Propagating probability distributions of stand variables using sequential Monte Carlo methods

JEFFREY H. GOVE

United States Department of Agriculture Forest Service, Northern Research Station, 271 Mast Road, Durham, NH 03824, USA
E-mail: jgove@fs.fed.us

Summary

A general probabilistic approach to stand yield estimation is developed based on sequential Monte Carlo filters, also known as particle filters. The essential steps in the development of the sampling importance resampling (SIR) particle filter are presented. The SIR filter is then applied to simulated and observed data showing how the 'predictor–corrector' scheme employed leads to a general probabilistic mechanism for updating growth model predictions with new observations. The method is applicable to decision making under uncertainty, where uncertainty is found in both model predictions and inventory observations.

Introduction

Forest managers are frequently faced with the prospect of making decisions using imperfect data. Measurements and model predictions can contain error from numerous sources, including sampling, measurement, classification, model estimation and misspecification, to name but a few. These errors may act in ways that are often difficult to quantify and may affect the decision variables in a variety of ways both linearly and non-linearly. In the formulation of optimal planning models for decision making, the manager must also determine whether the sources of uncertainty may accumulate additively or multiplicatively in the system under consideration. This need to incorporate uncertainty into the decision-making process has been recognized and numerous methods have been developed to accommodate it into optimal planning models (Hof *et al.*, 1988; Pickens and Dress, 1988; Gove and Fairweather, 1992; Kangas and Kangas,

1999). One of the major sources of uncertainty in the decision-making process comes from the error in the estimation of stand variables from inventories and model projections. Other sources of uncertainty often include those concerning future prices, interest rates, and changing ownership patterns. While these other sources of error are important, here we concentrate on the evolution of uncertainty in estimates of stand variables through time. That is, the quantification of the probability distribution of stand variables through the combination of model predictions and inventory adjustments over time. Such information not only plays an important role in optimal planning efforts but also can be used in a variety of other less rigorous decision processes such as the determination of stand stocking levels for silvicultural activities and the scheduling of stand examinations.

Much effort has gone into the problem of the estimation of forest stand variables and other quantities in a way that allows for an assessment

of the error in the estimates to be approximated as well. Traditional design-based sampling methods allow for the estimation of sampling error associated with means and totals in the form of standard errors (SEs). Similarly, simulation approaches might be used, for example to quantify the error in predictions from growth models. Too often, however, the steps of stand projection via models and sampling are thought of as disjoint events. And this is where the efforts of compositing and similar strategies have attempted to address merging both model and sampled estimates (Burk *et al.*, 1982; Fairweather and Turner, 1983; Green and Strawderman, 1988). The underlying theme of these efforts is to derive a melded estimate with associated error approximation resulting from the combination of the two different estimates.

In a more general approach to the quantification of error, the entire probability distribution of a stand variable could be estimated. This is the topic of probabilistic state estimation, which was recognized early on by engineers. In the evolution of the state through time, assumptions concerning Gaussianity and linearity are often made, and the methods that have been developed largely are used to propagate the mean and covariance of some vector of state variables. In so doing, these filtering methods seamlessly composite model predictions with incoming measurements in a two-step process. When the Gaussian assumption is tenable, the propagation of the mean and covariance establish the joint probability distribution of the state. When it is not tenable, there will be some associated bias in the estimates. It would make sense in such cases to actually estimate the form of the distribution itself, not just the mean and variance; other methods exist in such cases as will be detailed below.

Propagating the stand distribution through time has the added benefit that it can be used in Bayesian decision making. Ståhl *et al.* (1994) recognized the utility of such an approach by developing a probability-based technique for incorporating inventory planning into the silvicultural decision-making process. Underlying the planning component was a probabilistic framework for propagating the mean and variance of the state through time, where a Gaussian assumption was made on all distributions. These authors used Bayes' rule to incorporate inventory information

with assumed sampling error, into a prediction prior using a discretization process. In a related study, Nyström and Ståhl (2001) showed how different components of error could be incorporated into the state density propagation employing a Monte Carlo simulation approach. They also noted that measurements could be incorporated through Bayes' rule and illustrated the conceptual effect this can have in decision making. However, both of these studies appear to be limited in that the methods used may not generalize well. For example, Ståhl *et al.* (1994) note that using distributions other than the Gaussian would quickly lead to intractable calculations. Similarly, it is unclear how Bayes' rule might be employed to perform a measurement update on a prediction prior composed of thousands of Monte Carlo simulated predictions as in Nyström and Ståhl (2001), unless perhaps some parametric form were assumed and estimated. Moreover, as mentioned by Ståhl *et al.* (1994), the generalization of these methods to more than one state variable may present computational challenges.

In this paper, a general state space framework for estimating forest growth and yield in a probabilistic context is introduced that allows for nonlinear, non-Gaussian assumptions, as well as integration of new inventory information with model predictions. With regard to forest stands, the state of the system refers to any variables that are recognized as being important to the quantification and classification of the stand under consideration for estimation purposes. The system state might include such stand-level variables as basal area, number of stems, volume and biomass over all species, or broken down by species; a more detailed state vector might also include parameters for quantifying the stand diameter distribution, for example. Stand-level projection systems might employ these state variables as both dependent and independent variables, and sampling methods would target these state variables to be estimated from an inventory, either directly or indirectly.

The state space approach allows the time course of measurements and model predictions to be incorporated together in a sequential manner, taking into account both model and sampling errors. In this type of system, the disconnect of inventories and models is not present, since the two get fused as part of the sequential march through time. The

general framework that allows this probabilistic data–model fusion stems largely from the engineering literature on filtering and smoothing. Engineering applications have long taken the probabilistic approach to state estimation. The most well-known of the filtering methods is the Kalman filter (Kalman, 1960), which sequentially propagates the mean and covariance of the system state through a two-step process of prediction and update. The prediction step applies the process models to the prior state estimate, while updating corrects this prediction when the new measurement arrives. This 'predictor–corrector' structure of the Kalman filter is found in many subsequent filters of this class. The Kalman filter does not specifically place a Gaussian assumption on the states; however, the fact that it propagates the mean and covariance of the state would suggest that the closer the underlying state to a Gaussian, the better. A more formidable restriction to the Kalman filter is that of linearity: both measurements and process model predictions have to be a linear function of the states.

Because many real-world problems are non-linear and non-Gaussian, numerous extensions or new filters have been developed over the years to better handle such problems. Some examples include the extended (Gelb, 1974, p. 182) and unscented Kalman filters (Julier and Uhlmann, 2004). Shortly after the Kalman filter's introduction, Ho and Lee (1964) showed how the general non-linear non-Gaussian filtering problem could be formulated in a Bayesian setting. Unfortunately, there are only a few special cases to this general filtering problem that have known solutions. This is because it entails knowing the exact form of the distributions involved in order to solve the resulting integrals. Recently, a new class of filters has been put forth that relies on a combination of importance sampling and Monte Carlo simulation to circumvent the integration difficulties encountered in such problems. This class of filters, collectively known as particle filters, provide a straightforward solution to the problem of propagating probability distributions of stand variables through time.

In this paper, we present the probabilistic formulation to the filtering problem and its solution using one particular form of particle filter known as the bootstrap, or sampling importance resampling (SIR) filter. Examples are provided showing the details of the filter's application to a problem concerning the estimation of basal area yield of eastern white pine (*Pinus strobus* L.).

