United States
Department of
Agriculture

Forest Service

Pacific Northwest
Research Station

Research Note
PNW-RN-532
June 2002

A Comparison of Two Estimates of Standard Error for a Ratio-of-Means Estimator for a Mapped-Plot Sample Design in Southeast Alaska

Willem W.S. van Hees

Abstract

Comparisons of estimated standard error for a ratio-of-means (ROM) estimator are presented for forest resource inventories conducted in southeast Alaska between 1995 and 2000. Estimated standard errors for the ROM were generated by using a traditional variance estimator and also approximated by bootstrap methods. Estimates of standard error generated by both traditional and bootstrap methods were similar. Percentage differences between the traditional and bootstrap estimates of standard error for productive forest acres and for gross cubic-foot growth were generally greater than respective differences for nonproductive forest acres, net cubic volume, or nonforest acres.

Keywords: Sampling, inventory (forest), error estimation.

Introduction

Between 1995 and 2000, the Pacific Northwest Research Station (PNW) Forest Inventory and Analysis (FIA) Program in Anchorage, Alaska, conducted a land cover resource inventory of southeast Alaska (fig. 1). Land cover attribute estimates derived from this sample describe such features as area, timber volume, and understory vegetation. Each of these estimates is subject to measurement and sampling error. Measurement error is minimized through training and a program of quality control. Sampling error must be estimated mathematically.

This paper presents a comparison of bootstrap estimation of standard error for a ratio-of-means estimator with a traditional estimation method. Bootstrap estimation techniques can be used when a complex sampling strategy hinders development of unbiased or reliable variance estimators (Schreuder and others 1993.) Bootstrap techniques are resampling methods that treat the sample as if it were the whole population. Repeated samples are taken, with replacement, from the original sample and the statistic of interest is calculated for each sample. The average of the statistic over all the samples is the bootstrap estimator.

This study was conducted to examine briefly whether or not traditional estimates of standard error differed in some meaningful way from the bootstrap approximation. It was felt that the two estimates would produce substantially similar results.

Willem W.S. van Hees is a research forester, Forestry Sciences Laboratory, 3301 C. St., Suite 200, Anchorage, AK 99503.

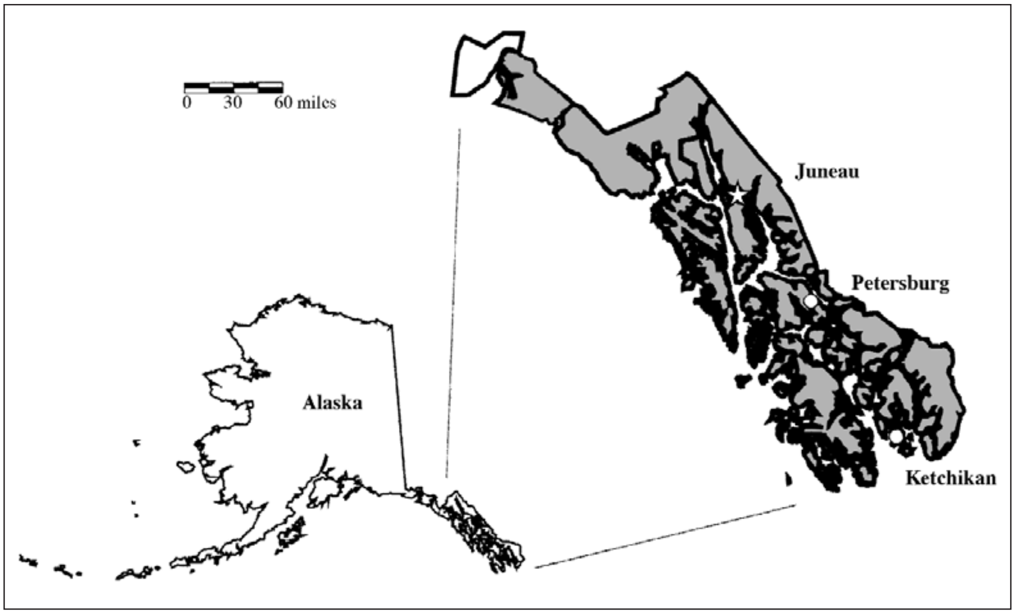

Figure 1—Southeast Alaska inventory region, 1995-2000.

