

biometrics

Toward Robust Estimation of the Components of Forest Population Change

Francis A. Roesch

Multiple levels of simulation are used to test the robustness of estimators of the components of change. I first created a variety of spatial-temporal populations based on, but more variable than, an actual forest monitoring data set and then sample those populations under a variety of sampling error structures. The performance of each of four estimation approaches is evaluated when the temporal scale of the estimand of interest is 1 year while the temporal scale of observation is t years. Three approaches for estimating the individual components of forest change are compared over five simulated populations under four sets of sampling error structure. The performance of a modification to these approaches is shown when extraneously obtained information indicates that a deviation to the assumed population model exists. Finally, the extraneous information is incorporated into a mixed estimator, combining each of three general transition models with a single compatibility model. The first three approaches, without the incorporation of extraneous information, are compatible with large monitoring efforts that require intervention-free results. The mixed-estimation approach accounts for model assumptions that sometimes remain latent in other approaches and is amenable to the incorporation of the extraneously obtained information and to ensuring estimator compatibility. All four approaches are shown to work well when the sampling error structure is unbiased, while some notable differences in performance were observed at the temporal extremities of observation in the presence of temporal anomalies and in the presence of biased sampling error structures.

Keywords: sampling forest change, spatial-temporal sample design, components of change

An estimator is robust when it provides significant revelations about the conditions being investigated even in the presence of violations of population and sampling error assumptions. This paper uses two levels of simulation, the first to create a variety of spatial-temporal populations based on actual forest monitoring data and the second to sample those populations under a variety of sampling error structures. I examine the robustness of three estimators for the annual components of forest change below when the temporal scale of the population estimand of interest is finer than the scale of observation in a temporally rotating panel sample design. The approaches are similar but differ in their underlying trend assumptions, leading to differences in how measurement intervals that are not centered on a target year inform the estimate for the target year. A modification to these approaches is shown when extraneously obtained information indicates that a deviation to the assumed trend model by each method has occurred. Mixed estimation is then used to package the extraneous information with each of three general transition models and a single compatibility model.

The components of forest change that are of interest when monitoring a dynamic landscape include:

1. Land-use change during the period of interest in the form of reversion from some other land use to forest and diversion away from forest to some other land use, including the growth on trees that occurred prior to a land diversion and growth on trees that occurred subsequent to a land reversion but prior to the first observation.
 2. Change of trees on land that is forested during the entire period of interest, which can be further categorized into:
 - a. Fully observable forest tree change, that is the growth, such as in value or volume, on trees that survive in the population from one observation to the next, (also known as survivor growth), a component of live tree growth.
 - b. Partially observable forest tree change: The value of trees entering into the population, the value growth on entry trees prior to the first observation (another component of live tree growth), and the value loss due to the death or harvest of trees prior to the next observation.
 - c. Unobserved forest tree change to include the growth subsequent to the final measurement on trees prior to death or harvest.
- For a dynamic forest inventory to be fully informative, it is important to acknowledge the distinction between land-use change

Manuscript received September 4, 2013; accepted December 16, 2013; published online January 23, 2014.

Affiliations: Francis A. Roesch (froesch@fs.fed.us), USDA Forest Service, Asheville, NC.

Acknowledgments: The author was funded by the USDA Forest Service Forest Inventory and Analysis Program. An associate editor and two anonymous reviewers provided comments that helped to improve the manuscript.

Table 1. Distribution statistics for population 1.

Component	Statistic	Year												
		1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Volume	Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0
	1st quartile	4.68	11.1	16.78	21.82	26.3	30.81	34.43	36.86	39.3	40.89	42.37	43.28	44.62
	Median	56.14	64.25	71	76.6	82.02	85.6	88.27	90.41	92.59	95	98.34	102.38	106.3
	Mean	83.06	88.69	94.05	98.69	103.0	106.94	110.14	112.71	114.87	117	119.88	123.33	127.69
	3rd quartile	127	134.71	141.72	147.24	152.21	156.87	161.05	164.54	167.4	170.31	173.11	178.22	184.87
	Maximum	599.56	602.77	620.56	636.28	650.25	661.39	669.52	674.7	679.09	679.78	676.46	684.32	831.11
Live growth	Minimum	-5.66	-6.86	-43.60	-21.13	-10.62	-13.44	-22.80	-32.27	-24.60	-10.21	-10.94	-6.57	-12.45
	1st quartile	0	0.46	1.12	1.58	1.95	1.931	1.78	1.29	0.51	0	0	0	0
	Median	2.91	3.72	4.30	4.68	4.96	4.88	4.51	4.02	3.53	2.95	2.408	2.13	1.99
	Mean	6.49	6.84	7.06	6.99	6.79	6.44	6.02	5.61	5.37	5.32	5.36	5.56	5.76
	3rd quartile	9.10	9.50	9.80	9.63	9.45	8.95	8.49	8.18	7.89	7.91	8.06	8.42	8.67
	Maximum	70.95	67.54	70.22	66.27	74.08	63.73	56.33	83.58	110.91	134.04	157.17	180.30	203.43
Entry	Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0
	1st quartile	0	0	0	0.09	0.13	0.12	0.10	0.02	0	0	0	0	0
	Median	0.19	0.29	0.35	0.39	0.40	0.39	0.35	0.28	0.20	0.13	0.043	0	0
	Mean	0.99	0.98	0.90	0.86	0.84	0.78	0.71	0.63	0.60	0.60	0.62	0.65	0.69
	3rd quartile	0.82	0.90	0.95	0.95	0.94	0.91	0.84	0.76	0.71	0.67	0.64	0.65	0.68
	Maximum	65.92	67.49	69.45	43.39	73.56	33.3	32.30	11.83	13.74	15.31	16.88	18.44	20.01
Mortality	Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0
	1st quartile	0	0	0	0	0	0	0	0	0	0	0	0	0
	Median	0	0	0	0.04	0.11	0.10	0.06	0	0	0	0	0	0
	Mean	1.48	1.33	1.19	1.07	0.97	0.88	0.82	0.78	0.79	0.82	0.87	0.92	0.97
	3rd quartile	0.56	0.63	0.72	0.84	0.90	0.87	0.77	0.61	0.47	0.40	0.37	0.32	0.31
	Maximum	162.27	125.27	88.28	63.89	57.18	50.48	43.77	37.07	30.36	29.23	32.39	34.55	37.63
Harvest	Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0
	1st quartile	0	0	0	0	0	0	0	0	0	0	0	0	0
	Median	0	0	0	0	0	0	0	0	0	0	0	0	0
	Mean	0.37	1.14	2.132	2.47	2.73	3.14	3.34	3.3	3.06	2.23	1.66	0.93	0.35
	3rd quartile	0	0	0	0	0	0	0	0	0	0	0	0	0
	Maximum	461.93	493.15	478.92	486.63	493.17	495.13	497.93	575.7	591.30	590.59	602.60	589.36	628.30

and the change that occurs on forestland. Both are important factors when evaluating the state of the forests in an area or nation, but the full implications of the distinction are often lost in aggregations of national forest inventory (NFI) data. That is, (1) and (2) above are sometimes confounded. The discussion below assumes that there exists a method of subsetting land by the temporal period of forestland classification, within the temporal period of interest. The methods discussed here can then be applied to the subsets individually. Van Deusen and Roesch (2009) and Roesch and Van Deusen (2012) explored estimation of the change in forestland classification in the context of the NFI in the United States, while the current paper concentrates on estimation of the change in the tree population on land that remains forested throughout the period of interest.

Historically, the definition of the components of growth was sample dependent. The resulting quantities could easily be calculated from remeasured samples but were not truly estimates of population parameters. Eriksson (1995) addressed this deficiency with the presentation of a set of definitions for the continuous components of change. In this paper, I subscribe to the definitions of Roesch (2007b), which presented a discrete version of the Eriksson definitions and discussed the distinctions between the discrete population components of change and the traditional components of growth.

Consider a sample design that consists of g mutually exclusive, spatially disjoint temporal panels in which, subsequent to a random areal start, one panel per year is measured, in turn, for g consecutive years. After each cycle, the panel measurement sequence reinitiates. Such a design is discussed in Bechtold and Patterson (2005) and Roesch (2007b). For instance, in a five-panel system, the five panels

are measured for growth over a 10-year period, that is panel 1 is measured in years 1 and 6, panel 2 is measured in years 2 and 7, etc. Several philosophies have emerged as to how data resulting from this design should be applied to estimates of growth and change because remeasurement of the panels provides observations that are spatially disjoint but temporally overlapping, and the temporal scale of the population of interest is finer than the scale of observation. Roesch (2007b) argued that the average annual growth within each individual panel is best applied to the center of the measurement interval, which is analogous to an assumption of linear change between observations. Presumably, this would be a reasonable first approximation in lieu of contradictory evidence. All analytical methods proposed to date, for this class of sample designs for forest monitoring, have been predicated on this or similar assumptions. This is also true for most of the discussion below, however, I do look at the effects of a simple nonconforming population trend, which can be either latent or overt.

The population components of change are compatible, that is

$$Y_{t+1} = Y_t + L_{t,t+1} + E_{t,t+1} - M_{t,t+1} - H_{t,t+1}, \quad (1)$$

where

Y_t = the value of interest at time t .

$L_{t,t+1}$ = growth in the value of interest on live trees between time t and time $t + 1$.

$E_{t,t+1}$ = the value of interest on live trees as they enter the population between time t and time $t + 1$.

$M_{t,t+1}$ = the value of interest on trees as they die between time t and time $t + 1$.

Table 2. Distribution statistics for population 2.

Component	Statistic	Year												
		1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Volume	Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0
	1st quartile	4.68	11.13	16.78	21.74	26.16	30.62	34.15	36.32	38.42	39.84	41.33	42.10	43.36
	Median	56.13	64.25	71.04	76.52	81.91	85.37	87.90	89.73	91.90	94.25	97.37	101.10	104.79
	Mean	83.07	88.72	94.05	98.63	102.91	106.77	109.87	112.23	114.24	116.25	119.01	122.40	126.71
	3rd quartile	126.99	134.77	141.73	147.21	152.17	156.79	160.95	164.30	167.08	169.98	172.70	177.70	184.12
	Maximum	595.09	603.74	622.28	638.03	651.04	661.42	668.94	673.73	677.89	678.48	675.11	691.20	828.74
Live growth	Minimum	-5.71	-6.83	-43.71	-21.18	-10.77	-13.48	-22.79	-32.38	-24.33	-10.20	-11.00	-6.54	-12.55
	1st quartile	0	0.46	1.12	1.58	1.95	1.93	1.78	1.29	0.50	0	0	0	0
	Median	2.91	3.71	4.30	4.68	4.96	4.88	4.51	4.02	3.52	2.95	2.41	2.13	1.99
	Mean	6.49	6.84	7.06	6.99	6.79	6.44	6.02	5.61	5.37	5.32	5.36	5.56	5.76
	3rd quartile	9.10	9.49	9.80	9.63	9.45	8.96	8.49	8.18	7.88	7.91	8.06	8.42	8.67
	Maximum	71.19	67.80	70.52	67.14	74.87	63.18	56.71	83.66	111.09	134.23	157.36	180.49	203.62
Entry	Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0
	1st quartile	0	0	0	0.09	0.13	0.12	0.10	0.02	0	0	0	0	0
	Median	0.19	0.29	0.35	0.39	0.40	0.39	0.35	0.28	0.20	0.13	0.04	0	0
	Mean	1.00	0.98	0.90	0.86	0.84	0.78	0.71	0.63	0.60	0.60	0.62	0.65	0.69
	3rd quartile	0.82	0.90	0.95	0.95	0.94	0.91	0.84	0.76	0.71	0.67	0.64	0.65	0.68
	Maximum	66.43	67.37	69.43	42.65	73.55	33.63	32.12	11.68	13.49	15.06	16.63	18.20	19.76
Mortality	Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0
	1st quartile	0	0	0	0	0	0	0	0	0	0	0	0	0
	Median	0	0	0	0.04	0.11	0.10	0.06	0	0	0	0	0	0
	Mean	1.48	1.33	1.19	1.07	0.97	0.87	0.82	0.78	0.79	0.82	0.87	0.92	0.97
	3rd quartile	0.56	0.63	0.72	0.84	0.90	0.87	0.76	0.61	0.47	0.40	0.37	0.32	0.31
	Maximum	162.00	125.00	88.01	63.83	57.13	50.42	43.71	37.01	30.30	29.03	32.19	34.57	37.65
Harvest	Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0
	1st quartile	0	0	0	0	0	0	0	0	0	0	0	0	0
	Median	0	0	0	0	0	0	0	0	0	0	0	0	0
	Mean	0.34	1.17	2.19	2.51	2.81	3.23	3.55	3.45	3.18	2.35	1.72	0.99	0.38
	3rd quartile	0	0	0	0	0	0	0	0	0	0	0	0	0
	Maximum	459.38	481.46	485.90	497.54	505.01	504.85	514.68	577.39	586.59	589.27	599.83	594.19	676.00

$H_{t,t+1}$ = the value of interest on trees as they are harvested between time t and time $t + 1$.

As such, it is often argued that estimators of these components should also be compatible. In this article, I compare four general approaches for estimating the components of forest change utilizing data from an annually rotating five-panel sample design, with 5-year change observations made for each panel. In the first three approaches, each component is estimated independently of the others. In the fourth approach, the estimates are constrained to be compatible. For clarity and without loss of generality, assume that the estimands of interest are the cubic meter volume per hectare of all live trees in a fixed area in each component of change category during each year of a multidecadal period.

One estimation approach uses a centralized difference estimator (CDE), the second uses the exponentially weighted difference estimator (EWD) of Roesch (2007a), another uses the (semicentralized) moving-window mean of ratios estimator (MWMOR) in Roesch and Van Deusen (2013), and the fourth uses the mixed-estimator variant in Roesch (2007a). Three general transition models were each combined with a single compatibility model for the mixed-estimation approach. Because we know that compatibility will have a cost in terms of squared error loss for one or more of the estimated components, the initial results compare only the first three approaches. The fourth approach is discussed in a situation in which the first three approaches are shown to give unfavorable results in lieu of the incorporation of extraneous information.

