United States Department of Agriculture

Forest Service

Northern Research Station

General Technical Report NRS-53

FIA National Assessment of Data Quality for Forest Health Indicators

James A. Westfall, Editor

Abstract

The Forest Inventory and Analysis (FIA) program of the U.S. Forest Service collects vast amounts of field data to assess the condition and trends of the nation's forest resources. A quality assurance (QA) program is implemented to assure that data are collected accurately with consistent protocols. A random subset of field plots is chosen to receive an additional, independent measurement by another field crew. This 'blind check' approach allows for comparison of measurements between the two crews. The measurement differences are evaluated against measurement quality objectives (MQO), which specify a level of measurement precision for each attribute. In this report, differences in repeated measurements from blind check data are analyzed to assess the quality of forest health data nationwide.

For the understory vegetation indicator, the ground cover variables attained the MQO or were slightly below the desired level. Variables related to canopy and species identification were substantially below the MQO.

Tree crown attributes that effectively met the MQO include crown light exposure, crown dieback, and foliage transparency. The uncompacted crown ratio, vigor class, and crown density measurements did not attain the MQO standards.

Quality assessment for the lichen indicator is evaluated using a second measurement by an expert and not an independent measurement by another field crew. Overall, the MQO was not attained, although compliance improved over time and western regions were close to achieving the standard.

Most variables measured for the down woody material indicator were below MQO standards.

Due to the sampling protocols for the ozone indicator, direct analyses of MQO attainment are not possible. Mean biosite index was not significantly different between crews, although there were some large discrepancies on individual plots. There was agreement on presence/absence of ozone damage for about 80 percent of the plots.

For field measurements of forest soils, 10 of the 12 variables attained the MQO. Although MQOs are not specifically stated for laboratory analyses of soil properties, comparisons were performed to assess the variability of lab measurements.

This information should be useful to data collection experts, as variables having poor measurement repeatability can be identified and examined for potential resolution. The results may also be of interest to analysts and researchers wanting to evaluate whether the repeatability of measurements is sufficient for their respective studies.

Cover photos, clockwise from top left:

Lichen identification; photo by Scott Bates, University of Colorado, used with permission.

Understory vegetation quadrat; photo by U.S. Forest Service.

Soil core; photo by U.S. Forest Service.

Coarse woody debris diameter measurement; photo by U.S. Forest Service.

Crown density/foliage transparency; photo by U.S. Forest Service.

Manuscript received for publication February 2009

Published by:

USDA FOREST SERVICE
11 CAMPUS BLVD., SUITE 200
NEWTOWN SQUARE, PA 19073

October 2009

For additional copies:

USDA Forest Service
Publications Distribution
359 Main Road
Delaware, OH 43015-8640
Fax: 740-368-0152

Visit our homepage at: **http://www.nrs.fs.fed.us/**

FIA National Assessment of Data Quality for Forest Health Indicators

James A. Westfall, Editor

Contributing Authors

Michael C. Amacher
William A. Bechtold
David Gartner
Mark H. Hansen
Olaf Kuegler
Paul L. Patterson
Charles H. Perry
KaDonna C. Randolph
Bethany Schulz
Marie T. Trest
James A. Westfall
Susan Will-Wolf
Christopher W. Woodall

Author Affiliations

Michael C. Amacher - Research soil scientist, U.S. Forest Service, Rocky Mountain Research Station, Logan, UT.

William A. Bechtold - Research forester, U.S. Forest Service, Southern Research Station, Asheville, NC.

David Gartner - Mathematical statistician, U.S. Forest Service, Southern Research Station, Knoxville, TN.

Mark H. Hansen - Research forester, U.S. Forest Service, Northern Research Station, St. Paul, MN.

Olaf Kuegler - Mathematical statistician, U.S. Forest Service, Pacific Northwest Research Station, Portland, OR.

Paul L. Patterson - Mathematical statistician, U.S. Forest Service, Rocky Mountain Research Station, Fort Collins, CO.

Charles H. Perry - Research soil scientist, U.S. Forest Service, Northern Research Station, St. Paul, MN.

KaDonna C. Randolph - Mathematical statistician, U.S. Forest Service, Southern Research Station, Knoxville, TN.

Bethany Schulz - Research forester/ecologist U.S. Forest Service, Pacific Northwest Research Station, Anchorage, AL.

Marie T. Trest - FIA lichen research specialist, Department of Botany, University of Wisconsin, Madison, WI.

James A. Westfall - Research forester, U.S. Forest Service, Northern Research Station, Newtown Square, PA.

Susan Will-Wolf - FIA lichen indicator advisor, Department of Botany, University of Wisconsin, Madison, WI.

Christopher W. Woodall - Research forester, U.S. Forest Service, Northern Research Station, St. Paul, MN.

Contents

Introduction

The Forest Inventory and Analysis program (FIA) of the U.S. Forest Service is responsible for inventorying and monitoring forests within the United States. FIA employs a three-phase sampling design. In the first phase (P1), strata are developed for use in estimation. The second phase (P2) entails collection of data for standard mensuration variables, such as amount of forest area, tree species, diameters, and heights on all sample plots. Phase three (P3) consists of a subset of these plots, where forest health information is obtained for indicators such as down woody material, tree crowns, lichens, ozone damage, soils, and vascular vegetation composition and diversity.[1]

As part of a comprehensive quality assurance (QA) program, a number of plots are randomly selected for 'blind check' remeasurement. On these plots, a second set of measurements are taken by a different crew (QA crew) who have no knowledge of the results obtained by the field crew (production crew). The two sets of data are then compared to evaluate how repeatable the measurements are. The differences between the two measurements are indicative of measurement uncertainty, not measurement error, since the true values are unknown (Lischke 2001). Most variables have a Measurement Quality Objective (MQO), which contains two parts: a measurement tolerance and a compliance rate.

The acceptable difference between independent measurements of the production crew and the QA crew is defined by the tolerance (e.g., ±10 percent). The expected proportion of measurement differences that are within the tolerance is specified by the compliance rate (e.g., 90 percent of the time). For this example, data quality is deemed acceptable when 90 percent or more of the measurement differences are within ±10 percent. The number of differences within measurement tolerance divided by the total number of paired observations gives the percentage of measurements having acceptable repeatability. If the percentage is less than the compliance rate, the variable has failed the MQO.

Some blind check remeasurements cannot be appropriately analyzed within the tolerance/compliance rate framework. Thus, measurement quality objectives for certain variables are stated in a substantially different way. For the lichen indicator, blind check remeasurements are performed by a QA specialist and repeatability is judged on attainment of a minimum standard relative to the QA specialist rather than the differences between two independent measurements. Similarly, data collected for the ozone indicator defy the standard analytical method due to data collection protocols that prohibit matching of observations between crews. In this case, differences in site-level variables are compared.

This report focuses on measurement precision and does not address measurement bias, nor does the report examine sampling, prediction, and nonresponse errors.

---

[1]We use the term compliance rate to refer to the values described as MQOs in FIA field guides, available at http://www.fia.fs.fed.us/library/field-guides-methods-proc/ .

It should be noted that measurement precision is an important component of estimation error, because the estimates are based on sample measurements. For all P3 indicators, estimates of change are of ever-increasing importance. As P3 plots are remeasured, FIA will be able to provide estimates of change based on two independent observations of the same individual (plots, subplots, trees, quadrats, biosites, etc.) taken at two points in time. One component of the error in those estimates will be the measurement precision of those two independent observations. The blind check process emulates this independent remeasurement process in the absence of time change. Inherent in the measures of change for a given attribute will be measurement variation at least as large as those observed on blind check plots, making the results presented here useful as predictors of our ability to detect change. To increase the power of detecting change over time (the probability of detecting a statistically significant change, if such a change has truly taken place), paired measurements with good measurement precision are required (Yanai et al. 2003).

The purpose of this report is to analyze the repeatability of forest health indicator variables measured on P3 sample plots. The report is organized into sections, with each section pertaining to a specific forest health indicator. Within each section are assessments that focus on attainment of the MQO for important indicator attributes. Also, some sections contain the results of additional analyses that provide insight into various other aspects of repeatability. However, specific recommendations for improvement are limited, as such recommendations must be made with consideration to a host of additional factors.

The data used in this report were collected from all FIA regions across the United States. The following abbreviations and their references are used throughout the report:

SRS: Southern FIA
NE: Northeast FIA
NC: North Central FIA
IW: Interior West FIA
PNW: Pacific Northwest FIA

The information contained herein should be useful to data collection experts, as variables having measurements that are below the MQO can be identified and appropriately addressed. Analysts and researchers may also use the results to determine if the repeatability of measurements is adequate for their analyses.

Section 1: Tree Crown Indicator

James A. Westfall
William A. Bechtold
KaDonna C. Randolph

Introduction

Tree crown conditions provide valuable insight into the health of forest ecosystems. The U.S. Forest Service has been collecting tree crown data since 1990. Between 1990 and 1999, these data were obtained by the Forest Health Monitoring (FHM) program. In 2000, the FHM plot network and the Forest Inventory and Analysis (FIA) program merged, which resulted in crown data collection being undertaken by FIA. A national quality assurance (QA) program was implemented by FIA (USDA Forest Service 2004a) to continuously monitor the quality of inventory data collection. One aspect of this QA program entails independent remeasurement of a randomly selected subset of sample plots. These data are used to evaluate measurement repeatability by comparing the QA measurements to those obtained during the visit by the production plot crew.

The purpose of this study is to analyze measurement repeatability for the suite of crown condition variables collected by FIA to assess the health of tree crowns.

Methods

Field Data Collection

FIA phase 2 (P2) plots occur at an intensity of approximately 1 plot per 6,000 acres in a spatially distributed sampling grid across the United States (Bechtold and Patterson 2005). Crown condition data are obtained from FIA Phase 3 (P3) sample plots, which comprise one-sixteenth subset of FIA P2 sample plots. Thus, each P3 plot represents roughly 96,000 acres (Reams et al. 2005).

For all attributes measured on a sample plot, a repeatability criterion (measurement quality objective [MQO]) is specified in terms of a tolerance and compliance rate. The acceptable range of differences between independent measurements is defined by the tolerance. The expected proportion of differences that are within the tolerance is specified by the compliance rate (see the introductory text). The FIA P3 field guide (USDA Forest Service 2004b) provides in-depth descriptions of field data collection procedures and repeatability standards for crown indicator variables.

This assessment is based on data from trees measured on 147 FIA P3 plots that were randomly chosen for QA inspection between 2002 and 2004. The distribution of QA plots by region were as follows: SRS = 49; NE = 20; NC = 35; IW = 37; PNW = 6. There are two sets of data from these plots. The first set of observations was collected during the production inventory visits. The second set of data is from ‘blind check’ remeasurement of these same plots by an independent field crew within two weeks after the production inventory measurement. The measurement protocols were identical for both crews.

Matching Procedures

Due to the ‘blind check’ nature of the plot remeasurement, there was no guarantee that the same tree number was assigned by both crews to the same tree. There are a number of reasons this may occur, including missed trees and differences in sample inclusion determination. Thus, the trees measured by each crew need to be matched to compute valid measurement differences. To accomplish this, a data-matching algorithm was used to ensure that both measurements were observations of the same tree.

The matching algorithm employed weighted distance functions based on distance and azimuth from plot center and tree diameter at breast height (d.b.h.) to determine appropriate pairings of data from the independent measurements. Although species are assigned numeric codes, there is no intrinsic meaning in terms of differences, e.g., how ‘far away’ is an oak from a maple. As such, there was no clear method to quantify species differences and species was not used in tree matching. This process was designed to only match trees where the location and size attributes were within prespecified ranges of agreement. The remaining unmatched trees were individually evaluated and either added to the matched data or set aside as extra trees. This same procedure was employed by Pollard et al. (2006) to match trees on FIA P2 plots.

Analytical Techniques

The repeatability of the various crown indicator variables was assessed via the differences between observed values from production and QA crews. Differences were computed by subtracting the production inventory measurement from the QA measurement.

$$d_{ij} = x_{ij} - y_{ij}$$

where: d_{ij} = difference between QA crew and production crew for tree i, crown variable j

x_{ij} = observation from QA crew for tree i, crown variable j

y_{ij} = observation from production crew for tree i, crown variable j

The percentage of observations falling within the range of the specified tolerance level was determined:

$$p_j = \frac{\sum w_{ij}}{n_j}$$

where: p_j = estimated proportion of observations within tolerance t_j for crown variable j

t_j = specified tolerance level for crown variable j

$$w_{ij} = \begin{cases} 1 \text{ if } |d_{ij}| \le t_j \\ 0 \text{ if } |d_{ij}| > t_j \end{cases}$$

n_j = number of observations for crown variable j

Exact binomial 95 percent confidence intervals for the percentage of the observations within tolerance also were computed (Balakrishnan and Nevzorov 2003). If we denote the number of observations within tolerance as $m_j = \sum w_{ij}$ and solve

$$0.025 = \sum_{m_{Uj}}^{n_j} \binom{n_j}{m_{Uj}} p_j^{m_{Uj}} (1 - p_j)^{n_j - m_{Uj}}$$

and

$$0.025 = \sum_{0}^{m_{Lj}} \binom{n_j}{m_{Lj}} p_j^{m_{Lj}} (1 - p_j)^{n_j - m_{Lj}}$$

for m_{Uj} and m_{Lj} respectively, then we can compute the upper (p_{Uj}) and lower (p_{Lj}) confidence limits for p_j as

$p_{Uj} = \frac{m_{Uj}}{n_j}$ and $p_{Lj} = \frac{m_{Lj}}{n_j}$, respectively.

These computations were performed using SAS® statistical software (SAS Institute, Inc. 2003).

The confidence intervals were applied to the calculated proportion of observations within tolerance for each variable. If the MQO compliance rate was within the interval, the measurement repeatability was assumed to have met the established standards. Other reported statistics that describe the distribution of measurement differences include mean difference and root mean squared difference (RMSE).

Results and Discussion

Uncompacted Crown Ratio

The length of the tree section between the live crown top and the live crown base defines the uncompacted crown length. To obtain uncompacted crown ratio, the uncompacted crown length is divided by the length from the base of the tree to the live crown top. Observations are recorded to the nearest 1-percent, with values ranging from 0 to 100.

Overall, independent measurements of uncompacted crown ratio were within the ±10 percent tolerance 79.1 percent of the time, which is below the 90 percent MQO compliance rate (Table 1.1). Regionally, the values ranged from 67.9 percent (NE) to 88.1 percent (NC). NC was the only region that met the target compliance rate (i.e., the confidence interval includes 90 percent). The magnitude of the mean difference suggests bias should be of little concern. RMSE values show that the average discrepancy between independent measurements ranged from roughly 10 percent to 15 percent. As shown in Tables 1.2 and 1.3, separating the data into hardwood and softwood groups indicated that repeatability was poorer for hardwood trees (76.7 percent) than softwood trees (81.9 percent). This result was not unexpected as the deliquescent crown form for many hardwoods makes it difficult to clearly establish the tree top and crown base. These results are comparable to the results reported by Pollard and Smith (1999, 2001). In their evaluation of QA data collected by the FHM Program in 1998 and 1999, repeatability was also poorer for hardwoods than for softwoods.

A more detailed analysis was conducted to determine if measurement repeatability problems may be related to crown ratio values. The uncompacted crown ratio data were thus grouped into 5 percent classes (based on the measurement recorded by the QA crew) and repeatability statistics were generated for each class (Table 1.4). For classes having reasonable sample sizes (≥20 trees), the repeatability was relatively consistent (70 to 80 percent), with perhaps the exception of the 90 percent and higher categories where repeatability increased notably. However, a trend is evident in the mean differences. For trees having low crown ratios, the QA measurement is, on average, smaller than the production inventory measurement. Conversely, there is a tendency for the QA measurement to be higher than the production inventory measurement for ratios above 55 percent. QA crews are more likely than production crews to assign extremely low and extremely high values. Further investigation is needed to ascertain the underlying cause(s) of this pattern.

Subsequent to the data used for this QA analysis, procedures were modified in 2006 for estimating uncompacted crown ratio for leaning and down trees and trees with dead tops. Uncompacted crown ratio was previously estimated based on tree height above the ground (not tree length). Currently, uncompacted crown ratio for all trees (leaning or not) is based on tree length. In addition, the denominator of the ratio has been redefined to include actual tree length (as defined by FIA), rather than the top of the live crown. Uncompacted crown ratio is thus expressed as a percentage of actual tree length. The result is that uncompacted crown ratio now includes dead tops (but not broken or missing tops). It is not yet apparent if or how these changes affect measurement repeatability.