In lieu of extensive simulations, multiple bootstrap estimates of standard error, along with traditional estimates, were made for each of five population parameters in three sampling units. These parameters, for unreserved USDA Forest Service lands, were productive forest-land acres, nonproductive forest-land acres, nonforest acres, net cubic-foot volume, and gross cubic-foot growth in the Chatham, Stikine, and Ketchikan inventory units (van Hees 2001a, 2001b, 2001c).

Methods

Sample Design

Land cover of the southeastern panhandle of Alaska (fig. 1) was sampled with a single-phase, unstratified, systematic grid sample (grid spacing was 4.8 kilometers). Ground plots (1 hectare) were established at each grid intersection and were subsampled by a cluster of four, 7.3-meter, fixed-radius subplots. Three subplots were sited, one each at 63.4 meters north, southeast, and southwest from the first, centrally located subplot. Each subplot was mapped for land cover (Scott and Bechtold 1995.)

Ratio-of-Means Estimator

Estimation of resource attributes such as area and timber volume used ratio-of-means estimators. The ratio-of-means estimator and associated traditional variance estimate are defined as (Cochran 1977):

$$\hat{R} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i}, \qquad (1)$$

where

y_i = the variable of interest on plot i,
x_i = an auxiliary variable (such as area) on plot i that is correlated with y_i, and
n = number of plots selected from the population,

with variance

$$V(\hat{R}) = \frac{1}{n(n-1)\left(\frac{\sum x_i}{n}\right)^2}\left(\sum y_i^2 + \hat{R}^2 \sum x_i^2 - 2\hat{R}\sum y_i x_i\right). \qquad (2)$$

Zarnoch and Bechtold (2000) demonstrate use of this ratio estimator to generate various population estimates from data collected by using a sampling design similar to that described above. Essentially, the ratio estimate defined by equation (1) is multiplied by some total, such as area, to estimate the amount of the total in the condition of interest (y_i).

Table 1—Ratio-of-means (ROM) population estimates, traditional estimate of standard error, and bootstrap estimates of standard error by number of iterations, sampling unit, and population attribute

			Estimated standard error									
				Bootstrap								
Sampling unit/ population attribute	ROM population estimate	No. of plots	Traditional[a] estimate	500 iterations						1,000 iterations			3,000 iterations
Ketchikan:		550											
Productive forest, acres	1,405,362		58,036	61,359	57,410	56,880	59,002			58,449			
Nonproductive forest, acres	851,262		53,754	56,895	55,148	53,482	54,649			54,689			
Nonforest, acres	449,053		41,962	43,615	43,416	41,269	45,148			41,961			
Net cubic-foot volume	7,293,763,469		483,113,144	504,230,788	462,260,072	479,479,823	463,481,657			468,461,285	470,675,742		463,202,816
Gross cubic-foot growth	63,725,479		3,250,717	3,516,024	3,378,918	3,461,131	3,328,462			3,414,326	3,446,209		3,400,647
Stikine:		520											
Productive forest, acres	1,085,711		60,713	62,704	58,648	63,772	57,690			62,517			
Nonproductive forest, acres	708,385		53,341	53,910	51,318	54,061	53,152			55,388			
Nonforest, acres	1,077,131		60,469	62,760	59,758	62,718	58,038			61,135			
Net cubic-foot volume	5,437,872,712		456,690,704	462,024,168	457,544,019	450,882,246	429,174,042			453,095,411			453,530,368
Gross cubic-foot growth	40,659,537		2,652,600	2,660,205	2,779,894	2,580,381	2,568,629			2,730,045			2,697,633
Chatham:		979											
Productive forest, acres	1,302,152		72,813	74,370	69,716	74,637	70,059	71,547		71,726			
Nonproductive forest, acres	709,450		56,555	56,060	60,671	56,542	56,008	58,272		59,047			
Nonforest, acres	3,368,739		82,340	79,570	83,406	84,722	79,248	79,016		82,717			
Net cubic-foot volume	7,561,291,112		540,235,099	534,119,374	519,603,815	541,402,469	558,017,539	515,248,933	535,130,404	536,971,963	551,209,556	568,524,342	
Gross cubic-foot growth	59,953,509		3,803,595	3,856,939	3,654,326	3,786,265	4,034,644	3,728,864	3,856,871	3,863,259	3,929,279	4,036,473	

[a] Traditional estimate of standard error is from Cochran (1977).