Roesch and Van Deusen (2013) discussed the effects of ignoring differences in the temporal aspects of a realized sample from the intended design. Specifically, they showed the effects of two (usually

tactic) assumptions in NFI designs. The first was that variation in the time of observation for an individual areal sample is ignorable. The second was that variation in the remeasurement period lengths between individual plots in successive areal samples was ignorable. They explored the effects of these assumptions and discussed how inference can be improved by a judicious accounting of these sampling disparities. They showed the remeasurement period assumption to be especially problematic, that is, plots in NFI systems are never remeasured on exact temporal intervals, and large biases can be introduced when there is little effort made to restrict the distribution of temporal interval lengths. They concluded that further research was needed to determine what restrictions should be placed on the distribution of temporal intervals to achieve specific objectives.

Again without loss of generality, in this paper, I define: entry as the cubic meter volume (or value) of trees as they attain the entry criterion; live growth as the annual growth in volume that occurs on trees after a defined entry criterion has been achieved; mortality as the volume of trees as they die; and harvest as the volume of trees as they are harvested. In addition, without loss of generality I use the entry criterion of 12.7 cm at 1.3716 m above the ground (dbh). A reader interested in how definitions of the components of change differ from the traditional components of growth, as in Meyer (1953), can refer to Eriksson (1995) or Roesch (2007b).

Estimators for the Components of Change

I first present three estimators that can be expected to yield equivalent results until they are differentially affected when an underlying assumption becomes tenuous. I then present a reweighting scheme that can be applied to each of the three estimators to incorporate

Table 3. Distribution statistics for population 3.

Component	Statistic	Year												
		1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Volume	Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0
	1st quartile	4.95	11.35	16.92	21.96	26.34	30.84	33.85	36.19	38.53	39.87	41.12	42.26	43.58
	Median	56.68	64.91	71.49	76.92	82.19	85.65	86.53	88.33	90.98	93.18	96.48	100.53	103.69
	Mean	84.02	89.49	94.64	99.10	103.25	106.97	108.10	110.65	112.81	114.99	117.83	121.32	125.72
	3rd quartile	127.27	135.26	142.04	147.53	152.35	156.86	157.92	161.43	164.42	167.13	170.33	175.53	181.97
	Maximum	667.96	609.09	628.10	644.38	658.01	668.73	667.17	672.10	676.24	676.75	673.31	688.97	822.51
Live growth	Minimum	-5.68	-6.76	-43.76	-21.12	-10.72	-13.35	-22.78	-32.17	-24.85	-10.08	-11.08	-6.56	-12.57
	1st quartile	0	0.47	1.12	1.58	1.95	1.93	1.78	1.29	0.50	0	0	0	0
	Median	2.93	3.73	4.31	4.68	4.96	4.88	4.51	4.02	3.52	2.95	2.41	2.13	1.99
	Mean	6.53	6.85	7.06	6.99	6.79	6.44	6.02	5.61	5.37	5.32	5.36	5.56	5.76
	3rd quartile	9.16	9.51	9.80	9.63	9.45	8.95	8.49	8.18	7.88	7.91	8.06	8.42	8.67
	Maximum	71.03	67.42	70.88	66.71	74.77	63.07	56.87	84.05	111.66	134.79	157.92	181.05	204.18
Entry	Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0
	1st quartile	0	0	0	0.09	0.13	0.12	0.10	0.02	0	0	0	0	0
	Median	0.19	0.29	0.35	0.39	0.40	0.39	0.35	0.28	0.20	0.13	0.04	0	0
	Mean	0.99	0.99	0.91	0.86	0.84	0.78	0.71	0.63	0.60	0.60	0.62	0.65	0.69
	3rd quartile	0.82	0.90	0.95	0.95	0.95	0.91	0.84	0.76	0.71	0.67	0.64	0.65	0.68
	Maximum	66.93	67.97	69.15	43.37	72.30	44.83	32.57	11.77	13.61	15.17	16.74	18.31	19.88
Mortality	Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0
	1st quartile	0	0	0	0	0	0	0	0	0	0	0	0	0
	Median	0	0	0	0.04	0.11	0.38	0.06	0	0	0	0	0	0
	Mean	1.68	1.53	1.39	1.27	1.17	2.96	0.81	0.78	0.78	0.81	0.84	0.89	0.94
	3rd quartile	0.58	0.65	0.75	0.87	0.97	3.13	0.76	0.60	0.46	0.40	0.36	0.32	0.30
	Maximum	161.99	132.67	124.81	116.96	109.11	101.25	43.73	37.02	30.31	28.49	31.57	34.65	37.73
Harvest	Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0
	1st quartile	0	0	0	0	0	0	0	0	0	0	0	0	0
	Median	0	0	0	0	0	0	0	0	0	0	0	0	0
	Mean	0.37	1.16	2.11	2.43	2.74	3.13	3.36	3.31	3.02	2.27	1.65	0.91	0.37
	3rd quartile	0	0	0	0	0	0	0	0	0	0	0	0	0
	Maximum	455.44	477.77	483.53	484.98	489.37	501.90	506.03	581.44	578.92	584.22	582.56	601.94	628.32

extraneously obtained information when an anomaly occurs that can be expected to affect these estimators under the given sampling design. Subsequently, the extraneously obtained weights are incorporated directly into a mixed estimator, and all four estimators are tested in the anomalous situation.

Centralized Difference Estimator

The CDE is a moving average estimator applied to a series of within-panel differences (i.e., a series of change component values). The CDE gives equal weights to the interval observations surrounding the interval of interest. In the CDE, a panel difference is applied to the center interval (or year) and combined with adjoining panel differences.

Let $\bar{d}_{h+t,h+1}$ the annual mean of a remeasured panel difference, such as $t^{-1}\hat{L}_{h+t,h+1}$. $r = (m-1)/2$, where m is odd and is the number of remeasured panels used in the estimator.

The CDE for component C in year k is then

$$C_k^{CDE} = m^{-1} \sum_{i=-r}^r \bar{d}_{k+r+i,k-r+i} \tag{2}$$

A trivial method of assuring compatibility when using this estimator would be to estimate entry, live growth, mortality, and harvest separately and then use the sum of these component vectors to estimate the total value vector. The CDE provides no estimates for m years on each end of the time string. In practice, an ad hoc variation would have to be incorporated to provide some or all of these estimates. Here, five panels are used in the estimator when

there are differences from at least two panels available before and after the central panel. Three panels are used in the estimator when there is only one additional panel difference available on either side of the central panel (e.g., for the second and penultimate estimates). A single panel difference is used for the first and the final estimates.

Exponentially Weighted Difference Estimator

An estimator introduced in Roesch (2007a) is the EWD, similar in concept to the exponentially weighted moving average (EWMA) estimator common in the quality control literature (i.e., Chandra 2000) and the econometrics literature (i.e., West and Harrison 1989, p. 55). In the EWD, a series of differences (i.e., a series of change component values) within panels is calculated. The EWD gives larger weights to the interval observations closest to the interval of interest, allowing more local variation than if equal weights are used. In the EWD, the panel difference is applied to the center interval and combined with the $m-1$ adjoining interval differences. The supporting panels are down-weighted exponentially with each step away from the central interval.

In addition to the notation above for the CDE, let

$$\alpha = \left(\frac{r}{r+1} \right).$$

The EWD for component C in year y is then

$$C_y^{EWD} = \sum_{i=-r}^r \frac{(1-\alpha)}{(1+\alpha-2\alpha^{r+1})} \alpha^{|i|} \bar{d}_{y+r+i,y-r+i} \tag{3}$$

As with the CDE, a trivial method of assuring compatibility

Table 4. Distribution statistics for population 4.

Component	Statistic	Year												
		1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Volume	Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0
	1st quartile	7.55	13.69	18.64	21.62	25.20	27.87	30.13	31.57	32.68	33.18	33.70	34.31	35.19
	Median	60.42	67.19	72.32	76.17	79.53	80.75	80.88	80.38	80.30	80.88	81.91	83.56	85.78
	Mean	86.40	91.11	95.22	98.23	100.65	102.47	103.42	103.70	103.65	103.61	104.24	105.51	107.70
	3rd quartile	130.65	137.25	143.28	146.07	148.62	150.60	151.03	151.21	151.33	151.43	152.03	152.77	155.92
	Maximum	608.21	602.18	615.03	625.44	634.34	640.85	645.19	647.47	649.17	648.63	645.73	640.58	674.22
Live growth	Minimum	-4.73	-5.33	-32.68	-15.17	-7.41	-9.00	-14.86	-20.59	-15.24	-6.27	-6.59	-3.93	-7.37
	1st quartile	0	0.49	0.92	1.20	1.37	1.33	1.17	0.83	0.31	0	0	0	0
	Median	2.47	3.05	3.31	3.43	3.49	3.33	2.97	2.58	2.22	1.83	1.46	1.28	1.15
	Mean	5.69	5.83	5.64	5.29	4.89	4.44	3.97	3.61	3.39	3.34	3.40	3.59	3.81
	3rd quartile	7.69	7.87	7.65	7.08	6.71	6.12	5.60	5.25	4.97	4.94	4.99	5.24	5.50
	Maximum	67.08	60.22	60.14	52.60	57.23	45.53	37.36	53.90	69.99	93.12	116.25	139.38	162.51
Entry	Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0
	1st quartile	0	0	0	0.08	0.11	0.10	0.08	0.01	0	0	0	0	0
	Median	0.19	0.27	0.31	0.33	0.33	0.32	0.28	0.22	0.15	0.10	0.03	0	0
	Mean	0.94	0.84	0.80	0.76	0.70	0.61	0.58	0.50	0.47	0.47	0.49	0.52	0.56
	3rd quartile	0.78	0.81	0.84	0.81	0.80	0.74	0.67	0.60	0.56	0.52	0.50	0.51	0.54
	Maximum	66.89	62.86	49.51	53.90	24.77	13.09	31.69	9.36	10.71	12.28	13.85	15.41	16.98
Mortality	Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0
	1st quartile	0	0	0	0	0	0	0	0	0	0	0	0	0
	Median	0	0	0	0.05	0.12	0.11	0.07	0	0	0	0	0	0
	Mean	1.56	1.42	1.28	1.17	1.07	0.99	0.94	0.92	0.93	0.96	1.00	1.03	1.07
	3rd quartile	0.63	0.70	0.78	0.93	1.01	0.99	0.88	0.71	0.55	0.48	0.44	0.38	0.34
	Maximum	167.89	130.90	93.90	69.67	62.96	56.25	49.55	42.84	36.13	34.56	37.64	40.72	43.80
Harvest	Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0
	1st quartile	0	0	0	0	0	0	0	0	0	0	0	0	0
	Median	0	0	0	0	0	0	0	0	0	0	0	0	0
	Mean	0.36	1.15	2.14	2.46	2.71	3.11	3.33	3.24	2.97	2.22	1.62	0.89	0.36
	3rd quartile	0	0	0	0	0	0	0	0	0	0	0	0	0
	Maximum	474.55	479.35	482.29	486.63	485.76	501.16	496.55	542.87	559.86	553.63	552.81	554.08	618.84

when using this estimator would be to estimate each component separately and then use the sum of the component vectors to estimate the total value vector. The EWD also does not provide estimates for m years on each end of the time string. Here, as with the CDE, five panels are used in the estimator when there are differences from at least two panels available before and after the central panel. Three panels are used in the estimator when there are differences from at least one panel available before and after the central panel. A single panel difference is used otherwise, rendering the EWD equal to the CDE when there is not at least one panel difference available before and after the central panel.

Moving-Window Mean of Ratios Estimator

Roesch and Van Deusen (2013) proposed an estimator that arose from a different perspective than the two estimators above that could also be applied to the current objective. In the temporal dimension, the idea was simple. One stacks the observations on a temporal scale (or a function of the temporal scale) and then slices through the stack (say to create annual segments) to determine how much of each observation contributes to the estimate for each year. For this problem, as in Roesch and Van Deusen (2013), I use the general three-dimensional selection model given in Roesch (2008) with the exception that time will be rescaled relative to the proportion of the growing season elapsed within each year. Assign to each observation of variable x labels for plot i and a superscript representing the beginning value and ending value as x_i^b and x_i^e , respectively. Because there are no observations between x_i^b and x_i^e , the distribution of the volume growth between the two observations must be modeled. Initially, I make two simplifying assumptions, both of

which can be refined by an appropriate model, as needed. In the first, I assume that the growing season does not vary within the area of interest. In the second, I assume that growth for each plot is uniform throughout the growing season. I can then temporally order each observation by the year of observation plus the proportion of the growing season that has elapsed (i.e., in the format $year.p$), and use s_i to represent the temporal span between the beginning and ending observations. I then allocate the proportion of growth observed over s_i to the proportion of each year spanned by s_i , (thereby accounting for the marginal probability of the time dimension). Modeling growth between observations allows me to allocate growth within components to the years the growth occurred. A simple time-adjusted estimator for annual volume growth (within growth component) is the MWMOR for component C in year y .

$$C_y^{MWMOR} = \frac{1}{n_y} \sum_{i=1}^{n_y} \frac{c_{i,y}}{p_{i,y}}, \tag{4}$$

where

n_y = the number of plots observing growth in year y .

$p_{i,y}$ = the product of portion of year y growing season observed by plot i and the portion of plot i area within the area of interest.

$c_{i,y}$ = the value of component C observed on plot i , assignable to year y .

The general statistical properties of ratio estimators are well known by now and can be found in such early works as Raj (1968), Walton and DeMars (1973), Cassel et al. (1977), and Cochran (1977).

