Model-Based Time-Series Analysis of FIA Panel Data Absent Re-Measurements

Raymond L. Czaplewski and Mike T. Thompson

United States Department of Agriculture / Forest Service
Rocky Mountain Research Station
Research Paper RMRS-RP-102WWW
May 2013

Czaplewski, Raymond L.; Thompson, Mike T. 2013. **Model-based time-series analysis of FIA panel data absent re-measurements.** Res. Pap. RMRS-RP-102WWW. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 13 p.

ABSTRACT

An epidemic of lodgepole pine (*Pinus contorta*) mortality from the mountain pine beetle (*Dendroctonus ponderosae*) has swept across the Interior West. Aerial surveys monitor the areal extent of the epidemic, but only Forest Inventory and Analysis (FIA) field data support a detailed assessment at the tree level. Dynamics of the lodgepole pine population occur at a more rapid rate than the FIA 10-year re-measurement cycle. A model-based approach links population-level estimates from each annual FIA panel estimate. A simple multivariate model predicts the statewide rates of annual change among live uninfected trees, live infected trees, mortality trees, and standing dead trees. A multivariate weighted sum of panel estimates and model predictions of the same attributes improve estimates for each year. Biological structure incorporated into the model improves logical consistency among the various categories of tree-level estimates and smooths annual fluctuations caused by random sampling error. We present concepts in simple terms and illustrate results with FIA data from 2002 to 2008.

AUTHORS

Raymond L. Czaplewski, Emeritus Scientist, Interior West Forest Inventory and Analysis Program, Fort Collins, Colorado

Mike T. Thompson, Forester, Interior West Forest Inventory and Analysis Program, Ogden, Utah

CONTENTS

Introduction

A regional epidemic of tree mortality from mountain pine beetles (*Dendroctonus ponderosae*) is sweeping through the montane forests of the western United States. The Forest Inventory and Analysis (FIA) program designed an annual system, in part, to monitor such episodic changes (Gillespie 1999). We present one case study that compares conventional analyses of annual FIA data to a model-based statistical approach. The objective is to compute efficient and reliable estimates of the state of the lodgepole pine (*Pinus contorta*) tree population in Colorado each year from 2002 to 2008.

The model-based estimator considered here is the Kalman filter (e.g., Maybeck 1979). Multitudes of engineering applications rely on this mathematical tool, with the common GPS system being one everyday example. Other mature applications abound in electrical engineering, econometrics, physical oceanography, and atmospheric sciences.

The Kalman filter estimates the multivariate "state" of a system over time. An example is the number of lodgepole pine trees in the State of Colorado within various categories of live trees, mortality trees, and standing dead trees (Czaplewski and others 2012). The Kalman filter separates a time-series (e.g., 2002 to 2008) into a sequence of discrete time intervals. In the case of FIA, that interval is one year. Initial conditions are established with the best available vector estimate of the multivariate state of the system at the beginning of the year. A simple linear model predicts the state-transition from the beginning to the end of the year (e.g., Table 1). The Kalman filter estimate at the end of the year is the weighted vector sum of the predicted state-vector and an independent measurement vector. A multivariate estimate from a single annual FIA panel is an example of the latter. The vector of weights, which are computed as known constants, minimizes the estimated covariance matrix for the vector sum. The best estimate of the state-vector at the end of the time-increment serves as initial conditions for the next time-increment. This sequential estimation cycle repeats until the full time-series is complete. Czaplewski and Thompson (2009) provided a tutorial example.

Monitoring Design

In the western United States, FIA measures a 10 percent interpenetrating sub-sample (i.e., "annual panel") of all Phase-2 field plots each year (Patterson and Reams 2005). It requires 10 years to measure the full sample of FIA field plots. During these 10 years, there are no re-measured plots to estimate components of change at the tree or plot levels. Re-measurements of plots begin during the 11th year, and all plots are fully re-measured by the 20th year. Unlike periodic inventories produced by FIA between 1950 and 2000, which measured all plots in a large region once every 10 to 20 years, the annual design produces new design-based sample survey estimates each and every year. However, each annual estimate uses the 10 percent sub-sample, which introduces relatively large sampling error and inferential uncertainty.

To increase sample size and reduce uncertainty from random sampling error, FIA uses the "temporally indifferent" estimator, which is a moving

Table 1. Multivariate state-transition model used in Kalman filter to predict the state of the Colorado lodgepole pine forest at the end of a one-year time-increment (i.e., time *t*) using the best estimate of that population at time *t*-1 as initial conditions. Our best estimates for each parameter c_i appear in Table 2.