Table 2—Differences between bootstrap and traditional estimates of standard errors as a percentage of the bootstrap estimate, by number of bootstrap iterations, sampling unit, and population attribute

Sampling unit/ population attribute	Number of bootstrap iterations									
	500						1,000			3,000
	Percent									
Ketchikan:										
Productive forest, acres	5.42	-1.09	-2.03	1.64			0.71			
Nonproductive forest, acres	5.52	2.53	-0.51	1.64			1.71			
Nonforest, acres	3.79	3.35	-1.68	7.06			0.00			
Net cubic-foot volume	4.19	-4.51	-0.76	-4.24			-3.13	-2.64		-4.30
Gross cubic-foot growth	7.55	3.79	6.08	2.34			4.79	5.67		4.41
Stikine:										
Productive forest, acres	3.18	-3.52	4.80	-5.24			2.89			
Nonproductive forest, acres	1.06	-3.94	1.33	-0.35			3.70			
Nonforest, acres	3.65	-1.19	3.59	-4.19			1.09			
Net cubic-foot volume	1.15	0.19	-1.29	-6.41			-0.79			-0.70
Gross cubic-foot growth	0.29	4.58	-2.80	-3.27			2.84			1.67
Chatham:										
Productive forest, acres	2.09	-4.44	2.44	-3.93	-1.77		-1.52			
Nonproductive forest, acres	-0.88	6.78	-0.02	-0.98	2.95		4.22			
Nonforest, acres	-3.48	1.28	2.81	-3.90	-4.21		0.46			
Net cubic-foot volume	-1.15	-3.97	0.22	3.19	-4.85	-0.95	-0.61	1.99	4.98	
Gross cubic-foot growth	1.38	-4.08	-0.46	5.73	-2.00	1.38	1.54	3.20	5.77	

Bootstrap Estimate

Bootstrap estimation of the variance was accomplished by iterative random sampling, with replacement, of the plot list to generate a new plot list. Within each plot, subplots were randomly selected, with replacement, to generate a new subplot list for each plot. The ROM estimates of population totals were then recomputed. The variance of the estimated totals provides an estimate of the sample variance. This process was repeated 500 to 3,000 times. At least four bootstrap estimates, for each of the five population parameters, in all three sampling units, were made by using 500 iterations. Greater numbers of iterations (1,000 and 3,000) were used fewer times and for increasingly fewer population parameters.

Although bootstrap samples are not independent, variance estimates resulting from them are considered conservative approximations (Schreuder and others 1993).

Results

Bootstrap and traditional estimates of standard errors are presented in table 1. In no case did the bootstrap estimate differ from the traditional estimate by more than 7.55 percent (table 2). Eleven of 92 observations were more then 5 percent different; the remaining 81 bootstrap estimates were less than 4.99 percent different from the ROM estimate. The larger percentage differences (greater than $\pm$ 5.5 percent) are not specific to a particular inventory area or estimated population parameter (figs. 2 through 6).