Population 1 - Sample Error Structure 1

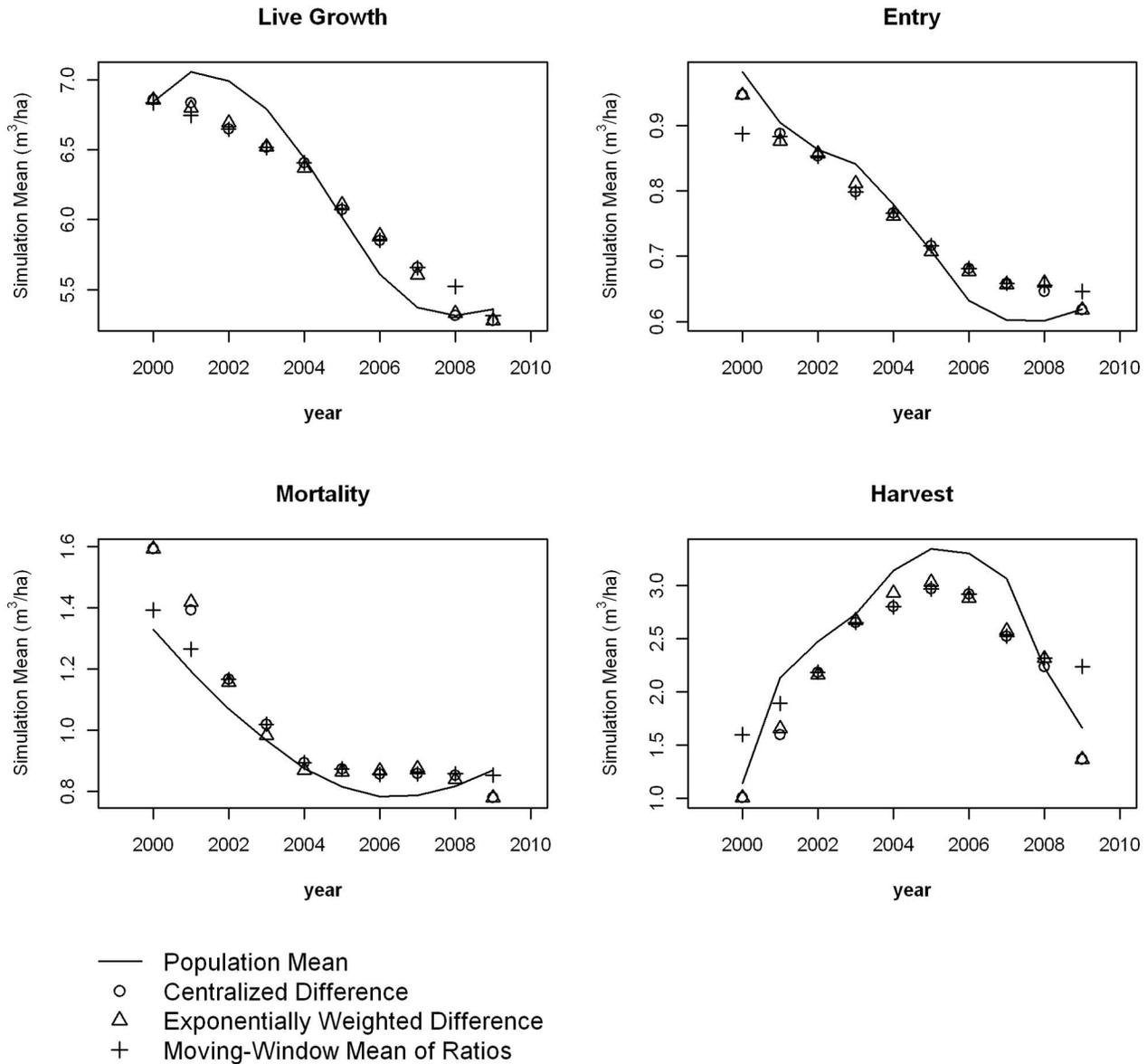


Figure 1. The mean over 1,000 iterations of 1,000 samples each from population 1 under sampling error structure 1 for the CDE, EWD, and MWMOR by growth component and estimation year.

Incorporating Outside Information

Situations arise within the scope of national-scale forest monitoring efforts for which the data obtained from the sample design are inadequate. Many approaches to incorporating extraneous information in forest inventories have been proposed and proven useful for particular applications. The example below demonstrates the utility of a weighting method that can be used in conjunction with each of the estimators above and with the mixed estimator below.

Assume that there exists strong external information that suggests that expected value of X at time t , $E(X_t)$ and only at time t differs from the expected value of a previous estimate of X , $E(X_t^p)$, at time t by a factor kw . For the estimators above, this suggests a

reweighting of previous estimates for all estimates that had used the previous estimate for time t .

To accomplish this reweighting, let

$$kv = kw + n_t - 1, \quad (5)$$

where n_t is the number of years used for each annual estimate, and let

$$kt^- = n_t/kv, \quad (6)$$

and

$$kt = (n_t * kw) / kv. \quad (7)$$

Population 1 - Sample Error Structure 1

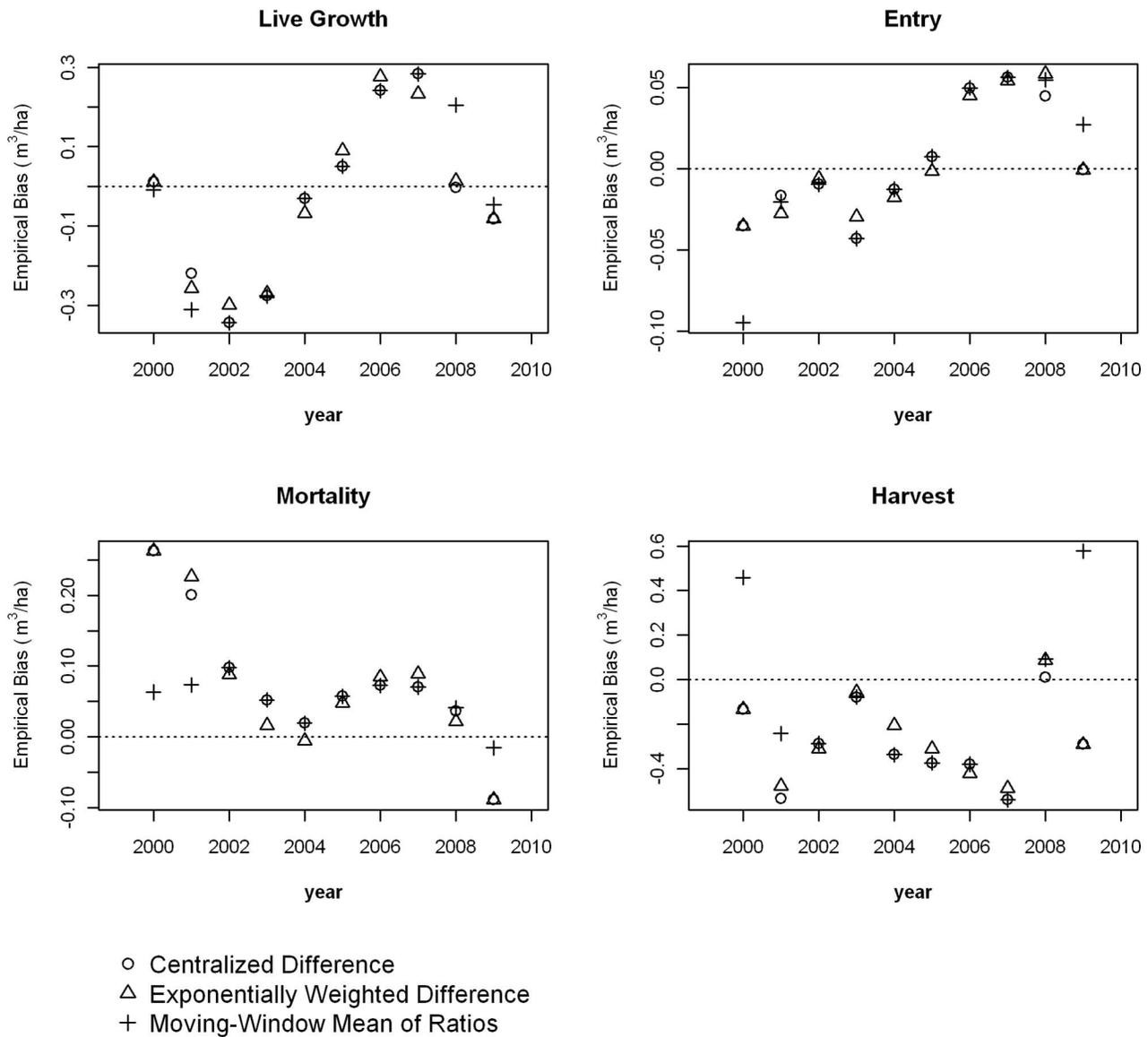


Figure 2. The EB over 1,000 iterations of 1,000 samples each from population 1 under sampling error structure 1 for the CDE, EWD, and MWMOR by growth component and estimation year.

Then weight the previous estimates at times other than time t that used the previous time t estimate by kt^{-} , and weight previous estimates at time t by kt to form the reweighted estimates.

Mixed Estimation

The mixed estimator (Theil 1963) can be used to draw strength from overlapping panels and easily incorporate extraneous information into forest monitoring efforts. Mixed estimation was first proposed for use in forestry when Korhonen (1993) used the method for calibrating tree volume functions. Van Deusen (1996, 1999, and 2000) developed mixed estimators for annual forest inventory designs, and Roesch (2007a) used it for components of change estimation. Here, I use mixed estimation as a convenient way to incorpo-

rate both (1) a belief in how the individual growth components should be related and transition from year to year and (2) extraneous information that suggests that a modification to those beliefs is appropriate. To achieve these goals, I start with the three transition models below and then adapt those models to incorporate the extraneous information. The three base models all assume compatibility of the total annual change with the components of change; that is, for each year t

$$\delta_{i,t+1} = \hat{V}_{t+1} - \hat{V}_t = \hat{L}_{i,t+1} + \hat{E}_{i,t+1} - \hat{M}_{i,t+1} - \hat{H}_{i,t+1}. \quad (8)$$

Initially, Model 1 assumes that for each component $C = L, E, M,$ or H , at each time t

Population 1 - Sample Error Structure 1

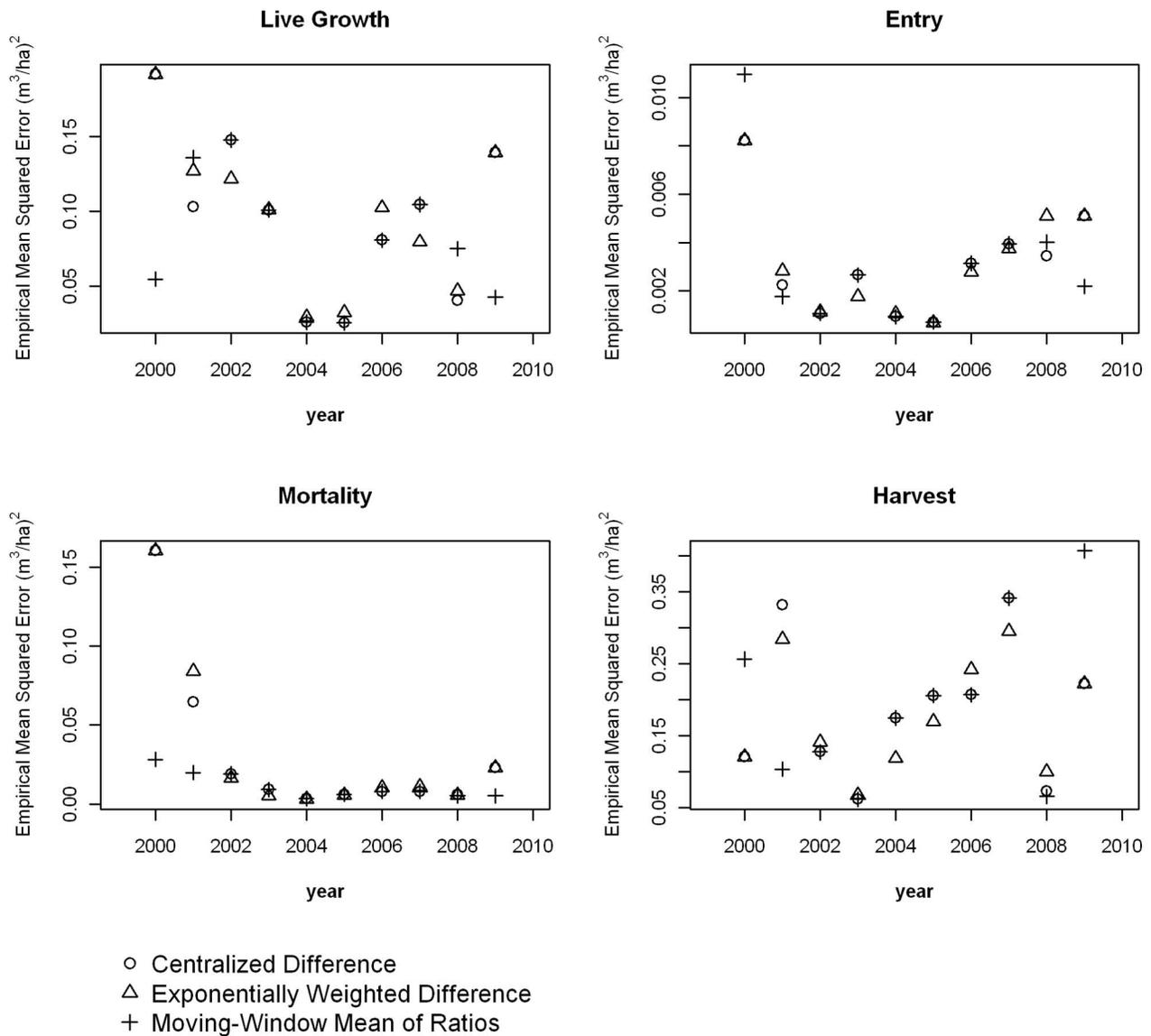


Figure 3. The EMSE over 1,000 iterations of 1,000 samples each from population 1 under sampling error structure 1 for the CDE, EWD, and MWMOR by growth component and estimation year.

$$\hat{C}_{t,t+1} - \hat{C}_{t+1,t+2} = \varepsilon_{C,t} \quad (9)$$

Model 2 assumes that for each component, at each time t

$$\hat{C}_{t,t+1} - 2\hat{C}_{t+1,t+2} + \hat{C}_{t+2,t+3} = \varepsilon_{C,t} \quad (10)$$

while Model 3 assumes that for each component, at each time t

$$\hat{C}_{t,t+1} - 3\hat{C}_{t+1,t+2} + 3\hat{C}_{t+2,t+3} - \hat{C}_{t+3,t+4} = \varepsilon_{C,t} \quad (11)$$

Formulation of the constraints under each of these models is straightforward.