Particle filtering

Particle filtering is a general state space method for the sequential estimation problem of assimilating model predictions with incoming measurements. Let the unobserved system state be given by the vector $\mathbf{x}_t$ with dimension $(n_x \times 1)$, for all time $t = 0, \ldots, T$. The sequence of states is assumed to be a first-order Markov process, which implies that the entire past history of information in $\mathbf{x}_t$ is contained in $\mathbf{x}_{t-1}$. More formally, the distribution of the state $\mathbf{x}_t$ conditional on the entire history of states $\mathbf{x}_{0:t-1} = \mathbf{x}_0, \ldots, \mathbf{x}_{t-1}$ is completely determined by $\mathbf{x}_{t-1}$. Assume that the process dynamics are given by $\mathbf{f}_t$, which can be non-linear, then the model governing the evolution of the states is

$$\mathbf{x}_t = \mathbf{f}_{t-1}(\mathbf{x}_{t-1}, \mathbf{v}_{t-1}), \qquad (1)$$

where $\mathbf{v}_t$ is the process noise. The process noise allows, for example for process model misspecification, estimation error or other disturbances in the process model.

Similarly, let $\mathbf{y}_t$ $(n_y \times 1)$ be the measurements or observations at time t. Note that in general, $\mathbf{y}_t$ does not have to be present at each time period, nor does it have to arrive in the estimation process at periods of equal intervals. This will become clear in the next section. If the measurement model is given by $\mathbf{h}_t$, which may also be non-linear, then the state is related to the measurements via the measurement equation

$$\mathbf{y}_t = \mathbf{h}_t(\mathbf{x}_t, \mathbf{n}_t), \qquad (2)$$

where $\mathbf{n}_t$ is the measurement noise. The measurement noise not only might be composed of measurement error but also additionally could include components of sampling, or classification error, to name a few.

Equations (1) and (2) make up the dynamic state space formulation of the problem. Note that the only assumptions are that both of these noise processes are white (a white noise process is one in which the random variables are mutually independent (Gelb, 1974, p. 42)) and that their probability density functions (PDFs) are known

and can be sampled from. Note particularly that the noise random variables are not required to be Gaussian.

The random noise components to the models above allow us to state the general state space model in probabilistic terms as

$$p(\mathbf{x}_t|\mathbf{x}_{t-1}) \text{ transition density} \qquad (3)$$

$$p(\mathbf{y}_t|\mathbf{x}_t) \text{ measurement density (likelihood)}, \quad (4)$$

with $p(\mathbf{x}_0) = p(\mathbf{x}_0|\mathbf{y}_0)$ as the initial condition. Since both the process dynamics and measurement equations are stochastic, it is not possible to infer the exact state from the measurements. Therefore, the goal is to estimate the posterior distribution $p(\mathbf{x}_t|\mathbf{y}_{1:t})$ at time t from all the measurements $\mathbf{y}_{1:t}$.

The general Bayesian solution

The general Bayesian filtering problem can now be formulated. As with the Kalman filter, it consists of prediction and update steps. First, the prediction step yields an estimate of the dynamic prior, or prediction density at time t, based on all of the measurements through time $t - 1$ (Gordon et al., 1993)

$$p(\mathbf{x}_t \mid \mathbf{y}_{1:t-1}) = \int p(\mathbf{x}_t \mid \mathbf{x}_{t-1}) p(\mathbf{x}_{t-1} \mid \mathbf{y}_{1:t-1}) \mathrm{d}\mathbf{x}_{t-1}. \quad (5)$$

In the update step, the measurement is assimilated with the prediction density through the application of Bayes' rule, yielding the posterior at time t. It can be shown that (e.g. Simon, 2006, p. 464)

$$p(\mathbf{x}_t \mid \mathbf{y}_{1:t}) = \frac{p(\mathbf{y}_t \mid \mathbf{x}_t) p(\mathbf{x}_t \mid \mathbf{y}_{1:t-1})}{p(\mathbf{y}_t \mid \mathbf{y}_{1:t-1})}, \qquad (6)$$

where the normalizing density in the denominator is often termed the data evidence and depends on the likelihood (4), and the dynamic prior (5); that is $p(\mathbf{y}_t \mid \mathbf{y}_{1:t-1}) = \int p(\mathbf{y}_t \mid \mathbf{x}_t) p(\mathbf{x}_t \mid \mathbf{y}_{1:t-1}) \mathrm{d}\mathbf{x}_t$.

In the linear, Gaussian case ($\mathbf{f}_t$ and $\mathbf{h}_t$ linear; $\mathbf{v}_{t-1}$, $\mathbf{n}_t$ additive Gaussian), it can be shown that the Kalman filter is the solution to this filtering problem given in equations (5) and (6). Unfortunately, a general solution to the above Bayesian filtering problem is not available because the integrals involved are generally high dimensional and

complex; this has spawned a number of Bayesian algorithms that are suboptimal approximations (Arulampalam et al., 2002; Ristic et al., 2004, pp. 6–8). A more general approach is to use Monte Carlo integration methods to evaluate the intractable integrals in equations (5) and (6). The Monte Carlo approach has the advantage that it can be applied in non-linear, non-Gaussian state space models.

Sequential importance sampling

In what follows, the general idea of the presentation is to motivate the concept of a particle filter through its development from the Bayesian filtering problem above using sequential importance sampling (SIS). The details are lengthy, and therefore are omitted; the interested reader should consult such references as Doucet et al. (2001), Ristic et al. (2004, Chapter 3), and Simon (2006, Chapter 15).

In order to establish an approximate solution to the conceptual Bayesian filtering problem in equations (5) and (6), importance sampling is used. Importance sampling is a Monte Carlo integration method that facilitates the approximation of general integrals (e.g. Rubenstein, 1981, p. 122). When the true distribution, $p(\mathbf{x}_t)$, is difficult to sample, a proposal distribution, $q(\mathbf{x}_t)$, is chosen whose support covers that of the true distribution, and from which it is easier to sample. The objective in filtering is to estimate equations (5) and (6), and expectations based on these densities sequentially, by propagating the appropriate PDFs through time.

We begin with the general importance sampling approach, and then show how this approach can be made sequential for filtering. In general, we would like to approximate the following expectation

$$\mathrm{E}\big[\mathbf{g}_t(\mathbf{x}_{0:t})\big] = \int \mathbf{g}_t(\mathbf{x}_{0:t}) p(\mathbf{x}_{0:t} \mid \mathbf{y}_{1:t}) \mathrm{d}\mathbf{x}_{0:t}, \qquad (7)$$

where $\mathbf{g}_t(\cdot)$ is a function, possibly non-linear, of the states. Employing importance sampling with the proposal density $q(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$, which covers the support of $p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$, yields

$$\mathrm{E}[\mathbf{g}_t(\mathbf{x}_{0:t})] = \int \mathbf{g}_t(\mathbf{x}_{0:t}) p(\mathbf{x}_{0:t} \mid \mathbf{y}_{1:t}) \frac{q(\mathbf{x}_{0:t} \mid \mathbf{y}_{1:t})}{q(\mathbf{x}_{0:t} \mid \mathbf{y}_{1:t})} \mathrm{d}\mathbf{x}_{0:t},$$

Applying Bayes' rule to $p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$ and some manipulation eventually yields the importance weights

$$w_t(\mathbf{x}_{0:t}) = \frac{p(\mathbf{y}_{1:t} \mid \mathbf{x}_{0:t})p(\mathbf{x}_{0:t})}{q(\mathbf{x}_{0:t} \mid \mathbf{y}_{1:t})}, \qquad (8)$$

allowing the expectation to be written as

$$E[\mathbf{g}_t(\mathbf{x}_{0:t})] = \frac{E_q[\mathbf{g}_t(\mathbf{x}_{0:t})w_t(\mathbf{x}_{0:t})]}{E_q[w_t(\mathbf{x}_{0:t})]},$$

where in each case $E_q[\cdot]$ means that we are taking the expectation with respect to the proposal distribution $q(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$.

