Kalman Filter to Update Forest Cover Estimates

Raymond L. Czaplewski
Mathematical Statistician
USDA Forest Service
Rocky Mountain Forest and Range Experiment Station
240 West Prospect Road
Fort Collins, Colorado 80526 U.S.A.

ABSTRACT

The Kalman filter is a statistical estimator that combines a time-series of independent estimates, using a prediction model that describes expected changes in the state of a system over time. An expensive inventory can be updated using model predictions that are adjusted with more recent, but less expensive and precise, monitoring data. The concepts of the Kalman filter are explained with a simple, hypothetical example of estimating percent forest cover over time, using remote sensing and field plots from a forest inventory.

INTRODUCTION

The Kalman filter is a composite estimator, which combines two independent estimates at a time, each of which is weighted inversely proportional to its variance. Gregoire and Walters (1988) note that composite estimators are widely used in forestry, including sampling with partial replacement (e.g., Ware and Cunia 1962). Green and Strawderman (1986) and Thomas and Rennie (1987) show how a composite estimator may be used to combine independent estimates of stem density, basal area, or wood volume.

In the Kalman filter, one of the independent estimates is a current estimate or monitoring measurement (e.g., remotely sensed data or midcycle update); the other is a previous estimate, or forest inventory, that is updated for expected changes over time using a deterministic prediction model. Variance for this updated estimate includes effects of (1) errors in the previous forest inventory that are propagated over time, and (2) model prediction errors between the previous and current estimates. Errors in a composite estimate are typically less than errors in either prior estimate alone.

EXAMPLE OF A COMPOSITE ESTIMATE

First, consider the problem of estimating percent forest cover in Fig. 1. Make an ocular estimate of the percent forest in Fig. 1, now. Then record your estimate in Table 1.

Nine other photointerpreters independently made ocular estimates for this same image, and their results are recorded in Table 1. Assume each ocular estimate has no bias or sampling error, and each has an identical distribution of measurement errors (i.e., Fig. 1 is imperfectly censused, with nine independent replicates). The mean of these ocular measurements (x_1=56.0% from Table 1) is the first estimate of percent forest cover in Fig. 1, with variance of the mean Var(x_1)=3.11%% (variance units are $\%^2$, denoted as %%)

Next, consider a second estimate of percent forest cover in Fig. 1 using error-free classification of 400 temporary plots. There are 204 forested plots, producing the estimate y_1=51.0%. Using the binomial distribution, the estimated sampling variance is Var(y_1)=(51.0%)(49.0%)/400=6.25%%, which produces the approximate 95% confidence interval $CI_{95\%}(y_1)$=51.0±4.9%.

Figure 1. Hypothetical example requiring an estimate of the percent forest cover (shaded area in this image) at time $t=1$. You are requested to make an ocular estimate of percent forest, and record your estimate in Table 1. Subsequently, 400 randomly located plots are independently classified into forest or nonforest.

Table 1. Independent ocular estimates (photointerpretations) of percent forest in Fig. 1 by nine (or more) different, but unbiased and equally skilled, observers. These are pooled into a group estimate. Recompute the mean, variance among interpretations, and variance of the mean, using your ocular estimate of percent forest cover in Fig. 1 ($n=10$).

Observer	Ocular Estimate (x)	x^2	Estimator Calculations ($n=9$)
1	60	3600	Group mean x_I =504%/9
2	55	3025	x_I =56%
3	50	2500	
4	55	3025	Variance among photointerpretations
5	52	2704	=[28448%%-(504%)^2/9]/8
6	50	2500	=28%%
7	65	4225	
8	62	3844	Variance of mean=28%%/9
$n=9$	55	3025	Var(x_I^J)=3.11%%
Subtotal	504%	28448%%	
			Approximate 95% confidence interval
Yours	___	___	$x_I \pm 1.96(3.11)^{1/2}$=56.0±3.5%
Total($n=10$)			