Let $\mathbf{Y} = \mathbf{an}(n_c^* n_t)$ row \times n_t column response matrix, where $n_c = 5$ is the number of growth components plus 1, n_t is the number of

years in the estimation interval, and n_s is the sample size for all n_p panels. The columns of \mathbf{Y} are arranged in n_t successive ordered five-tuples of (1) the MWMOR estimate of annual change, (2) observed annual live growth, (3) entry, (4) mortality, and (5) harvest, for each year in the observation interval for a plot and zeroes otherwise.

A user's level of belief in the underlying model can be incorporated into the mixed estimator in a number of ways. In a similar application, Van Deusen (1999) showed that the level of belief in the constraints could be incorporated into the estimation process by choice or estimation of the value of a parameter p in the mixed estimator. Here, I choose to preprocess \mathbf{Y} by first reweighting according to the extraneous information, adapting the constraint

Population 1 - Sample Error Structure 4

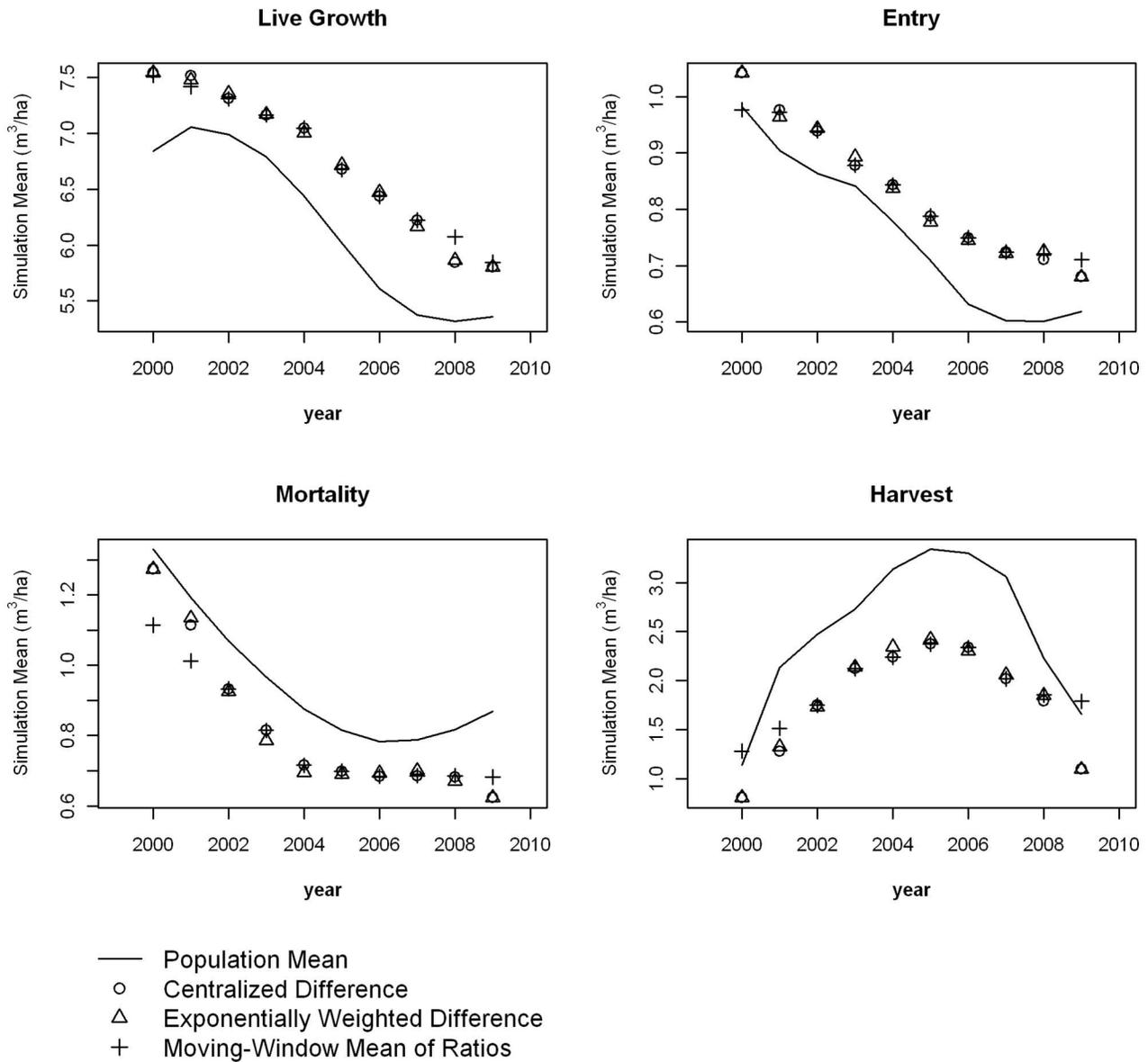


Figure 4. The mean over 1,000 iterations of 1,000 samples each from population 1 under sampling error structure 4 for the CDE, EWD, and MWMOR by growth component and estimation year.

matrix, and then strictly applying the constraints to ensure growth component compatibility. To accomplish this weighting is simple. For estimates at times other than time t , that use estimates for time t , weight the time t estimate by kw^{-1} prior to combining with the nontime t estimates. For estimates at time t , weight the nontime t estimates by kw prior to combining with the time t estimate. Indicate the outside information weighted response matrix as \mathbf{Y}_{OI} . Then, let

$\Sigma = \text{an}(n_c * n_t) \text{ row} \times (n_c * n_t) \text{ column variance/covariance matrix of } \mathbf{Y}_{OI}$.

$\mathbf{R} = \text{an}(n_R) \text{ row} \times (n_c * n_t) \text{ column constraint matrix, appropriate for a given model, where } n_R \text{ is the number of constraint rows.}$

$\hat{\beta} = \text{an}(n_c * n_t) \text{ row} \times 1 \text{ column predictive coefficient vector for strict constraints.}$

$$\hat{\beta} = \mathbf{Y}_{OI} - [\Sigma \mathbf{R}' (\mathbf{R} \Sigma \mathbf{R}')^{-1} \mathbf{R} \mathbf{Y}_{OI}]. \quad (12)$$

Let \mathbf{M} be an $(n_c * n_t) \text{ row} \times n_s$ column matrix, with each column consisting of n_t repetitions of the vector $(n_s^{-1}, n_s^{-1} n_p, n_s^{-1} n_p, n_s^{-1} n_p, n_s^{-1} n_p)$.

Then

$$\hat{\mathbf{Y}}_{mix} = \hat{\beta} \mathbf{M}'. \quad (13)$$

Population 1 - Sample Error Structure 4

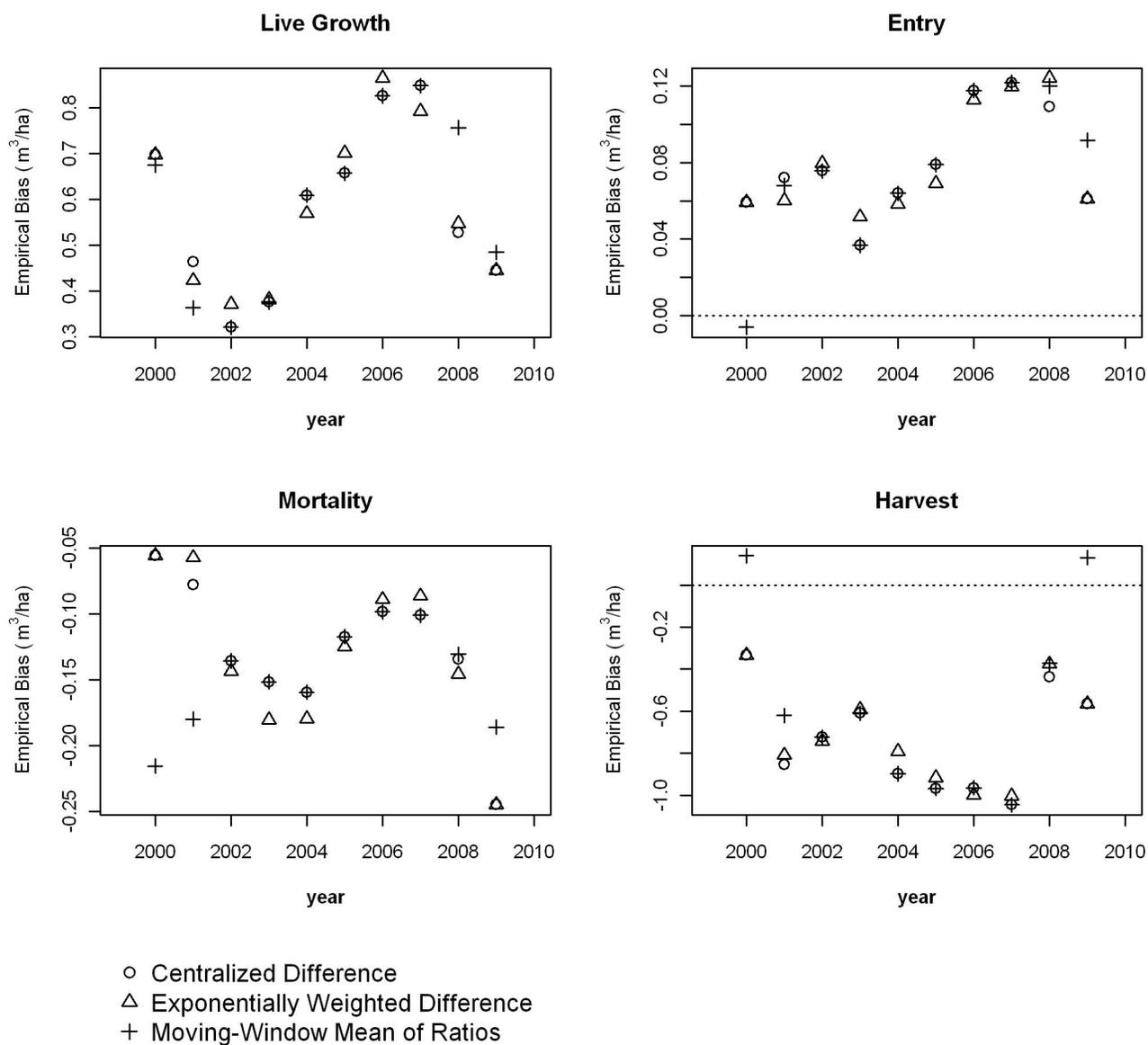


Figure 5. The EB over 1,000 iterations of 1,000 samples each from population 1 under sampling error structure 4 for the CDE, EWD, and MWMOR by growth component and estimation year.

Population Simulations

To motivate the discussion, in this application, I use data from the USDA Forest Service's Forest Inventory and Analysis Program (FIA) collected in South Carolina to construct five simulated populations. Although all of the populations are plausible, the intention was for the first population (population 0) to be a seed population utilizing the simplest possible model for deriving annual values from multiyear observations. The seed population allows the simulation of a series of populations, some of which we might assume to be like the one from which the sample data were drawn and others that might arise from a wider diversity of conditions. As explained more fully in the ensuing paragraphs, the next four populations differed

from population 0 as follows. For population 1, a mild (latent) nonlinear trend was introduced into each of the components of population 0. For population 2, a mild nonlinear trend was introduced into the components of live growth, entry, and mortality, and a stronger nonlinear trend was introduced into the harvest component. For population 3, a mild nonlinear trend was introduced into the components of live growth, entry, and harvest, and a catastrophic high mortality event of four times the mortality rate of population 1 was introduced for the year 2004. Population 4 was initially as in population 1 and then postulated climate change effects were simulated by increasing mortality and decreasing growth and recruitment, with harvest levels remaining the same. Each of the

Population 1 - Sample Error Structure 4

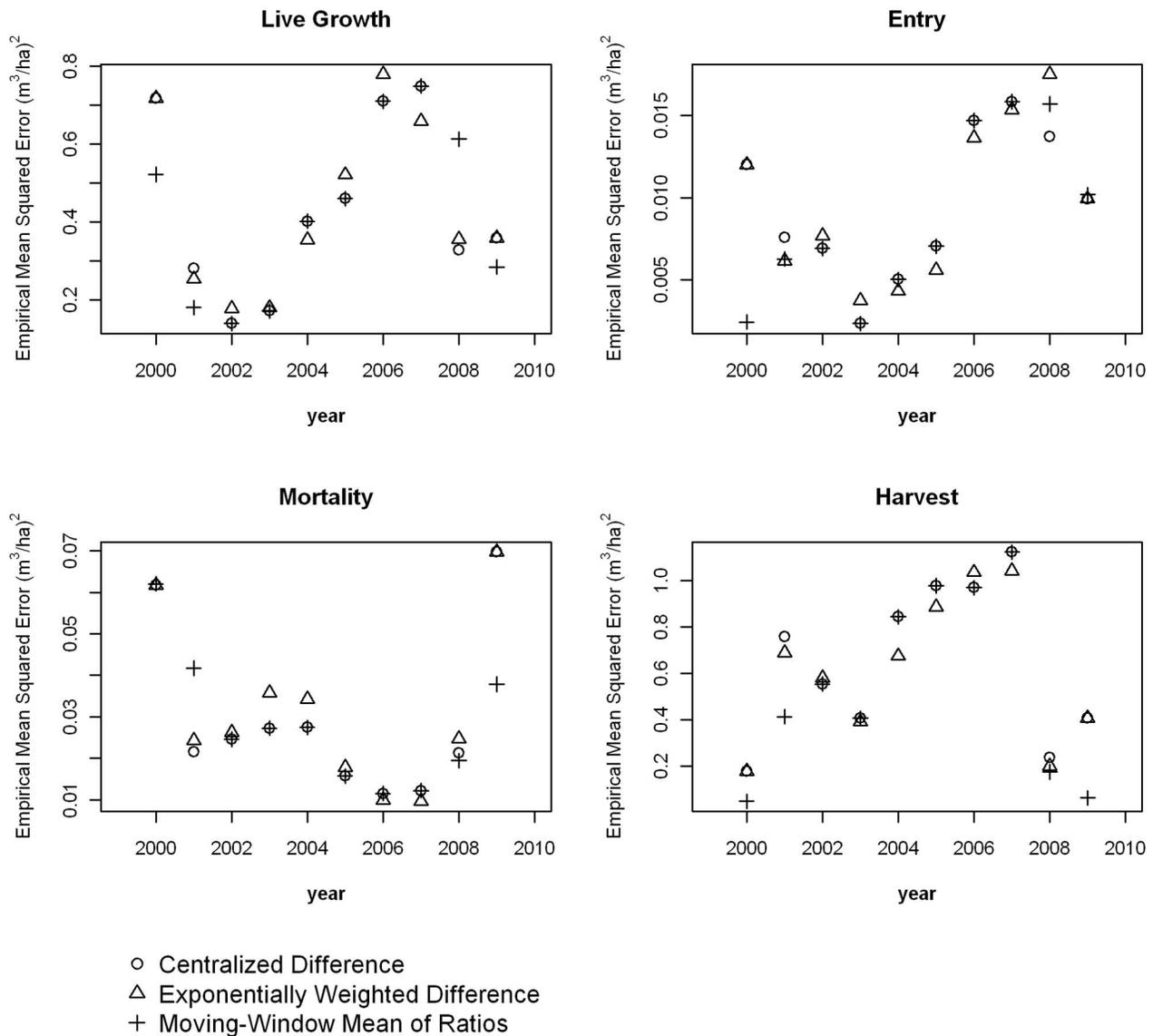


Figure 6. The EMSE over 1,000 iterations of 1,000 samples each from population 1 under sampling error structure 4 for the CDE, EWD, and MWMOR by growth component and estimation year.

five populations consists of 1,458,000 forested “hectares” or “elements” with measurable cubic meter volume at some time in the 14-year period from 1998–2011.