Table 1.1—Repeatability statistics for tree crown attributes by FIA region.

Variable	Tolerance	MQO compliance rate	FIA region	Percent within tolerance	95% CI	Mean difference (QA crew – production crew)	RMSE	Records
Uncompacted crown ratio	±10%	90	SRS	75.3	71.8 - 78.6	2.6	14.9	648
			NE	67.9	63.9 - 71.8	1.9	15.4	558
			NC	88.1	85.5 - 90.4	-0.4	9.7	712
			IW	83.0	80.3 - 85.4	-0.8	10.1	881
			PNW	68.1	57.7 - 77.3	2.8	13.9	94
			ALL	79.1	77.6 - 80.6	0.7	12.5	2893
Crown light exposure	±1 class[a]	85	SRS	87.2	84.4 - 89.7	-0.1	0.9	648
			NE	83.5	80.2 - 86.5	-0.1	1.0	558
			NC	89.5	87.0 - 91.6	-0.1	0.8	712
			IW	84.4	81.9 - 86.8	-0.2	1.2	881
			PNW	92.6	85.3 - 97.0	0.2	0.8	94
			ALL	86.4	85.1 - 87.6	-0.1	1.0	2893
Crown position[b]	No tolerance	85	SRS	90.0	87.4 - 92.2	-	-	647
			NE	81.2	77.7 - 84.3	-	-	558
			NC	90.3	87.9 - 92.4	-	-	712
			IW	70.7	67.6 - 73.7	-	-	881
			PNW	77.7	67.9 - 85.6	-	-	94
			ALL	82.1	80.6 - 83.5	-	-	2892
Vigor class	No tolerance	90	SRS	80.0	72.8 - 86.0	-0.1	0.5	155
			NE	82.2	77.2 - 86.5	0.0	0.4	281
			NC	70.5	61.9 - 78.1	-0.2	0.6	132
			IW	63.0	52.3 - 72.9	0.3	0.7	92
			PNW	25.0	0.6 - 80.6	-0.8	0.9	4
			ALL	76.4	72.9 - 79.5	0.0	0.5	664
Crown density	±10%	90	SRS	78.9	75.0 - 82.4	1.0	11.7	493
			NE	60.7	54.6 - 66.6	-2.8	14.9	270
			NC	78.6	75.0 - 81.9	-0.5	11.0	579
			IW	67.8	64.4 - 71.1	0.0	13.4	789
			PNW	67.8	57.1 - 77.2	-3.1	13.0	90
			ALL	72.2	70.3 - 74.1	-0.4	12.6	2221
Crown dieback	±10%	90	SRS	97.4	95.5 - 98.6	-0.2	6.3	493
			NE	90.4	86.2 - 93.6	-1.2	8.5	270
			NC	96.9	95.1 - 98.1	0.0	7.4	579
			IW	95.9	94.3 - 97.2	0.9	7.6	789
			PNW	96.7	90.6 - 99.3	-0.9	6.7	90
			ALL	95.9	94.9 - 96.6	0.1	7.4	2221
Foliage transparency	±10%	90	SRS	86.6	83.3 - 89.5	-0.9	10.3	493
			NE	77.8	72.3 - 82.6	-1.6	15.1	270
			NC	94.8	92.7 - 96.5	-1.3	8.7	579
			IW	95.9	94.3 - 97.2	-0.1	7.2	789
			PNW	96.7	90.6 - 99.3	-1.4	7.9	90
			ALL	91.4	90.2 - 92.5	-0.8	9.6	2221

[a] ±1 class when exposure > 0; no tolerance when exposure = 0
[b] Mean difference and RMSE not reported for categorical variables

Table 1.2—Repeatability statistics for tree crown attributes from 1543 hardwood trees across the U.S.

Variable	Tolerance	MQO compliance rate	Percent within tolerance	95% CI	Mean difference (QA crew – production crew)	RMSE	Records
Uncompacted crown ratio	±10%	90	76.7	74.5 - 78.8	1.1	13.8	1543
Crown light exposure	±1 class[a]	85	87.4	85.7 - 89.0	-0.1	0.9	1543
Crown position[b]	No tolerance	85	86.3	84.5 - 88.0	-	-	1542
Vigor class	No tolerance	90	73.1	68.5 - 77.3	0.0	0.6	405
Crown density	±10%	90	75.8	73.1 - 78.2	0.1	12.1	1130
Crown dieback	±10%	90	94.1	92.5 - 95.4	0.0	9.0	1130
Foliage transparency	±10%	90	87.4	85.4 - 89.3	-1.2	11.6	1130

[a] ±1 class when exposure > 0; no tolerance when exposure = 0
[b] Mean difference and RMSE not reported for categorical variables

Table 1.3—Repeatability statistics for tree crown attributes from 1350 softwood trees across the U.S.

Variable	Tolerance	MQO compliance rate	Percent within tolerance	95% CI	Mean difference (QA crew – production crew)	RMSE	Records
Uncompacted crown ratio	±10%	90	81.9	79.7 - 83.9	0.2	10.8	1350
Crown light exposure	±1 class[a]	85	85.2	83.2 - 87.0	-0.1	1.1	1350
Crown position[b]	No tolerance	85	77.3	74.9 - 79.5	-	-	1350
Vigor class	No tolerance	90	81.5	76.2 - 86.0	0.0	0.5	259
Crown density	±10%	90	68.6	65.7 - 71.3	-0.9	13.2	1091
Crown dieback	±10%	90	97.7	96.6 - 98.5	0.2	5.1	1091
Foliage transparency	±10%	90	95.5	94.1 - 96.7	-0.4	7.0	1091

[a] ±1 class when exposure > 0; no tolerance when exposure = 0
[b] Mean difference and RMSE not reported for categorical variables

Table 1.4—Repeatability statistics for crown ratio measurements by crown ratio class.

QA crew observation[a]	Tolerance	Percent within tolerance	95% CI	Mean difference (QA crew – production crew)	RMSE	Records
0%	±10%	50.0	6.8 - 93.2	-10.3	14.2	4
5%	±10%	66.7	9.4 - 99.2	-25.0	46.3	3
10%	±10%	80.0	44.4 - 97.5	-11.7	22.4	10
15%	±10%	62.5	40.6 - 81.2	-9.7	20.7	24
20%	±10%	72.9	58.2 - 84.7	-6.4	14.2	48
25%	±10%	71.0	58.1 - 81.8	-4.7	14.6	62
30%	±10%	75.6	64.9 - 84.4	-5.0	14.4	82
35%	±10%	81.3	73.8 - 87.4	-1.2	12.0	139
40%	±10%	83.1	77.1 - 88.1	-1.5	10.8	195
45%	±10%	79.9	74.0 - 85.0	-2.5	10.5	219
50%	±10%	77.6	72.0 - 82.5	-1.5	12.8	263
55%	±10%	79.6	73.5 - 84.9	1.3	11.2	206
60%	±10%	74.0	67.7 - 79.6	2.0	13.1	223
65%	±10%	68.8	61.7 - 75.2	1.2	12.7	192
70%	±10%	71.9	64.4 - 78.5	3.2	11.9	167
75%	±10%	70.8	63.5 - 77.3	4.0	15.0	178
80%	±10%	76.4	69.0 - 82.8	2.6	14.3	157
85%	±10%	79.6	72.0 - 85.9	3.6	13.0	142
90%	±10%	87.4	81.7 - 91.9	3.5	12.1	183
95%	±10%	88.8	82.7 - 93.3	3.9	12.2	152
100%	±10%	95.1	91.6 - 97.4	2.3	7.6	244

[a] QA crew observations were grouped into 5-percent classes: 0% = 0%, 5% = 1-5%, 10% = 6-10%, …100% = 96-100%.

Crown Light Exposure

A measure of the amount of direct sunlight a tree receives is given by the crown light exposure variable. The measurement protocol requires the tree to be divided vertically into four equal sides. A count of the number of sides receiving direct light if the sun were directly above the tree is made. One is added to this count if the tree receives direct light from the top. The total is the value entered for crown light exposure, with acceptable values ranging from 0 to 5.

The repeatability assessment for crown light exposure measurements indicated 86.4 percent of the observations were within the tolerance (Table 1.1). The attained repeatability (86.4 percent) slightly exceeded the MQO compliance rate of 85 percent. Evaluations by FIA region showed a high level of consistency across regions and the confidence intervals for each region included the MQO value. The small values for mean difference indicate negligible bias, while RMSE were generally near 1 class. The results also were very consistent between hardwood and softwood species (Tables 1.2 and 1.3), suggesting that evaluation of crown light exposure is unaffected by tree crown characteristics and foliage type.

Crown Position

Crown position information is collected for each tree to indicate where an individual tree crown is located in relation to the overstory trees. Three codes are used in stands where the crowns are closed (i.e., ≥50 percent crown cover at the stand level). These codes denote superstory, overstory, and understory crown positions. An additional code is used for stands lacking crown closure (<50 percent crown cover). In this case, all trees in the stand are assigned a single code of open grown.

The comparison of matched crown position measurements showed exact agreement (tolerance level is 0) between independent measurements 82.1 percent of the time with a 95 percent confidence interval of 80.6 percent - 83.5 percent (Table 1.1). The desired compliance rate is 85 percent. There was a notable amount of regional variation, from 70.7 percent (IW) to 90.3 percent (NC). Based on the confidence intervals, the NE region was slightly out of compliance, but the IW region failed to achieve the MQO compliance rate by a substantial amount. Repeatability was somewhat less for softwood trees (77.3 percent) as compared to hardwood trees (86.3 percent) for all regions combined (Tables 1.2 and 1.3). This difference probably contributed to the low compliance rate in the IW since the region contains a high proportion of softwoods. The fact that repeatability for hardwood trees slightly exceeded the MQO compliance rate indicates that improved repeatability for softwoods is needed.

A confounding factor in the analysis of crown position repeatability was the conditional use of codes depending on whether the stand-level crown closure attained 50 percent or more. In situations where differences in determining whether crown closure has occurred, all of the crown position codes will differ between crews. Given that the tolerance is zero, this lowers the overall repeatability statistic. For these data, 210 trees on 17 plots were affected by differences in crown closure determination. Excluding these data from the analysis increased agreement between the independent crown position measurements to 88.5 percent, which exceeds the MQO compliance rate of 85 percent.

Additional analyses indicated that the crown closure differences primarily occurred in the IW region where there is a high level of forest/woodland interface. Of the 210 trees affected by crown closure differences, 193 occurred in the IW region. These 193 trees occurred over 14 conditions, where open crown conditions were observed nine times by the QA crew and five times by the production crew. Thus, there is a tendency for the QA crew to determine lack of crown closure more often than the production crew.

IW also had the lowest repeatability statistic (70.7 percent) for crown position among all regions (Table 1.1). Excluding the data where crown closure determinations differed resulted in an increase to 90.6 percent repeatability between crews in IW. Decreasing differences on degree of crown closure between crews, especially in the IW region, will increase the repeatability of crown position observations.

Crown Vigor Class

Tree saplings (1.0 inch ≤ d.b.h. ≤4.9 inches) are assigned a crown vigor class for the purpose of identifying very good or extremely poor crowns. The ratings are based on crown dieback, uncompacted crown ratio, and amount of missing or damaged foliage. Vigor is recorded in three classes: 1 = vigorous, 2 = moderate vigor, and 3 = poor vigor. Most saplings are in class 1.

The repeatability of crown vigor class measurements is evaluated with no tolerance for differences. Results of the analysis show the overall repeatability of 76.4 percent (95 percent confidence interval of 72.9 percent to 79.5 percent) did not attain the MQO compliance rate of 90 percent (Table 1.1). All regions failed to attain the desired compliance rate, although SRS and NE regions were fairly close with upper confidence limits near 86 percent. Statistics presented for the PNW region should be viewed with caution, as only four trees contributed to the analysis. Comparisons between hardwood and softwood species (Tables 1.2 and 1.3) show that agreement between measurements was better for softwoods (81.5 percent) than hardwoods (73.1 percent).

Repeatability of crown vigor measurements was also independently assessed for each of the three classes. The results shown in Table 1.5 clearly indicate that repeatability is poor when the QA crew observes crown vigor other than class 1. Analysis of the mean difference and RMSE values indicate that the poor repeatability for classes 2 and 3 are the result of the QA crew recording a higher vigor class (lower tree vigor) than the production inventory crew, on average.

Table 1.5—Repeatability statistics for crown vigor measurements by crown vigor class.

QA crew observation	Tolerance	Percent within tolerance	95% CI	Mean difference (QA crew – production crew)	RMSE	Records
1	No tolerance	83.5	79.9 - 86.7	-0.2	0.4	491
2	No tolerance	56.7	48.6 - 64.6	0.4	0.7	157
3	No tolerance	50.0	24.7 - 75.3	0.8	1.1	16

Crown Density

Crown density is a measure of how much light is blocked by the tree crown. The measurement protocol entails defining tree crown shape and using a reference card to determine the percentage of light being obscured. Observations are recorded in 5-percent classes, with a value of 0 for trees with no crown (i.e., epicormic branches only).

Overall, crown density repeatability was the lowest among all the crown indicator variables. Evaluated with a tolerance of ±10 percent, the attained repeatability for crown density was 72.2 percent with an associated 95 percent confidence interval of 70.3 to 74.1 percent (Table 1.1). This was an undesirable realization, as the MQO standard is 90 percent. Regional results ranged from 60.7 to 78.9 percent for the NE and SRS units, respectively. No individual regions achieved the MQO compliance rate. The mean differences show no practical bias. The average deviation between independent measurements (based on RMSE) ranged from approximately 11 to 15 percent. Consistency was higher for hardwood species than for softwood species (75.8 percent and 68.6 percent, respectively) (Tables 1.2 and 1.3).

Pollard and Smith (1999, 2001) also found that between-crew agreement was more consistent for hardwoods than softwoods and that both species groups were below the MQO standard of 90 percent. Repeatability has declined since 1999. Pollard and Smith (1999, 2001) observed 83 percent agreement for hardwoods in 1998 and 89 percent in 1999; and 75 percent for softwoods in 1998 and 86 percent in 1999.

Further analyses of crown density repeatability shows that the most consistent agreement between measurements occurred when observed values were 60 percent or less (Table 1.6). In this range, measurements are within ±10 percent of each other, on average, more than 70 percent of the time. Repeatability decreased with increasing crown density when crown densities became greater than 60 percent. A precipitous decline occurs beyond the 60 percent point, e.g. only one-third of the production crew measurements were within tolerance when the QA crew recorded a value of 80 percent. There were no matched observations within ±10 percent for crown density of 85 percent. No values were recorded by the QA crew above 85 percent. This poor repeatability is probably due in part to the rarity of trees with high crown density. Crown densities above 75 percent are rare for most species unless they are open-grown. FHM data from the South (1995-1999) had a maximum crown density of 85 percent. The QA data used by Pollard and Smith had a maximum density of 90 percent (1 tree). Caro et al. (1979) noted that rare events are less likely to have strong inter-observer agreement simply because observers do not assess these conditions frequently.

An interesting bias in crown density ratings was observed (Table 1.6). Generally, the difference between crews becomes larger as crown density moves away from the 45 to 50 percent range. For trees having low crown density, the QA measurement is, on average, smaller than the production inventory measurement. Conversely, there is a tendency for the QA measurement to be higher than the production inventory measurement for density ratings above 50 percent. The RMSE values depict a similar trend. This indicates that production crew values tend to gravitate toward the mean, which was also observed with the crown ratio data.

Table 1.6—Repeatability statistics for crown density measurements by crown density class.