The estimate x_1=56.0% from Table 1 can be combined with the sample estimate from Fig. 1 (y_1=51.0%) to produce a new composite estimate x^*_1, as shown in Fig 2. x_1 is weighted more heavily than y_1 because $Var(x_1)$=3.11%% for replicated ocular measurement error is less than $Var(y_1)$=6.25%% for sampling error from 400 point plots:

$$x^*_1 = [A_1\, y_1] + [(1 - A_1)\, x_1]$$
$$= [(0.33)\ 51.0\%] + [(0.67)\ 56.0\%]$$
$$= 54.4\%$$

$$A_1 = Var(x_1)/[Var(x_1)+Var(y_1)]$$
$$= 3.11\%\% \;/\; (3.11\%\% + 6.25\%\%)$$
$$= 0.33$$

Weight (or shrinking coefficient) A_1 can be derived using maximum likelihood, minimum variance, or empirical Bayes theory. The expected variance of the composite estimate $Var(x^*_1)$ is based on an elementary theorem in mathematical statistics for the variance of a linear transformation Y of independent random variables X_a and X_b:

$$Y = (a\, X_a) + (b\, X_b)$$

$$Var(Y) = a^2\, Var(X_a) + b^2\, Var(X_b)$$

Applying this theorem to the composite estimator x^*_1:

$$Var(x^*_1) = [A_1^2\, Var(y_1)] + [(1 - A_1)^2\, Var(x_1)]$$
$$= [(0.1089)\ 6.25\%\%] + [(0.4489)\ 3.11\%\%]$$
$$= 2.08\%\%$$

This estimator of $Var(x^*_1)$ is also given by Gregoire and Walters (1988). The variance of the composite estimate is smaller than either of the two independent estimates (Green and Strawderman 1986), as illustrated in Fig. 2; the approximate 95% bounds on estimation error for the composite estimate x^*_1 are ±2.8%, compared to ±3.5% for the mean ocular estimates, and ±4.9% for the sample estimate using 400 plots.

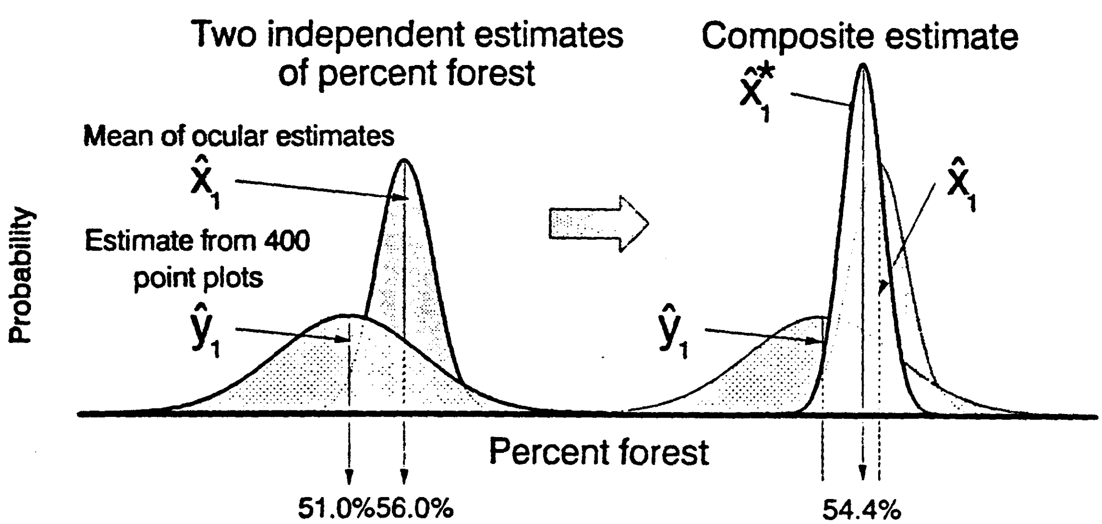

Figure 2. Probability densities for estimates of percent forest cover from mean ocular estimates (x_1=56.0%) and 400 point plots (y_1=51.0%). These are weighted inversely proportional to their variances, and combined into the composite estimate (x^*_1=54.4%).