The Monte Carlo approach to the evaluation of this expectation is to draw N independent, identically distributed samples from the proposal distribution to obtain a weighted approximation to this integral; viz.

$$\widehat{E}[\mathbf{g}_t(\mathbf{x}_{0:t})] = \sum_{i=1}^{N} \mathbf{g}_t(\mathbf{x}_{0:t}^i)\, \widetilde{w}_t^i, \qquad (9)$$

where the normalized weights are given as

$$\widetilde{w}_t^i = \frac{w_t^i(\mathbf{x}_{0:t})}{\sum_{i=1}^{N} w_t^i(\mathbf{x}_{0:t})} \qquad (10)$$

It should be clear from the form of equation (9) that the weights (10) form a discrete representation of the posterior $p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$, since the entire quantity is an approximation to equation (7); this will be formalized momentarily. However, the general importance sampling procedure outlined thus far is not a sequential, or recursive, filtering solution. To make this outline of the general procedure sequential, assume that the proposal can be factored as follows

$$q(\mathbf{x}_{0:t}|\mathbf{y}_{1:t}) = q(\mathbf{x}_{0:t-1} \mid \mathbf{y}_{1:t-1})q(\mathbf{x}_t \mid \mathbf{x}_{0:t-1}, \mathbf{y}_{1:t})$$

where the rightmost density on the right-hand side represents an adjustment to the existing path. Then it is straightforward to show by iteration that

$$q(\mathbf{x}_{0:t} \mid \mathbf{y}_{1:t}) = q(\mathbf{x}_0)\prod_{k=1}^{t}q(\mathbf{x}_k \mid \mathbf{x}_{0:k-1}, \mathbf{y}_{1:k}).$$

In addition, the other PDFs in equation (8) can be factored similarly. From these results, it can be shown that the importance weights can also be written recursively (Arulampalam et al., 2002) as

$$w_t^i \propto w_{t-1}^i \frac{p(\mathbf{y}_t \mid \mathbf{x}_t^i)p(\mathbf{x}_t^i \mid \mathbf{x}_{t-1}^i)}{q(\mathbf{x}_t^i \mid \mathbf{x}_{t-1}^i, \mathbf{y}_t)} \qquad (11)$$

The sketch of derivation given above yields the components necessary for a particle-based approximate solution to the sequential filtering problem. In the Monte Carlo particle context, the state space is represented by N particles, or support points $\{\mathbf{x}_t^i, i = 1, \ldots, N\}$ at time t. The normalized weights $\{\widetilde{w}_t^i, i = 1, \ldots, N\}$ represent the contribution to the discretization of the posterior of each support point at time t. The set of support points and associated weights determines a random measure $\{\mathbf{x}_t^i, \widetilde{w}_t^i\}_{i=1}^{N}$. The posterior, therefore, is represented by (Arulampalam et al., 2002; Ristic et al., 2004, p. 39)

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) \approx \sum_{i=1}^{N} \widetilde{w}_t^i\, \delta(\mathbf{x}_t - \mathbf{x}_t^i), \qquad (12)$$

where $\delta(\cdot)$ is Dirac's delta (Maybeck, 1979, p. 84) and the weights, $\widetilde{w}_t^i$, have been normalized via

$$\widetilde{w}_t^i = \frac{w_t^i}{\sum_{i=1}^{N} w_t^i}, \qquad (13)$$

such that $\sum_{i=1}^{N} \widetilde{w}_t^i = 1$. Furthermore, it can be shown that as $N \to \infty$, the posterior approximation (12) approaches the true posterior density (Crisan and Doucet, 2002).

The general SIS algorithm consists of simply (a) drawing samples $\mathbf{x}_t^i \sim q(\mathbf{x}_t^i \mid \mathbf{x}_{t-1}^i, \mathbf{y}_t)$, (b) calculating the weights via equation (11), and (c) normalizing the weights via equation (13). Steps (a)–(c) are conducted for each particle at each time step, and the entire algorithm is applied sequentially through time. Therefore, the SIS filter algorithm can be seen to propagate the support points and associated weights (i.e. the random measure $\{\mathbf{x}_t^i, \widetilde{w}_t^i\}_{i=1}^{N}$), and thus the discrete estimate of the posterior (12), through time.

The SIR filter

The SIS algorithm has been known for some time. Unfortunately, it has also been known that in application, the variance of the importance weights will increase over time in the SIS algorithm, leading to degeneracy. The result of degeneracy is that

one importance weight will be nearly equal to one, while all of the others will be approximately zero, yielding a poor representation of the posterior. However, a resampling step can be judiciously inserted into the SIS algorithm when certain criteria are met in order to counteract degeneracy (Ristic *et al.*, 2004, pp. 40–41). The main idea of resampling is to multiply particles with high importance weights and eliminate those with small importance weights.

Gordon *et al.* (1993) developed the SIR filter as an extension to the SIS algorithm by resampling at every time step and using the transition density as the importance proposal density. This algorithm was originally called the 'bootstrap' filter because it employed a multinomial resampling scheme, which is equivalent to drawing a weighted bootstrap resample, with weights given by the normalized weights (13). There are many different ways to resample the particles (Hol *et al.*, 2006); however, in this paper, we use the original multinomial scheme. Conceptually, in the resampling step, an integer number of particles, N_t^i, are drawn based on the magnitude of the weight at each support point: if $N_t^i = 0$, the particle is pruned, while   connotes replication. More precisely, the weighted measure   is replaced by the uniform measure   such that   ; therefore, the weights are all equal post-resampling.

Table 1 presents the SIR filter algorithm in pseudo-code. Notice in particular, that when the transition density is used as the proposal density, the importance weights (11) become simply   . Note also that the weights are all uniform after resampling, so   does not appear in this equation. Since the form of the likelihood is known by assumption, calculation of the weights in the importance sampling step is straightforward. In the sampling step, a sample   may be drawn as follows: (a) draw a sample   from the process noise distribution, then (b) apply the process model (1) as   (Ristic *et al.*, 2004, p. 48). This SIR algorithm is used as the non-linear filter in the yield analysis that follows.

Finally, approximating integral quantities such as expectations based on the posterior at any given time period becomes simple under the SIR

Table 1: SIR particle filter algorithm

for i = 1:N   *Initialize the filter*
$\sim p(\mathbf{x}_0)$
for t = 1:T {   *For each time period*
for i = 1: N {
Sampling step…
Draw
Importance sampling step…
Calculate
}
for i = 1: N
Normalize the weights via (13)
Resampling selection step…
Resample from
}

filter. For example, to estimate the general expectation at time t,

$$\tag{14}$$

one would use the approximation based on equation (9)

$$.\tag{15}$$

Note that such quantities are calculated based on the weighted approximation to the posterior (12) prior to resampling (Ristic *et al.*, 2004, p. 44).

Filtering stand yield

In this section we show how the SIR particle filter can be applied to the estimation of stand yield. In this problem, the system state is the yield at time t, and the process dynamics are stand growth. Measurements enter the system from an inventory. The SIR filter propagates the joint posterior distribution in the form of equation (12), through time. It is important to note that the system state can be multivariate, as outlined in the problem statement. For example, it could be a simple compatible growth and yield system in the spirit of Clutter (1963), or something more complicated with more state variables. However, because it is helpful to the understanding of the filter implementation, here we restrict attention to one state variable, basal area.