QA crew observation[a]	Tolerance	Percent within tolerance	95% CI	Mean difference (QA crew – production crew)	RMSE	Records
0%	±10%	66.7	22.3 - 95.7	-17.5	27.6	6
5%	±10%	83.3	35.9 - 99.6	-5.8	11.4	6
10%	±10%	66.7	38.4 - 88.2	-9.0	16.3	15
15%	±10%	61.3	42.2 - 78.2	-9.7	17.0	31
20%	±10%	68.2	52.4 - 81.4	-10.1	16.8	44
25%	±10%	67.5	55.9 - 77.8	-7.0	13.5	77
30%	±10%	73.4	66.3 - 79.8	-5.2	12.1	177
35%	±10%	77.7	72.9 - 82.1	-2.6	11.1	337
40%	±10%	74.8	70.3 - 78.8	-2.7	11.7	420
45%	±10%	78.7	74.0 - 82.9	-0.2	10.7	343
50%	±10%	71.8	65.7 - 77.5	0.7	12.2	238
55%	±10%	70.9	64.0 - 77.1	3.6	11.8	199
60%	±10%	71.6	63.6 - 78.7	5.7	13.6	148
65%	±10%	57.3	46.4 - 67.7	8.2	15.4	89
70%	±10%	52.1	37.2 - 66.7	9.8	16.0	48
75%	±10%	38.5	20.2 - 59.4	16.7	21.2	26
80%	±10%	33.3	9.9 - 65.1	17.9	21.7	12
85%	±10%	0.0	0.0 - 52.2	26.0	27.0	5

[a] QA crew observations were grouped into 5-percent classes: 0% = 0%, 5% = 1-5%, 10% = 6-10%, …100% = 96-100%.
Note: There were no observations above 85% from the QA crews.

Crown Dieback

An estimate of the recent mortality of branches with fine twigs occurring in the upper and outer portions of the tree is provided by the crown dieback variable. Percent dieback is measured as a percentage of the total crown area. Observations are recorded in 5 percent classes, with a value of 0 for trees with no dieback.

Both these results and those reported by Pollard and Smith (1991, 2001) indicate that crown dieback is the most repeatable measurement within the crowns indicator suite. Overall repeatability was assessed at 95.9 percent with a tolerance of ±10 percent (Table 1.1). This exceeds the MQO compliance rate of 90 percent. All FIA regions met the MQO standard and RMSE values ranged from roughly 6 to 8 percent. As expected with a measurement exhibiting high levels of consistency, evaluations by hardwood and softwood species groups both exceeded the specified compliance rate. Softwood trees attained a statistic of nearly 98 percent, while hardwood trees were somewhat lower at 94.1 percent (Tables 1.2 and 1.3).

Most trees had little or no crown dieback. In the data used for this study, 95 percent of the trees had dieback of 10 percent or less according to the QA crew (Table 1.7). The small amount of dieback occurring on a large proportion of the trees resulted in high repeatability statistics. However, most classes between 15 to 40 percent did not attain the MQO standard. It is also in this range where sample sizes decreased substantially, suggesting that this may be a ‘threshold’ area where crews began to measure trees with amounts of dieback that were not commonly encountered. Minimal experience in measuring trees with these attributes could contribute to increased measurement variability. There were

Table 1.7—Repeatability statistics for crown dieback measurements by crown dieback class.

QA crew observation[a]	Tolerance	Percent within tolerance	95% CI	Mean difference (QA crew – production crew)	RMSE	Records
0%	±10%	98.1	97.2 - 98.7	-1.7	4.5	1357
5%	±10%	98.2	96.8 - 99.1	0.9	6.7	605
10%	±10%	100.0	97.6 - 100	4.5	6.9	150
15%	±10%	72.2	54.8 - 85.8	7.6	9.9	36
20%	±10%	39.1	19.7 - 61.5	5.7	15.1	23
25%	±10%	60.0	14.7 - 94.7	8.0	12.2	5
30%	±10%	22.2	2.8 - 60.0	18.9	22.5	9
35%	±10%	0.0	0.0 - 45.9	20.8	27.2	6
40%	±10%	40.0	5.3 - 85.3	23.0	27.7	5
45%	±10%	33.3	0.8 - 90.6	26.7	34.9	3
50%	±10%	50.0	6.8 - 93.2	-1.3	28.6	4
55%	±10%	100.0	2.5 - 100	0.0	0.0	1
60%	±10%	0.0	0.0 - 97.5	35.0	35.0	1
65%	±10%	50.0	1.3 - 98.7	27.5	38.9	2
70%	±10%	100.0	2.5 - 100	5.0	5.0	1
75%	±10%	33.3	0.8 - 90.6	23.3	30.8	3
80%	±10%	50.0	1.3 - 98.7	-7.0	13.9	2
85%	±10%	-	-	-	-	0
90%	±10%	50.0	6.8 - 93.2	10.0	18.4	4
95%	±10%	-	-	-	-	0
100%	±10%	50.0	6.8 - 93.2	49.5	70.0	4

[a] QA crew observations were grouped into 5-percent classes: 0% = 0%, 5% = 1-5%, 10% = 6-10%, …100% = 96-100%.

relatively few trees beyond 40 percent crown dieback, although wide confidence intervals caused by small sample sizes in this data range consistently included the MQO compliance rate of 90 percent. Mean differences between crews are generally positive, indicating that QA crews are more likely to assign higher dieback values than production crews, again indicating that production crews tend to gravitate toward the mean.

Foliage Transparency

Foliage transparency describes the amount of light coming through the live portion of the crown (i.e., where normal, damaged, or partly missing foliage occurs). Crown dieback and other dead branches are not included. Once the foliated area of the crown is determined, a reference card is used to estimate transparency in 5-percent classes, with a value of 99 percent for trees with no crown (i.e., epicormic branches only).

With a tolerance of ±10 percent between independent measurements, the repeatability of foliage transparency (91.4 percent) exceeded the MQO compliance rate of 90 percent; however, the regional analyses show that the SRS slightly missed the specified standard and the NE was substantially out of compliance (Table 1.1). As with the other variables, the mean differences were trivial. The RMSE values reflected the regional differences, with a minimum of 7.2 percent (IW) and a maximum of 15.1 percent (NE). There was a marked difference in repeatability between hardwood and softwood species. The measurements had 95.5 percent repeatability for softwoods, but only 87.4 percent agreement for hardwoods (Tables 1.2 and 1.3). Pollard and Smith (1999, 2001) also found a difference in repeatability between hardwood and softwood species, though not to the same extent. Repeatability was higher for softwoods than for hardwoods but only by 3 to 4 percent. The characteristics of hardwood crowns probably contribute to this discrepancy, as exclusion of light gaps between branches must be taken into account.

An analysis of repeatability by foliage transparency classes showed some trends that were similar to other crown indicator variables (Table 1.8). Most notably, the same bias was evident that was found for both crown ratio (Table 1.4) and crown density (Table 1.6), i.e., the mean differences indicated the QA measurement was smaller than the production crew for low values of foliage transparency and were higher than the production crew when foliage transparency was high. The transition from negative to positive mean differences occurred near the 20 percent level, which is in the region where most of the data were concentrated.

Another discerned trend was a drop in repeatability outside the 5-35 percent foliage transparency range. Again, this seems to coincide with a precipitous decline in sample size and a progression into amounts of foliage transparency that are not often observed.

Table 1.8—Repeatability statistics for crown foliage transparency measurements by foliage transparency class.

QA crew observation[a]	Tolerance	Percent within tolerance	95% CI	Mean difference (QA crew – production crew)	RMSE	Records
0%	±10%	10.0	0.3 - 44.5	-20.4	21.0	10
5%	±10%	80.0	28.4 - 99.5	-8.0	9.5	5
10%	±10%	92.1	88.4 - 94.8	-5.3	8.7	302
15%	±10%	96.6	95.1 - 97.8	-2.1	7.9	746
20%	±10%	95.3	93.4 - 96.7	-0.1	7.3	69
25%	±10%	87.3	82.8 - 91.0	1.0	10.7	276
30%	±10%	77.1	67.9 - 84.8	4.3	10.7	105
35%	±10%	61.4	45.5 - 75.6	5.5	15.2	44
40%	±10%	40.0	12.2 - 73.8	7.0	22.6	10
45%	±10%	25.0	3.2 - 65.1	13.1	19.3	8
50%	±10%	40.0	5.3 - 85.3	4.2	28.5	5
55%	±10%	33.3	0.8 - 90.6	20.0	21.6	3
60%	±10%	50.0	1.3 - 98.7	17.5	19.0	2
65%	±10%	-	-	-	-	0
70%	±10%	100.0	15.8 - 100	5.0	5.0	2
75%	±10%	50.0	1.3 - 98.7	25.0	35.4	2
80%	±10%	-	-	-	-	0
85%	±10%	-	-	-	-	0
90%	±10%	-	-	-	-	0
95%	±10%	0.0	0.0 - 97.5	80.0	80.0	1
100%	±10%	40.0	5.3 - 85.3	43.4	58.3	5

[a] QA crew observations were grouped into 5-percent classes: 0% = 0%, 5% = 1-5%, 10% = 6-10%, …100% = 96-100%.

Conclusions

Crown attributes that effectively meet the MQO compliance rate include crown light exposure, crown dieback, and foliage transparency. Current training methods and measurement protocols for these variables are producing acceptable levels of measurement repeatability.

Crown position was slightly below the MQO compliance rate when evaluated at the national level due to repeatability problems with the open-grown classification in the IW region. More training in this region, especially with respect to 50 percent crown closure, should help to achieve compliance both nationally and regionally.

The uncompacted crown ratio, vigor class, and crown density measurements did not attain the MQO compliance rate. Experts in field data collection of crown attributes should re-evaluate the standards for these variables to determine if they are realistic. If additional training cannot correct the problem, the MQO should be modified to be in accordance with observed repeatability statistics. If an MQO adjustment (compliance rate and/or tolerance) is implemented, the utility of each indicator needs to be re-evaluated with respect to the revised MQO. Additionally, analysts and researchers who utilize these data should assess whether the repeatability of the crown measurements is sufficient for their respective analyses.

Finally, measurement protocols for uncompacted crown ratio and crown density have changed since the data used in this analysis were collected. As data collected under these revised protocols becomes available, additional analyses are warranted. These protocol changes are expected to yield at least some improvement in measurement repeatability.

Section 2: Down Woody Materials Indicator

James A. Westfall
Christopher W. Woodall

Introduction

Efficient and accurate sampling of down woody materials (DWM) in forests is critical for estimation of national carbon stocks and forest fuel loadings. The U.S. Forest Service's Forest Inventory and Analysis program (FIA) has been conducting a national inventory of DWM since 2001. In an effort to maintain and improve the quality of inventory data collection nationwide, a continuous quality assurance (QA) program is in effect (USDA Forest Service 2004a). Under this program, a portion of inventory sample plots are independently remeasured by a qualified field crew each year. These QA measurements can then be compared to the measurements collected during the regular plot visit to evaluate measurement repeatability.

The goal of this study was to assess the repeatability of all variables sampled by the FIA program's national DWM inventory using remeasurement data.

Methods

Field Data Collection

FIA phase 2 (P2) sample plots occur in a spatially distributed design across the United States at an intensity of approximately 1 plot per 6,000 acres. DWM data are obtained from FIA phase 3 (P3) sample plots, which at a national core sample intensity comprise a one-sixteenth subset of FIA P2 plots. Thus, each P3 plot represents roughly 96,000 acres (Reams et al. 2005). DWM data are only collected on forested portions of sample plots.

Attributes collected for the DWM sample are broadly grouped into five categories: coarse woody debris (CWD), fine woody debris (FWD), shrub and herb fuels, residual piles, and depth of forest floor components. Attributes of CWD, FWD, and forest floor depth are sampled along transects within the plot area (Figure 2.1). CWD diameters are measured (to nearest inch) at the point of intersection with the transect and at the small end and large end of each sampled piece. The length of each piece is also recorded to facilitate computation of population estimates (e.g., CWD pieces per acre). Additional CWD variables include slope distance along the transect from subplot center, decay class, species, whether or not the piece is hollow, and its history (natural cause or harvest). CWD residue piles (i.e., slash piles) are measured using specific pile sampling protocols if their centers intersect the 24-foot fixed radius subplot.

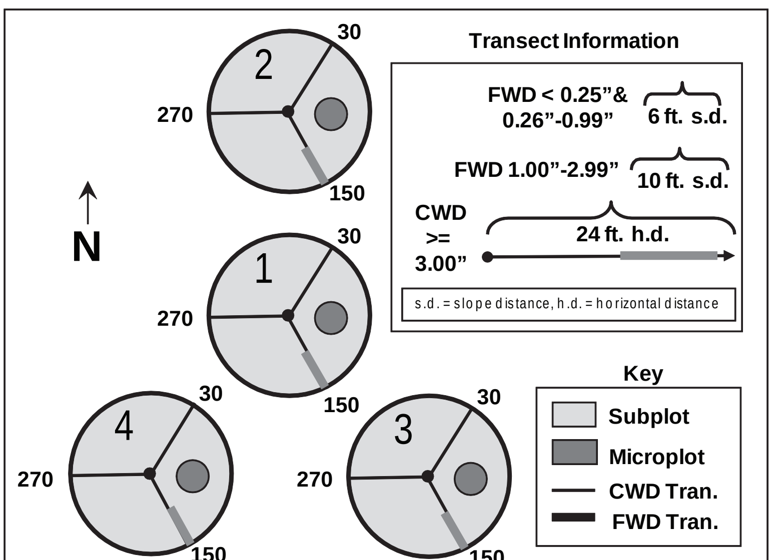
FWD is sampled only on a portion of the 150-degree transect for each subplot where the number of FWD intersecting the sampling transect are tallied by diameter class. The length of the transect used for sampling depends on which of three size classes is being counted: 6 feet for small FWD (0.01-0.24 inch diameter) and medium FWD (0.25-0.99 inch diameter), and 10 feet for large FWD (1.00-2.99 inch diameter). The presence/absence of a residue pile on the FWD segment is also observed.

Measurements of duff, litter, and fuelbed depth are taken at 24-foot slope distance of each transect line. If there is an obstruction (rock, log, etc.) at this point, duff and litter measurements are not taken and fuelbed data is only taken if the obstruction is a log. Duff and litter depths are measured to the nearest 0.1 inch; fuelbed depths are recorded to the nearest 0.1 foot.

Shrub and herb fuels are assessed on the 6.8-foot radius microplot (Figure 2.1). The primary fuel loading variables are percent cover and height of live and dead herbs and shrubs. Percentage cover of forest floor litter is also assessed. Percentage cover is recorded in 10-percent classes and heights are measured to the closest 0.1 foot up to 6 feet for herbs and 99.9 feet for shrubs. For further information regarding the field protocols and estimation/analytical procedures for the DWM indicator please refer to Woodall and Monleon (2008).

For each variable measured, a repeatability criterion (measurement quality objective [MQO]) is specified in terms of a tolerance and compliance rate. The tolerance defines the range of differences between independent measurements that are considered acceptable. The MQO compliance rate states the desired proportion of differences that are within the tolerance (see the introductory chapter). Detailed explanations of field data collections procedures and repeatability standards for DWM variables can be found in the FIA P3 field guide (USDA Forest Service 2004b).

Figure 2.1—Plot design for sampling DWM attributes



Distance between subplots (2, 3, and 4) and subplot center (1): 120 ft at angles (deg.) 0, 120, and 240 respectively, distance between subplot center and microplot center: 12 ft; shrubs/herbs sampled on microplot. Duff/litter sampled at 24 ft slope distance on each CWD transect.

The data used for these analyses arise from FIA P3 sample plots that were randomly chosen for QA inspection. All 118 plots that were collected from 2002 to 2004 were used in this analysis. Specifically, these were blind check plots, where the plot was remeasured by an independent field crew within 2 weeks after the plot was measured as part of the regular inventory. The measurement protocols were identical for both crews.

Matching Procedures

All DWM measurements, except for CWD, could be uniquely identified and matched based on the unique identifiers of microplot, transect, subplot, plot, and state, or some combination thereof.

CWD measurements differ from all other DWM attributes in that individual pieces are identified and measured along each transect. Thus, the pieces measured by each crew need to be matched together in order to compute measurement differences. A matching algorithm was developed that used the location on the transect and size of each piece of CWD to determine appropriate pairings of data from the independent measurements. The process was designed to be conservative and only match pieces where the location and size attributes were in close agreement. As such, there were a number of unmatched pieces remaining. These pieces were individually evaluated and either added to the matched data or set aside as extra pieces. This procedure was similar to that developed by Pollard et al. (2006) to match trees on FIA P2 plots. About 17 percent of the all matched data was obtained via this manual process. There were 72 QA unmatched pieces and 111 regular inventory pieces unmatched, resulting in nearly 17 percent of the data remaining unpaired. Westfall and Woodall (2007) discuss reasons why transect location and piece sizes may be inconsistent between measurements.

Analytical Techniques

The differences between observed values from each crew were used to assess the repeatability of the various DWM measurements. For each attribute, a distribution of differences was generated by pooling data from all available QA plots nationwide. To assess the adequacy of repeatability, the percentage of observations falling within the range of the specified tolerance levels was determined and compared to the MQO compliance rate for each variable. Additionally, the percentage of observations within two, three, or four times the original tolerance level was determined for evaluation of tolerance adjustment effects. Some attributes have zero tolerance, and thus no additional information is provided.

