+1}=\mathbf{H}_{t+1}\mathbf{x}_{t+1}$, where $\mathbf{H}_{t+1}$ is a $\underline{m}x\underline{k}$ matrix. For example, coarse-resolution satellite data might be able to discriminate well-stocked forest (undisturbed and plantation forests) from all other cover types (degraded, cleared, and fallow forests, and non-forest), in which case $\underline{m}$=2 and $\underline{k}$=6. The $\underline{k}x\underline{m}$ matrix weight for the composite estimator is $\mathbf{K}_{t+1}=[\hat{\mathbf{P}}_{t+1|t}\mathbf{H}'_{t+1}+\hat{\mathbf{C}}_{t+1}][\mathbf{H}_{t+1}\hat{\mathbf{P}}_{t+1|t}\mathbf{H}'_{t+1}+\hat{\mathbf{R}}_{t+1}+\mathbf{H}_{t+1}\hat{\mathbf{C}}_{t+1}+\hat{\mathbf{C}}'_{t+1}\mathbf{H}'_{t+1}]^{-1}$, where $\hat{\mathbf{C}}_{t+1}$ is the $\underline{k}x\underline{m}$ covariance matrix between model prediction errors and the estimation errors for $\hat{\mathbf{y}}_{t+1}$. The sum of each proportion vector is $\mathbf{1}\hat{\mathbf{x}}=\mathbf{1}\hat{\mathbf{y}}=1$, their covariance matrices sum to $\mathbf{1}\hat{\mathbf{P}}\mathbf{1}'=\mathbf{1}\hat{\mathbf{R}}\mathbf{1}'=0$, and dimensions $\underline{k}$ and $\underline{m}$ can be decreased by 1 to avoid singularity without loss of information. The multivariate composite estimate is $\hat{\mathbf{x}}_{t+1|t+1}=[\mathbf{I}-\mathbf{K}_{t+1}]\hat{\mathbf{x}}_{t+1|t}+\mathbf{K}_{t+1}\hat{\mathbf{y}}_{t+1}$, with covariance matrix $\hat{\mathbf{P}}_{t+1|t+1}=\hat{\mathbf{P}}_{t+1|t}-\mathbf{K}_{t+1}[\mathbf{H}_{t+1}\hat{\mathbf{P}}_{t+1|t}+\hat{\mathbf{C}}'_{t+1}]$. The vector of residual differences between the estimates and predicted estimates acquired at time $\underline{t}$+1 is $\mathbf{r}_{t+1}=\hat{\mathbf{y}}_{t+1}-\mathbf{H}_{t+1}\hat{\mathbf{x}}_{t+1|t}$, where $E[\mathbf{r}_{t+1}]=\mathbf{0}$ and covariance matrix $\hat{E}[\mathbf{r}_{t+1}\mathbf{r}'_{t+1}]=\mathbf{H}_{t+1}\hat{\mathbf{P}}_{t+1|t}\mathbf{H}'_{t+1}+\hat{\mathbf{R}}_{t+1}+\mathbf{H}_{t+1}\hat{\mathbf{C}}_{t+1}+\hat{\mathbf{C}}'_{t+1}\mathbf{H}'_{t+1}$.

The Kalman filter repeats this procedure recursively for each time period to produce a time-series of composite estimates. If any underlaying assumptions are invalid (e.g., model prediction errors not distributed with mean vector **0** and covariance matrix $\hat{\mathbf{Q}}_{t+1|t}$), then the Kalman filter can produce inaccurate estimates. Diagnostic procedures must be used to detect quantitative symptoms that suggest invalid assumptions.

Statistical estimates for proportions of cover in k categories can be produced for a sub-continental region each year as follows. A random sample of high-resolution satellite scenes (e.g., Landsat) is imaged at time t=0, and each Landsat pixel in the sample is classified into 1 of k=6 categories; Landsat classification error is assumed negligible. An estimate $\hat{\mathbf{x}}_{0|0}$ for the proportion of the k categories in the population is made using the proportions of the k categories in the scene, and the sample covariance matrix used for $\hat{\mathbf{P}}_{0|0}$. Coarse-resolution satellite data (e.g., AVHRR) is acquired for the entire population at t=0, and each AVHRR pixel is classified into 1 of m=2 simplified categories. The proportion of classified AVHRR pixels in this census is a direct estimate of population proportions. Misclassification error (i.e., measurement error) will bias this estimate. However, AVHRR proportion estimates can be calibrated, using AVHRR pixels registered to the random sample of Landsat scenes, to produce an unbiased estimate ($\hat{\mathbf{y}}_1$) of the