Specifically, I created the five simulated populations by first selecting all remeasured forested plots spanning the 14-year period from the South Carolina data. This resulted in 2,430 forested plots (set 1), most of which had three measurement times (i.e., two observed growth intervals for each component). Because exact harvest times were unknown, harvested volume was randomly allocated to a year within each observation interval. Linear interpolation and extrapolation were used to obtain an initial value for the live growth, entry, and mortality change components for each year from 1998 through 2011 (set 1), as well as a starting cubic-meter per hectare value in the beginning of 1998, with temporal adjustments made as

necessary for high levels of harvest and mortality. (Note that I am not attempting to reconstruct the plot but rather a reasonable facsimile to the forested condition from which it was drawn.) Each population has five (1,458,000 row by 14 column) matrices, one for each change component and one for initial annual volume. Construction of population 0 then proceeded with 600 variance-interjected copies of set 1, resulting in 1,458,000 hectares. Variance was interjected at two levels, in step 1, to keep trend but add variance to the seed, by multiplying all values for each component on each hectare by a random variate, unique for that hectare, drawn from an $N(1, 0.025)$ distribution. The second level of variance was introduced temporally by multiplying the result of step 1 for each annual value for each component on each hectare by a unique random variate drawn from an $N(1, 0.0025)$ distribution.

Population 3 - Sample Error Structure 4

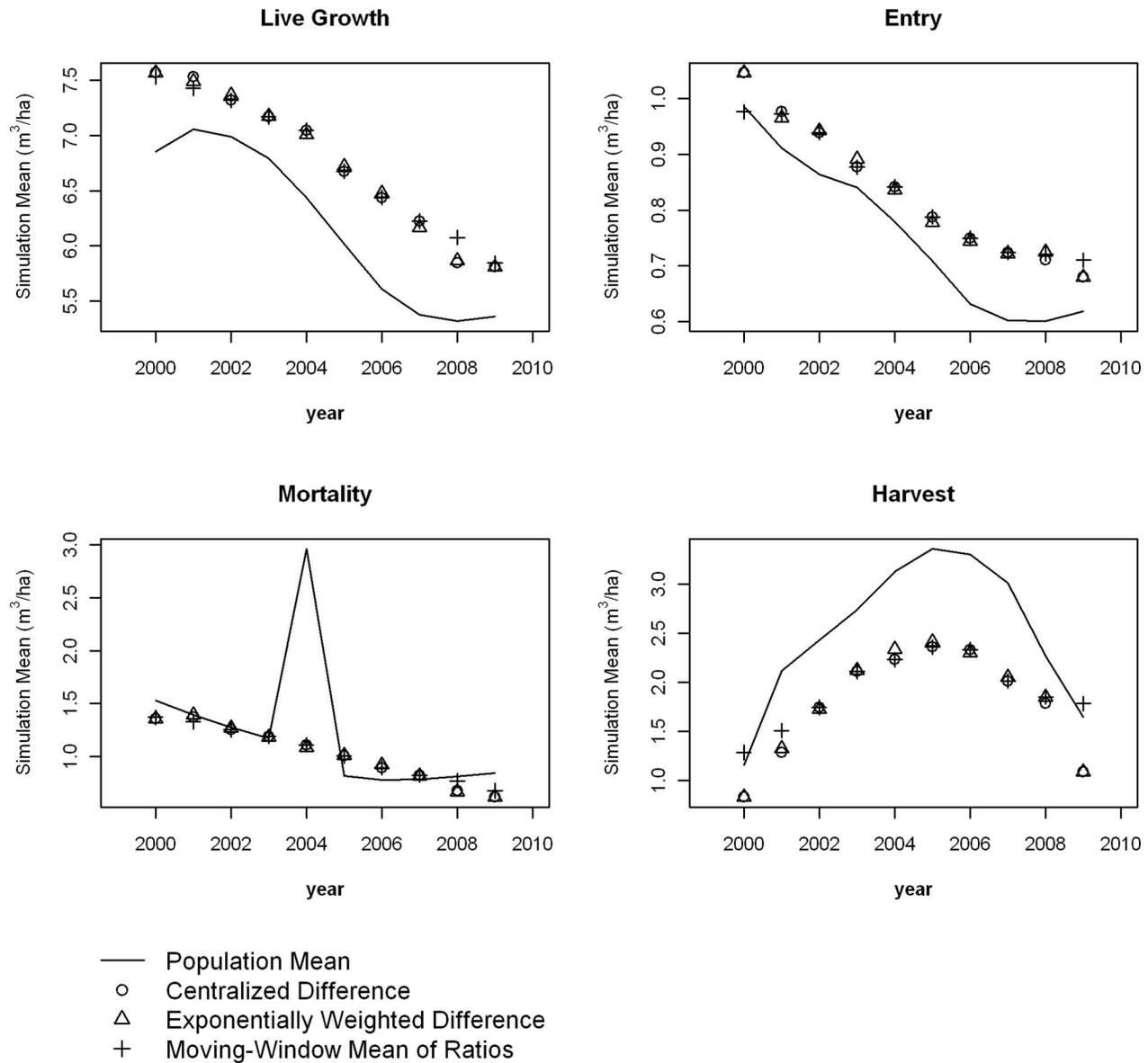


Figure 7. The mean over 1,000 iterations of 1,000 samples each from population 3 under sampling error structure 4 for CDE, EWD, and MWMOR by growth component and estimation year.

Population 1 was initiated in the same manner as population 0 and then multiplied by a mild nonlinear trend

$$T1_i = [0.95 + (.05 \ln(i - 1997))], \quad (14)$$

where each value in each year $i = 1998$ to 2011 is multiplied by $T1_i$. Table 1 gives the population 1 distribution statistics for 1999–2011.

Population 2 had increasing harvesting pressure introduced into the cut component as

$$T2_i = [0.90 + (.10 \ln(i - 1997))], \quad i = 1998 \text{ to } 2011. \quad (15)$$

That is, each value of the cut component in each year i is

multiplied by $T2_i$, while all other components are as in population 1. Table 2 gives the population 2 distribution statistics for 1999–2011.

Population 3 was constructed as in population 1, except with an introduced catastrophic event of four times the amount of mortality in of population 1 in 2004. Table 3 gives the population 3 distribution statistics for 1999–2011.

Population 4 was initially constructed in the same manner as population 0 and then postulated climate change effects were simulated by increasing mortality and decreasing growth and recruitment, with harvest levels remaining the same as in population 1. Specifically, each value of live growth in year i is multiplied by

Population 3 - Sample Error Structure 4

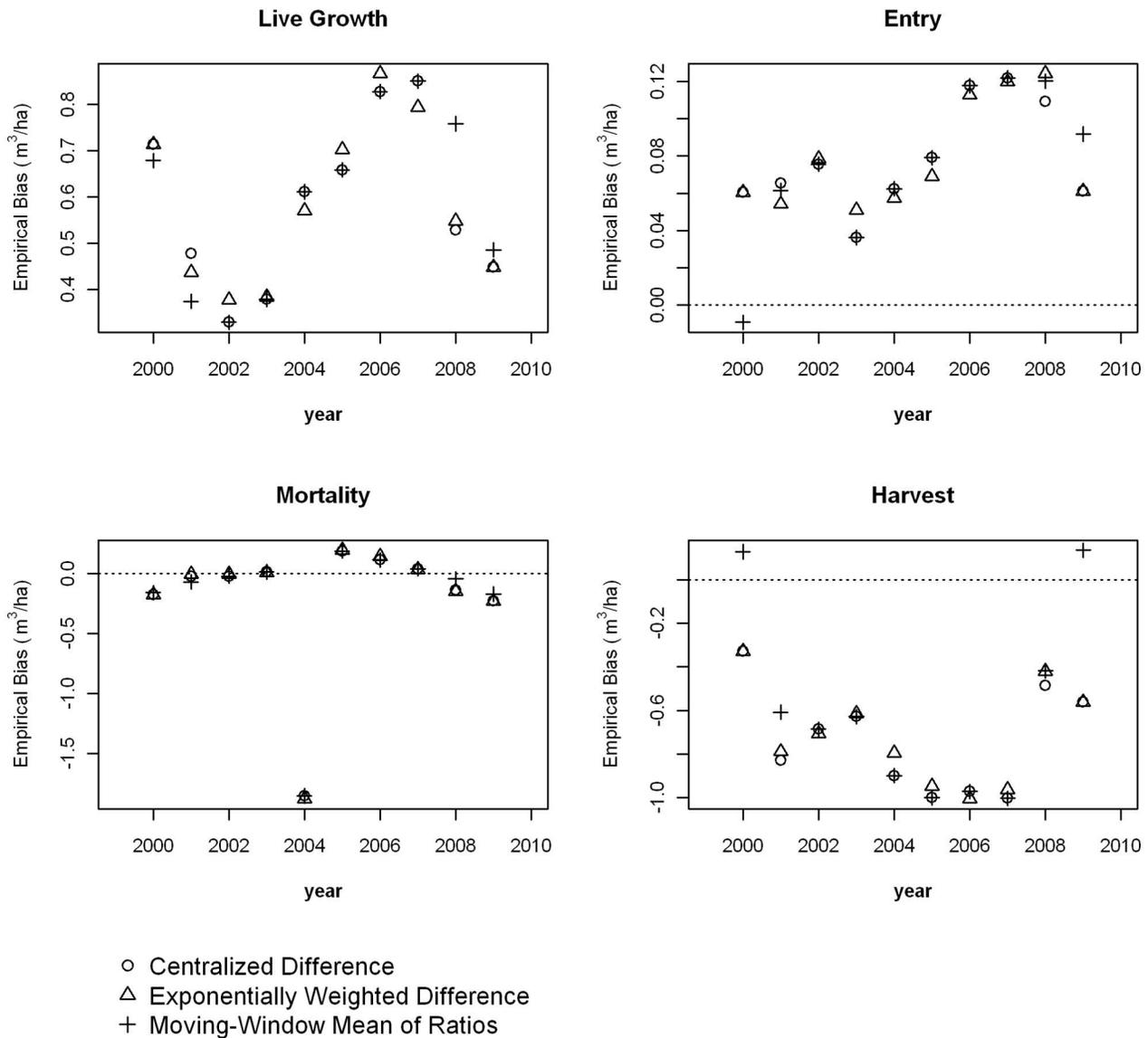


Figure 8. The EB over 1,000 iterations of 1,000 samples each from population 3 under sampling error structure 4 for the CDE, EWD, and MWMOR by growth component and estimation year.

$$T4_i = [0.90 - (.10 \ln(i - 1997))], i = 1998 \text{ to } 2011, \quad (16)$$

while each value of entry in year i is multiplied by

$$T5_i = [0.95 - (.05 \ln(i - 1997))], i = 1998 \text{ to } 2011; \quad (17)$$

and each value of mortality in year i is multiplied by

$$T6_i = [0.80 + (.20 \ln(i - 1997))], i = 1998 \text{ to } 2011. \quad (18)$$

Table 4 gives the population 4 distribution statistics for 1999–2011.

The five populations in this study were constructed to examine estimator performance in the presence of plausible suboptimal population characteristics, in the form of nonlinear trends and a fine-scale anomaly for the given sample design.