Frothingham (1914) presents several yield tables for fully stocked eastern white pine stands

in different quality classes. A Chapman–Richards yield model (Pienaar and Turnbull, 1973) was fitted to the basal area yields in the lowest quality class given (Frothingham, 1914, Table 5). The model is

where t_0 is the initial time and A, k and m are parameters. The fit of the Chapman–Richards yield model to the Frothingham yield data (not shown) was nearly perfect. The differential form of this yield equation is given as

where the parameters   and are determined from the parameter estimates of the fitted yield curve; viz.   ,   and   .

The state space formulation of the dynamics equations given in equations (1) and (2), can be applied to this model as

where the measurement equation is a simple linear random walk and the vector state notation has been retained for consistency, even though the current problem is scalar. To complete the probabilistic formulation of the state space model as given in (3) and (4), the errors are assumed to be normally distributed, yielding

$$\tag{16}$$

$$\tag{17}$$

where $\mathbf{Q}_{t-1}$ and $\mathbf{R}_t$ are the associated variances (or, more generally covariance matrices when $n_x > 1$ or $n_y > 1$).

Simulated yield example

In order to apply the particle filter with the model given by equations (16) and (17), the variances in both densities as well as the initial condition require quantification. The SE from the non-linear least squares fit of the Chapman–Richards model can be used for the process noise variance in equation (16); that is $\mathbf{Q}_t = 32$, $t = 1, …, T$. The measurement variance could come from the SE of

a forest inventory, such as a periodic stand assessment, and need not be from permanent plots. For this example, we simply set $\mathbf{R}_t = 100$, $t = 1, …, T$, yielding an approximate error of 10 ft^2 ac^{-1} (2.3 m^2 ha^{-1}). In both instances, we have made the simplifying assumption that the variance is constant through time; however, this is not necessary as will be made clear later, but is only done for expedience. The initial condition was drawn from a density according to the yield equation with appropriate age, corrupted by a Gaussian disturbance with the variance equal to the process noise variance.

Figure 1 presents a single 50-year simulation generated using the state space model described above. The circles denote the true system state in terms of basal area, the quantity we are trying to estimate and is normally (outside of simulation) unknown. Also shown are the estimates from the SIR filter using $N = 1000$ particles. The particle approximation of the posterior given in (12) is used with equation (15) to calculate filter-based estimates such as the mean and variance. In Figure 1, the weighted mean estimate is shown as the solid line, while the maximum *a posteriori* (MAP) estimate is represented by the dot–dash line. The MAP estimate is seen to be almost indistinguishable from the mean largely due to the Gaussian assumptions and the mildly non-linear form of the process dynamics. The MAP estimate was calculated by fitting a weighted kernel density estimator with Gaussian kernel (Silverman, 1986, p. 43) to the posterior representation (12) at each time period. The shaded area shows the 95% posterior or credible intervals, as estimated from the approximation (12). As mentioned earlier, Ristic *et al.* (2004, p. 44) note that all such quantities should be calculated prior to resampling, as done here.

The top panel in Figure 1 shows the SIR filter results when yearly measurements are available. The simulated states are seen to wander about the model yield line due to the addition of process noise. The measurements behave similarly, with the addition of measurement (i.e. sampling) variability. The objective in filtering is to estimate the unknown states using noisy measurements. We consider the filter successful if the posterior intervals capture the unknown states. In this example, this happened 90% of the time (five were missed). The mean estimates from the filter are 'corrected'
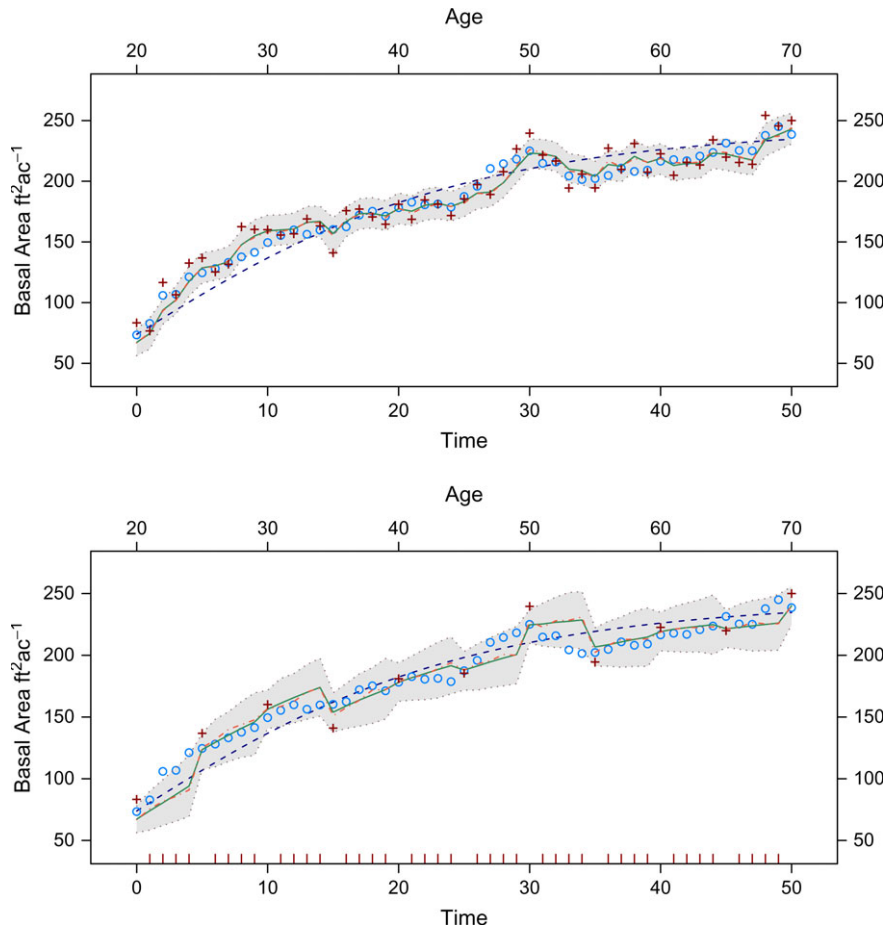
Figure 1. Trajectory of simulated white pine basal area yield. In both panels, the measurements are denoted by the plus, while the true states are circles. The shaded area marks the 95% credible intervals. The solid line is the SIR filter estimate of the mean and the dot–dash line is the MAP estimate based on $N = 1000$ particles. The deterministic yield equation B_t is shown as the long-dash line. In the top panel, measurements are yearly, while in the bottom they are periodic at 5-year intervals, with bars denoting missing measurements. (Note: 1 ft^2 ac^{-1} equals 0.2296 m^2 ha^{-1}.)

by the measurements, but only to a certain extent, which is based on the uncertainty in the likelihood. Consider, for example that what appears to be an aberrant measurement at $t = 15$ pulls the trajectory down, but does so only slightly relative to the measurement. This shows the filter's robustness to poor inventory information in such circumstances.