$\underline{m}$=2 simplified cover categories; covariances of errors propagated from the calibration model are estimated by $\hat{\mathbf{R}}_1$. The Kalman filter combines population estimates $\hat{\mathbf{x}}_{0|0}$ and $\hat{\mathbf{y}}_1$ into the composite estimate $\hat{\mathbf{x}}_{1|1}$, where times $\underline{t}$=0 and $\underline{t}$=1 are identical (i.e., $\hat{\mathbf{\Phi}}_{0|1}$=**I**, $\hat{\mathbf{Q}}_{0|1}$=**0**), and $\hat{\mathbf{C}}_1$ estimates the covariance between $\hat{\mathbf{x}}_{0|0}$ and $\hat{\mathbf{y}}_1$ caused by shared use of the same Landsat sample scenes. $\hat{\mathbf{C}}_1$ might be difficult to estimate, and the AVHRR data might not substantially improve precision unless the calibration sample of Landsat scenes is large (i.e., $\hat{\mathbf{R}}_1$ is small). If Landsat classifications contain errors, then estimates can be further calibrated using a different sample of higher-resolution data.

A probability transition matrix $\hat{\mathbf{\Phi}}_{2|1}$ can estimate the proportions ($\hat{\mathbf{x}}_{2|1}$) of $\underline{k}$=6 cover types at time $\underline{t}$=2 given the previous composite estimate ($\hat{\mathbf{x}}_{1|1}$) at $\underline{t}$=1. The transition matrix can be estimated with models that predict change in forest cover caused by environmental and anthropogenic forces, or empirical models based on past changes among cover types on permanent sample plots. If no direct estimates ($\hat{\mathbf{y}}_2$) are acquired at time $\underline{t}$=2, then $\hat{\mathbf{x}}_{2|2}$=$\hat{\mathbf{x}}_{2|1}$; otherwise, new ancillary estimates are combined with the Kalman filter, as described above. This procedure is repeated each time period. A time-series of known residual vectors

($\mathbf{r}_t$, $\underline{t}$=0,1,2,...) can be standardized with spectral decompositions of their expected residual covariance matrices; standardized residuals are expected to be independent and identically distributed, with mean 0 and variance 1. If the realized distribution of standardized residuals from all time periods fails to meet this expectation, then a problem is detected. The most likely causes are biased model predictions ($\hat{\mathbf{\Phi}}_{t+1|t}\hat{\mathbf{x}}_{t|t}$) or estimates of model prediction error covariance ($\hat{\mathbf{Q}}_{t+1|t}$), but there might also be unrecognized biases in the direct estimates ($\hat{\mathbf{y}}_{t+1}$) or their covariance matrices ($\hat{\mathbf{R}}_{t+1}$). Good technical judgement is required to diagnose the probable cause of the problem, and prescribe corrective actions.

This estimation strategy assumes direct estimates are based upon a random sample of Landsat scenes, and estimates for sub-populations are not required. Both assumptions might be poor because national mapping priorities might produce a non-random sample of classified Landsat scenes, and sub-population estimates are often requested. One solution is to (1) consider the population equivalent to all Landsat scenes that image the region, (2) estimate of the state of each Landsat scene during each time period, and (3) consider this a census of the entire population each time period, with varying degrees of multivariate measurement error for each Landsat scene.

Measurement errors are caused by imperfect model predictions. One model estimates current state of each Landsat scene based on past measurements and expected changes, and the calibration model estimates the current state of each Landsat scene based on imperfectly classified AVHRR pixels that fall within each Landsat scene. As in the Kalman filter, model predictions and AVHRR estimates can be combined for each Landsat scene. If a Landsat scene has an unusually large residual difference between model predictions and AVHRR estimates, then measurement error for this scene might be high, and new Landsat data could be obtained for this scene to reduce measurement error. However, the multivariate measurement errors among Landsat scenes will not be independent or identically distributed. This presents formidable challenges to prudent statistical estimation. Spatial statistics and propagation of error techniques are needed to estimate the covariance among and within Landsat scenes. An estimator analogous to the multivariate composite estimator might be needed to efficiently combine state estimates for all Landsat scenes into a population or sub-population estimate.