Error Structures

Usually, the overriding criterion for selection of an estimator in forest inventories is the minimum mean squared error for the candidate unbiased estimators. Rarely, is the effect of sampling error in the form of bias on the robustness of theoretically unbiased estimators considered. Some notable exceptions have been Gertner (1987), Thomas and Roesch (1990), and more recently Eastaugh and Hasenauer (2013). Eastaugh and Hasenauer (2013) gave the results of a particularly thorough investigation into the potential bias that can be introduced into theoretically unbiased estimators under plausible assumptions concerning common sampling and measurement errors in remeasured forest inventories. In the current investigation, estimator robustness was tested in a simulation by sampling each population under four different assumptions of total

Population 3 - Sample Error Structure 4

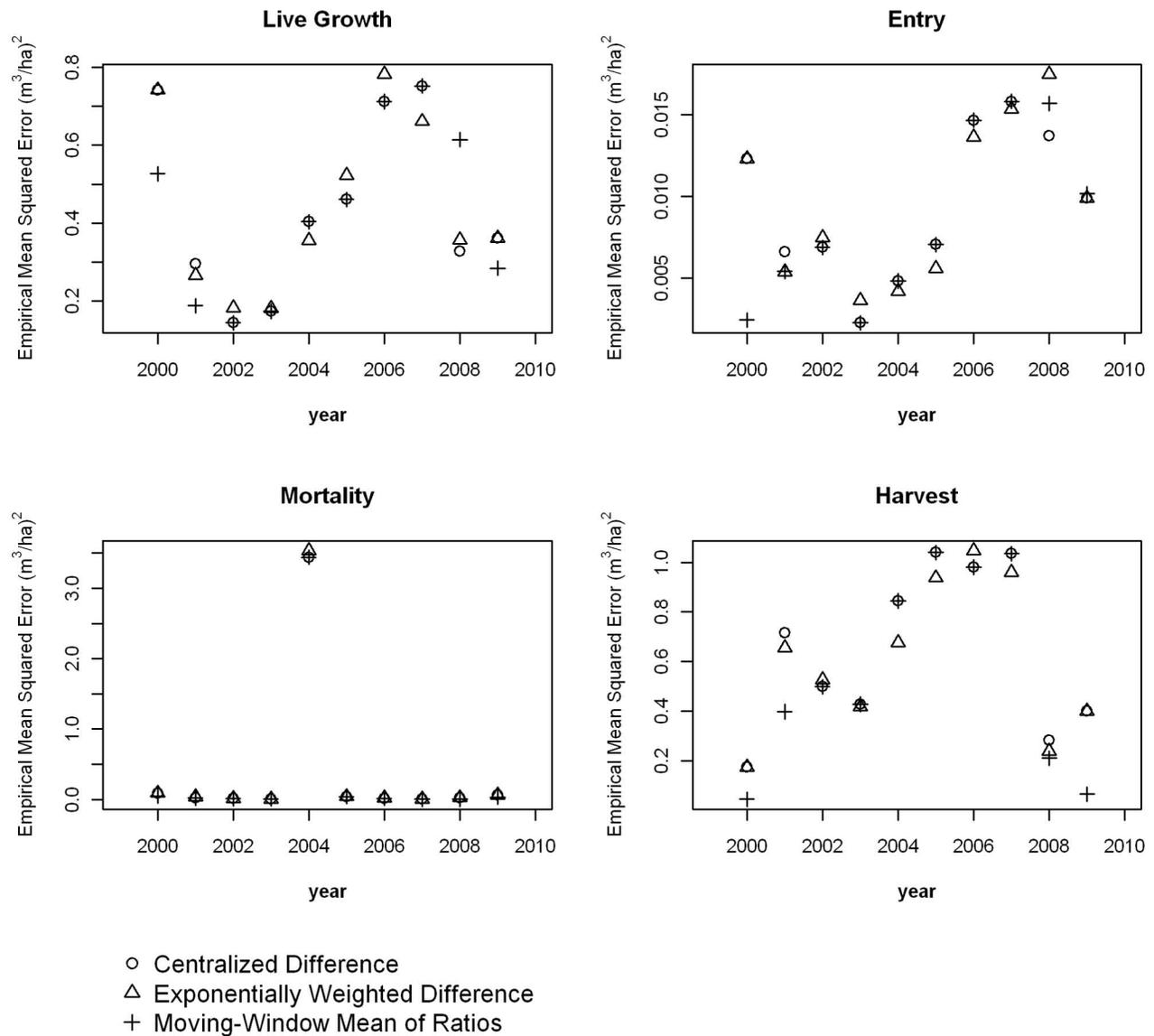


Figure 9. The EMSE over 1,000 iterations of 1,000 samples each from population 3 under sampling error structure 4 for the CDE, EWD, and MWMOR by growth component and estimation year.

sampling error structure. The models are intended to represent all error that would not be addressed by the use of an unbiased sampling simulation. The unbiased simulation might be the closest one can come to the pure sampling error inherent in a perfectly observed and measured sample, while in a realized sample there could also be “item observation” errors, frame errors, and measurement errors to say the least. That is, the error structure models used here are intended to represent all of the ways that a realized sample might differ from its perfectly observed theoretical counterpart.

Each simulation consisted of 1,000 iterations of 1,000 plots each (without replacement) from each population, under each of the four sampling error structures.

For error structure 1, sampling error was assumed to consist exclusively of a small variance, effected by multiplying a unique

random normal deviate of mean 1 and standard deviation (SD) 0.025 by each sampled observation of each component.

For error structure 2, I assumed that sampling error consisted of a small variance and a positive bias on all change components. I effected error structure 2 by multiplying a unique random normal deviate of mean 1.05 and SD of 0.025 by each observation of each component.

For error structure 3, I assumed that sampling error consisted of variance and positive bias on volume, live growth, and entry and a negative bias on harvest and mortality. I effected error structure 3 as follows.

Each observation of live growth and entry was multiplied by a unique random normal deviate of mean 1.05 and SD of 0.025. Each observation of mortality and harvest was multiplied by a unique

Population 4 - Sample Error Structure 3

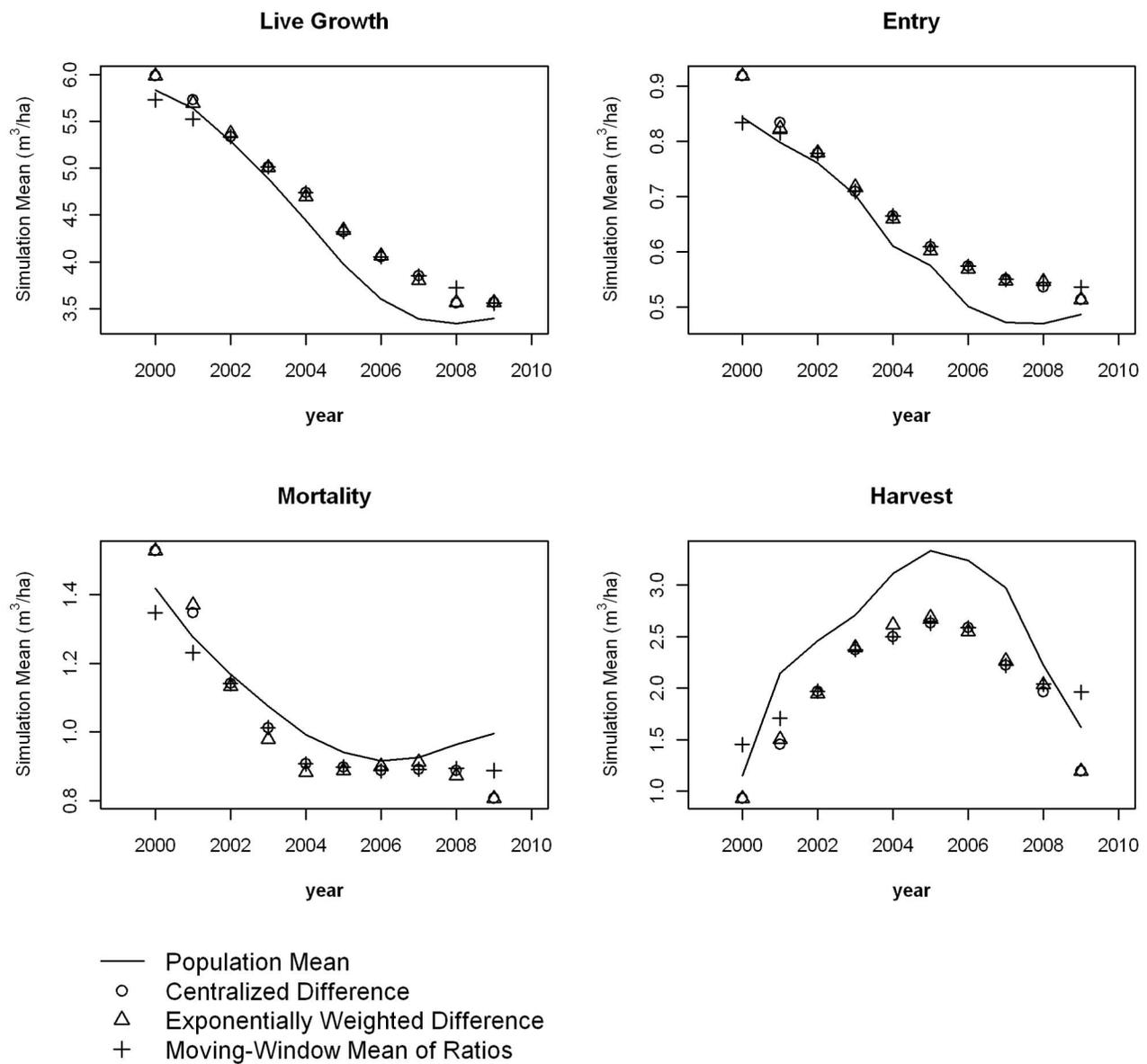


Figure 10. The mean over 1,000 iterations of 1,000 samples each from population 4 under sampling error structure 3 for the CDE, EWD, and MWMOR by growth component and estimation year.

random normal deviate of mean 0.90 and SD of 0.025. Although the level of simulated error is somewhat arbitrary, errors of approximately these magnitudes seem reasonable based on the work in Thomas and Roesch (1990) and the results in Eastaugh and Hasenauer (2013).

Error structure 4 was similar to error structure 3, but I assumed that sampling error consisted of a higher variance and greater bias, as follows: each observation of live growth and entry was multiplied by a unique random normal deviate of mean 1.1 and SD of 0.05. Each observation of mortality and harvest was multiplied by a unique random normal deviate of mean 0.80 and SD of 0.05.

For population 3, I also simulated extraneously obtained information that mortality during 2004 was about 4× higher than in

surrounding years by drawing a random variate from a normal distribution of mean 4 and SD of 0.05 and setting it equal to kw for each observation. Then estimates for each estimator, the CDE, the EWD, and the MWMOR, were reweighted to obtain the CDE-OI, the EWD-OI, and the MWMOR OI estimators. For instance, for ($t = 2002, 2003, 2005, \text{ and } 2006$), $\text{CDE-OI}_{mort,t} = \text{CDE}_{mort,t} * kt^-$, and for $t = 2004$, $\text{CDE-OI}_{mort,t} = \text{CDE}_{mort,t} * kt$. The input values and constraints for the mixed estimator models were reweighted analogously.

For each iterate, for each year, I calculated the empirical bias (EB) and the empirical mean squared error (EMSE) over the 1,000 iterations between each estimator and the true population values under each of the four error structures.

Population 4 - Sample Error Structure 3

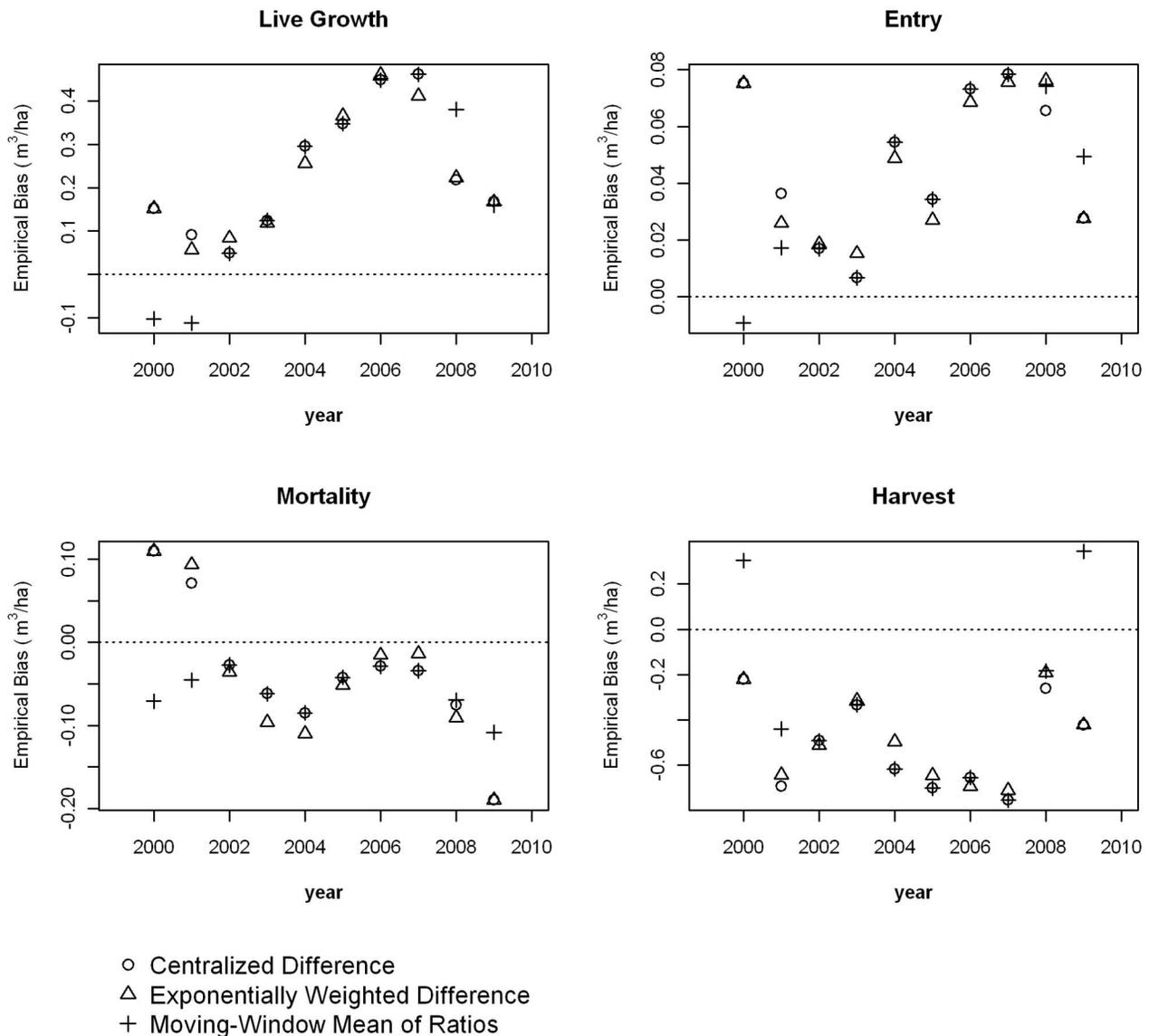


Figure 11. The EB over 1,000 iterations of 1,000 samples each from population 4 under sampling error structure 3 for the CDE, EWD, and MWMOR by growth component and estimation year.