Periodic forest inventories, inventory updates, or forest monitoring programs seek estimates of forest condition after changes have occurred. Fig. 3 shows the condition of forest cover at time $t=2$, after clearcuts have changed the hypothetical example in Fig. 1 at time $t=1$. Make another ocular estimate of percent forest cover in Fig. 3, now. Then, record your answer next to Fig. 3.

In Fig. 3, 200 temporary plots are independently classified to estimate percent forest cover at $t=2$; 90 plots have forest cover, for an estimate $y_2=45.0\%$, with $Var(y_2)=(45.0)(55.0)/200=12.38\%\%$, and $CI_{95\%}(y_2)=45.0\pm6.9\%$.

Another estimate of percent forest x_2 at time $t=2$ can be made from prior estimate $x^*_1=54.4\%$, given an estimated rate of change. It is predicted, from historical trends or a deterministic model, that 5% of all forest cover in Fig. 1 is clearcut between $t=1$ and $t=2$; therefore, the expected percent forest at $t=2$ is $x_2=0.95(x^*_1)=0.95(54.4\%)=51.7\%$.

If this linear transformation x_2 of x^*_1 is a perfect prediction model, variance of the propagated estimation error at $t=2$ is simply $Var(x_2)=(0.95)^2 var(x^*_1)=(0.90)2.08\%\%=1.88\%\%$. However, models are imperfect; prediction errors, denoted w, occur; and linear transformation $x_2=0.95x^*_1+w$ is a more realistic prediction model. If prediction error is unbiased, i.e., the expected value $E[w]$ is zero, then the updated estimate x_2 is unaffected.

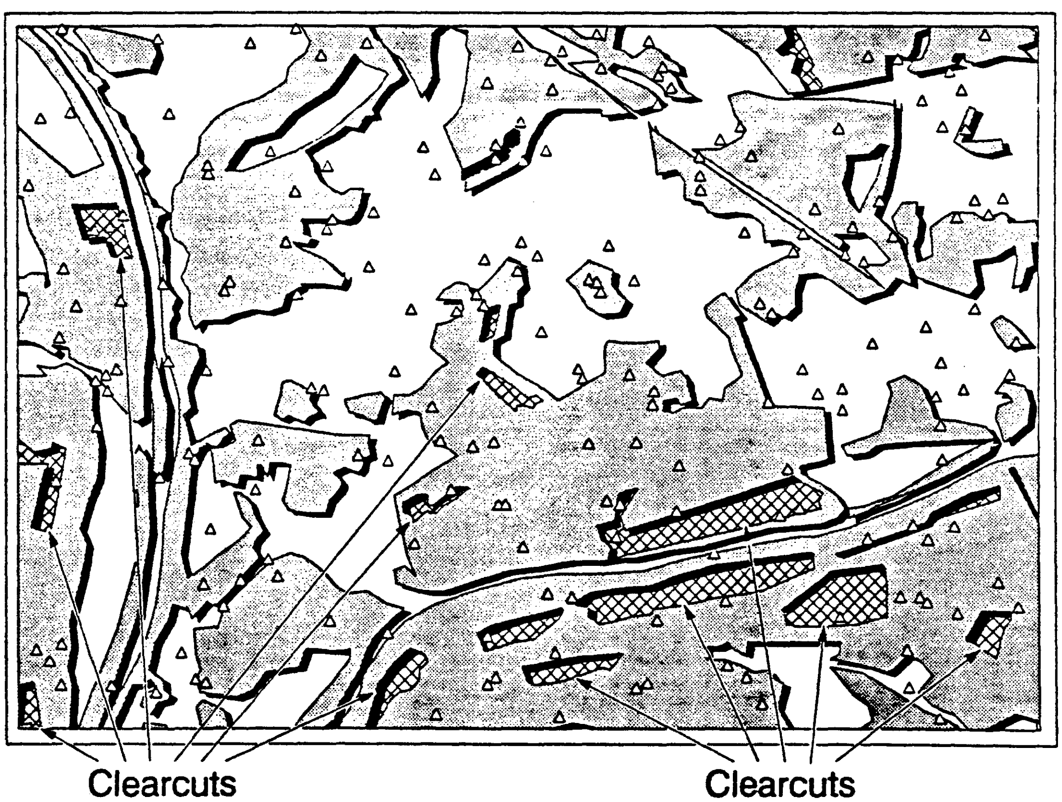

Figure 3. Between times $t=1$ and $t=2$, clearcuts reduce forest cover. In a monitoring system, 200 point plots are used at $t=2$ to independently estimate percent forest cover.