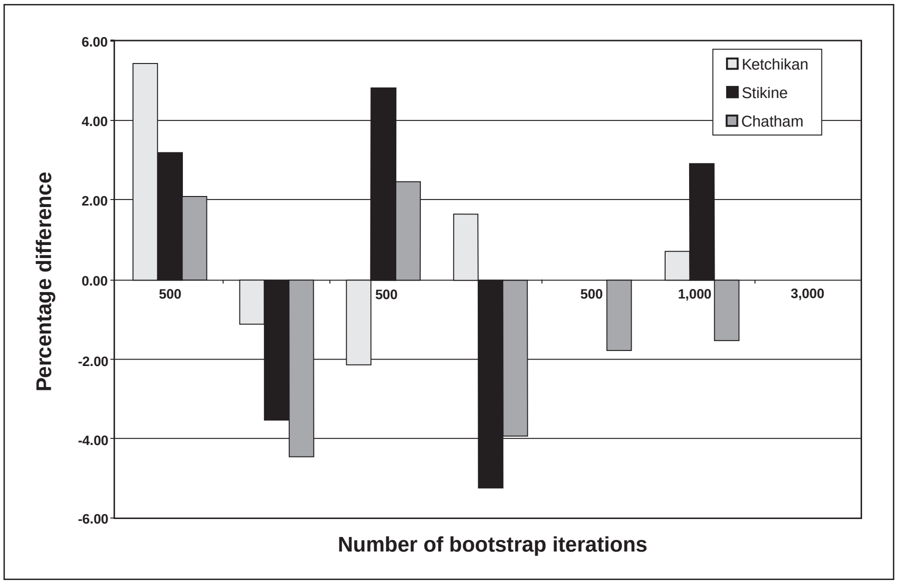

Figure 2—Differences between bootstrap and traditional estimates of standard errors as a percentage of the bootstrap estimate, by number of bootstrap iterations, for estimates of productive forest-land area within unreserved national forest lands, southeast Alaska.

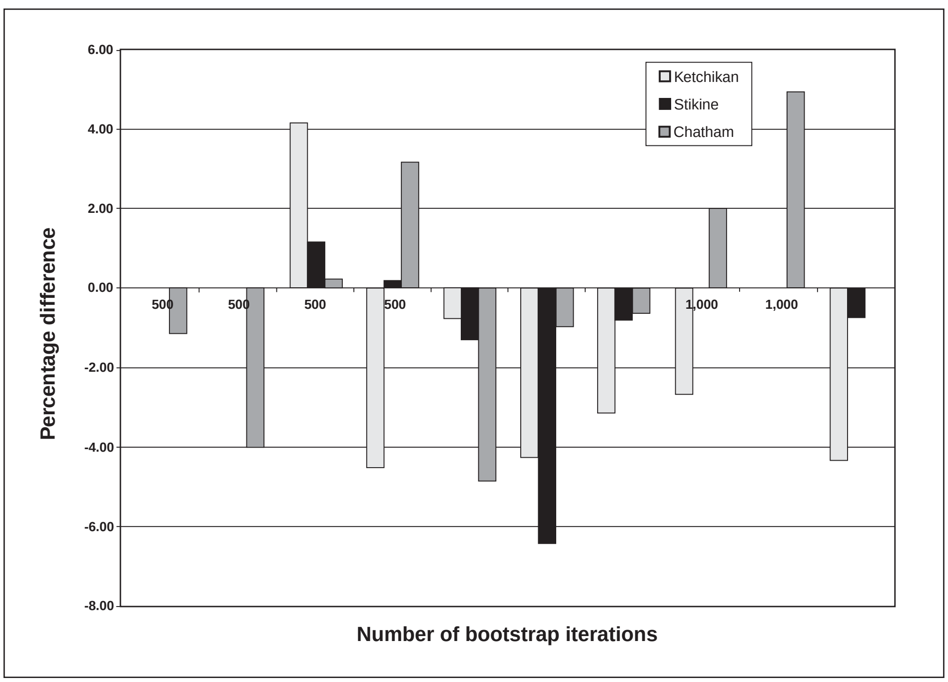

Figure 3—Differences between bootstrap and traditional estimates of standard errors as a percentage of the bootstrap estimate, by number of bootstrap iterations, for estimates of net cubic-foot volume on productive forest-land area within unreserved national forest lands, southeast Alaska.