That is

$$EB_{PES} = \frac{1}{1000} \sum_{i=1}^{1000} (\hat{x}_{PESi} - X_P), \quad (19)$$

where \hat{x}_{PESi} is the sample estimate of X in population P for estimator E under error structure S for iterate i . Likewise,

$$EMSE_{PES} = \frac{1}{1000} \sum_{i=1}^{1000} (\hat{x}_{PESi} - X_P)^2. \quad (20)$$

Usually, the overriding criterion for selection of an estimator in forest inventories is the minimum mean squared error for the can-

didate unbiased estimators. This might suggest that the emphasis in the presentation of these results should focus on $EMSE_{PES}$. Because the four error structures were devised to examine estimator performance over these populations in presence of plausible differentially biased sampling error, I will pay special attention to EB_{PES} in the presentation of results.

Results

In the interest of parsimony, I am providing a subset of the results that can be used to demonstrate the salient points. The results for all populations under all sampling error structures are available from the author on request. Figures 1–3 give the empirical mean, bias, and mean squared error, respectively, over 1,000 iterations of 1,000

Population 4 - Sample Error Structure 3

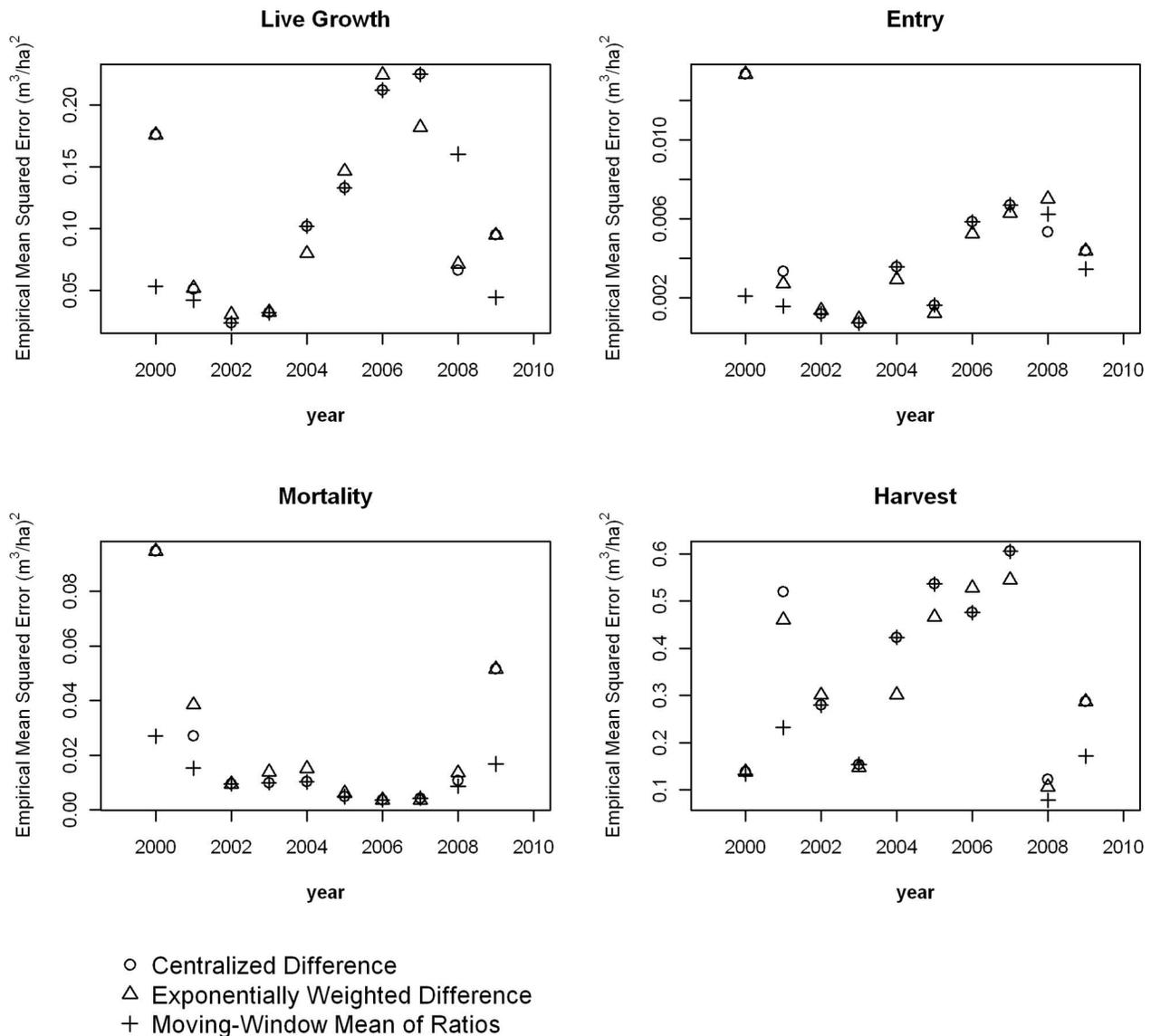


Figure 12. The EMSE over 1,000 iterations of 1,000 samples each from population 4 under sampling error structure 3 for the CDE, EWD, and MWMOR by growth component and estimation year.

samples each from population 1 under sampling error structure 1 for the CDE, the EWD, and the MWMOR, by growth component for estimation years 2000–2009. These are the years that are the center of at least one observation window (or panel) for a sample drawn under this design from a population spanning 1998–2001. By definition, the CDE and EWD use the observations from a single panel’s remeasurement for years 2000 and 2009 while the noncentralized MWMOR uses the observations from three panels. The CDE and EWD use the observations from the remeasurement of three panels for years 2001 and 2008 while the semicentralized MWMOR uses the observations from four panels for these years. Five panels are used for all of the estimators for estimates of years 2002–2007. All three estimators, in conjunction with the design are

shown to be general smoothers. That is the design observes 5-year windows, which provides an average annual change or “smooths” the actual annual change, and the estimators, while drawing strength from overlapping panels, provide further smoothing. In Figure 1, this effect can be seen quite clearly in all four components under this simplest sampling error structure. As expected, owing to the nonlinear latent trends, in the central years of the estimation span, the EWD is sometimes slightly closer to the population value than the other estimators are. In the temporal extremes, the results for the MWMOR are mixed, with the expected bias due to the use of the off-center panels sometimes, but not always, being overridden by the lower variance from the larger sample. This can be verified by comparing the harvest and mortality results in Figures 2 and 3.

Population 3 - Mortality

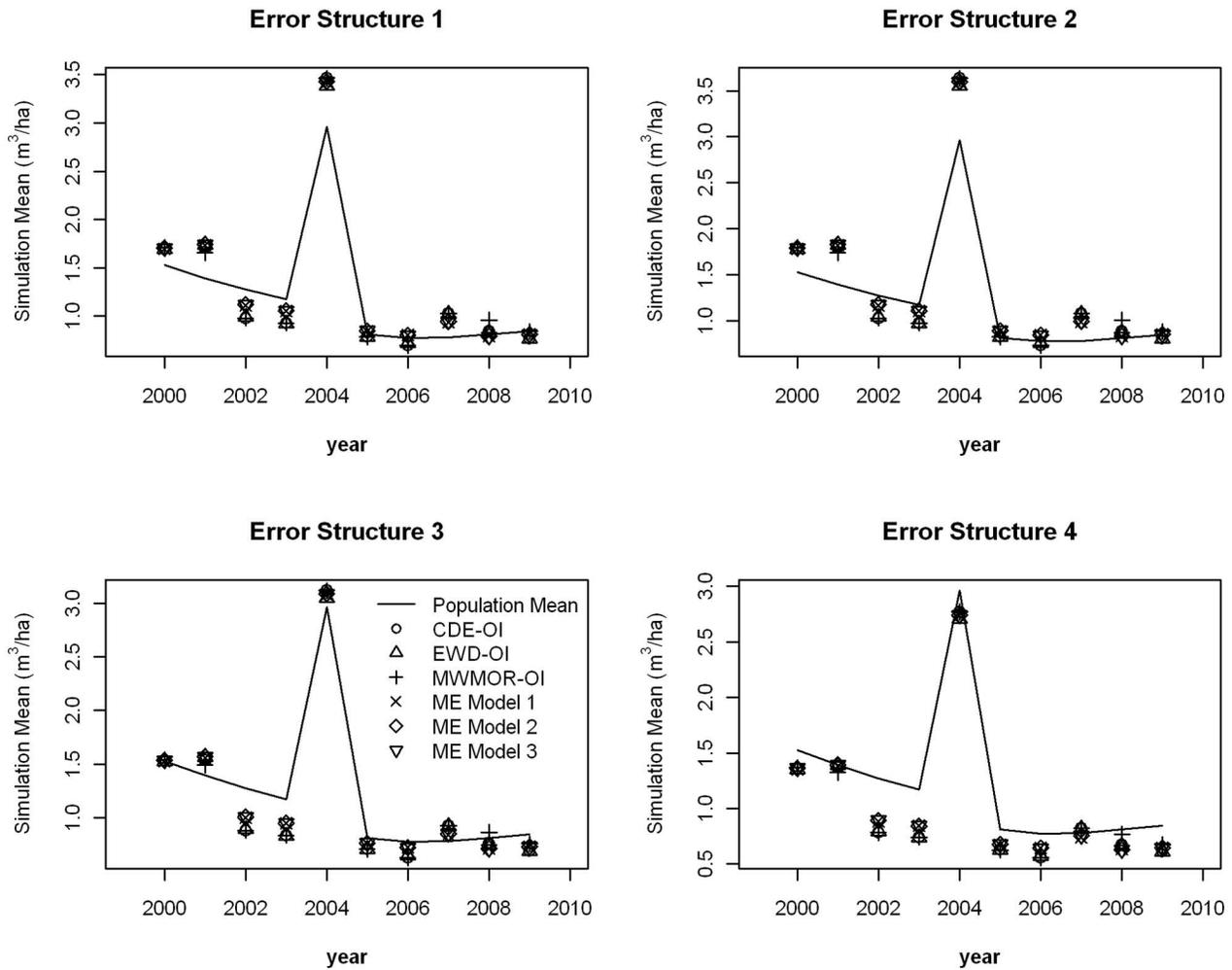


Figure 13. The mean over 1,000 iterations of 1,000 samples each from population 3 under sampling error structures 1–4, for the estimators incorporating outside information for the mortality component by estimation year.

Figures 4–6 give the empirical mean, bias, and mean squared error, respectively, over 1,000 iterations of 1,000 samples each from population 1 under sampling error structure 4 for the CDE, the EWD, and the MWMOR by growth component and estimation year. In general, the three figures reflect the bias and greater variance of error structure 4 over error structure 1 for this population. The estimators do not appear to be differentially affected by the bias in error structure 4. I do note that occasionally the bias in the temporal extremities shown in Figure 2 for the MWMOR is offset in Figure 5 by the counteracting bias in the sampling error structure, resulting in lower EMSE in Figure 6 relative to Figure 3. This reinforces a point made in Eastaugh and Hasenauer (2013) that being that often the implementation of theoretically unbiased sample designs will result in biased samples for a myriad of reasons, and estimators should, therefore, be evaluated in consideration of that possibility. With respect to robustness, note that if we were examining only the outcome for one or more of these estimators and did not know the true population values, that is, the population line was missing from

Figures 1 and 4, we would draw the same conclusions about the trend in each of the components. This suggests that all of these estimators are, at least in this regard, robust.

Figures 7–9 give the empirical mean, bias, and mean squared error, respectively, for 1,000 iterations of 1,000 samples each from population 3 under sampling error structure 4, for the CDE, the EWD, and the MWMOR by growth component and estimation year. Although I show the results for error structure 4, the most notable outcome for population 3 under all error structures is that all three estimators give no indication of the extreme anomaly for mortality in 2004 due to the smoothing effects of the both the sample design and the estimators applied to the outcomes of the design. The anomaly in 2004 is spread out over the estimates for the surrounding years, so the estimators reflect very high EB and EMSE for mortality in 2004 in Figures 8 and 9, respectively. The sample design itself makes this single-year anomaly particularly difficult to evaluate and impossible to definitively separate it from a possible spatial effect.