If prediction errors w are independent of errors for x^*_1, then $Var(x_2)=(0.95)^2 Var(x^*_1)+Var(w)$. If $Var(w)$ is assumed to be 1.00%%, then $Var(x_2)$=1.88%%+1.00%%=2.88%%, which yields an updated estimate x_2 with $CI_{95\%}(x_2)=51.7\pm3.3\%$.

The Kalman filter is merely a different linear transformation (x^*_2) of the estimates y_2 and x_2, with weights (A_2) inversely proportional to their variances (Fig. 4):

$$\begin{aligned} A_2 &= Var(x_2)/[Var(x_2)+Var(y_2)] \\ &= 2.88\%\% \,/\, (2.88\%\% + 12.38\%\%) \\ &= 0.19 \end{aligned}$$

$$\begin{aligned} x^*_2 &= [A_2\, y_2] + [(1 - A_2)\, x_2] \\ &= [(0.19)\, 45.0\%] + [(0.81)\, 51.7\%] \\ &= 50.4\% \end{aligned}$$

$$\begin{aligned} Var(x^*_2) &= [A_2^2\, Var(y_2)] + [(1 - A_2)^2\, Var(x_2)] \\ &= [(0.0361)\, 12.38\%\%] + [(0.6561)\, 2.88\%\%] \\ &= 2.34\%\% \end{aligned}$$

Approximate 95% bounds on estimation errors for the Kalman composite estimate are $\pm3.0\%$, compared to $\pm6.9\%$ for the sample estimate y_2, and $\pm3.3\%$ for model-updated estimate x_2.