Results and Discussion

Coarse Woody Debris

Several measurements are taken for each tallied CWD piece. Variables that attained the specified repeatability goals are decay class, hollow, history, small-end diameter, and transect diameter for pieces less than 20.0-inch diameter (Table 2.1). Large-end diameter for pieces less than 20.0-inch diameter was very close to meeting the MQO compliance rate. Attributes not meeting the specified compliance rate include slope distance, species, total length, and small-end, transect, and large-end diameter for pieces larger than 20.0-inch diameter. An MQO compliance rate for total piece count by plot is not specified; however, the fact that counts are within the tolerance for almost 90 percent of the plots measured indicates that little concern is warranted.

The diameter measurements for pieces larger than 20.0-inch diameter have relatively few observations, so conclusions regarding repeatability for these variables should be made with caution. Although the percent agreement (69 percent) on species identification is below the MQO compliance rate (80 percent), the results are not particularly noteworthy given the various stages of decay that are encountered for CWD. Slope distance and total length are variables that may deserve further scrutiny. Given the propensity of CWD pieces to be easily trampled or possibly decay (especially in the southeastern U.S.) between field and QA measurements, it can be suggested that the measurement repeatability of dead and downed wood may never reach the levels attained by that of live trees.

Although CWD pieces attained adequate repeatability, it has been observed on non-QA plots and noted by field staff that the difference in diameter measurement precision for live (tenth of an inch) and down dead trees (nearest inch) causes problems with data entry during field sampling. Field crews accustomed to entering live tree diameters to the nearest tenth of an inch will inadvertently do the same for CWD. This relatively rare phenomena can result in a 3-inch CWD piece being recorded as 30 inches. For further information on this measurement error, please refer to Woodall and Westfall (2008).

Table 2.1—Repeatability statistics for CWD attributes based on 523 observations across the U.S.

		MQO compliance	Percent within tolerance				Number of times data exceeded tolerance				
Variable	Tolerance	rate	@1x	@2x	@3x	@4x	@1x	@2x	@3x	@4x	Records
Slope distance	±1 ft	90%	78.8%	89.1%	92.7%	93.9%	111	57	38	32	523
Decay class	±1 class	90%	93.1%	100.0%			36	0			523
Species	No Tolerance	80%	69.0%				152				491
Intersection diameter <20 in	±3 in	90%	98.6%	100.0%			7	0			510
Intersection diameter => 20 in	±20%	90%	61.5%	61.5%	61.5%	61.5%	5	5	5	5	13
Small diameter <20 in	±2 in	90%	94.9%	98.8%	100.0%		25	6	0		487
Small diameter => 20 in	±10%	90%	75.0%	75.0%	75.0%	75.0%	1	1	1	1	4
Large diameter <20 in	±2 in	90%	89.4%	96.9%	99.0%	99.8%	51	15	5	1	482
Large diameter => 20 in	±15 %	90%	66.7%	88.9%	88.9%	88.9%	3	1	1	1	9
Total length	±20%	90%	72.5%	86.6%	91.6%	93.3%	144	70	44	35	523
Hollow	No Tolerance	90%	98.0%				10				490
History	No Tolerance	90%	91.2%				43				491
Total plot count	±2 pcs/5 %	N/A	89.5%	97.7%	98.8%	98.8%	9	2	1	1	86

Fine Woody Debris

The repeatability of the FWD counts was poor for all three size classes (Table 2.2). The small FWD count fared the worst, with slightly less than 25 percent of the counts being within the tolerance. However, percent agreement climbed steadily as tolerance levels increased, suggesting that perhaps a more lenient tolerance specification is appropriate. Similar, but slightly better results (35 percent) were obtained for counts of medium FWD. Agreement between counts of large FWD was the best, likely due to the size of the pieces being observed. A little more than 50 percent of the comparisons were within tolerance for these larger pieces. Interestingly, increasing tolerance levels had a relatively minor effect on count agreement for the large FWD, indicating that wider tolerances will not appreciably improve repeatability. This suggests that there are large differences in counts between crews and/or other types of errors (e.g., data entry) are occurring. There was only one instance where there was disagreement on the existence of a residue pile on the FWD transect segment.

Clearly, the current methods for FWD counts do not provide repeatable measurements. The lack of consistency may be attributed to several causes. Due to the small size of FWD, very slight variations in the establishment of the sampling transect line can result in different FWD pieces being counted. The small size also results in FWD becoming intermingled with organic litter on the forest floor, which makes identification of individual pieces difficult and subjective. Field implementation of FWD sampling protocols also results in some unavoidable trampling of the sample location. This disturbance may alter counts obtained in a remeasurement. Finally, natural phenomena, such as wind and rain storms, may substantially affect the amount of FWD. If such an event occurs between the initial measurement and the subsequent QA assessment, FWD counts are likely to be dissimilar. The above factors likely affect small FWD the most and large FWD to a lesser extent.

Duff, Litter, and Fuelbed Depth

All three depth measurements failed to attain the desired MQO compliance rate (Table 2.3). For duff and litter depths, the lack of repeatability primarily stems from difficulty in determining where the duff/litter breakpoint occurs and the overall depth. A more detailed analysis shows that the deeper duff and litter layers found in the NC and NE regions play some role in measurement repeatability (Table 2.4), as the percentage of data within tolerance is generally lower for these areas in comparison to other regions where the layers are

Table 2.2—Repeatability statistics for FWD attributes based on 385 observations across the U.S.

Variable	Tolerance	MQO compliance rate	Percent within tolerance				Number of times data exceeded tolerance				Records
			@1x	@2x	@3x	@4x	@1x	@2x	@3x	@4x	
Small FWD count	±20 vary	90%	23.6%	42.9%	59.5%	72.2%	294	220	156	107	385
Medium FWD count	±20%	90%	34.5%	44.2%	55.1%	61.6%	252	215	173	148	385
Large FWD count	±20%	90%	56.4%	59.0%	64.7%	67.0%	168	158	136	127	385
Residual piles	No tolerance	90%	99.7%				1				378

Table 2.3—Repeatability statistics for duff, litter, and fuelbed depth based on 1220 observations across the U.S.

Variable	Tolerance	MQO compliance rate	Percent within tolerance				Number of times data exceeded tolerance				Records
			@1x	@2x	@3x	@4x	@1x	@2x	@3x	@4x	
Duff depth	±0.5 in	90%	74.8%	86.2%	92.0%	95.4%	291	160	93	53	1157
Litter depth	±0.5 in	90%	72.6%	86.9%	93.2%	97.6%	317	151	79	28	1157
Fuelbed depth	±20%	90%	38.9%	43.9%	54.8%	59.8%	745	685	552	491	1220

Table 2.4—Repeatability statistics and median depths for duff and litter by FIA region.

Region	Duff depth % w/in tolerance	Median duff depth (in)	Litter depth % w/in tolerance	Median litter depth (in)
NC	69.6%	0.5	55.4%	1.4
NE	68.4%	0.6	41.9%	1.4
PNW	73.0%	0.1	54.0%	0.0
SRS	77.2%	0.1	67.0%	0.3
IW	78.2%	0.1	76.0%	0.5

relatively thin. This suggests that tolerance levels for duff and litter depth might be better specified in relation to observed depth than as a constant value across all conditions.

Among the three depth measurements, fuelbed depth had the lowest percentage (39 percent) of data within the specified tolerance. Fuelbed depth is determined from the top of the duff layer to the top of the CWD/FWD fuel complex and is designed to measure the continuum of fuels. This can produce measurements that vary widely because determination of where these points occur can be unclear. It is particularly difficult to ascertain the top of the CWD/FWD layer due to complexity in defining what constitutes a reasonable continuum of fuels.

Shrub and Herb Fuels

The dead shrub cover and dead herb cover variables exceeded the specified MQO compliance rate, while the remaining variables were notably lower than the desired value (Table 2.5). Live herb height was worse (39 percent) than any of the other variables and height and cover measurements for live vegetation exhibited poorer repeatability than the same measurements for dead vegetation. The percentage of litter cover measurements that were within tolerance (61 percent) was a little less than the values obtained for live vegetation cover.

Often, a thorough understanding of the data collection protocols is necessary to provide an accurate interpretation of repeatability statistics. An example relevant to this analysis would be dead vegetation cover and height observations. When no dead herbs or shrubs are found, a value of zero is recorded for both cover and height. Often, both crews agree no dead herbs or shrubs were present and no ‘measurements’ were actually taken. When zero-valued agreements compose a significant portion of the data, it can result in high levels of repeatability. One approach to evaluating this phenomenon is to re-compute the statistics after removing observations where both crews recorded a zero value.

Table 2.5—Repeatability statistics for shrub and herb fuel attributes based on 433 observations across the U.S.

Variable	Tolerance	MQO compliance rate	Percent within tolerance @1x	@2x	@3x	@4x	Number of times data exceeded tolerance @1x	@2x	@3x	@4x	Records
Live shrub cover	±1 class	85%	67.9%	86.8%	91.7%	95.2%	139	57	36	21	433
Live shrub height	±0.5 ft	90%	57.0%	69.1%	75.5%	82.0%	186	134	106	78	433
Dead shrub cover	±1 class	85%	92.4%	98.6%	99.5%	99.8%	33	6	2	1	433
Dead shrub height	±0.5 ft	90%	63.7%	73.2%	79.2%	83.8%	157	116	90	70	433
Live herb cover	±1 class	85%	66.5%	80.4%	87.1%	90.8%	145	85	56	40	433
Live herb height	±0.2 ft	90%	38.8%	52.2%	61.2%	68.6%	265	207	168	136	433
Dead herb cover	±1 class	85%	86.1%	94.7%	95.8%	96.8%	60	23	18	14	433
Dead herb height	±0.2 ft	90%	53.3%	64.4%	76.0%	80.4%	202	154	104	85	433
Litter cover	±1 class	85%	60.5%	75.3%	79.4%	83.4%	171	107	89	72	433

Roughly 46 percent of the observations for dead shrub cover/height were agreements on the absence of dead shrubs. Similarly, observations of absence of dead herb cover/height comprised nearly 35 percent of all observations. Table 2.6 shows the impact of removing these records and only comparing repeatability when non-zero measurements were taken. There was a precipitous drop in repeatability for the dead vegetation height measurements, indicating the absence of dead vegetation was having a large affect on assessments of height measurement consistency. The repeatability of cover estimates changed only slightly, suggesting the absence of dead vegetation had little overall impact on the results.

A similar analysis for live cover/height (not shown) indicated that the removal of zero-valued observations had a relatively small effect on the percentage of data within tolerance.

Residue Piles

Residue piles are a rare forest attribute compared to that of CWD or FWD, therefore there were few repeated measures of residue piles for QA assessment. Table 2.7 provides the analytical results for the various field measurements associated with residual piles. These results are provided for information purposes only, as no meaningful conclusions regarding repeatability can be drawn due to the small number of observations.

Table 2.6—Repeatability statistics for dead vegetation heights when both crews agree dead vegetation is present.

Variable	Tolerance	Percent within tolerance @1x	@2x	@3x	@4x	Number of times data exceeded tolerance @1x	@2x	@3x	@4x	Records
Dead shrub cover	±1 class	85.8%	97.4%	99.1%	99.6%	33	6	2	1	233
Dead shrub height	±0.5 ft	32.6%	50.2%	61.4%	70.0%	157	116	90	70	233
Dead herb cover	±1 class	78.8%	91.9%	93.6%	95.1%	60	23	18	14	283
Dead herb height	±0.2 ft	28.6%	45.6%	63.3%	70.0%	202	154	104	85	283

Table 2.7—Repeatability statistics for residual pile attributes based on two observations across the U.S.

			Percent within tolerance				Number of times data exceeded tolerance				
Variable	Tolerance	MQO compliance rate	@1x	@2x	@3x	@4x	@1x	@2x	@3x	@4x	Records
Pile shape	No tolerance	90%	50.0%				1				2
Pile azimuth	±10º	90%	100.0%				0				2
Pile length 1	±10%	90%	0.0%	100.0%			2	0			2
Pile length 2	±10%	90%									
Pile width 1	±10%	90%	0.0%	0.0%	0.0%	50.0%	2	2	2	1	2
Pile width 2	±10%	90%	0.0%	100.0%			1	0			1
Pile height 1	±10%	90%	0.0%	0.0%	0.0%	0.0%	2	2	2	2	2
Pile height 2	±10%	90%	0.0%	100.0%			1	0			1
Density	±20%	75%									

Conclusions

For the national DWM indicator, numerous measurement variables had poor repeatability as indicated by MQO attainment results. Experts in field data collection of DWM attributes should evaluate these results and determine what steps can be taken to bring the MQO compliance rates and realized field results closer together. Training procedures and data collection protocols should be examined to ascertain whether changes in these areas might result in more repeatable field measurements. Also, the tolerance and MQO compliance rate specifications should be revisited and potentially revised where warranted. In some cases, the achieved repeatability may be near what can reasonably be accomplished under actual field conditions. For instance, it might not be possible to notably increase agreement between crews regarding species identification due to the various stages of decay encountered on CWD pieces. Finally, other factors, such as field crew experience, measurement instrument calibration, and seasonal variability of environmental and forest conditions, can also influence measurement repeatability.

Specific attention should be given to problem areas already identified in this report. For example, further improvements in CWD diameter measurement repeatability may be achieved by altering the measurement protocol to be more compatible with standing tree measures. Increasing measurement repeatability for FWD may be attained by placing more emphasis on accurate re-establishment of the transect line. Also, the time lapse between measurements should be minimized to avoid actual changes in FWD resource conditions. Changes in tolerance specifications that reflect the relative difficulty of obtaining accurate measurements are suggested for duff and litter depth variables. These and other suggestions for improvement contained within the report should be evaluated, refined as needed, and implemented in the near future.

Section 3: Soil Quality Indicator

Mark H. Hansen
Charles H. Perry
Michael C. Amacher

Introduction

Beginning in the early 1990s, the Forest Health Monitoring program of the U.S. Forest Service developed, tested, and implemented a soil quality monitoring program to address indicators of forest soil health identified by the Montreal Process (O'Neill et al. 2005). The Agricultural Research, Extension, and Education Reform Act (PL 105-185) of 1998 transferred the responsibility for measuring these plots to the Forest Inventory and Analysis (FIA) program beginning with the 2001 field season, and the field protocols have been essentially stable since that time (O'Neill et al. 2005). Complete field protocols for the soil quality indicator are available for review (USDA Forest Service 2007). Additional laboratory protocols are documented by Amacher et al. (2003).

The introductory material at the front of this report identifies and outlines the QA process for FIA's phase 3 (P3) data. Here we present an analysis of the measurement quality objective (MQO) data for the field forest soils variables as well as an analysis of the measurement errors observed on the laboratory analysis of the samples collected on these plots.

Methods

Data Collection

Plot data were collected during the summers of 2000 to 2005. Some minor changes in procedures and attributes observed have been made over this period and these are summarized by O'Neill et al. (2005). We included in this analysis only those attributes that are collected in the most recent phase 3 field guide (USDA Forest Service 2007). Data attributes collected under previous manuals, which have been dropped from the soils collection protocol, are not included in this analysis, and all soil depth measurements that were collected in centimeters have been converted to inches.

The soil measurement and sampling procedures are divided into three parts: soil erosion, soil compaction, and soil chemistry. The soil erosion and soil compaction measurements are made during the field crew visit to the standard FIA plot. Forest floor samples are collected at three subplots (2, 3, and 4) on each plot and two mineral soil samples are collected only at subplot one from the 0-4 inch layer and the other from the 4-8 inch layer.

Table 3.1 lists the data items collected by the soil quality indicator that are analyzed in this report. Details of the methods used to measure the field data items and procedures used to collect the samples are presented in the field guide (USDA Forest Service 2007). The soil chemistry measurements are made under laboratory conditions on samples collected by the field crews during their visit. Local laboratories processed samples from individual regions using standard laboratory procedures (Amacher et al. 2003).

Southern (SRS)
MU Soil Characterization Lab
E 2509 Engineering Bldg East
University of Missouri, Columbia
Columbia, MO 65211

Northern (NC & NE)
Forestry Sciences Lab
1831 Highway 169
Grand Rapids, MN

Western (IW & PNW)
Forestry Sciences Lab
169 East 860 N. 1200 E.
Logan, UT 84321

Blind check remeasurements, described in detail in the Introduction of this report, of the observed percentage cover of bare soil, percent compacted area, and type of compaction on each subplot are not destructive and reflect a second observation taken on the same area. The interpretation of MQO compliance for these data items is similar to that for other data items in this report; they reflect the ability of two different field crews to make the same observation when observing the same resource.