	State variables
X_t	Number of live trees in Colorado with no insect damage at time *t*
Y_t	Number of live trees in Colorado with insect damage at time *t*
$(Z_1)_t$	Number of mortality trees in Colorado that died between times *t*-1 and *t*
$(Z_2)_t$	Number of mortality trees in Colorado that died between times *t*-2 and *t*-1
$(Z_3)_t$	Number of mortality trees in Colorado that died between times *t*-3 and *t*-2
$(Z_4)_t$	Number of mortality trees in Colorado that died between times *t*-4 and *t*-3
$(Z_5)_t$	Number of mortality trees in Colorado that died between times *t*-5 and *t*-4
$(Z_{6+})_t$	Number of standing dead trees in Colorado that died before time *t*-6
A_t	Area of lodgepole pine forest in Colorado at time *t*
	t = 2002, 2003, 2004, 2005, 2006, 2007, 2008

Model		Components of change
$X_t =$	$+ X_{t-1}$	Live trees with no insect damage at time *t*-1
	$+ (c_5 \times X_{t-1})$	Annual regeneration rate
	$- (c_6 \times X_{t-1})$	Annual non-insect mortality rate
	$+ (c_3 \times Y_{t-1})$	Infected live trees at time *t*-1 that are uninfected live trees at *t*
	$- [(X_{2002}-X_{t-1}) \times c_1] \times Y_{t-1}$	Uninfected live trees at time *t*-1 that are infected live trees at *t*
$Y_t =$	$+ [(X_{2002}-X_{t-1}) \times c_1] \times Y_{t-1}$	Density-dependent increase live uninfected trees at time *t*-1 infected at time *t*
	$+ (1 - c_2 - c_3) \times Y_{t-1}$	Infected live trees at time *t*-1 that remain infected live trees at *t*
	$- (c_2 \times Y_{t-1})$	Infected live trees at time *t*-1 that are mortality trees at time *t*
	$- (c_3 \times Y_{t-1})$	Infected live trees at time *t*-1 that are uninfected live trees at time *t*
	$+ (c_4 \times X_{t-1})$	Annual non-insect mortality rate
$(Z_1)_t =$	$+ (c_2 \times Y_{t-1})$	Infected live trees at time *t*-1 that are mortality trees at time *t*
	$- (Z_1)_{t-1}$	1st-year mortality trees at time *t*-1 that are 2nd-year mortality trees at time *t*
$(Z_2)_t =$	$+(Z_1)_{t-1} - (Z_2)_{t-1}$	2nd-year mortality trees at time *t* that were 1st-year mortality trees at time *t*-1
$(Z_3)_t =$	$+(Z_2)_{t-1} - (Z_3)_{t-1}$	3rd-year mortality trees at time *t* that were 2nd-year mortality trees at time *t*-1
$(Z_4)_t =$	$+(Z_3)_{t-1} - (Z_4)_{t-1}$	4th-year mortality trees at time *t* that were 3rd-year mortality trees at time *t*-1
$(Z_5)_t =$	$+(Z_4)_{t-1} - (Z_5)_{t-1}$	5th-year mortality trees at time *t* that were 4th-year mortality trees at time *t*-1
$(Z_{6+})_t =$	$+(Z_{6+})_{t-1}$	Standing dead trees at time *t*-1 (mortality trees 6 years and older)
	$+ (Z_5)_{t-1}$	5th-year mortality trees at time *t*-1 that become standing dead trees at *t*
	$- c_6\ (Z_{6+})_{t-1}$	Standing dead trees at time *t*-1 that are fallen coarse woody debris at time *t*
$A_t =$	A_{t-1}	No change in lodgepole pine forest area over time

average that essentially treats plots measured during the preceding five years as though they were measured during a single year (Patterson and Reams 2005). This estimator is design-unbiased if estimation of the five-year moving average is the analysis objective, although that metric is rarely an analyst's first choice. The temporally indifferent estimator is design-unbiased for conditions during a single year if there are no net changes in the population during these five years, although that situation requires strong assumptions by the analyst that cannot be rigorously tested. Regardless, this assumption is implausible during epidemic tree mortality, and the temporally indifferent estimator is biased whenever applied to rapidly changing populations (Patterson and Reams 2005).

Implementation of the FIA annual design in Colorado began during 2002. The Colorado State Forest Service and FIA measure approximately 1100 forest and non-forest Phase-2 plots each year across the State of Colorado. Between 2002 and 2008, about 60 plots per year had one or more tallied lodgepole pine trees, either live or dead. Less than half of these plots were affected by the bark beetle epidemic.

Thompson (2009) analyzed these data with conventional statistical methods. He treated annual panels as independent samples to estimate population-level conifer mortality in Colorado. Statistical tests rejected the null hypothesis that there was no change in mean number of mortality trees per acre during the 2002 to 2008 time periods. The most pronounced change during these six years occurred in the estimated number of insect-killed lodgepole pine, especially after 2005.

Design-Based Annual Panel Estimates

Figure 1 illustrates the multivariate trends in individual panel estimates. These fully estimate the multivariate state of the lodgepole pine system in Colorado between 2002 and 2008. Each annual panel supports independent estimates for a single year of acres that contain lodgepole pine trees and numbers of live trees, mortality trees, and standing dead trees. Transformations of those estimates further characterize the statewide lodgepole pine population as follows. The estimated mortality rate of infected live trees is the estimated number of first-year mortality trees in year *t* divided by the number of infected live trees in the preceding year *t*-1, although this metric is not exactly a mortality rate because live infected trees can survive for multiple years. The annual rate of net change in infected live trees is the number of infected live trees in year t divided by the number of infected live trees in the preceding year *t*-1. The overall rate of change (Figure 2) is estimated with a simple linear regression with measurement year as the independent variable; the dependent variable includes the seven annual population estimates (2002 to 2008) for numbers of live uninfected trees.