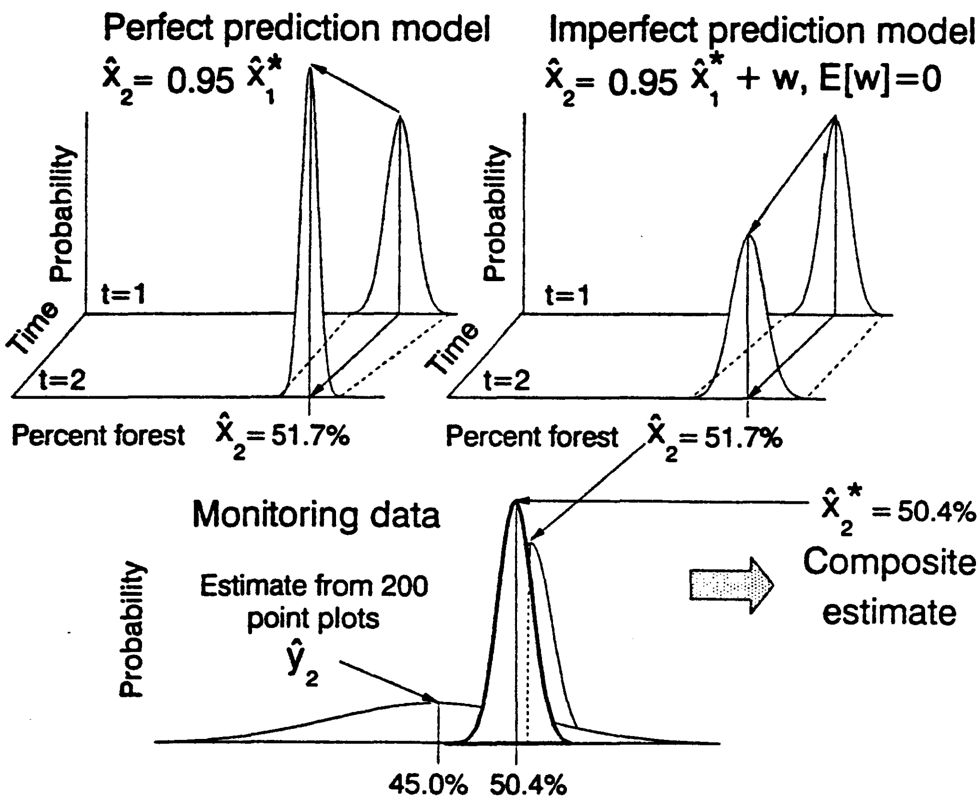

Figure 4. Probability densities for estimates of percent forest cover at time $t=2$, made from a prior estimate at time $t=1$, given model $x_2=(0.95)x^*_1=51.7\%$ (5% of the forest is expected to be clearcut between $t=1$ and $t=2$). Given a perfect prediction model, only the estimation error at $t=1$ is propagated to $t=2$. More realistically, the model is imperfect, and an unbiased prediction error (w) also occurs. An independent estimate y_2 is available from the 200 plots in Fig. 3. The Kalman filter combines these two independent estimates into a composite estimate (x^*_2), with weights inversely proportional to their variances.

TIME-SERIES OF MONITORING DATA

The Kalman filter is usually applied to a time series of monitoring measurements. With each new monitoring measurement, a composite estimate is made, which serves as new initial conditions for the next deterministic prediction (e.g., Fig. 5, year 4). Monitoring measurements can adjust predictions from a simple linear model for trends that are truly nonlinear, but are not well quantified. Precise data from the past can improve current estimates using less precise, but more recent, monitoring data. The Kalman filter can combine monitoring data from many sources, e.g., remote sensing or severance tax records.

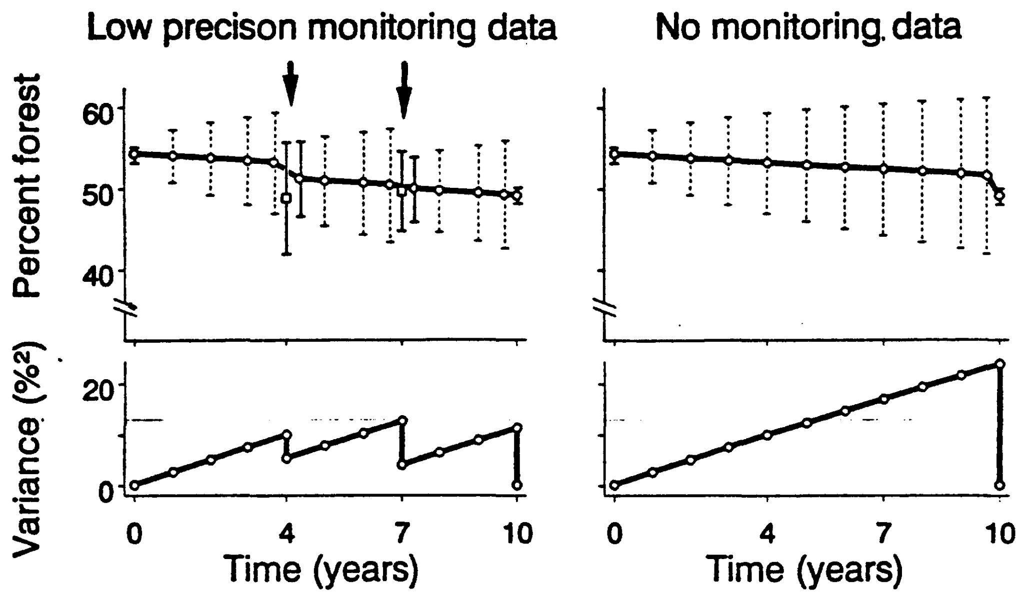

Figure 5. Kalman estimates and approximate 95% confidence intervals. Forest inventories were conducted in years 0 and 10; monitoring data were gathered in years 4 and 7. A time series of relatively imprecise (i.e., inexpensive) monitoring data can prolong utility of a previous, more expensive forest inventory.

TUTORIAL

To better understand the basic functions of the Kalman estimator, recompute the estimates above, using your ocular estimate of percent forest in Fig. 1, which you recorded in Table 1, as follows:

1. Compute a new mean of the ocular estimates x_1, and its variance of the mean $\mathrm{Var}(x_1)$, using $n=10$ in Table 1.