The disruptive/destructive nature of the soil sampling process and measurements of soil thickness, soil texture, and depth to restrictive horizon does not allow the blind checks be taken at exactly the same points as the production measurements. The soils field guide allows the original sampling point to be relocated to any location within 5 feet of the prescribed sampling point if the soil cannot be sampled. Blind checks were taken within this 5-foot circle at a randomly selected undisturbed location. Also, the blind checks were not taken at the same time as the production measurement; blind checks were made during the same growing season but often several weeks after the production measurement. These procedural decisions mean any analysis of the agreement between production and blind check of the five other field data items (thickness of the litter layer, forest floor thickness, soil texture in the 0- to 4-inch layer, soil texture in the 4- to 8-inch layer, and depth to restrictive horizon) and all of the soil chemistry measurements reflect the combined effects of measurement error and variation in soil properties over relatively short periods of time (a few weeks) and space (a few feet). For measurements such as water content, the variation over time or space may be much greater than those for other data items, and these factors must be considered in the interpretation of the results presented here.

Table 3.1—Forest soils data items.

Data items measured in field[a]		
Data item	**Where collected**	**Units**
Percent cover of bare soil	All 4 subplots	Percent
Percent compacted area	All 4 subplots	Percent
Type of compaction - rutted trail	All 4 subplots - if COMPCPCT > 0	Code
Type of compaction - compacted trail	All 4 subplots - if COMPCPCT > 0	Code
Type of compaction - compacted area	All 4 subplots - if COMPCPCT > 0	Code
Type of compaction - other	All 4 subplots - if COMPCPCT > 0	Code
Type of compaction - any	All 4 subplots - if COMPCPCT > 0	Code
Thickness of the Litter Layer	4 points (N, S, E & W) on subplots 2, 3, & 4	Inches
Forest floor thickness	4 points (N, S, E & W) on subplots 2, 3, & 4	Inches
Soil texture in the 0-4 inch layer	Subplot 2	Code
Soil texture in the 4-8 inch layer	Subplot 2	Code
Depth to restrictive horizon	Subplots 2, 3, & 4	Inches
Data items measured in laboratory[b]		
Data item	**Where collected**	**Units**
Total soil sample weight - wet	Forest floor and mineral soil samples	Grams
Total soil sample weight - oven dry	Forest floor and mineral soil samples	Grams
Water content	Forest floor and mineral soil samples	Percent
Coarse fragment [>0.08-in] content	Forest floor and mineral soil samples	Percent
Bulk density	Forest floor and mineral soil samples	g per cc
pH in water	Mineral soil samples	pH
pH in 0.01 M $CaCl_2$	Mineral soil samples	pH
Total carbon	Forest floor and mineral soil samples	Percent
Total inorganic carbon	Mineral soil samples	Percent
Total organic carbon	Mineral soil samples	Percent
Total nitrogen	Forest floor and mineral soil samples	Percent
Extractable phosphorus (Bray 1 meth.)	Mineral soil samples (pH < 6)	mg per kg
Extractable phosphorus (Olsen meth.)	Mineral soil samples (pH > 6)	mg per kg
Exchangeable cations - Na	Mineral soil samples	mg per kg
Exchangeable cations - K	Mineral soil samples	mg per kg
Exchangeable cations - Mg	Mineral soil samples	mg per kg
Exchangeable cations - Ca	Mineral soil samples	mg per kg
Exchangeable cations - Al	Mineral soil samples	mg per kg
Estimated cation exchange capacity	Mineral soil samples	cmol per kg
Exchangeable cations - Mn	Mineral soil samples	mg per kg
Extractable - Fe	Mineral soil samples	mg per kg
Extractable - Ni	Mineral soil samples	mg per kg
Extractable - Cu	Mineral soil samples	mg per kg
Extractable - Zn	Mineral soil samples	mg per kg
Extractable - Cd	Mineral soil samples	mg per kg
Extractable - Pb	Mineral soil samples	mg per kg
Extractable - S	Mineral soil samples	mg per kg

[a] All field data items are only collected on forested portions of FIA plots.
[b] Forest floor and mineral soil samples are only collected on forested portions of FIA plots.

Data Matching

Matching the QA data with the production data was easily accomplished using the plot identification information (to ensure that each QA plot was matched to the appropriate production plot) and the subplot numbers (to ensure that measurements came from the same subplot). The layer identifier was also used in matching the soil chemistry measurement data to ensure that the soil measurements came from the same soil layer, and the litter layer and forest floor thickness were matched by direction (N, S, E, and W) to ensure that they were made at the same location relative to the sampling point. We collected every available soils QA observation from the five FIA regions. A total of 225 plots from 36 states are included in the analysis of blind checks of field forest soils measurements. Eighty-nine of these plots distributed across 25 states included the soil chemistry set of laboratory measurements.

In cases where matched data items did not both contain valid measurements—that is, one crew measured the data item and the other crew did not measure the item or observed and recorded an invalid measurement—the paired observations were considered invalid and not included in the analysis. However, other observations on that same plot were included in the analysis if they met the matching criteria and both measurements were valid observations. Unmatched or invalid measurements could occur for a number of reasons. In some cases, crews did not agree that a specific condition or portion of the plot was forested and would not measure the item on that portion of the plot. In the laboratory measurements there could be equipment failures and other problems that invalidate one of the measurements. The matching process thus matched valid pairs of observations—where both crews successfully measured the same attribute—and ignored any differences due to equipment failures, lost samples, or disagreements that the measurement was appropriate or could be made. Nearly every variable matched greater than 99 percent of the time. The Olsen method (Kuo 1996) of phosphorus testing was one notable exception. Testing of soil phosphorus is dependent upon the pH of the sample. Samples with pH values greater than 6.0 are analyzed with the Olsen method; the other samples are analyzed with the Bray-1 method (Kuo 1996). The observed 32 percent of mismatches occurred when the pH of the QA sample was not in the same pH range as the production sample.

Type of compaction (rutted trail, compacted trail, compacted area, and other) is recorded as a set of binary values (0/1). In this analysis, matched data pairs were those observed on subplots that contain forest land and had an observed percentage compacted area greater than zero. In the analysis of these four data items, we included all subplots where both crews observed forest land and either crew observed a percent compacted area greater than zero, but we did not include pairs where both crews said the compacted area was zero. Type of compaction (any) is a computed binary data item (0/1) indicating the observation of some type of compaction on the sub-plot. The matched data pairs for this item included all subplots where both crews observed forest land.

Data Analysis

Our analysis of forest soils data follows methods similar to those used in the other sections of this report. Each field measurement has an MQO consisting of a measurement tolerance and a required percentage of observations that need to be within the measurement tolerance. Please review the introductory material at the front of this report for additional details. MQOs have not been established for the laboratory measurements due the unknown temporal and spatial variability of these measurements. To set a baseline and report on the laboratory measurements, a tolerance of ±25 percent of the mean of each pair of production and QA observations was arbitrarily set for all laboratory measurements. We did not specify a target percentage compliance. For each data item, a distribution of differences was generated by pooling data from all available QA plots nationwide. To assess the potential repeatability of lab measurements, the percentage of observations falling within the range of the specified tolerance levels was determined and compared to the MQO for each variable. Additionally, the percentage of observations within 2, 3, or 4 times the original tolerance level was determined for evaluation of tolerance adjustment effects. No additional information is provided for those attributes with zero tolerance.

Results and Discussion

Field Measurements

Of the 12 field measurements, 10 have observed compliance rates that meet or exceed the MQO (Table 3.2), indicating that measurement standards initially set in development of the forest soils are realistic and the field soils measurements are generally repeatable. The soil texture observations of both layers met established tolerances and MQOs while the other eight variables that were within tolerances generally did so to a much greater degree than expected.

Both the percent cover of bare soil and percent compacted area observations classify subplots into 5-percent classes—with two extra classes at the bottom end of the scale for absent and trace—with tolerances of 10 percent and 15 percent respectively. If the tolerance for these two items was changed to 5 percent, the observed compliance rate would be 82.2 percent for percentage cover bare soil and 88.8 percent for compacted area. These high compliance rates are due to the large portions of the sample where there is little or no bare soil or compacted area on the subplot. The majority of the discrepancies that exceed the tolerances are plots with mid-range observations (20-80 percent).

Consistent identification of subplots where compaction has occurred as well as the type of compaction present appears to be possible using the current field procedures. Only the identification of compacted trails did not meet the 75 percent MQO. Many of the disagreements for this variable occurred on subplots where one of the crews identified more than one type of compaction. There could be issues with the minimum width and length of compacted trail that require identification. The other types of compaction are well defined in the field guide. Ruts must be at least 2 inches deep in the mineral soil or 6 inches deep from the undisturbed litter surface to be tallied. Examples in the field guide of compacted areas include the junctions of skid trails and landings. A compacted trail, by contrast, is defined simply as “usually the result of many passes of heavy machinery, vehicles, or large animals”, and prototypical sizes are not provided.

The measurements of thickness of the litter layer and the forest floor (measured to one-tenth of an inch) have MQOs of ±2 inches 90 percent of the time, and the tolerance was achieved 98.5 and 94.3 percent, respectively. The MQOs for these two items should be reconsidered. It is counterintuitive to record a measurement to 1/10 of an inch and specify a tolerance of ±2 inches. A tolerance of 0.5 inches with an MQO of 75 percent is approximately what is being achieved. Assuming a tolerance of 0.5 inches, 84.2 percent of the litter layer and 76.2 percent of the forest floor measurements were within tolerance.

Table 3.2—Repeatability statistics for forest soils measurements.

Data items	Tolerance	MQO%	% within tolerance	Number over tolerance	Total number of paired obs.
Field measurements					
Percent cover of bare soil	±10%	75%	89.2%	88	813
Percent compacted area	±15%	75%	94.0%	49	813
Type of compaction - rutted trail	No tolerance	75%	88.1%	32	269
Type of compaction - compacted trail	No tolerance	75%	69.1%	83	269
Type of compaction - compacted area	No tolerance	75%	84.0%	43	269
Type of compaction - other	No tolerance	75%	97.0%	8	269
Type of compaction - any	No tolerance	75%	89.8%	83	813
Thickness of the Litter Layer	±2 in	90%	98.5%	40	2,708
Forest floor thickness	±2 in	90%	94.3%	155	2,708
Soil texture in the 0-4 inch layer	±1 class	80%	82.7%	32	185
Soil texture in the 4-8 inch layer	±1 class	80%	80.5%	36	185
Depth to restrictive horizon	±6 in	90%	65.3%	193	556
Lab measurements					
Total soil sample weight - wet	±25%	Not set	42.0%	213	367
Total soil sample weight - oven dry	±25%	Not set	41.7%	214	367
Water content	±25%	Not set	29.7%	258	367
Coarse fragment [>0.08-in] content	±25%	Not set	34.5%	95	145
Bulk density	±25%	Not set	72.3%	33	119
pH in water	±25%	Not set	99.3%	1	144
pH in 0.01 M $CaCl_2$	±25%	Not set	98.6%	2	143
Total carbon	±25%	Not set	53.7%	168	363
Total inorganic carbon	±25%	Not set	89.8%	9	88
Total organic carbon	±25%	Not set	53.7%	168	363
Total nitrogen	±25%	Not set	58.7%	150	363
Extractable phosphorus (Bray 1 meth.)	±25%	Not set	39.5%	75	124
Extractable phosphorus (Olsen meth.)	±25%	Not set	24.3%	28	37
Exchangeable cations - Na	±25%	Not set	37.5%	90	144
Exchangeable cations - K	±25%	Not set	43.1%	82	144
Exchangeable cations - Mg	±25%	Not set	53.5%	67	144
Exchangeable cations - Ca	±25%	Not set	46.5%	77	144
Exchangeable cations - Al	±25%	Not set	38.2%	89	144
Estimated cation exchange capacity	±25%	Not set	65.3%	50	144
Exchangeable cations - Mn	±25%	Not set	33.3%	96	144
Extractable - Fe	±25%	Not set	35.9%	84	131
Extractable - Ni	±25%	Not set	54.2%	60	131
Extractable - Cu	±25%	Not set	71.0%	38	131
Extractable - Zn	±25%	Not set	35.9%	84	131
Extractable - Cd	±25%	Not set	61.8%	50	131
Extractable - Pb	±25%	Not set	33.6%	87	131
Extractable - S	±25%	Not set	50.0%	72	144

The depth to restrictive soil horizon is the only other field measurement that does not meet the measurement quality objectives. This measurement involves the use of a tile probe at five locations to obtain the average depth to a restrictive horizon. In 193 of the 556 times both crews made these measurements, the difference between the two averages was greater than 6 inches. Two possible explanations include the high local spatial variation of stones in soils and different uses of the tile probe. The probe itself is a simple metal rod, and it may be driven into the soil with varying amounts of force. Different crews may make the measurement differently since there is no means of gauging the amount of pressure being applied. Training materials suggest that field crews should “stop pushing when you reach a resistant layer.” The definition of resistant layer may be sufficiently ambiguous to make the procedure less repeatable. The procedures and training for this variable should be re-evaluated.

Laboratory measurements

MQOs have not been set for the soil chemistry measurements that are made at the three regional soil labs. Using a tolerance of ±25 percent, there was a wide range in observed measurements within tolerance, with extractable phosphorus (Olsen method) and water content having less than 30 percent of the samples within tolerance and the two pH measurements having over 98 percent of the samples within tolerance. It is important to emphasize again the destructive nature of soil sampling. This combined with spatial and temporal variability in the soil properties that are being measured explains most of the disagreements observed here. By contrast, measurement errors are not a large component of the observed differences between QA and production lab measurements for most of the data items. Many of these lab measurements are reported to three or four significant digits indicating very low measurement error.

Rather than a detailed discussion of each measurement item, we present scatter-plots of the paired measurements in Figures 3.1 through 3.27 in Appendix A, beginning on page 31. These show the distribution of the laboratory analyses of the production and blind checks.

Appendix A. Plots of Paired Measurements for Soils Lab Measurements

Figure 3.1—Scatterplot of soil sample wet weight paired measurements.

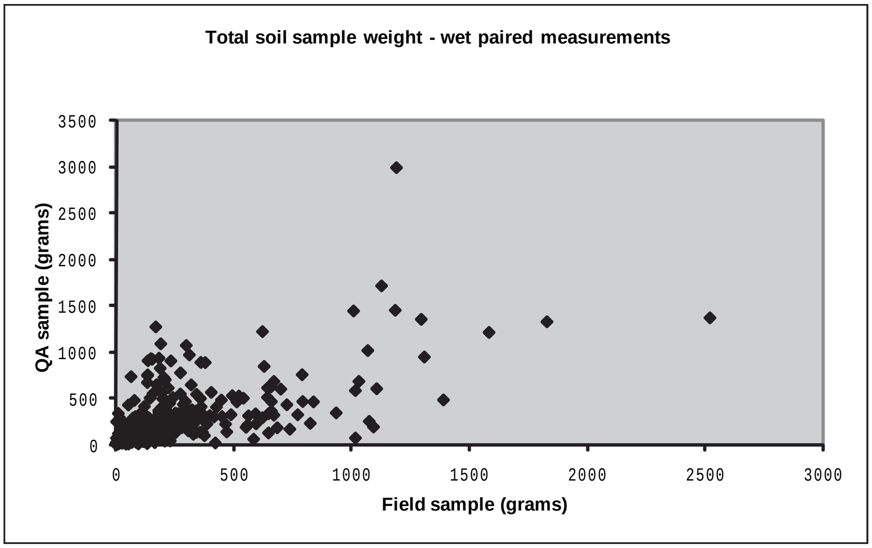


Figure 3.2—Scatterplot of soil sample dry weight paired measurements.