As expected, the estimated numbers of infected live trees and mortality trees increased rapidly over time (Figure 1). However, other estimates, if taken literally, describe unexpectedly volatile fluctuations. The estimated area of lodgepole pine forest and the estimated number of uninfected live

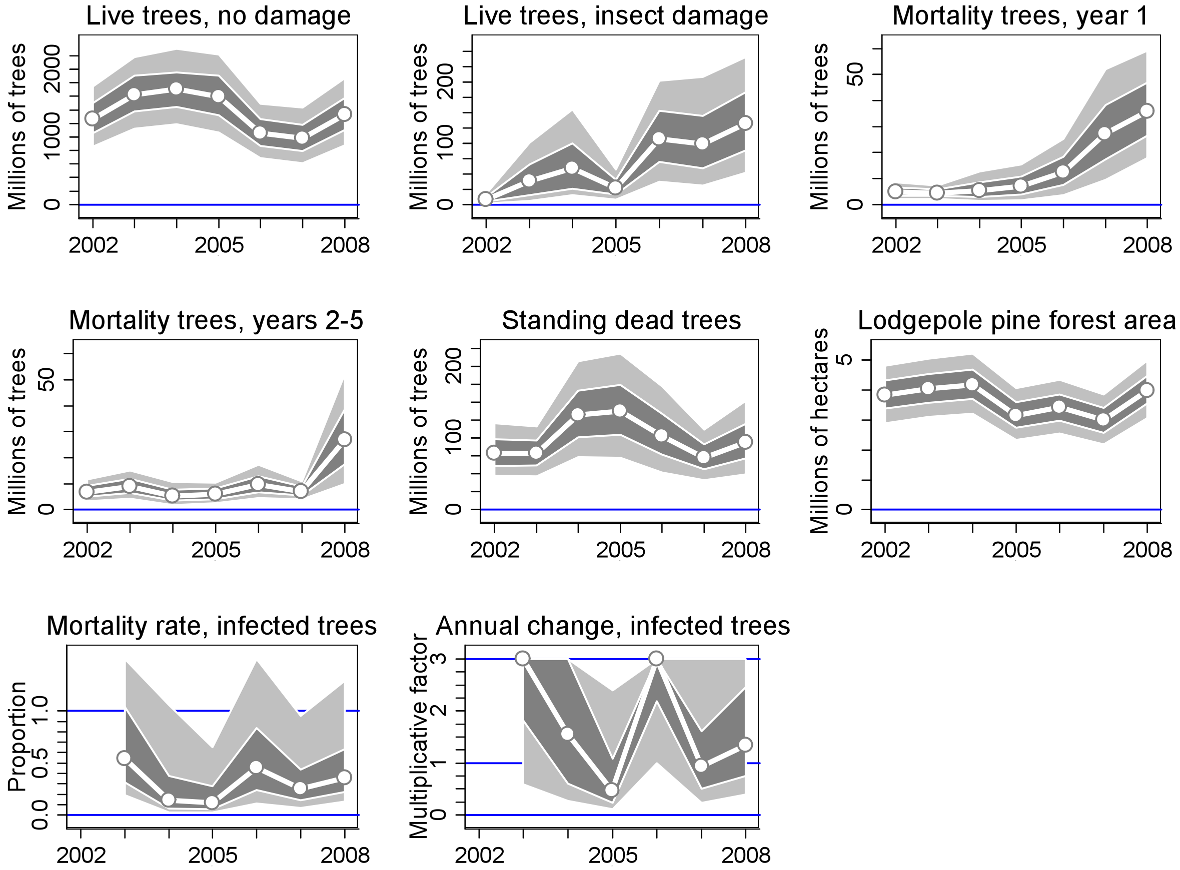

Figure 1. Annual panel estimates for the lodgepole pine population in the State of Colorado with individual annual panel data. Medians of the bootstrap distribution appear as the white center lines. The dark gray region portrays the range between the 0.16 and 0.84 quantiles of the bootstrap distribution. This asymmetrical region approximately corresponds to ±1 standard deviation units if the bootstrap distribution were Gaussian, while the light gray region represents the 0.02 and 0.98 quantiles, roughly corresponding to ±2 standard deviation units.