2. Recompute x^*_1, which combines this new x_1 with the sample estimate from the 400 point plots: $y_1=51.0\%$, $\mathrm{Var}(y_1)=6.25\%\%$.

3. Apply the prediction model $x_2=(0.95)x^*_1$, and the estimator for $\mathrm{Var}(x_2)$, where $\mathrm{Var}(w)$ is given as $1.00\%\%$.

4. Combine this updated estimate (x_2) with the sample estimate at $t=2$ [$y_2=45.0\%$, $\mathrm{Var}(y_2)=12.38\%\%$] to compute x^*_2.

The resulting composite estimate (x^*_2) is also the Kalman filter estimate. Compute an estimate of percent forest at $t=3$, given that 470 plots were forested in a sample of 1000 point plots at $t=3$, and using the same model for change (i.e., 5% of forest cover is clearcut between $t=2$ and $t=3$, with no regeneration). Repeat 1 to 4 above using only your ocular estimate (x_1) of Fig. 1, with $\mathrm{Var}(x_1)=28\%\%$ from Table 1, and vary the estimate of model prediction error dispersion, e.g., $\mathrm{Var}(w)=5\%\%$ or $10\%\%$. Assume ocular estimates are biased, and the true percent forest (x_1) is related to your ocular estimate (x_0) by the known linear model $x_1=0.85(x_0)+2\%$; recompute the composite estimate x^*_1 using only your ocular estimate (x_0) for Fig. 1. Assume the model that predicts x_1 from your x_0 is not precisely known, but is estimated by linear regression using a finite sample of reference sites ($n=100$) from a calibration experiment, in which residual mean square error is 30%%, mean of the 100 ocular estimates x_0 is 50.5%, and sum of $(x_0\text{-}50.5\%)^2$ in the experiment is 84000%%; recompute composite estimate x^*_1 using your calibrated ocular estimate x_1 of x_0. The true percent forest (usually unknown) is 54.68% at $t=1$ (Fig. 1), not including shadows, 49.65% at $t=2$ (Fig. 3), and 45.08% at $t=3$.

VERIFICATION

It is possible that two independent estimates disagree, or "diverge", in that neither estimate is likely given the other (Fig. 6). Contradictory estimates can be combined, but the result can be biased. Contradictions are probably caused by problems in estimating the error distribution of (1) the current monitoring measurement, or (2) the past estimate that is updated by a deterministic prediction model.

Concerning (2) above, it is very difficult to estimate the variance of model prediction errors $\mathrm{Var}(w)$; the estimate $\mathrm{Var}(w)=1.00\%\%$ in the hypothetical example above is quite arbitrary. Accurate estimates of $\mathrm{Var}(w)$ 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, estimates of $\mathrm{Var}(w)$ until disagreements are within acceptable bounds (Fig. 6), often using a time series of residuals $y_t\text{-}x_t$ (Sorenson 1985). Some Bayesian techniques assume model prediction errors are biased, and choose a weight or shrinking coefficient that minimizes a risk function.

It is also possible that bias exists in the measurement model. For example, different photointerpreters are rarely unbiased, and never have identical error distributions. Calibration experiments for each interpreter are needed to more realistically model such measurement errors.

The composite estimator assumes the two prior estimates are independent; however, their errors might be correlated. If your estimate of percent forest in Fig. 3 was influenced by any estimate of Fig. 1, then your measurement errors are correlated. The Kalman filter can treat correlated errors, but covariance estimates are required. Temporal or other patterns in standardized residuals from the Kalman filter can be used to detect such problems.

Negative variance estimates and asymmetrical covariance matrices can occur with the Kalman filter, especially when multivariate state vectors x_t are estimated and there are large differences between covariance matrices for multivariate x_t and y_t. Solutions to these numerical problems abound in the engineering literature, with the "square root" filter being frequently employed (Maybeck 1979).