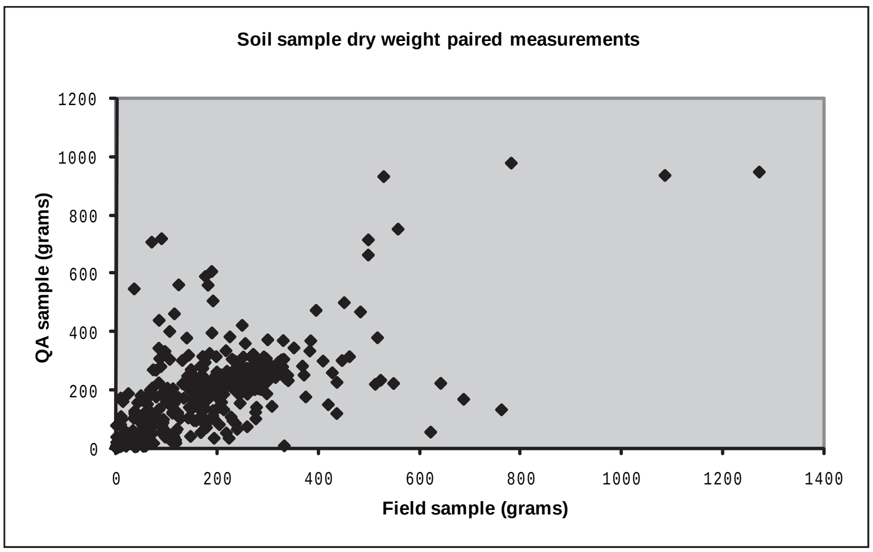


Figure 3.3—Scatterplot of coarse fragment content paired measurements.

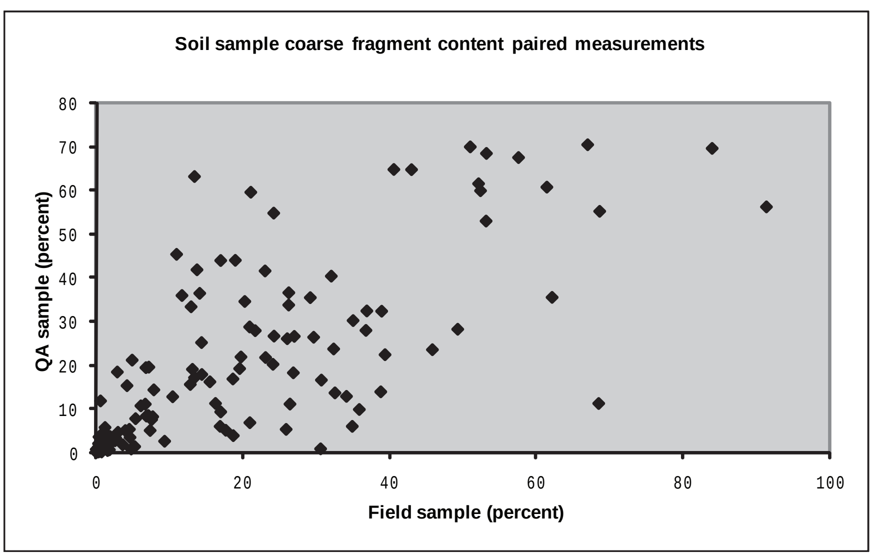


Figure 3.4—Scatterplot of water content paired measurements.

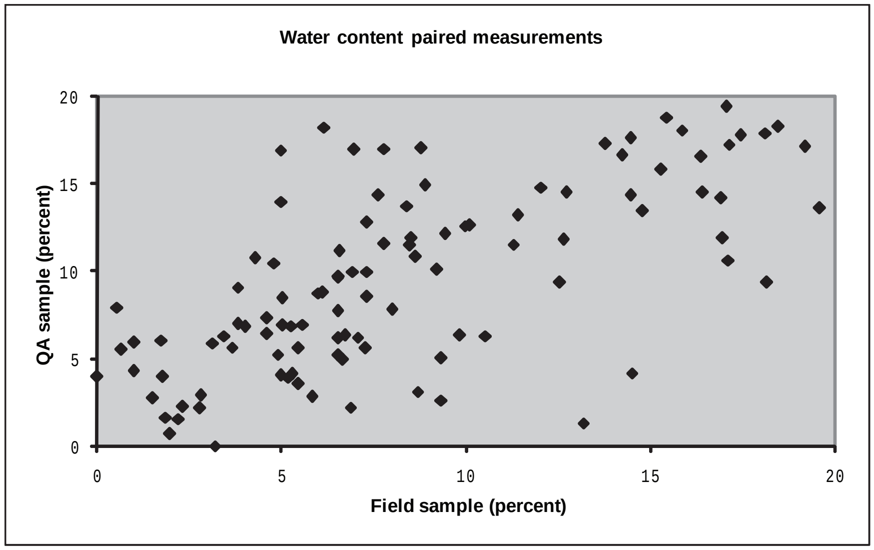


Figure 3.5—Scatterplot of bulk density paired measurements.

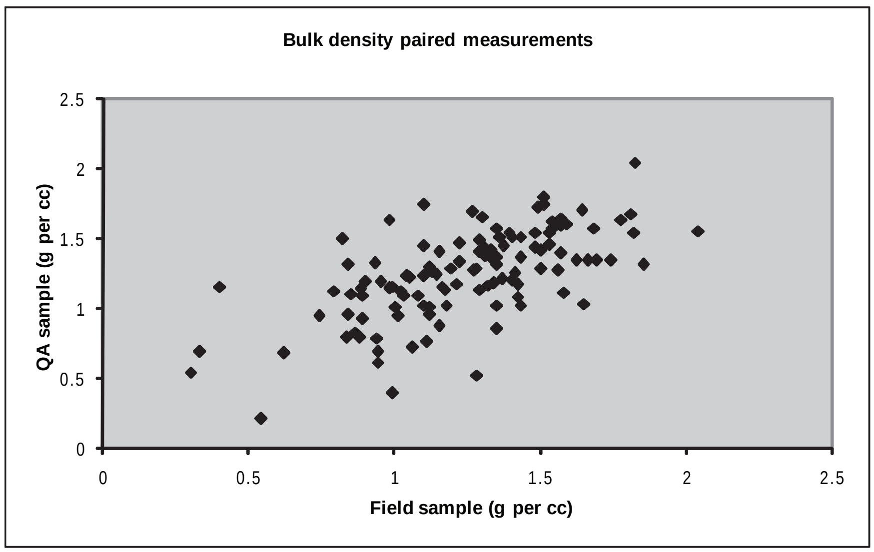


Figure 3.6—Scatterplot of soil pH (H_2O) paired measurements.

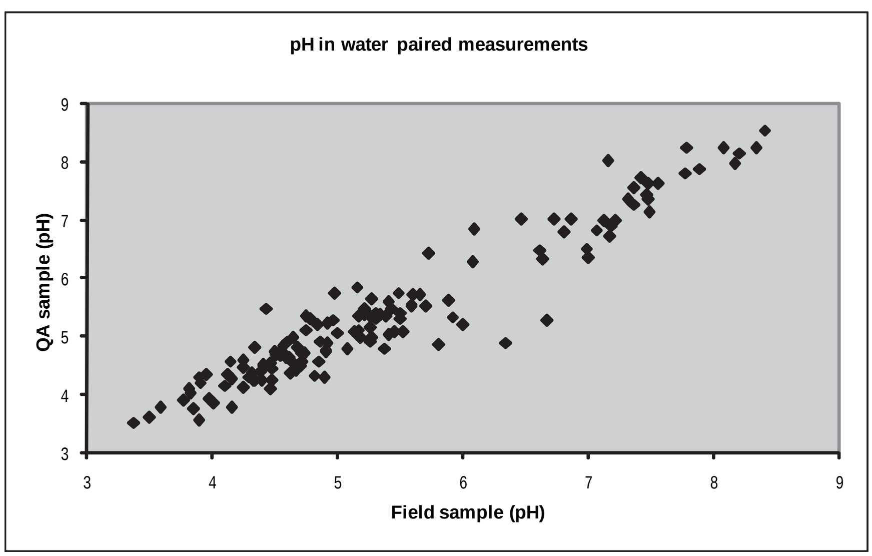


Figure 3.7—Scatterplot of soil pH ($CaCl_2$) paired measurements.

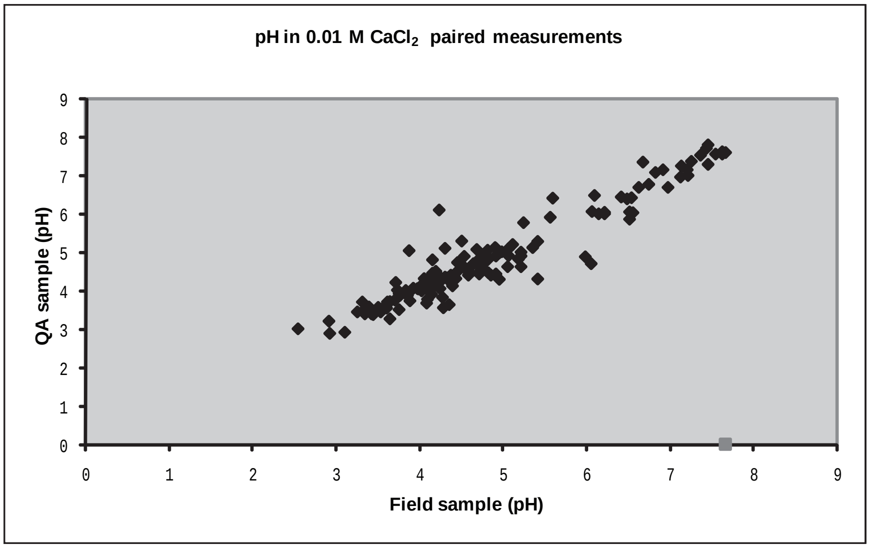


Figure 3.8—Scatterplot of total C paired measurements.

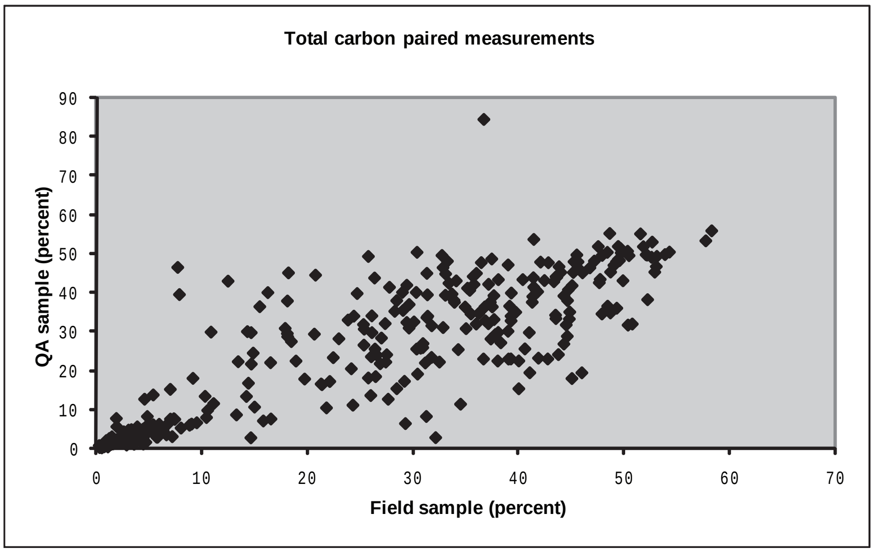


Figure 3.9—Scatterplot of inorganic C paired measurements.

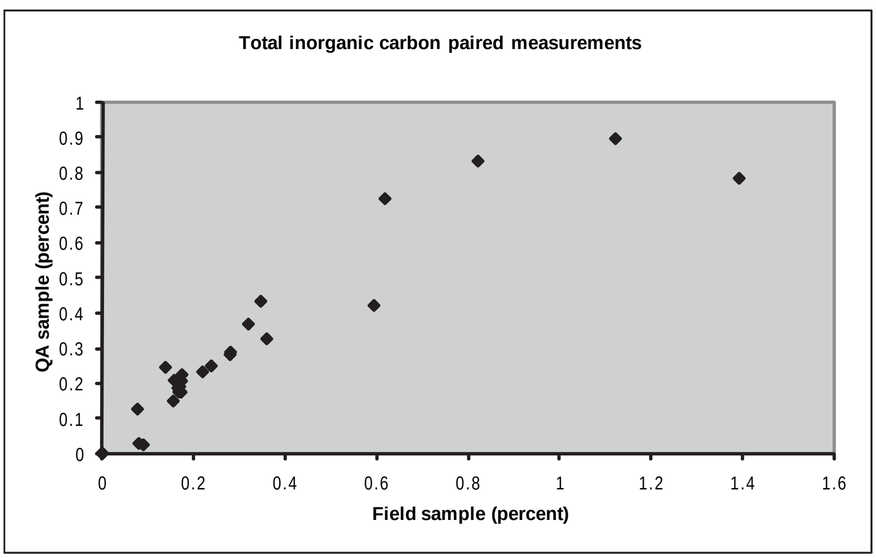


Figure 3.10—Scatterplot of organic C paired measurements.

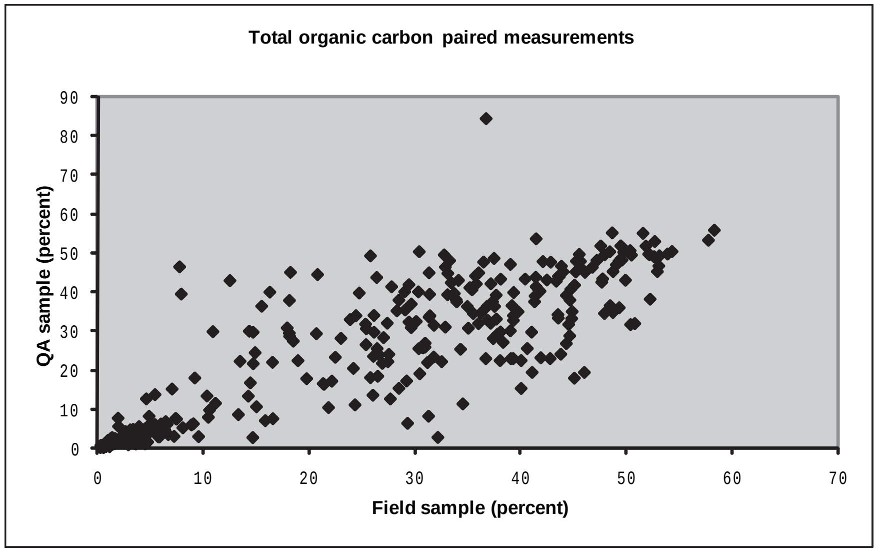


Figure 3.11—Scatterplot of total N paired measurements.

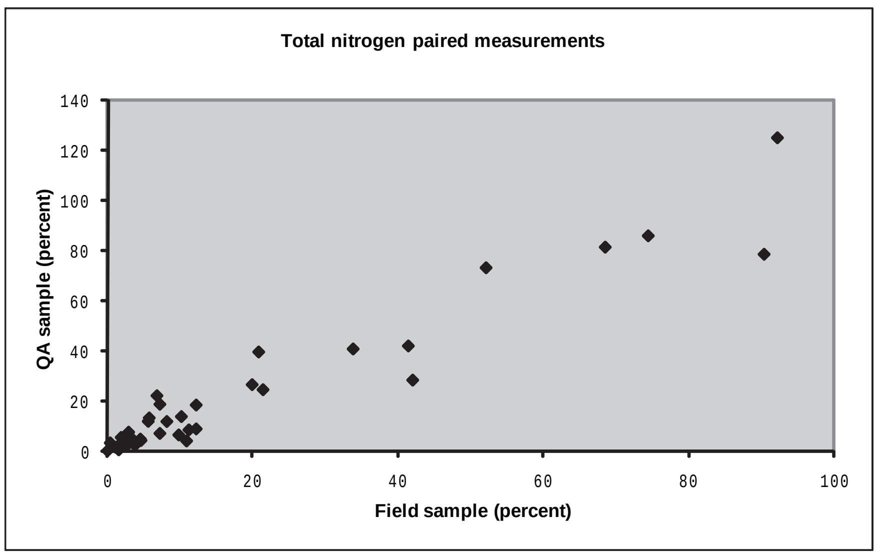
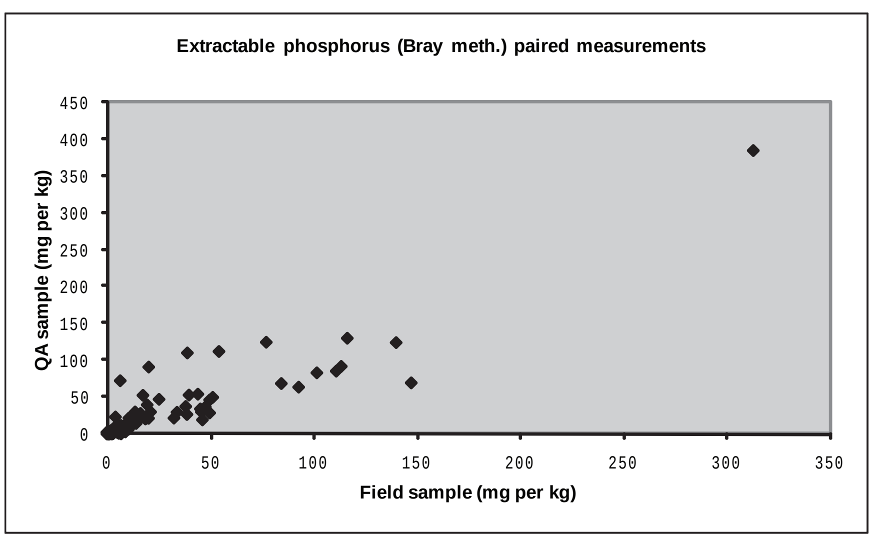


Figure 3.12—Scatterplot of extractable P (Bray 1 method) paired measurements.