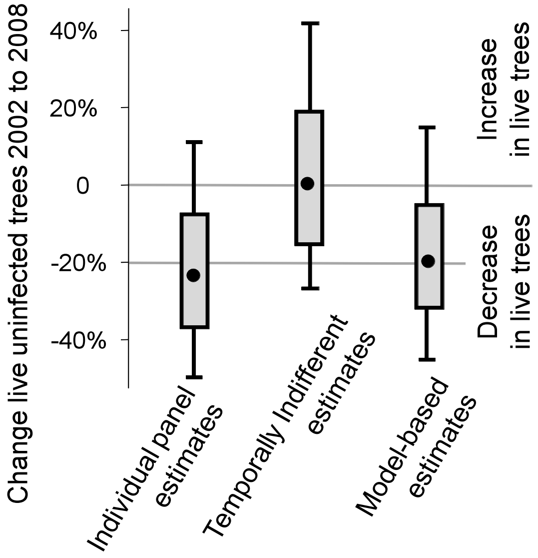

Figure 2. Overall rate of change in live trees with no insect damage estimated with a simple linear regression and the seven annual population estimates for 2002 through 2008. Spread of the estimated change during those seven years represents the bootstrap distribution. The temporally indifferent estimates indicate no change in number of uninfected live trees, in spite of the well documented epidemic of lodgepole pine tree mortality since 2002. Patterson and Reams (2005) anticipated this problem when they stated that the temporally indifferent estimator "dampens trends by obscuring annual fluctuations" and "makes changes appear smaller than they are." Unlike the temporally indifferent estimator, the FIA panel estimators and the model-based estimator both estimate an approximate 20 percent decrease in uninfected live trees between 2002 and 2008.

trees and standing dead trees can increase or decrease 30 percent or more from one year to the next. In 2005, the estimated number of infected live trees decreased markedly from the overall time trend, but the estimated number of mortality trees in 2006 did not exhibit a similar response. The estimated annual rates of change in infected live trees varied widely from a 50 percent decrease between 2004 and 2005 to a three-fold increase between 2002 and 2003 and from 2005 to 2006. The estimated mortality rates varied between 10 and 50 percent from one year to the next. While these latter estimates might be plausible, they might be misleading because ratios between estimates from two different years have relatively large estimation errors.

Most analysts would comfortably attribute these unexpected volatilities to random sampling error. This is most likely because an analyst's conceptual model of a statewide tree population envisions forest attributes that do not fluctuate wildly each year. This same conceptual model might further suggest that variables in Figure 1 should be related to each other through obvious demographic processes. For example, the number of mortality trees one year should be correlated with the number of infected live trees in the preceding year, and the number of standing dead trees should be related to the number of mortality trees in previous years. The next section describes one such model that we incorporated into the time-series analysis.

Model-Based Time-Series Estimates

Czaplewski and Thompson (2009) introduced the Kalman filter as an efficient and intuitive model-based estimator for time-series of annual FIA panel data. The model predicts annual changes in the statewide area of lodgepole forest and number of live trees, live infected trees, and mortality trees (Table 1). The analyst uses professional judgment to set the feasible range for each model parameter and selects the final set of model parameters aided by informal goodness-of-fit statistics. The results in our case are given in Table 2. While model parameters are treated as known constants in the Kalman filter, the model includes an additive vector of random prediction errors that accounts for imperfect model predictions (Maybeck 1979).

The Kalman filter estimate for year *t* is the weighted vector sum of model predictions at time *t*, which can be relatively precise but biased, and the independent design-based vector estimates from the FIA annual panel at time *t*, which are design unbiased but can be imprecise. The vector of weights is computed from the difference between the estimated covariance matrix for model prediction errors at time *t* and the covariance matrix for random sampling errors in the FIA panel estimate at time *t*. This computation is derived with the minimum variance optimality criterion, but it can also be viewed as a maximum likelihood estimator if the error distributions are multivariate Gaussian (Maybeck 1979). In addition, the Kalman filter qualifies as an empirical Bayes estimator (Maybeck 1979). More weight is placed on estimates with smaller variances, either model predictions or design-based panel estimates. Viewed differently, unbiased but imprecise

Table 2. Numerical parameters (c_i) used with the model-based estimator (Table 1). Ranges are estimated by the analyst through professional judgment, and the final parameter set selected by the analyst are aided by informal goodness-of-fit statistics. This provides estimates of certain components of change, which are not available without re-measurement data from the independent annual panel estimator or with the temporally indifferent estimator. However, the statistical properties of these parameter estimates are unknown.

		Sets of parameters evaluated			Best estimate
Model parameter (Table 1)		minimum	maximum	increment	
c_1	Baseline annual rate of increase in new infections without density dependence	1.9	2.7	0.1	2.1
c_2	Proportion of infected live trees that are mortality trees the following year	0.4	1.0	0.05	0.4
c_3	Proportion of surviving infected live trees that are uninfected the following year	0.1	0.7	0.1	0.7
c_4	Annual mortality rate for uninfected live trees from all causes except insect damage	0.0000	0.0004	0.0001	0.0001
c_5	Annual regeneration rate for live uninfected trees (assumes background regeneration and mortality rates are equal)	$= c_4$	$= c_4$	$= c_4$	$= c_4$
c_6	Annual rate of change from standing dead trees to fallen coarse woody debris trees (single value assumed in all trials)	0.01	0.01	0	0.01

estimates from each annual FIA panel help the simple demographic model more accurately predict the true population dynamics, which are rarely simple.

The exact mortality year for a tree is difficult to infer in the field. To reduce aberrations caused by nonrandom measurement errors, panel estimates for mortality trees that died during years *t*-2 and *t*-5 are merged into a single measure. Also, mortality trees that died between times *t* and *t*-1 are not separated by cause (i.e., insect damage or other causes). This is mathematically feasible because dimensions of the measurement vector from annual panel estimates need not conform to dimensions of the state-vector (Maybeck 1979). An indicator matrix of zeros and ones relates this abbreviated vector of panel estimates to the full state-vector in Table 1.