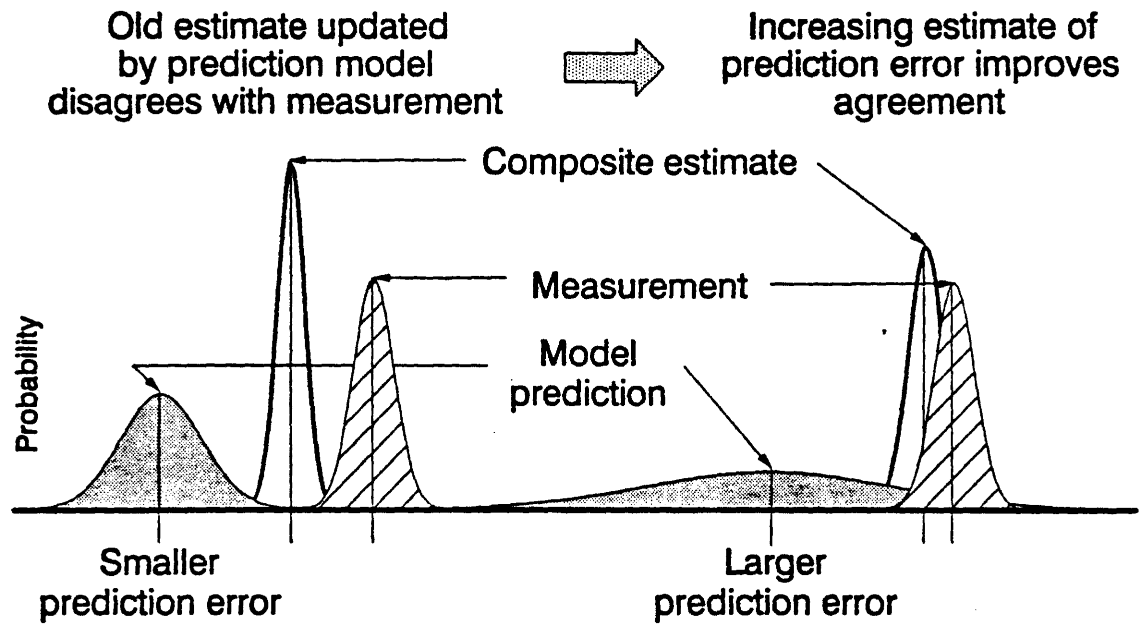

Figure 6. Predicted probability densities for two independent estimates (measurement y_t and model prediction x_t) that disagree; the residual difference between the two estimates is unlikely. This is probably caused by an unanticipated bias in the estimated error distribution for the measurement or model update. Adaptive filters assume that the estimated variance of model prediction errors Var(w) is inaccurate, and Var(w) is changed until the disagreement is within acceptable bounds.

CONCLUSIONS

The Kalman filter can be used to combine a time series of forest inventories, updates, or monitoring estimates using a model of expected change in forest condition over time. Many independent sources of data and knowledge can be combined, as discussed by Czaplewski et al. (1988). This can afford substantial efficiencies, and more timely estimates, compared to any one source of information by itself.

The Kalman filter is not complicated, but its application can be challenging. Accurate models for measurement and prediction errors are needed. Model formulation, parameter estimates, and assumptions must be verified; if residuals fail verification tests, then sound technical judgment is required to diagnose and cure any problems. However, solution of such problems can improve analysts' understanding of forest dynamics and the measurement processes.

REFERENCES

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.

Green, E.J. and Strawderman, W.E. 1986: Reducing sample size through the use of a composite estimator: an application to timber volume estimation. Can. J. For. Res., vol. 16, no 5, pp. 1116-1118.

Gregoire, T.G. and Walters, D.K. 1988: Composite vector estimators derived by weighting inversely proportional to variance. Can. J. For. Res., vol. 18, no. 2, pp. 282-284.

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

Sorenson, H.W. 1895. Kalman filtering: Theory and Applications. IEEE Press Inc., New York. 457pp.

Thomas, C.E. and Rennie, J.C. 1987: Combining inventory data for improved estimates of forest resources. South. J. Appl. For., vol. 11, no. 3, pp. 168-171.

Ware, K.D. and Cunia, T. 1962: Continuous forest inventory with partial replacement. For. Sci. Monogr. No. 3.