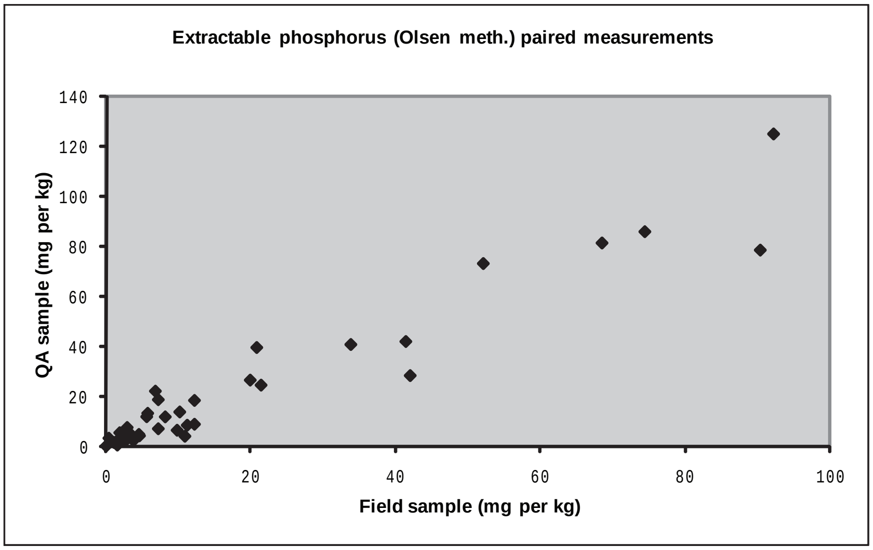


Figure 3.13—Scatterplot of extractable P (Olsen method) paired measurements.

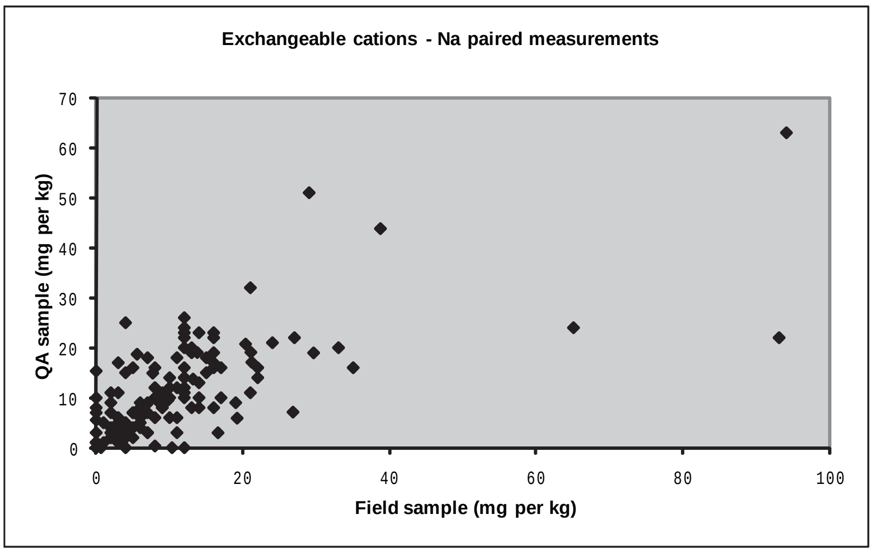


Figure 3.14—Scatterplot of extractable Na paired measurements.

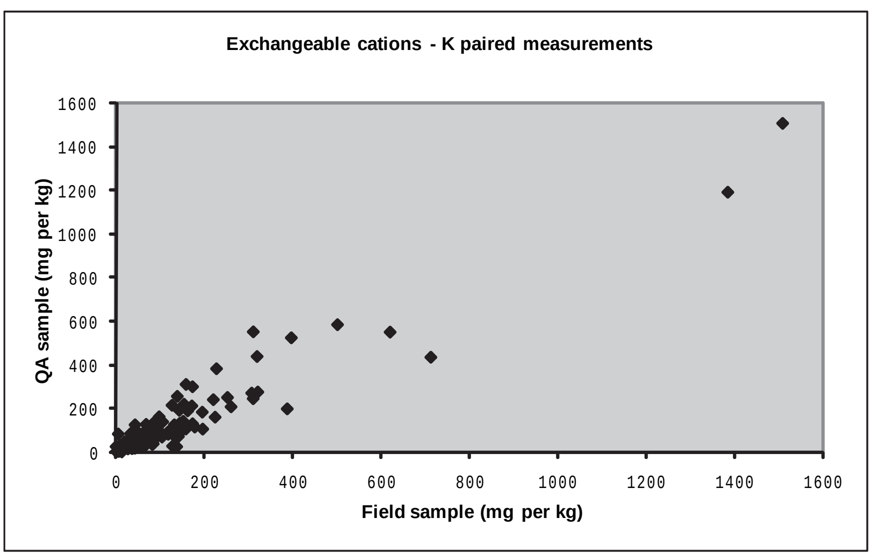


Figure 3.15—Scatterplot of extractable K paired measurements.

Figure 3.16—Scatterplot of extractable Mg paired measurements.

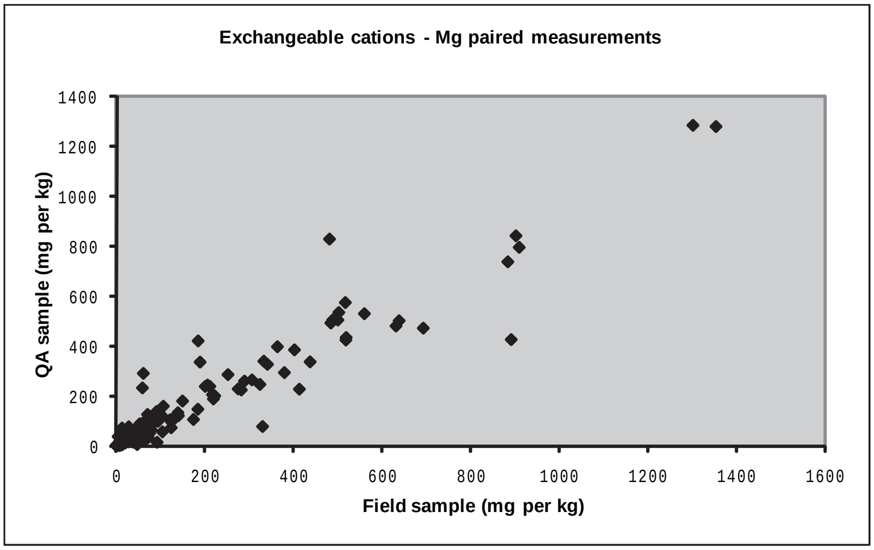


Figure 3.17—Scatterplot of extractable Ca paired measurements.

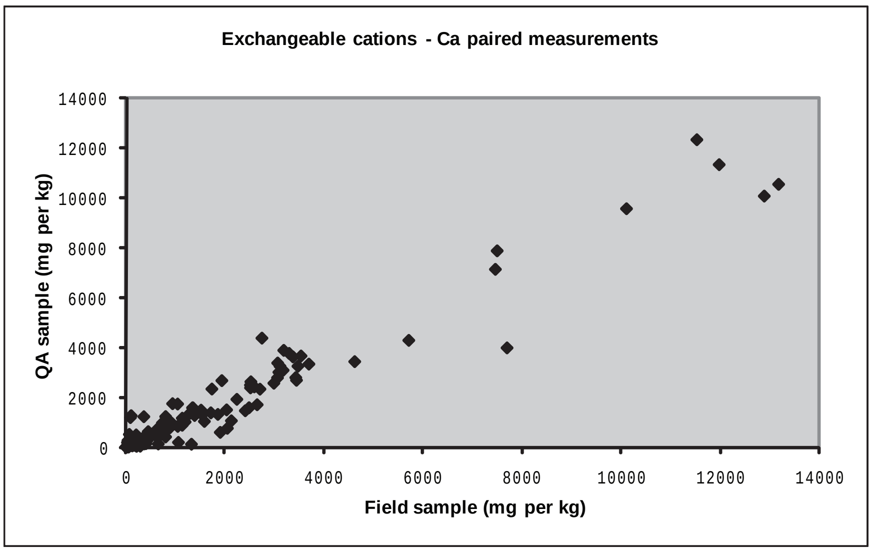


Figure 3.18—Scatterplot of extractable Al paired measurements.

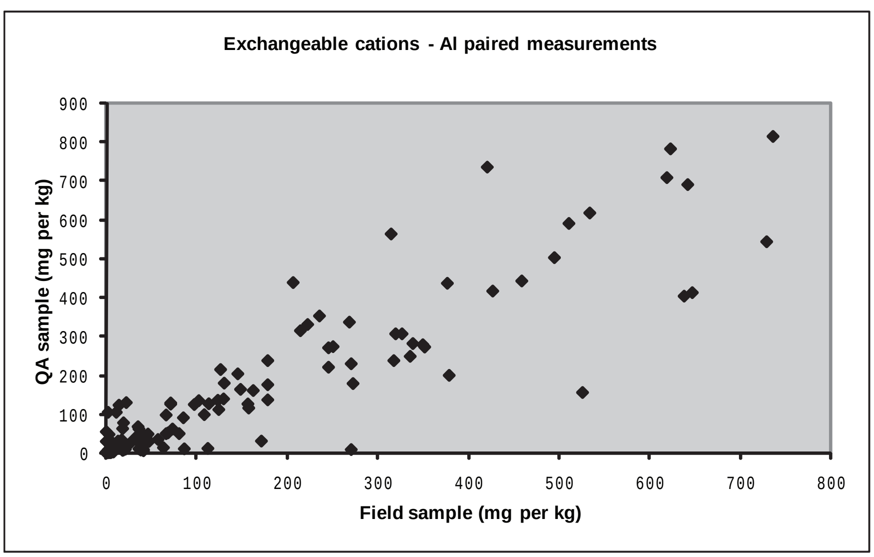


Figure 3.19—Scatterplot of extractable Mn paired measurements.

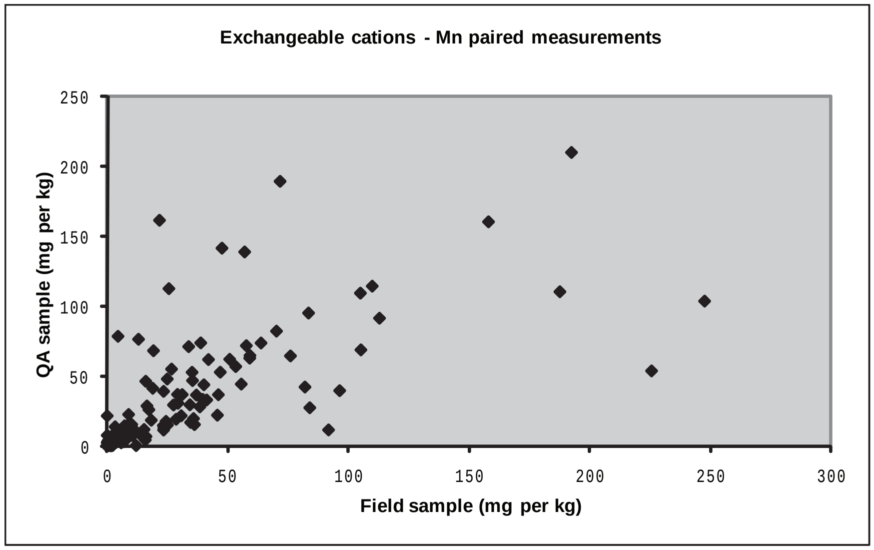


Figure 3.20—Scatterplot of effective cation exchange capacity (ECEC) paired measurements.

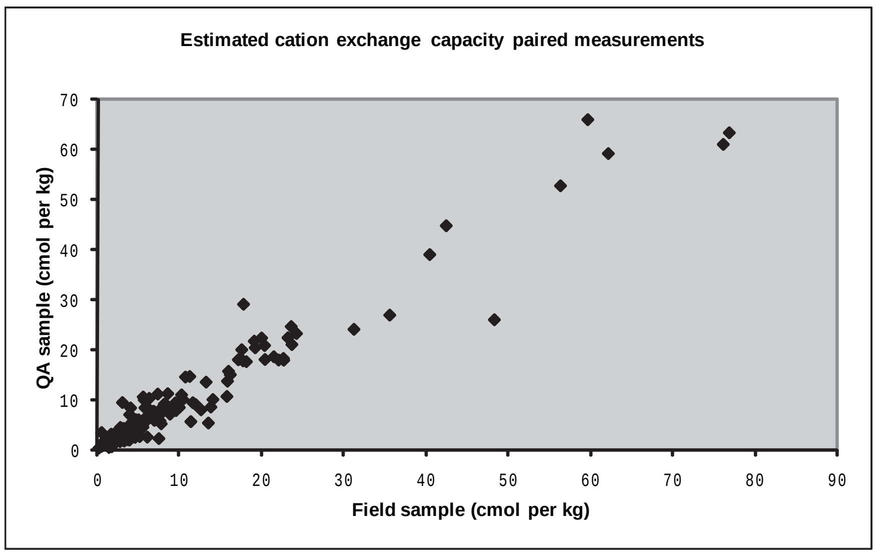


Figure 3.21—Scatterplot of extractable Fe paired measurements.

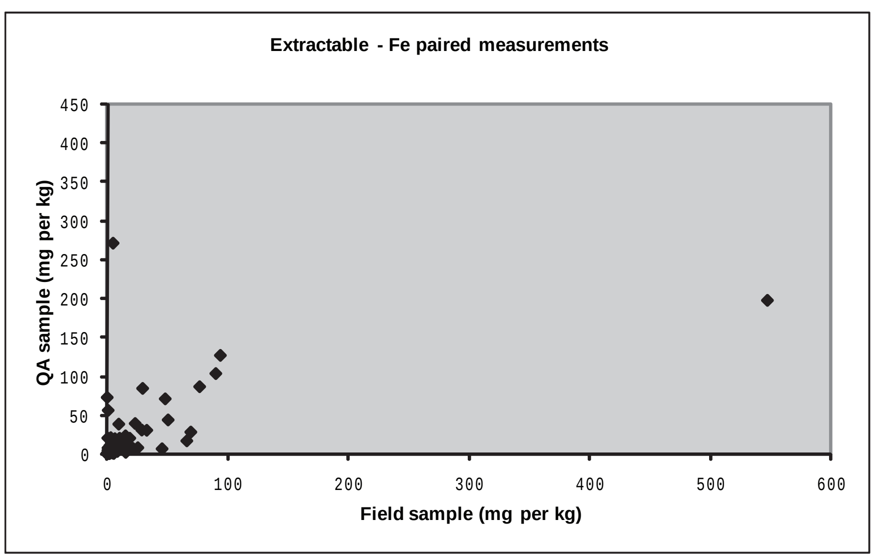


Figure 3.22—Scatterplot of extractable Ni paired measurements.