Bar-Shalom and others (2001) documented the **LDL**' square-root implementation of the Kalman filter, and all matrix computations in our study use this algorithm. The **LDL**' algorithm is an algebraic transformation of the matrix formulae in the standard Kalman filter. This algorithm substantively reduces numerical round-off error that is a common hazard in applications of the digital Kalman filter. Furthermore, a nonlinear optimization algorithm scaled the estimated covariance matrix for annual model prediction errors to reduce sub-optimal weighting of model predictions. The algorithm focuses on the time-series of multivariate residual differences between model estimates at time *t* and corresponding panel estimates at time *t*. There are seven vectors of residuals (2002 to 2008) and each measurement vector has six variables: X_t; Y_t; $(Z_1)_t$; $(Z_2)_t+(Z_3)_t+(Z_4)_t+(Z_5)_t$; $(Z_6+)_t$; and A_t (Table 1). The estimated covariance matrix for model errors is the basis for a linear transformation of these multivariate residuals into a time-series of 42 nominally orthogonal scalar values with unit variances (Maybeck 1979). If the true covariance matrix were known without error, then these 42 transformed residuals are expected to be mutually independent with mean equal to 0 and variance equal to 1. The optimization algorithm performs a multivariate scaling of the each of the seven estimated covariance matrices so that the realized mean squared errors of the standardized residuals conform to their expected distribution. The mean of the squared errors simultaneously addresses prediction bias (i.e., non-zero mean of residuals) and bias in estimated variance of prediction errors. This reduces risk from placing too much weight on potentially biased model predictions and too little weight on unbiased panel estimates. However, there is no known theoretical basis to assert that this estimator is unbiased.

Uncertainty Metric

Random sampling errors and model prediction errors cause uncertainty in population estimates. The standard-deviation statistic is one measure of uncertainty. The standard deviation accurately describes the spread of a Gaussian distribution. However, small sample sizes for non-negative and relatively rare attributes, such as tree mortality, are associated with skewed error distributions. Therefore, the bootstrap method (Efron and Tibshirani 1994) is used to measure uncertainty in the population estimates.

The bootstrap method independently re-samples measured FIA plots within each annual panel. Each of the three estimators is applied to each bootstrap sample. There were 2500 replicate bootstrap samples used to estimate the bootstrap distribution of random errors. The population estimates in Figures 1, 2, and 3 use the median of the bootstrap distribution to accommodate extreme but rare outliers in the bootstrap distributions. The uncertainty metric is the difference between the 16th and 84th percentiles of the bootstrap distribution. If the bootstrap distribution were Gaussian, then this uncertainty metric would approximately equal one standard deviation unit.

Results and Discussion

Figures 2 and 3 compare time-series estimates from individual annual panels, the temporally indifferent five-year moving average, and the model-based Kalman filter. All three estimators paint a similar picture of changes in the statewide population of lodgepole pine between 2002 and 2008. This is not surprising since all three use the same FIA field data. However, there are notable differences among estimators.

The temporally indifferent moving average estimates that there is no change in number of uninfected live trees between 2002 and 2008 (Figure 2), in spite of the well-documented epidemic of lodgepole pine tree mortality since 2002.The FIA panel estimators and the model-based estimator both estimate approximately 20 percent decrease in number of uninfected live trees between 2002 and 2008, which is far more plausible than estimates with the moving average.