Population 3 - Mortality

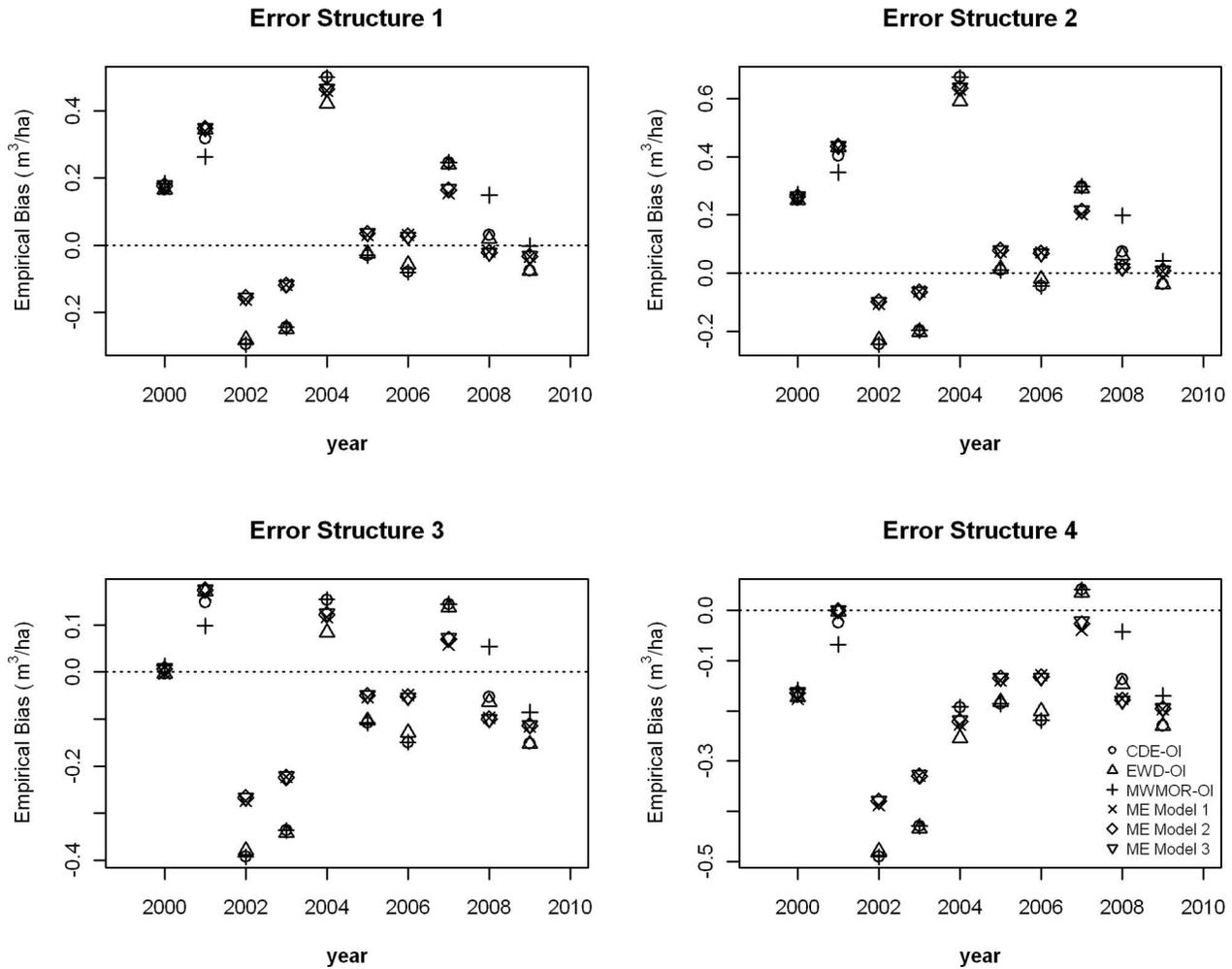


Figure 14. The EB over 1,000 iterations of 1,000 samples each from population 3 under sampling error structures 1–4 for the estimators incorporating outside information for the mortality component by estimation year.

The empirical mean, bias, and mean squared error, respectively, for 1,000 iterations of 1,000 samples each from population 4 under sampling error structure 3, for the CDE, the EWD, and the MWMOR by growth component and estimation year, are given in Figures 10–12. Although population 4 has a greater diversity of conditions than population 1, the differences in estimator performance between Figures 10–12, relative to the corresponding graphs in Figures 1–3, appear to correspond to those that could be expected by the more severe sampling error structure 3.

Figures 13–15 give the empirical mean, bias, and mean squared error, respectively, for 1,000 iterations of 1,000 samples each from population 3 under sampling error structures 1–4, for the estimators incorporating outside information for the mortality component by estimation year. Although it is true that the outside information was not perfect, all of the estimators incorporating the outside information benefited in the form of improved mortality estimates for the year 2004 and the surrounding years that used observations spanning 2004. Figure 13 shows that under each of the sampling

error structures each of the estimators exhibit a pattern very similar to the mortality trend for the population. The patterns differ from the population trend predictably by sampling error structure. Figure 14 shows more clearly than Figure 13 that the order of the estimates for each year remains constant through the different sampling error structures, indicating that sampling error structure did not differentially affect the estimators. Some interaction between the outside information and sampling error structure is indicated by the position of the group of estimators for 2004 relative to the groups of estimators for the surrounding years. Figure 15 shows that the EMSE results for the mixed estimator models are often, but not always, higher than they are for the other models. This effect could be viewed as the cost of compatibility.

Discussion and Conclusions

I explored some special problems that arise in estimation of the components of change when the temporal scale of the population estimand of interest is finer than the scale of observation under both

Population 3 - Mortality

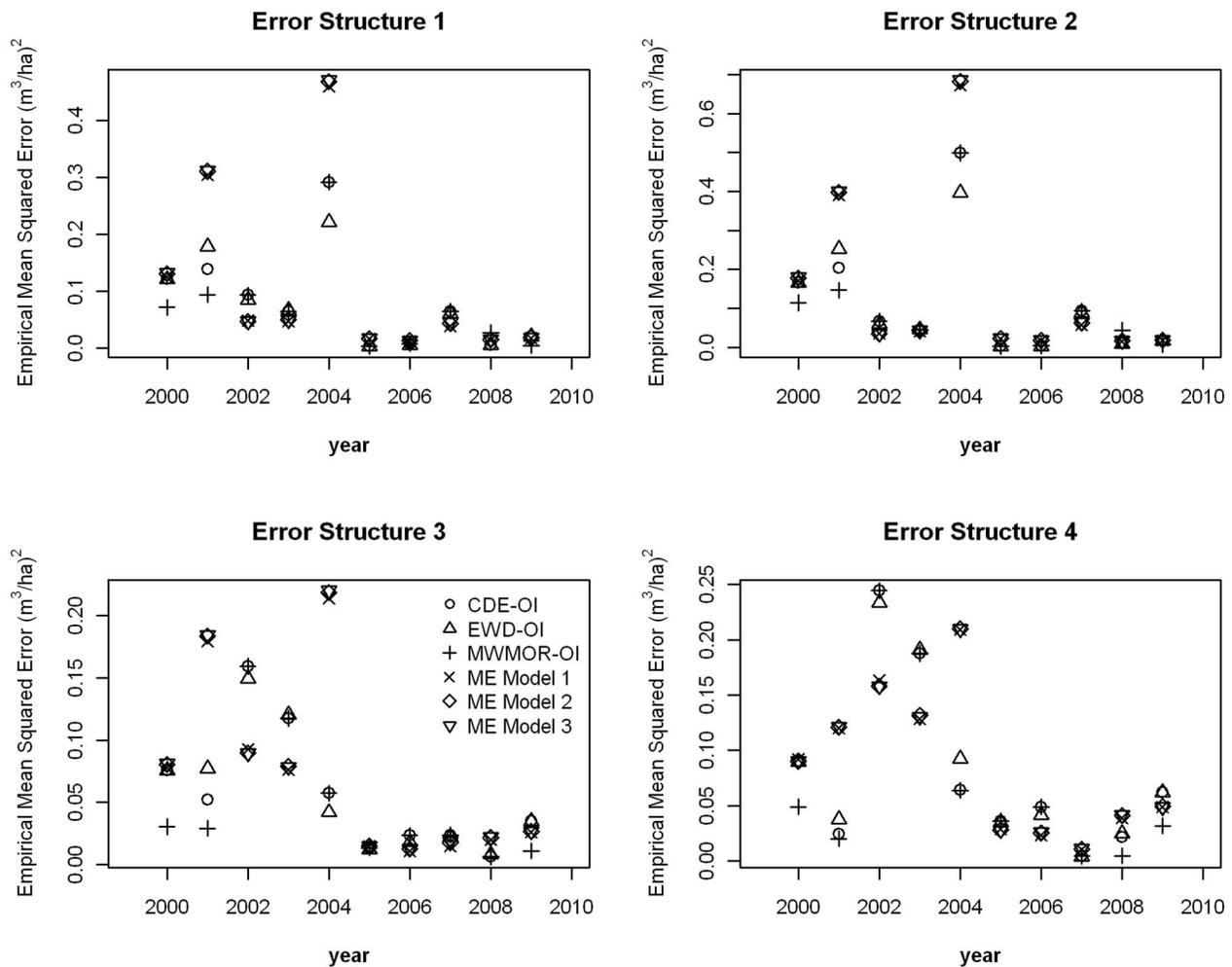


Figure 15. The EMSE over 1,000 iterations of 1,000 samples each from population 3 under sampling error structures 1–4, for the estimators incorporating outside information, for the mortality component by estimation year.

biased and unbiased sampling error structures. In the example simulations, the temporal scale of observation was 5 years while the temporal dimension of the population of interest was 1 year. All of the approaches worked well in the temporal midrange of observations in the presence of smooth population trends under unbiased sampling error structures. By interjecting a single year anomaly into population 3, I presented an especially difficult (but realistic) situation given the sampling frame. The results of these more thorough simulations support the simulations and conclusions of Roesch (2007a) with respect to comparisons between the EWD and this application of the mixed estimator. The simulations also showed the variance/bias tradeoff encountered when the MWMOR was used in the extremity years of observation. Although the MWMOR is sometimes biased in the presence of trend in the extremity years, the EMSE was often lower than it was for the other estimators. Of the four general approaches for estimating the components of forest change from this annually rotating five-panel sample design: the centralized moving average estimator, the EWD, the MWMOR,

and the mixed estimator, the first three approaches are very compatible with large monitoring efforts that require intervention-free results. These three simple approaches were shown to be amenable when outside information suggests an adaptation to the weighting scheme. The fourth approach, the mixed estimator, is also amenable to the incorporation of extraneously derived information and can easily incorporate complex models. No single estimation approach has (or could have) been shown to be a panacea and some notable differences in performance were observed at the temporal extremities of observation in the presence of temporal anomalies and in the presence of biased sampling error structures. With respect to trend, all of the estimators are robust, but the CWD, EWD, and MWMOR are somewhat nonresponsive to highly variable trends and all estimators are subject to the smoothing effect of the sample design considered here. This particular deficiency in the design was shown here to be readily corrected through the incorporation of outside information.

This study suggests that one should probably not attempt to

choose a single estimation approach to address the widest range of estimation objectives. The estimators considered here were shown to differ in their respective robustness to different but realistic, underlying error structures, aspects of which may not always be known. Rather, an investigator would be well served to embrace the philosophy behind the Thomas and Roesch (1990) and Eastaugh and Hasenauer (2013) articles that, when real data are involved, there is value in making estimates using as many theoretical approaches as possible. When different approaches produce similar results, there is strong evidence for those results. However, it can be even more informative when different (but defensible) approaches produce varied results. When this happens, it is incumbent on the analyst to figure out why the varied results have occurred. Quite simply, an analyst should look at the data from as many angles as she or he has the time and energy for in an attempt to understand fully the natural phenomena that are being imperfectly observed.

Literature Cited

- BECHTOLD, W.A., AND P.L. PATTERSON (EDS.). 2005. *The enhanced Forest Inventory and Analysis Program—National sampling design and estimation procedures*. USDA For. Serv., Gen. Tech. Rep. SRS-GTR-80, Asheville, NC. 85 p.
- CASSEL, C.M., C.E. SÄRNDAL, AND J.H. WRETMAN. 1977. *Foundations of inference in survey sampling*. John Wiley & Sons, New York. 192 p.
- COCHRAN, W.G. 1977. *Sampling techniques*, 3rd ed. John Wiley & Sons, New York. xvi + 428 p.
- CHANDRA, M.J. 2000. *Statistical quality control*. CRC Press, New York. 284 p.
- EASTAUGH, C.S., AND H. HASENAUER. 2013. Biases in volume increment estimates derived from successive angle count sampling. *For. Sci.* 59(1):1–14.
- ERIKSSON, M. 1995. Compatible and time-additive change component estimators for horizontal-point-sampled data. *For. Sci.* 41(4):796–822.
- GERTNER, G. 1987. Approximating precision in simulation projections: An efficient alternative to Monte Carlo methods. *For. Sci.* 33(1):230–239.
- KORHONEN, K.T. 1993. Mixed estimation in calibration of volume functions of Scots pine. *Silva Fen.* 27(4):269–276.
- MEYER, H.A. 1953. *Forest mensuration*. Penn Valley Publishers, Inc., State College, PA. 357 p.
- RAJ, D. 1968. *Sampling theory*. McGraw-Hill, New York. 302 p.
- ROESCH, F.A. 2007a. Compatible estimators of the components of change for a rotating panel forest inventory design. *For. Sci.* 53(1):50–61.
- ROESCH, F.A. 2007b. The components of change for an annual forest inventory design. *For.Sci.* 53(3) 406–413.
- ROESCH, F.A. 2008. An alternative view of continuous forest inventories. *For. Sci.* 54(4):455–464.
- ROESCH, F.A., AND P.C. VAN DEUSEN. 2012. Monitoring forest/non-forest land use conversion rates with annual inventory data. *Forestry: An Int. Jour. of For. Res.* 85(3):391–398.
- ROESCH, F.A., AND P.C. VAN DEUSEN. 2013. Time as a dimension of the sample design in national-scale forest inventories. *For. Sci.* Available online at <http://dx.doi.org/10.5849/forsci.12-075>; last accessed Nov. 25, 2012.
- THEIL, H. 1963. On the use of incomplete prior information in regression analysis. *J. Am. Stat. Assoc.* 58:401–414.
- THOMAS, C.E., AND F.A. ROESCH. 1990. Basal area growth estimators for survivor component: A quality control approach. *South. J. For.* 14(1):12–18.
- VAN DEUSEN, P.C. 1996. Incorporating predictions into an annual forest inventory. *Can. J. For. Res.* 26:1709–1713.
- VAN DEUSEN, P.C. 1999. Modeling trends with annual survey data. *Can. J. For. Res.* 29(12):1824–1828.
- VAN DEUSEN, P.C. 2000. Alternative sampling designs and estimators for annual surveys. P. 192–196 in *Integrated tools for natural resources inventories in the 21st century*, Hansen, M., and T. Burk (eds.). USDA For. Serv., NC-GTR-212, St. Paul, MN. 744 p.
- VAN DEUSEN, P.C., AND F.A. ROESCH. 2009. Estimating forest conversion rates with annual forest inventory data. *Can. J. For. Res.* 39(10): 1993–1996.
- WALTON, G.S., AND C.J. DEMARS. 1973. *Empirical methods in the evaluation of estimators*. USDA For. Serv., Res Pap. NE-272, Upper Darby, PA. 15 p.
- WEST, M., AND J. HARRISON. 1989. *Bayesian forecasting and dynamic models*. Springer-Verlag, New York. 704 p.