The correction mechanism just described can be illustrated by plotting all of the densities used in the filter at a given time step. This has been done for a more well-behaved measurement at $t = 30$ in

Figure 2. The likelihood is centred about the measurement at 238 ft^2 ac^{-1}, which is higher in basal area than almost the entire transition prior would represent. Indeed, the prior is centred in an area of very low likelihood. Because the likelihood establishes the weights in the SIR filter, the vast majority of the prior gets down-weighted by the low weights, while the upper tail of the prior receives higher weight. The result is the posterior $p(\mathbf{x}_t|\mathbf{y}_{1:t})$, which has shifted towards the observation. In this particular case, the shifting produced a correction in the direction of the true state (225 ft^2 ac^{-1}),
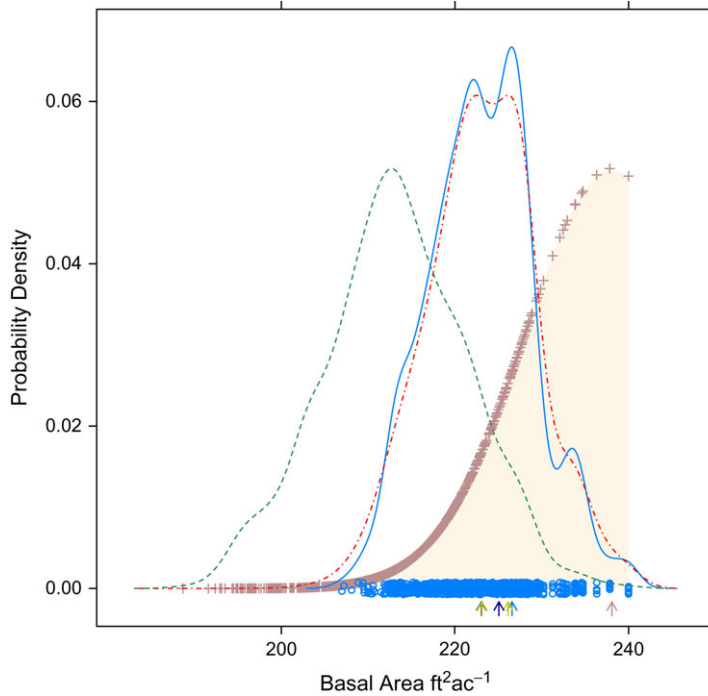
Figure 2. Illustration of the various SIR densities at time *t* = 30 from the top panel in Figure 1. A kernel density estimator was applied to the transition prior (dashed), posterior prior to resampling (dot–dashed) and the resampled posterior (solid), with the circles showing the resampled particles associated with this density. The likelihood is denoted as the shaded region and is plotted over the support of the prior (plus) and has been scaled to the prior for illustration. The leftmost two arrows (coincident) denote the means for both posterior representations; the next denotes the true state, with the following two denoting the weighted and resampled MAP estimates, respectively, and the rightmost arrow denoting the measurement. (Note: 1 ft^2 ac^{-1} equals 0.2296 m^2 ha^{-1}.)

which falls between the mean and MAP estimates for the posterior densities. Notice that resampling tends to exaggerate the modes and humps in the weighted representation of the posterior. Both representations of the posterior are approximately normal in overall shape, but have developed slight bimodality and humps due to the weighting of the prior. In cases where the likelihood and the prior are more closely aligned, the multiple modes will disappear and the posteriors will be approximately Gaussian in shape. Additionally, increasing the number of particles may tend to smooth the densities and more closely approach Gaussian in appearance. Notice also that the posterior has smaller variance than the prior: the standard deviation for the posterior is 6.4 ft^2 ac^{-1}while that for the prior is 7.4 ft^2 ac^{-1} (1.47

and 1.7 m^2 ha^{-1}, respectively). This is typically the case when the inventory provides relatively precise information.

Unfortunately, it would be a rare situation for forest managers to obtain yearly measurements from an annual inventory; more generally, measurements are available periodically. The bottom panel in Figure 1 shows a more realistic estimation run of the SIR filter, where periodic remeasurements enter the record every 5 years. In the absence of measurements, the SIR filter becomes a sequential reapplication of sampling from the prior (Doucet *et al.*, 2000). This is because the rest of the algorithm depends on the weights, which cannot be calculated when measurements are missing (Table 1). There are two salient points to be noticed in this figure. First, as

the time increases since the last received measurement, the credible intervals (now based only on the transition prior) become wider. When a new measurement is recorded – a new inventory taken – the likelihood adjusts the prior density as discussed above and the posterior intervals narrow. Second, notice that in the absence of measurements, the filter mean and MAP estimates tend to track approximately parallel to the underlying yield model. An example of the value of the measurement correction occurs at $t = 35$. At $t = 31$, the true state begins a slow decline in basal area, but the filter continues to estimate on the high side, as there has been no external input signalling that any change to the system has occurred. The widening of the intervals still catches the true state until $t = 33$, when basal area has declined to a point where the intervals fail to cover the true state at the 95% level for the next two periods. The new measurement at $t = 35$ has appropriately adjusted the prediction trajectory, resulting in the true state's inclusion within the posterior interval again. The credible intervals caught the true state 92% of the time (four missed) over all time periods in this particular example.

The rates at which the posterior intervals captured the true states are slightly lower in the simulation above than the nominal coverage rate of 95%. To determine whether the number of particles used in the above simulation was adequate, a Monte Carlo experiment was performed. Series of length $T = 50$ were simulated 100 times for each of $N = 50, 100, 250$ and 500 particles. The results were coverage rates of 90.9%, 93.1%, 94.4% and 94.9%, respectively. Therefore, for this particular set of models and variance parameters, somewhere between $N = 250$ and 500 particles should result in approximate nominal coverage. Root mean square error analysis based on the same set of simulations supported this conclusion. The fact that the rates in the example above were slightly lower is simply due to a chance selection of the particular trajectory, since the number of particles ($N = 1000$) was more than adequate.

Updating permanent growth plot estimates

In this section, the SIR filter is applied to a single permanent growth plot with records dating back over 40 years, to illustrate filter estimation on real data. The plot is 0.1 acre (0.4 ha) in size, and is part of an eastern white pine growth study established in the 1950s throughout southern New Hampshire known as the Hatch Study. The Frothingham growth and yield formulation fitted previously is used again to show how the filter will correct aberrant predictions when modelled growth is somewhat optimistic. The same correction mechanism is at work here through the SIR filter as in the simulation example, but the true state of the stand basal area is now unknown.

The top panel in Figure 3 presents the trajectory. The figure shows the somewhat erratic remeasurement schedule adopted with anywhere from 1–6 years lapsing before subsequent remeasurements were taken. Note that the stand measurements plot well below the Frothingham yield curve, implying that the growth predictions will indeed be optimistic. The SIR filter was run on these data with the same assumptions as in the previous simulation; however, since there is only one plot, the measurement variance, $\mathbf{R}_t$, is assumed to be composed solely of measurement error. Regarding the filter mean and MAP estimates, the optimism of the Frothingham yield curve is seen to be especially true early in the trajectory where the slope of the yield curve is more severe. However, after $t = 26$, the growth of the stand and model predictions are more closely aligned. Notice that when a prediction is corrected by a new measurement, the posterior estimate is drawn towards the measurement, but they do not coincide. Ideally, the correction would be less dramatic with a more suitable growth model (e.g. in the previous simulation example). However, this example illustrates that even an optimistic growth model can be used in this setting, because an estimate of the probability distribution of the state is being propagated, providing a degree of belief in our estimate, rather than simply the stand mean from a deterministic growth projection.