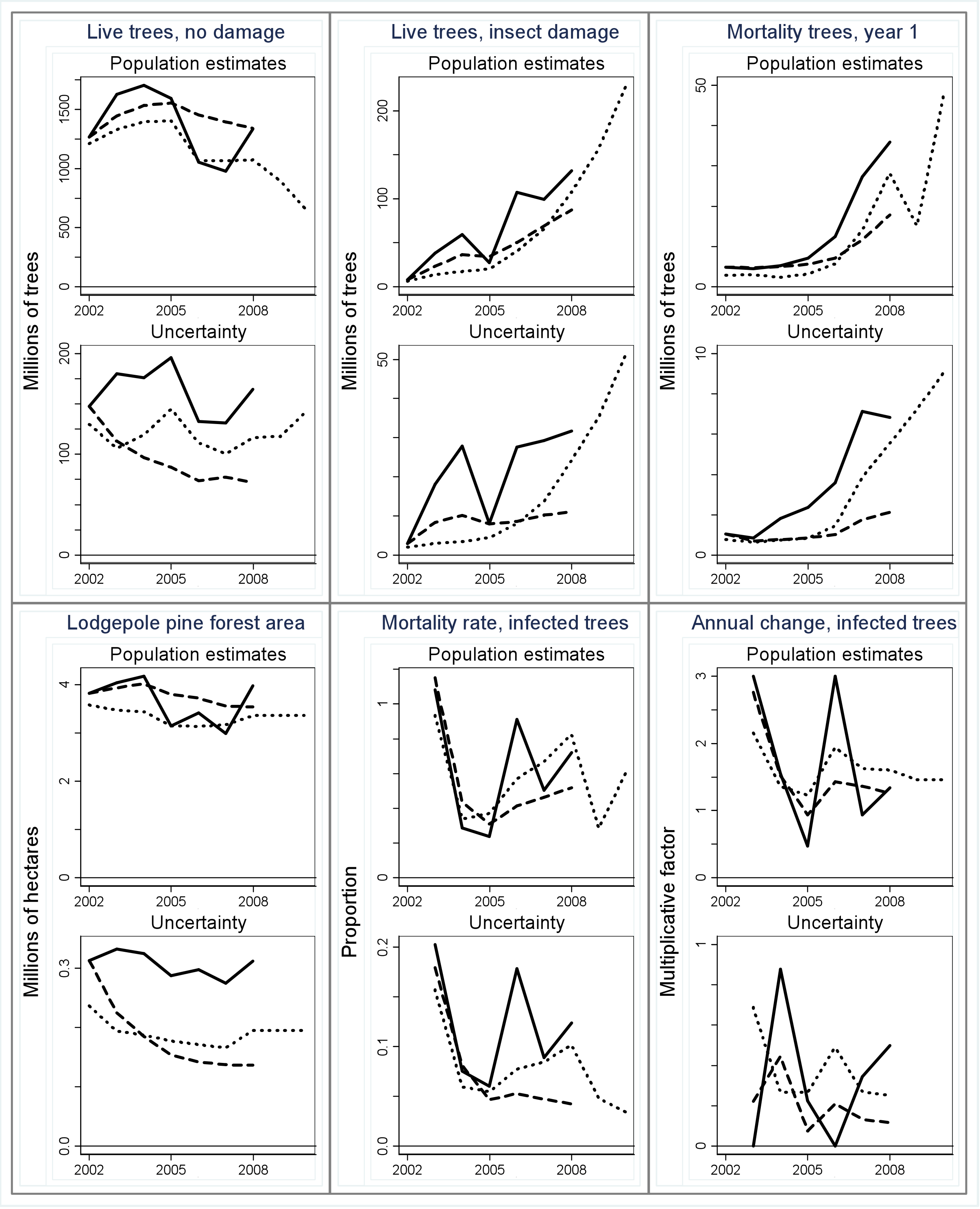

Figure 3. Comparison of results from three alternative estimators. The estimator based on individual annual panels (solid lines), which does not merge observations from different panels, tends to be more variable over time with higher uncertainty. The estimated time-series from the other two estimators are smoother over time and less uncertain. The temporally indifferent five-year moving average estimator (dashed lines) tends to lag behind the model-based Kalman filter estimator (dotted lines). Absent knowledge of the true population state, there are no definitive criteria available to judge which estimator is the most accurate.

The time-series estimates from individual annual panels (solid lines in Figure 3) fluctuate more from one year to the next than those with other estimators. Uncertainty, as gauged by the spread of the bootstrap distribution, is high compared to the other two estimators, both of which use multiple annual panels with larger numbers of FIA plots.

Estimates with the temporally indifferent five-year moving average (dashed lines in Figure 3) have the least variation from one year to the next and tend to have the lowest uncertainty metrics. However, Patterson and Reams (2005) cautioned that "older panels potentially create a lag bias (with the temporally indifferent estimator) when estimating current conditions." The uncertainty metric in Figure 3 ignores lag bias, the degree of which is unknown, and uncertainty with the temporally indifferent estimator is surely higher than portrayed with the bootstrap distribution.

Uncertainty metrics for the model-based estimates with the Kalman filter (dotted lines in Figure 3) are nearly as low as those for the temporally indifferent estimator. Model-based estimates more closely track variations in the panel estimates, but with less-extreme annual fluctuations. Like the temporally indifferent estimator, the Kalman filter "smoothes" variations in annual panel estimates. This is particularly true with estimates of live trees with no insect damage, first-year mortality trees, annual net mortality rates of infected live trees, and annual rate of increase in infected live trees. Unlike the alternative estimators, the model-based approach can forecast trends in the absence of FIA panel data, examples of which are the estimates for 2009 and 2010 in Figure 3.

Which Estimator Is Best?

The best estimator is expected to produce estimates with the least variance and greatest precision (i.e., least uncertainty) with negligible bias relative to that uncertainty. But bias and precision are expectations over all possible samples, and we have only one sample in this case study. Statistical theory can infer precision and bias for certain estimators, such as the design-based estimator separately applied to each annual panel. That same theory can prove the existence of lag bias with the temporally indifferent estimator, but the degree of bias is unique to each population and unknown in any practical application without strong but unverifiable assumptions. Mathematical theory can prove that the model-based estimator is unbiased if there is no bias in model predictions, but that too is unverifiable. We believe that our implementation of the model-based algorithm effectively reduces the magnitude of potential model bias, especially relative to the temporally indifferent estimator, but that has not been proven through theory. Therefore, we simply describe the differences we observe among the three sets of estimates given the sample survey data in hand. At best, this single case study offers weak evidence for judging the merits of alternative estimators.