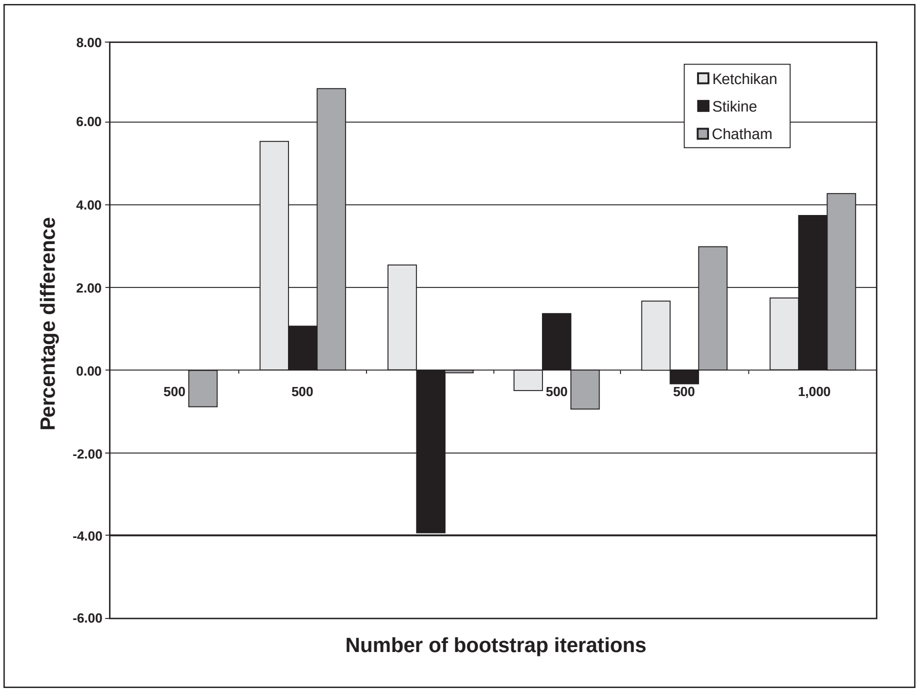

Figure 4—Differences between bootstrap and traditional estimates of standard errors as a percentage of the bootstrap estimate, by number of bootstrap iterations, for estimates of non-productive forest-land area within unreserved national forest lands, southeast Alaska.

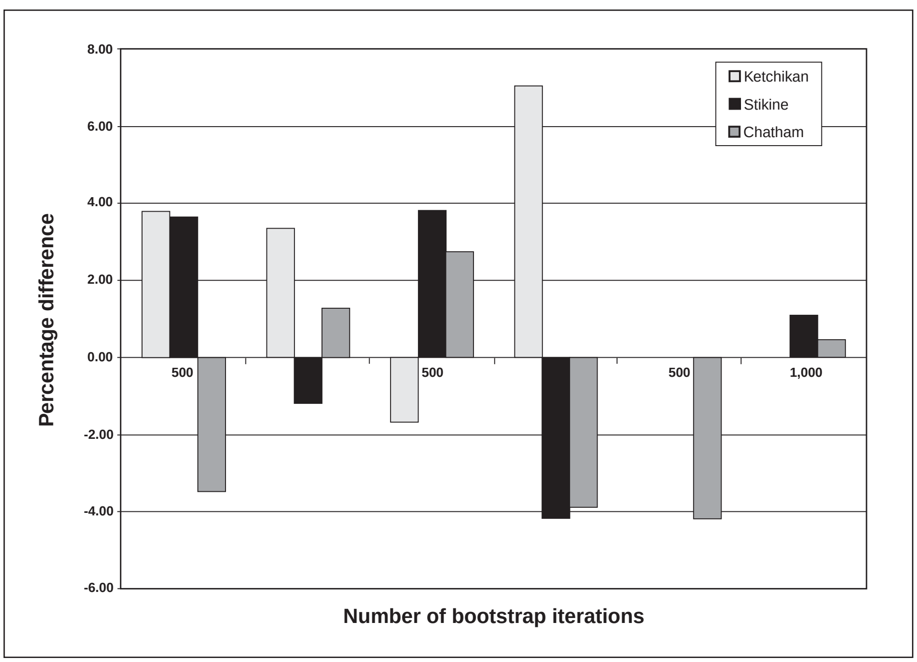

Figure 5—Differences between bootstrap and traditional estimates of standard errors as a percentage of the bootstrap estimate, by number of bootstrap iterations, for estimates of non-forest-land area within unreserved national forest lands, southeast Alaska.