Suppose, alternatively, that the perceived bias associated with the Frothingham growth equation is determined to be too extensive when applied to these data. In this case, the process noise component can be adjusted in an attempt to correct the bias. For example, the bottom panel in Figure 3 shows a run of the SIR filter with everything the same, except that the process error distribution is now assumed to be   . The variance
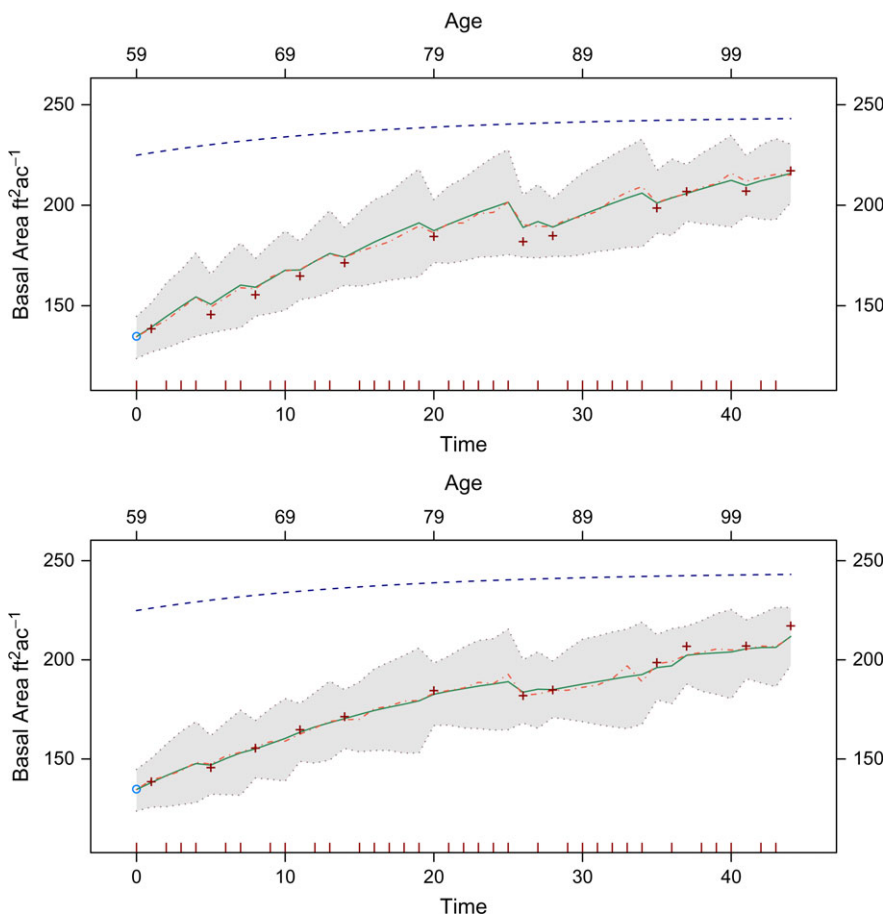
Figure 3. Single Hatch plot trajectory of white pine basal area yield from 40 years of remeasurements. Top panel trajectory has no bias correction; bottom panel trajectory has process model bias correction with   . The measurements are denoted by the plus with bars at the figure base denoting missing measurements; the initial state is shown as a circle. The shaded area marks the 95% credible intervals. The solid line is the SIR filter estimate of the mean and the dot–dash line is the MAP estimate using $N = 1000$ particles. The deterministic yield equation B_t is shown as the long-dash line for comparison. (Note: 1 ft^2 ac^{-1} equals 0.2296 m^2 ha^{-1}.)

is still the same; however, a 2-ft^2 ac^{-1} (0.46 m^2 ha^{-1}) downward adjustment on average has been made through the random error component. This adjustment could be approximated by comparing the recorded growth between two periods with that proposed by the Frothingham model and, in general, need not be constant over time; however, here it was simply estimated by eye for illustration. Notice how a simple small change has led to an adjustment in the particle trajectories such that the mean and MAP estimates no longer have

a 'sawtooth' form (as a result of the optimistic growth), but rather, more closely conform to the growth trajectory of the measurements. However, even though the filter trajectories align well with the measurement trajectory for a few periods, the intention was not to match the measurements, since they are assumed to be corrupted with noise, but simply to adjust the overly optimistic growth produced by the Frothingham model.

The non-uniform nature of the remeasurement periods in this example should serve to illustrate

that the filter does not require an exact periodic measurement schedule. Additionally, as in the simulation example, notice that the probability intervals widen the further out in time the prediction gets from the past measurement. One very simple method for planning a future remeasurement using such probabilistic estimates would be to schedule the inventory when the intervals get past some tolerable degree of error – this rate may be set somewhat higher with a mismatched growth model like the one used here if left uncorrected.

Figure 4 illustrates how the form of the estimated densities change over time. The density at $t = 28$ is the posterior density and is nearly normal. In the absence of measurements, the prediction densities distort, spread and are translated as they proceed through time (Ristic *et al.*, 2004, p. 4) due to compounding of the uncertainty in the predictions. When the new measurement is assimilated at $t = 35$, the posterior density is again nearly normal, more highly peaked, and has much smaller variance.

It should be clear from these examples that the value of sequential Monte Carlo methods lies not only in the probabilistic propagation of the state but also in the fact that such methods offer a general methodology for updating inventory

projections on permanent remeasured growth plots with the latest measurements. In the general filtering literature, this is often referred to equivalently as data assimilation or data–model fusion. A general inventory-updating scheme can be envisioned by running a sequential filter on individual permanent growth plots where the state variables might be, for example the basal area, number of trees, or volume, at each plot. In a large-scale inventory, filters could be run independently, one per plot, since the spatial covariances between plots would probably be negligible. Alternatively, the state vector might be composed of the state variables from many plots, and the covariances estimated, either as part of an adaptive scheme or outside the filter. There is no restriction in either case on when the measurements are recorded on the individual plots; indeed, only a portion of the plots might be remeasured in any particular year.

Discussion and conclusions

Filtering methods owe their genesis to applications such as object tracking, where measurements arrive sequentially in real time, often on the order of fractions of a second. The fact that the time scale is different in forest yield estimation does
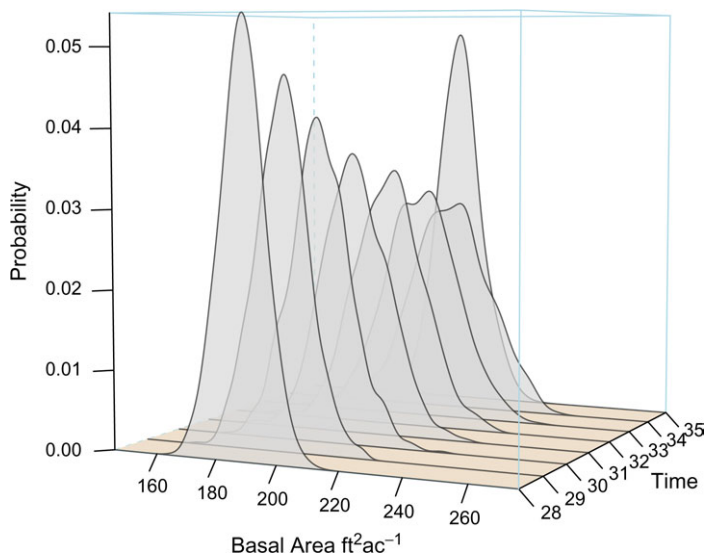

Figure 4. Estimated probability densities for time periods $t = 28$ through $t = 35$ corresponding to the trajectory in the top panel of Figure 3. (Note: 1 ft^2 ac^{-1} equals 0.2296 m^2 ha^{-1}.)

not matter, it simply gives forest managers the luxury of making adaptive decisions concerning the scheduling of events, such as the next inventory, in a more relaxed manner. Because the time frame between inventory updates will normally be measured in years, it enables the manager and biometrician to consider alternate models, or even to recalibrate existing models based on new sets of measurements in the intervening period. Indeed, the luxury of time might even allow propagating future state estimates based on some new filtering algorithm that has been developed in the interim.