Both the model-based and temporally indifferent estimators have less uncertainty than individual panel estimates (Figure 3). However, this empirical metric ignores lag bias and model bias. For example, the temporally indifferent estimator infers that there is no change in number of live trees between 2002 and 2008 (Figure 2). This is almost surely an artifact of lag bias. The uncertainty metrics in Figure 3 simply do not capture this type of uncertainty because the magnitude of bias, if any, remains a mystery. However, the model-based algorithm automatically places less weight on model predictions if those predictions show evidence of prediction bias during analyses of residuals (see the Uncertainty Metric section). This is no guarantee that the model-based estimator is unbiased, but the above-mentioned estimation method presumably reduces risk. In principal, the model-based estimator should lessen or eliminate lag bias and under-estimation of annual changes that are inherent with the temporally indifferent estimator.

All three estimators estimate overall net rates of annual change in both infected and uninfected live trees (Figure 3), but only the Kalman filter estimates the components of those changes (c_1, c_2, c_3, and c_4 in Tables 2 and 3). However, the statistical properties of these estimates are not well understood, so little can be generalized about their precision and accuracy. In addition, the Kalman filter incorporates a demographic model, which forecasts the biological state of the population into the future, including the estimated error distribution propagated through the model (Figure 3). This all contributes to analysts' quantitative understanding of annual demographic process within a dynamic tree population.

Future Improvements to Model-Based Approach

In this case study, we considered a relatively small suite of demographic models with the Kalman filter, and there is potential for improvement. For example, our best model has parameter values equal to the limits of those tested (Table 2). These are percent of infected live trees that are mortality trees following year (c_2 in Table 3) and percent of surviving infected live trees that are uninfected following year (c_3 in Table 3). Future analyses will extend the range of these parameter values in pursuit of more accurate estimates.

The Kalman filter algorithm includes a nonlinear optimization routine that adjusts weights on all model predictions so that the distribution of the 42 standardized residuals agrees with its expected mean and variance (see the Uncertainty Metric section). If residuals from any single variable or measurement year are unexpectedly large, then model predictions from all variables for all years receive less weight. This increases overall uncertainty in the model-based estimates (see Figure 3). For example, the annual panel estimate for number of live infected trees is strikingly low in 2005, but there is no corresponding phenomenon with the number of mortality trees in 2006 (Figure 3). This is inconsistent with our conceptual model for lodgepole pine population dynamics. Random sampling error could cause a spurious anomaly in the time-series estimates. Ten percent of the bootstrap samples manifested a similar anomaly in infected live trees with 2006 to 2008 panel data, and 60 percent of all bootstrap samples had a 50 percent decrease or two-fold increase in at least one variable between two consecutive years. Therefore, an anomaly caused by sampling error is not unusual given the relatively small sample size with annual panels in Colorado. Future analysis will consider censoring the 2005 panel estimate for number of infected live trees as part of our model, which should reduce the uncertainty in the model-based estimates for all variables over all years. (Recall that dimensions of the annual measurement-vector need not conform to those of the population state-vector so long as a linear relationship with known constants, e.g., zeros and ones, exists between the two vectors.)

A less striking anomaly is the estimated area of Colorado occupied by lodgepole pine trees, which varies by ±30 percent from one year to the next. This too is inconsistent with our conceptual model of statewide population dynamics. In an attempt to reduce uncertainty in estimates of the tree state-variables, future analyses will consider replacing annual estimates of lodgepole pine area with a single time-invariant state-variable for the entire 2002 to 2008 time-interval.

Estimates of mortality rate of infected trees and net rate of increase in infected trees exhibit wide fluctuations over time (Figure 3). This might be caused by random sampling error rather than actual processes in the lodgepole pine population. These variables are not included in the state-vector or state-transition model. Rather, they are separate nonlinear transformations of other estimates. However, these highly variable rates of change could be added to the model as additional state-variables in the Kalman filter. Expert judgment could constrain their values to a biologically feasible range, limit their annual variability, or serve as a different type of independent "measurement." This would push our model-based approach toward the realm of Bayesian estimation, which is another valid perspective on the Kalman filter (Maybeck 1979).

Bootstrap resampling uses all FIA plots, including plots that do not include lodgepole pine trees. Uncertainty would likely decrease if we analyzed only those plots (regardless of their forest condition) within the potential range of lodgepole pine in Colorado. This would require external geospatial data to supplement FIA plots. In the longer-term, geospatial data from forest health risk models and forest health aerial surveys and remotely sensed data from earth-observing satellites might further improve estimates (Czaplewski and others 2012). This would require expanding dimensions of the state-vector and state-transition model, but that would be no more challenging than developing the model in Table 1. These geospatial data need not be unbiased or share the same accuracy as the field data. Geospatial data merely need to be sufficiently correlated or associated with field measurements in order to improve accuracy of model-based population estimates. However, the potential effect of these changes is unknown. Regardless, success would be constrained by the small sample size of annual FIA plots within different geospatial categories. Success might improve with continuous geospatial variables, such as the remotely sensed proportion of mortality trees, rather than simply a dichotomous indicator of insect damage or no damage.