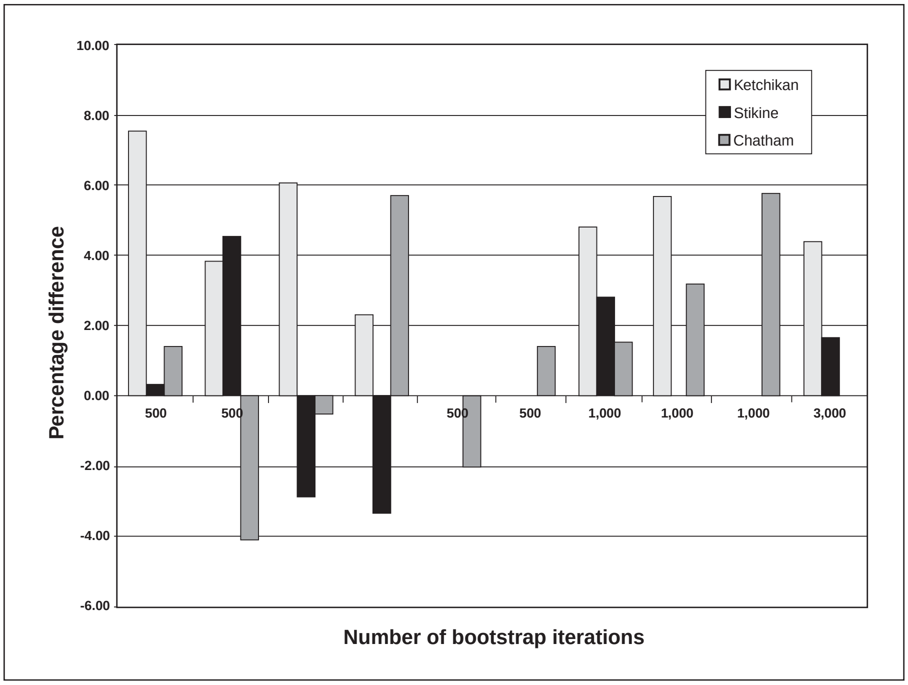

Figure 6—Differences between bootstrap and traditional estimates of standard errors as a percentage of the bootstrap estimate, by number of bootstrap iterations, for estimates of gross cubic-foot growth on productive forest-land area within unreserved national forest lands, southeast Alaska.

Discussion

For the response variables listed, there is little difference between the two estimates. The largest percentage differences between error estimates are in the Ketchikan unit. The forest resource of the Ketchikan unit is more variable than in the other two units; it is ecologically transitional (approaching the southern extent of the Sitka spruce/western hemlock zone), and logging has been more intensive and extensive resulting in greater variety of age and size classes. Greater variability between bootstrap and traditional estimates may indicate the number of iterations was low for those estimates.

Relative rapidity of data processing is a significant client service consideration. Rapid responses are important when providing clients with inventory results. Although bootstrap estimates can be generated for any of the population estimates FIA produces, the computational resources needed to generate those estimates for each cell in all the output tables typical for FIA reporting exceed those needed for ROM estimates. Use of the ROM estimator reduces production time dramatically.

Acknowledgments

Kevin Dobelbower, Team Leader for Information Management, Anchorage, Alaska, created and ran necessary computer programs for this study.

English Equivalents

0.3048 meter = 1 foot
0.0283 cubic meter = 1 cubic foot
0.4047 hectare = 1 acre
1.609 kilometers = 1 mile

Literature Cited

Cochran, W.G. 1977. Sampling techniques. 3rd ed. New York: John Wiley and Sons, Inc. 428 p.

Schreuder, H.T.; Gregoire, T.G.; Wood, G.B. 1993. Sampling methods for multiresource forest inventory. New York: John Wiley and Sons, Inc. 446 p.

Scott, C.T.; Bechtold, W.A. 1995. Techniques and computations for mapping plot clusters that straddle stand boundaries. Forest Science Monograph. 31: 46-61.

van Hees, W.W.S. 2001a. Summary estimates of forest resources on unreserved lands of the Chatham inventory unit, Tongass National Forest, southeast Alaska, 1998. Resour. Bull. PNW-RB-234. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 12 p.

van Hees, W.W.S. 2001b. Summary estimates of forest resources on unreserved lands of the Ketchikan inventory unit, Tongass National Forest, southeast Alaska, 1998. Resour. Bull. PNW-RB-233. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 8 p.

van Hees, W.W.S. 2001c. Summary estimates of forest resources on unreserved lands of the Stikine inventory unit, Tongass National Forest, southeast Alaska, 1998. Resour. Bull. PNW-RB-232. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 8 p

Zarnoch, S.J.; Bechtold, W.A. 2000. Estimating mapped-plot attributes with ratios of means. Canadian Journal of Forest Research. 30: 688-697.

This page intentionally left blank.

Pacific Northwest Research Station

Web site	http://www.fs.fed.us/pnw
Telephone	(503) 808-2592
Publication requests	(503) 808-2138
FAX	(503) 808-2130
E-mail	pnw_pnwpubs@fs.fed.us
Mailing address	Publications Distribution Pacific Northwest Research Station P.O. Box 3890 Portland, OR 97208-3890

U.S. Department of Agriculture
Pacific Northwest Research Station
333 S.W. First Avenue
P.O. Box 3890
Portland, OR 97208-3890