The SIR particle filter described here is only one in a class of general non-linear filters referred to earlier as suboptimal. While the appellation 'suboptimal' may sound less than desirable, especially when applied to a statistical estimation technique, as is explained more fully in Ristic *et al.* (2004, Chapter 2), it is used because there is no single overall solution method to the general non-linear non-Gaussian filtering problem. In this sense, suboptimality is not something to be shunned, it is the best one can do given the circumstances. In the case of linear Gaussian problems, the Kalman filter can be used and is optimal in the minimum mean square error sense. For the general non-linear filtering problem, particle filters are the most flexible because of their ability to handle various error distribution models. This flexibility does, however, come with an associated computational cost – for as the dimension of the state space increases, so too the number of particles must increase to adequately cover the support of the state space in this higher dimensional problem (Gordon *et al.*, 1993). Daum and Huang (2003) have shown that the computational cost of the SIR filter can indeed be high as the state dimension increases. However, Daum and Huang (2003) note that a well-designed filter, with conditional densities close to Gaussian and choice of a good proposal distribution (which is a component of filter design) can mitigate this cost substantially, an observation echoed by Ristic *et al.* (2004, p. 59) and elsewhere. Research into alternative proposal distributions has produced numerous versions of the particle filter and is a subject of continuing intense research that will not be discussed further here, the interested reader may consult Ristic *et al.* (2004, Chapter 3) for a recent overview.

The so-called curse of dimensionality probably will not be a major concern in most forest growth and yield applications for two reasons. First, the state space in stand-level problems is often small, made up of a few state variables characterizing the stand such as basal area, number of trees, volume and the like; and second, most problems encountered will undoubtedly have conditional densities that are nearly Gaussian, as in the examples above (assuming Gaussian noise). Moreover, if need be, algorithms exist for parallelizing particle filters, using specially designed resampling steps that can be run in parallel, to help ameliorate any burdens associated with large N (Hegyi *et al.*, 2007). Needless to say, when real-time sequential estimation is measured in years, rather than in seconds, computational time becomes somewhat of a moot point; even if millions of particles were used, estimation would still take a small fraction of the time available for forest planning.

The overall flexibility of the filtering approach owes much to its original formulation with linked state evolution and measurement equations. This allows the incorporation of two main sources of variability inherent in all coupled inventory and growth projection schemes: sampling error and model uncertainty. In theory, additive measurement errors $\mathbf{n}_t$ are the difference between the measurements and the true unknown value of the state as related through $\mathbf{h}_t$, which may be non-linear. In our examples, the measurement equation is a linear identity function, so that the measurement errors represent the difference between the measurements and the unknown state. The assumption was that these errors were zero mean, Gaussian. This assumption implies that there is no bias in the estimation of the states from the measurements. Inasmuch as the SE from a forest inventory provides information about the variability of the measurement estimate from the true unknown population mean (i.e. the unknown state), its use as the variance component seems reasonable. However, in the case where other errors (e.g. classification) might also be acting upon the measurements, possibly even in a biasing manner, the general error structure for the measurement equation can be used to incorporate these other sources and possible biases as well.

In addition, there is no consistency restriction on measurements, implying that they must always derive from an inventory with similar goals

in precision. For example, returning to the bottom panel of Figure 1 at t = 30, recall that mortality due to some event can be inferred, since the true state is decreasing. Perhaps this decline has been noticed by the forest manager. Then it is trivial to incorporate information from an auxiliary estimate during the interim, perhaps from a basal area walk-through count, or some other source like recently flown photography, with associated variance $\mathbf{R}_t$ applied. This variance estimate would likely be higher than that from a normal inventory, both due to the fact that the stand may be less homogeneous from the mortality and because the purpose of the inventory would be for a quick update rather than a full stand examination. The point is that measurement updates need not always come from a full inventory and can be done at any time, yielding a very flexible system.

Similar advantages apply to the process model component of the state space formulation. For example, bias in the process model can be accommodated by specifying a Gaussian or other distribution that is not assumed to have zero mean, as in the Hatch white pine example. In addition, notice in equation (1) that all components are time dependent, including both the model and the noise variance. This structure allows for different model formulations, if desired, at different time periods, with associated changes in variance. This could be as simple as external model recalibration based on the latest information or as extensive as a complete change in model formulation. Indeed, recalibration need not be external to the filter, as the application of filtering methods is not limited solely to state estimation but can include model parameters as well. In general, filters have been applied in various ways to the estimation of model parameters in a sequential manner. Depending upon the complexity of the models and their ability to match the process under consideration, filter-based parameter estimates may be approximately constant over time requiring little adjustment or be adjusted like the states, yielding time-varying parameter trajectories. An example where an unscented Kalman filter was used to estimated the parameters of a process model, in addition to the states, based on eddy covariance measurements (Baldocchi, 2003) is found in Gove and Hollinger (2006). The unscented Kalman filter is closely related to the particle filter described

here in that it is a general non-linear filter. They differ primarily in the underlying sampling mechanism. In the particle filter, sampling is based on many particles generated through Monte Carlo methods, while in the unscented filter, a small deterministic sample of the state space is taken at each time period structured in such a way that it captures the mean and covariance of the state. Relatedly, in linear Kalman filter applications to tree-ring studies, the differentiation between state variables and parameters is often blurred, because the models are of simple linear regression form with process mean and slope varying through time (Van Deusen 1991; Gove and Houston 1996).

State space models are known by various names including hidden Markov models (Cappé et al., 2005, p. 4). The latter derives from the fact that the value of the state may be hidden, or unknown, coupled with the Markov assumption discussed earlier. In the examples presented here, the true value of stand basal area per unit area is unknown, but is estimated from measurements on the same variable. This need not be so, the state itself can be unmeasurable (or hidden), for example in the case of stand biomass, and the measurement would then be related to this quantity at time t via the measurement equation $\mathbf{h}_t$.

The yield model used here was of simple structure for pure even-aged stands based on the Chapman–Richards function. More complicated growth and yield models could have been used. Systems of equations common in forest growth prediction on the stand level could easily be accommodated. In models where mortality is explicitly accounted for, it would also become a state variable. In inventory update situations based on temporary rather than permanent plots, which yield components of change, mortality would more naturally fit the description of a hidden state variable, to be inferred from the process model, but not necessarily measured directly in the inventory. In addition, rather than applying an *ad hoc* bias correction as was done here for illustration, a more appropriate growth model would serve to align the dynamic prior with the true posterior. Other filter-based methods that would also help accomplish this involve extensions to the basic SIR filter such as choosing different proposal distributions (Ristic et al., 2004, p. 55).

As detailed in Ståhl *et al.* (1994) and Nyström and Ståhl (2001), the impact of categorizing the stand in terms of its probable state values influences how decisions might be made with respect to scheduling of treatments or inventories. Both studies present decision criteria such as expected loss, and other quantities, based on posterior integrals. The integrals involved are naturally of the form (14); therefore, they may be approximated at any given time t from the particle filter state estimates using equation (15). Again, this particle approximation yields a tractable solution to problems where the posterior is intractable, or where the resulting integrals are difficult and may also be intractable.