We anticipate additional improvements. Large and small trees appear to have fundamentally different mortality processes during the pine beetle epidemic, and we will consider partitioning tree variables into two size classes. Since infected live trees can survive several years before mortality, we will consider subdividing the state-variable for infected live trees into two separate variables: live trees infected during the past year, and live trees infected two or more years previously. We will also add FIA panel data from 2009 through 2013. This could improve model-based estimates for 2002 to 2008 because estimates of current conditions improve estimates of past conditions with the Kalman filter. Future analyses might pool annual panel data from a larger ecoregion to increase sample size and decrease uncertainty. However, this would compromise the geographic specificity of the analyses. Finally, the density dependent parameter for rate of increase in infected live trees, i.e., $[(X_{2002}\text{-}X_{t\text{-}1}) \times c_1]$ in Table 1, is indexed relative to the estimated number of live uninfected live trees at the beginning of the time-series, i.e., X_{2002}. This estimate varies among bootstrap iterations, which might contribute unintended variability to the bootstrap distributions. The model will be changed so that the X_{2002} parameter is replaced with a similarly scaled constant across all bootstrap iterations.

The current research is strictly an exploratory proof-of-concept. Ultimately, any useful model-based approach must be implemented within FIA operational systems. This requires analysis tools that are intuitive, user-friendly, and do not need constant support from a mathematical statistician. If future results with this model-based approach are promising, significant development remains.

Conclusions

Can model-based time-series analysis of FIA panel data enhance estimates of mortality under an interpenetrating design? The annualized FIA inventory was implemented under the assumption that trends in forest attributes such as area, tree volume, tree growth, and tree mortality could be assessed in a timelier manner than periodic inventories. The penalty for more timely data is increased uncertainty from the small sample sizes inherent with independent panel estimates. Many proponents of the annual inventory system felt that supplementary data or modeling procedures could significantly reduce this uncertainty. The model-based estimator we described reduces conditional uncertainty and improves quantitative understanding of population dynamics and components of change during periods of rapid change. We envision several improvements to the model-based implementation, such as changes in the state-vector and state-transition model. Hopefully, these will further improve results. While these initial results appear promising, more years of annual data in Colorado might be needed before signatures of mortality events can be fully evaluated with confidence. Estimates derived from the re-measured (paired) plots, which will become available from 2012 to 2022, will substantially increase the power to infer significant effects related to tree mortality.

Acknowledgments

We appreciate the helpful reviews by Dr. KaDonna Randolph and Sara Goeking. Any remaining errors or omissions remain our sole responsibility.

Literature Cited

Bar-Shalom, Y.; Li, X.-R. Kirubarajan, T. 2001. Estimation with applications to tracking and navigation. New York: John Wiley and Sons. 584 p.

Czaplewski, R. L., Thompson, M. T. 2009. Opportunities to improve monitoring of temporal trends with FIA panel data. In W. McWilliams, G. Moisen, and R. Czaplewski (eds.), Forest Inventory and Analysis (FIA) Symposium. Proc. RMRS-P-56CD. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station: 1-55.

Czaplewski, R. L.; Thompson, M. T.; Moisen, G. G. 2012. An efficient estimator to monitor rapidly changing forest conditions. In R. Morin and G. Liknes (eds.), Forest Inventory and Analysis (FIA) Symposium. Gen. Tech. Rep. NRS-P-105. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station.

Efron, B., Tibshirani, R. J. 1994. An introduction to the bootstrap (Chapman & Hall/CRC Monographs on Statistics and Applied Probability). Chapman and Hall/CRC, 1st edition.

Gillespie, A. J. R. 1999. Rationale for a national annual forest inventory program. Journal of Forestry 97:16-20.

Maybeck, P. S. 1979. Stochastic models, estimation, and control, volume 141-1 of Mathematics in Science and Engineering. New York: Academic Press.

Patterson, P. L., Reams, G. A. 2005. Combining panels for forest inventory and analysis estimation. In W. A. Bechtold and P. L. Patterson (eds.), The enhanced Forest Inventory and Analysis program-national sampling design and estimation procedures, Gen. Tech. Rep. SRS-80. Asheville, NC: U.S. Department of Agriculture, Forest Service: 69-74.

Thompson, M. T. 2009. Analysis of conifer mortality in Colorado using Forest Inventory and Analysis's annual forest inventory. Western Journal of Applied Forestry 24:193-197.

Rocky Mountain Research Station

The Rocky Mountain Research Station develops scientific information and technology to improve management, protection, and use of the forests and rangelands. Research is designed to meet the needs of the National Forest managers, Federal and State agencies, public and private organizations, academic institutions, industry, and individuals. Studies accelerate solutions to problems involving ecosystems, range, forests, water, recreation, fire, resource inventory, land reclamation, community sustainability, forest engineering technology, multiple use economics, wildlife and fish habitat, and forest insects and diseases. Studies are conducted cooperatively, and applications may be found worldwide. For more information, please visit the RMRS web site at: www.fs.fed.us/rmrs.

Station Headquarters
Rocky Mountain Research Station
240 W Prospect Road
Fort Collins, CO 80526
(970) 498-1100

Research Locations

Flagstaff, Arizona
Fort Collins, Colorado
Boise, Idaho
Moscow, Idaho
Bozeman, Montana
Missoula, Montana

Reno, Nevada
Albuquerque, New Mexico
Rapid City, South Dakota
Logan, Utah
Ogden, Utah
Provo, Utah

To learn more about RMRS publications or search our online titles:

www.fs.fed.us/rm/publications

www.treesearch.fs.fed.us