The SIR filter is one example of a large class of similar filters for solving non-linear non-Gaussian sequential estimation problems. Its utility in problems concerning probabilistic state estimation of forest stand variables has been demonstrated using simple examples of basal area yield. Use of such methods allows the forest manager to combine model predictions and measurement updates of various forms into an integrated system. The Markovian assumption states that it is unnecessary to keep the past sequence of states beyond $t - 1$ when estimating $\mathbf{x}_t$. This assumption – which forms the basis for the majority of important filters – seems quite logical in forest dynamics as most growth models employ it implicitly; that is growth projection proceeds sequentially based on the most recent measurement (or prediction) only, and not based on estimates further back in time. In this sense, forest growth is envisioned in general as a short-memory process. Therefore, not only does the filtering approach make probabilistic sense but also it is in keeping with our current understanding and modelling of the biology of the system. Lastly, the particle approximation of the posterior yields an approximation to general expectations based on the posterior that would be difficult or intractable to solve analytically. The result is not only a general state estimation mechanism but also a methodology that provides the foundation for decision making under uncertainty as well.

Conflict of Interest Statement
None declared.

References

Arulampalam, S., Maskell, S., Gordon, N. and Clapp, T. 2002 A tutorial on particle filters for on-line non-linear/non-Gaussian Bayesian tracking. *IEEE Trans. Signal Process.* **50**, 174–188.

Baldocchi, D.D. 2003 Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future. *Global Chang. Biol.* **9**, 479–492.

Burk, T.E., Hansen, M.H. and Ek, A.R. 1982 Combining sources of information for improved in-place inventory statistics. In *In-place Inventories: Principles and Practices*. T.B. Brann, L.O. House IV and H.G. Lund (eds). Society of American Foresters, Bethesda, MD, 82-02, pp. 413–420.

Cappé, O., Moulines, E. and Rydén, T. 2005 *Inference in Hidden Markov Models*. Springer, New York.

Clutter, J.L. 1963 Compatible growth and yield models for loblolly pine. *For. Sci.* **9**, 354–371.

Crisan, D. and Doucet, A. 2002 A survey of convergence results on particle filtering methods for practitioners. *IEEE Trans. Signal Process.* **50**, 736–746.

Daum, F. and Huang, J. 2003 Curse of dimensionality and particle filters. In *Aerospace Conference 2003 Proceedings*. **volume 4**, IEEE, Big Sky, MT, pp. 1979–1993. http://ieeexplore.ieee.org/xpl/tocresult.jsp?isnumber=27670&isYear=2003

Doucet, A., de Freitas, N. and Gordon, N. 2001 An introduction to sequential Monte Carlo methods. In *Sequential Monte Carlo Methods in Practice*. A. Doucet, N. de Freitas and N. Gordon (eds). Springer, New York, pp. 3–14.

Doucet, A., Godseill, S. and Andrieu, C. 2000 On sequential Monte Carlo sampling methods for Bayesian filtering. *Stat. Comput.* **10**, 197–208.

Fairweather, S.E. and Turner, B.J. 1983 The use of simulated remeasurements in double sampling for successive forest inventory. In *Proceedings, Renewable Resource Inventories for Monitoring Changes and Trends*. In J.F. Bell and T. Atterbury (eds). Oregon State University, Corvallis, OR, pp. 609–612.

Frothingham, E.H. 1914 *White Pine under Forest Management*. Bulletin 13, U.S. Department of Agriculture, U.S. Government Printing Office, Washington, DC.

Gelb, A.E. (ed). 1974 *Applied Optimal Estimation*. MIT Press, Cambridge, MA.

Gordon, N.J., Salmond, D.J. and Smith, A.F.M. 1993 Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proc. F.* **140**, 107–113.

Gove, J.H. and Fairweather, S.E. 1992 Optimizing the management of uneven-aged forest stands: a stochastic approach. *For. Sci.* **38**, 623–640.

Gove, J.H. and Hollinger, D.Y. 2006 Application of a dual unscented Kalman filter for simultaneous state and parameter estimation in problems of surface-atmosphere exchange. *J. Geophys. Res.* **111**, 1–21.

Gove, J.H. and Houston, D.R. 1996 Monitoring the growth of American beech affected by beech bark disease in Maine using the Kalman filter. *Environ. Ecol. Stat.* **3**, 167–187.

Green, E.J. and Strawderman, W.E. 1988 Combining inventory data with model predictions. In *Forest Growth Modelling and Prediction*. A.R. Ek, S.R. Shifley and T.E. Burk (eds). **volume 2**, USDA Forest Service, North Central Forest Experiment Station, GTR NE-120, pp. 676–682.

Hegyi, A., Mihaylova, L., Boel, R. and Lendek, Z. 2007 Parallelizable particle filtering for freeway traffic state estimation. In *Proceedings of the European Control Conference*, Greece. pp. 2442–2449. http://www-sigproc.eng.cam.ac.uk/smc/papers.html (accessed on 13 March, 2008).

Ho, Y.C. and Lee, R.C.K. 1964 A Bayesian approach to problems in stochastic estimation and control. *IEEE Trans. Automat. Contr.* **AC-9**, 333–339.

Hof, J.G., Robinson, K.S. and Betters, D.R. 1988 Optimization with expected values of random yield coefficients in renewable resource linear programming. *For. Sci.* **34**, 634–646.

Hol, J.D., Schön, T.B. and Gustafsson, F. 2006 On resampling algorithms for particle filters. In *Nonlinear Statistical Signal Processing Workshop*. http://www-sigproc.eng.cam.ac.uk/smc/papers.html (accessed on 18 May, 2007), Cambridge, UK.

Julier, S.J. and Uhlmann, J.K. 2004 Unscented filtering and nonlinear estimation. *Proc. IEEE.* **92**, 410–422.

Kalman, R.E. 1960 A new approach to linear filtering and prediction problems. *Trans. ASME J. Basic Eng. Ser.* **82D**, 35–45.

Kangas, A.S. and Kangas, J. 1999 Optimization bias in forest management planning solutions due to errors in forest variables. *Silva Fenn.* **33**, 303–315.

Maybeck, P.S. 1979 *Stochastic Models, Estimation, and Control*. **Volume 1**, 1st edn. Academic Press, New York.

Nyström, K. and Ståhl, G. 2001 Forecasting probability distributions of forest yield for a Bayesian approach to management planning. *Silva Fenn.* **35**, 185–201.

Pickens, J.B. and Dress, P.E. 1988 Use of stochastic production coefficients in linear programming models: Objective function distribution, feasibility, and dual activities. *For. Sci.* **34**, 574–591.

Pienaar, L.V. and Turnbull, K.J. 1973 The Chapman-Richards generalization of Von Bertalanffy's growth model for basal area growth and yield in even-aged stands. *For. Sci.* **19**, 2–22.

Ristic, B., Arulampalam, S. and Gordon, N. 2004 *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Artech House, Boston, MA.

Rubenstein, R.Y. 1981 *Simulation and the Monte Carlo Method*. Wiley, New York.

Silverman, B.W. 1986 *Density Estimation for Statistics and Data Analysis*, 1st edn. Chapman and Hall/CRC, New York.

Simon, D. 2006 *Optimal State Estimation*. Wiley, Hoboken, NJ.

Ståhl, G., Carlsson, D. and Bondesson, L. 1994 A method to determine optimal stand data acquisition policies. *For. Sci.* **40**, 630–649.

Van Deusen, P.C. 1991 Trend monitoring with varying coefficient models. *For. Sci.* **37**, 1365–1375.