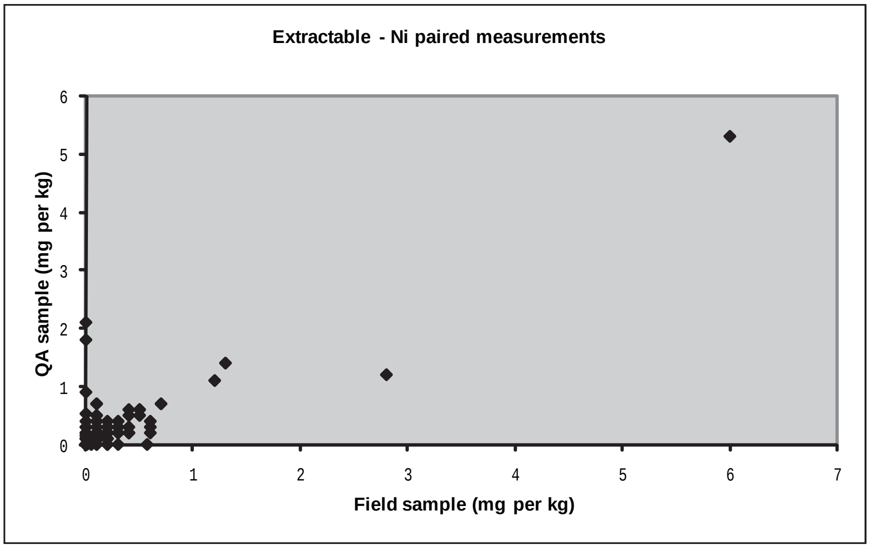


Figure 3.23—Scatterplot of extractable Cu paired measurements.

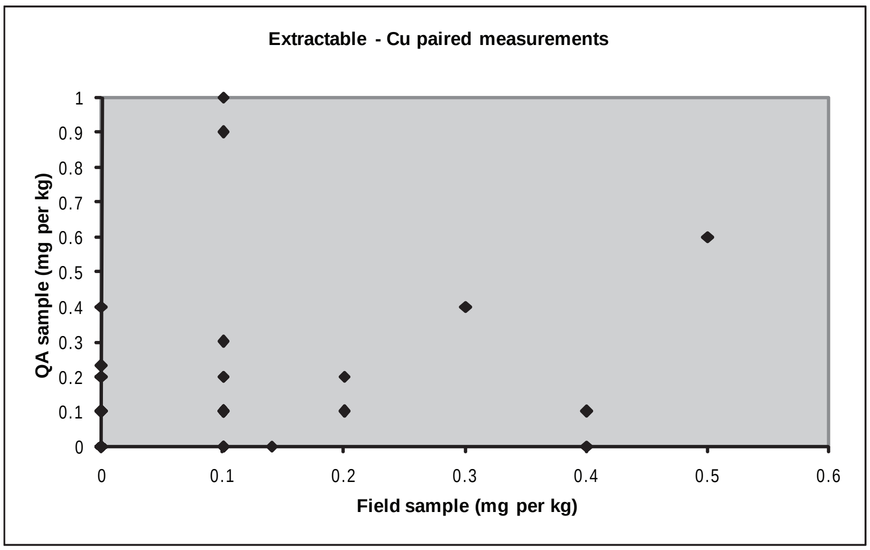


Figure 3.24—Scatterplot of extractable Zn paired measurements.

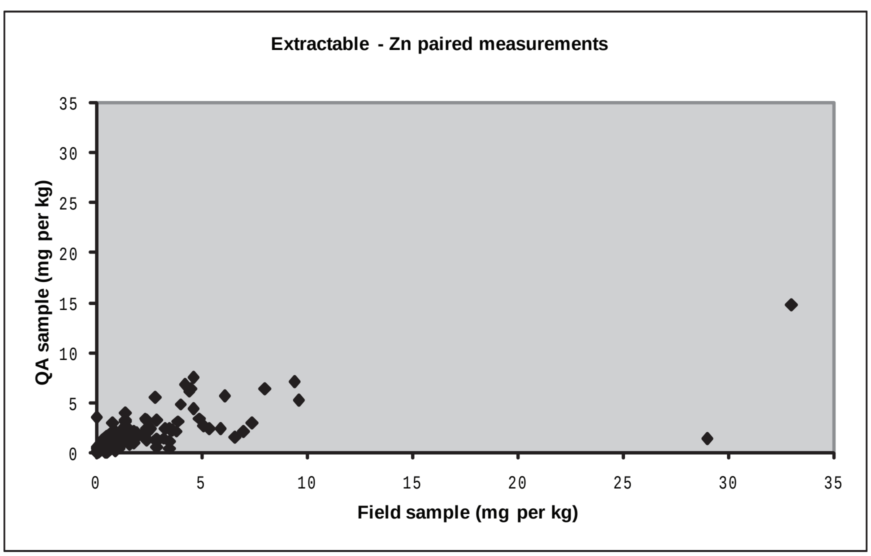


Figure 3.25—Scatterplot of extractable Cd paired measurements.

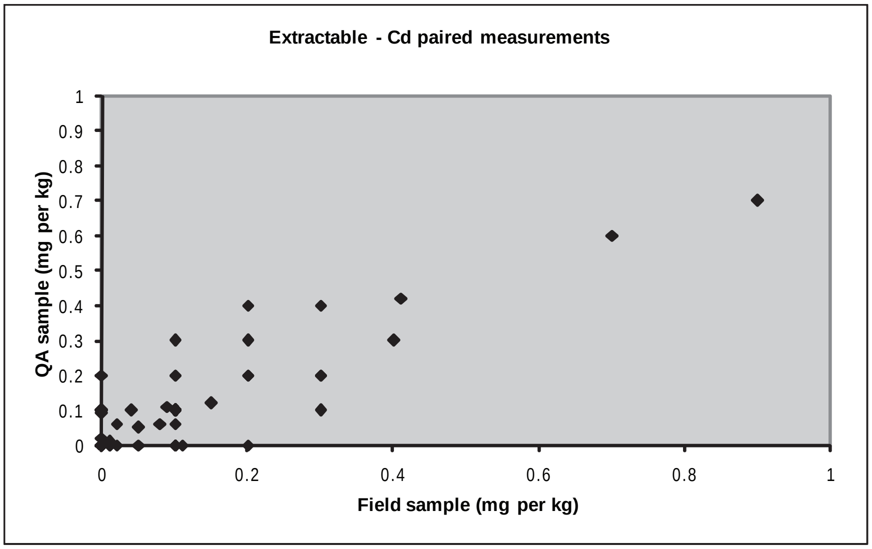


Figure 3.26—Scatterplot of extractable Pb paired measurements.

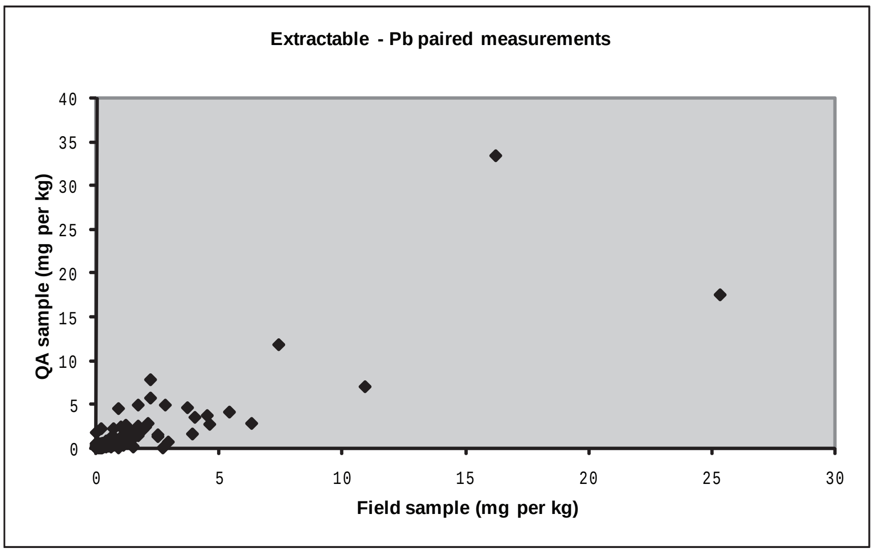
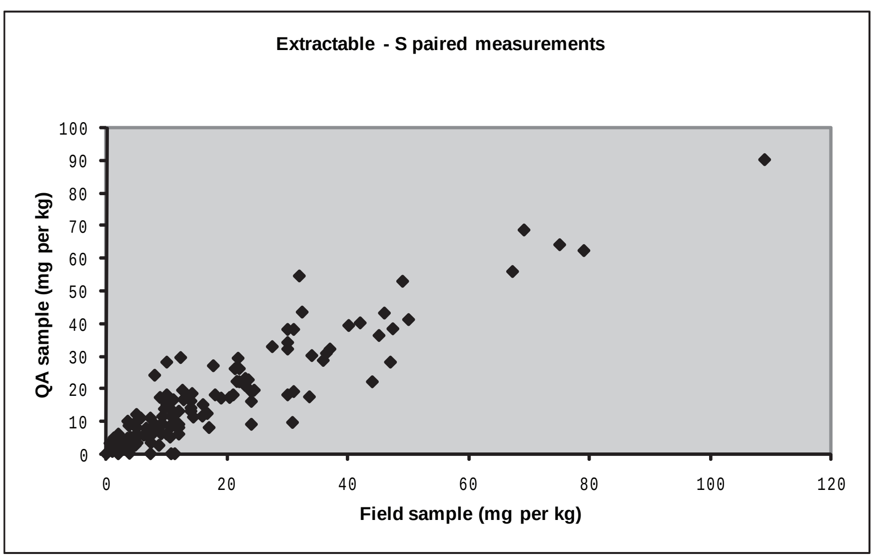


Figure 3.27—Scatterplot of extractable S paired measurements.



Section 4: Lichen Indicator

Paul L. Patterson
Susan Will-Wolf
Marie T. Trest

Introduction

Lichens are very responsive to environmental stressors in forests, including changes in forest structure, air quality, and climate. Each lichen species on a plot is an indicator of how lichen communities respond to ecological conditions. Individual lichen species occur erratically and even common species are often absent from plots with suitable habitat. The combined response of all species is a better indicator of ecological conditions than responses of any one or a few lichen species.

FIA is interested in developing ecological gradients; the main gradients of interest are for air pollution and climate. Plot values for these gradients are assigned based on lichen species data. Lichen data are evaluated to assess spatial patterns and trends in ecological conditions of the plots. For details on how gradients are developed and analysis of lichen data see Will-Wolf (in press).

The lichen indicator field procedures (USDA Forest Service 2007) require the production crew to collect samples from as many lichen species on a plot as possible in the allowed sample time and assign an abundance class to each species. The allowed sample time is variable; the instructions (found in section 10.1.4 of the phase 3 (P3) field guide; USDA Forest Service 2007) are: "Sampling continues for a maximum of two hours or until 10 minutes elapse with no additional species recorded. At least 45 minutes in the Northwest, North Central, South and West Coast including Alaska, and 30 minutes in the Intermountain West, must be spent searching the plot, even if few lichens are present." The production crew sends all collected samples to a regional lichen identification specialist to identify lichen species.

Lichen indicator quality analysis (QA) data and the structure of the management quality objective (MQO) are different from the standard FIA model of blind checks by a QA crew. A regional lichen QA specialist is considered an expert standard and provides measurements of variables. The lichen QA specialist visits the field plot and follows the same protocols as the production crew. A lichen QA specialist can be either a lichen identification specialist trained in FIA field protocols or a production crew member certified for lichens QA. Crew members that regularly achieve very high QA scores are candidates for certified lichen QA specialists.

The sampling protocol for the lichen indicator has been stable since 1994 and QA data have been collected since 1999.

As already stated, another difference from the standard FIA QA model is that lichen indicator data are from a timed sample rather than a complete inventory, and lichen species identification does not occur in the field. These differences have implications for how the MQO standards for the lichen variables are stated and evaluated.

Lichen data collected in the field can be divided into two sets. The first set contains three variables. Two of the variables, collection number and abundance score, are written on the packets containing the samples that are sent to the lichen specialist for species identification. The third variable, number of species, is determined after the lichen identification specialist identifies the species contained in the field samples. The second set of variables consists of all the remaining variables, which we will refer to as field-entered variables.

The collection number and abundance score (sections 10.3.8 and 10.3.9 in the lichen chapter of the P3 field guide; USDA Forest Service 2007) are recorded only on field sample packets. Collection numbers are unique to each field collection, so independent remeasurement of collection number is neither possible nor appropriate. The only logical MQO compliance factor is whether or not collection numbers were recorded. Currently there is no provision for entering into the database whether or not collection numbers were provided, so it would be difficult now to evaluate the MQO compliance rate and it is not considered further here.

What is important are the species of the lichen and the abundance score of each species. Since the species are not identified by the crew, the MQO is based on compliance rate for number of species found, stated as a minimum standard for 'successful performance of sample procurement procedure' (item 8 in section 10.2 in the lichen chapter of the P3 field guide; USDA Forest Service 2007). As will be explained in greater detail in the next section, what is of interest for MQO compliance is the number of lichen species found by the production crew in relation to the number of lichen species found by the lichen QA specialist, rather than the names of the species found. The current MQO compliance rate for abundance score only applies when the same species are tallied by both the production crew and lichen QA specialist, and the relevant comparison is between the final abundance score for each identified species as calculated by the lichen identification specialist, not the scores recorded in the field. Preliminary studies by lichen identification specialists have been conducted on how to evaluate the repeatability of abundance score. Further work is needed to determine the best MQO compliance rate that can be measured, given the way the lichen data are collected.

The field-entered lichen variables can be divided into two subsets. The first subset is the variables that identify the plot, i.e., P3 hex number, FIA plot number, state and county, and document the plot visit, i.e., date, etc. (variables 10.3.1 through 10.3.15 in the Lichen chapter of the P3 field guide; USDA Forest Service 2007). Except for variable 10.3.13, these variables are not of interest for this study. Variable 10.3.13 indicates whether the plot was searched for lichens and whether, if searched, any lichens were found. The values of this variable in the FIA database must completely match lichen lab data records (100 percent MQO compliance). Any mismatches are 100 percent resolved before lichen lab data can be successfully entered into the database. The incidence of such errors has been rare and has not been formally assessed.

The second subset (variables 10.3.16 through 10.3.24 in the Lichen chapter of the P3 field guide; USDA Forest Service 2007) are secondary variables that are intended for use by a lichen data analyst to understand any anomalies in the lichen species and abundance variables. On most QA plots, the lichen QA specialist has not collected these data electronically and the variables will not be analyzed in this study.

Summarizing, the only QA analysis that will be done in this study will be of the variable number of species. The number of species is determined after species identification of the field samples by a lichen identification specialist.

MQO for Number of Lichen Species

Studies of repeatability and reliability for the lichen indicator (McCune et al. 1997; and regional gradient models such as Jovan and McCune 2005, 2006; Geiser and Neitlich 2006) support establishment of the minimum QA standard. If the number of species the production crew finds is at least 65 percent of the number of species found by a lichen QA specialist, then repeat samples on the same plot reliably result in plot scores on gradients of interest that fall within the same 10 percent of the full range (length) of the original gradient. As long as the minimum QA standard of 65 percent is met, deviations of plot scores for repeat samples that are on average greater than 10 percent can be considered evidence of change in the relevant ecological condition. Deviations that are on average less than 10 percent are considered to be due to plot-level sampling error. If production crew samples on average achieve much higher than the minimum QA standard of 65 percent, it is likely that deviations due to sampling error might be smaller, thus smaller changes in response to gradients of interest might be detected.

Based on these studies, 65 percent of the number of species a QA lichen specialist finds is the minimum tolerance for lichen indicator lab data; the percentage, computed as 100 ∗ [(Number of species production crew) / (Number of species QA specialist)], is the QA score for a plot. Tolerance is expressed as a minimum QA score of 65 percent rather than a range because any higher percentage improves the reliability of the data; QA scores can be (and are, see Figs. 4.5-4.7) above 100 percent. The MQO compliance rate for the lichen indicator is that the minimum tolerance be achieved for 90 percent of the plots sampled for QA.

Methods

To evaluate each region’s performance, FIA established a goal of at least 10 QA plots per year in each region where lichen data are being collected. QA sampling intensity for lichens has varied by years and by regions. We analyzed QA data for PNW, IW, and NE FIA regions. There was only one QA plot for SRS, so this region is not included in the analysis of lichen indicator QA data. The number of QA plots by year and by region is given in Table 4.1. The total number of production plots by year and by region is also in Table 4.1. The datum for each plot consists of a QA score, where:

$$\text{QA score} = 100 * (\text{Number of Species production crew}) / (\text{Number of Species QA speciaalist})$$

Table 4.1—Number of lichen QA and production plots by year and region.

	IW		NE		PNW	
Year	QA	Production	QA	Production	QA	Production
1999	3	219	4	159	4	112
2000	3	45	9	257	10	165
2001	10	141	10	213	9	165
2002			11	226		
2003	12	185	10	234	12	165
2004	9	149	5	87		
2005	10	158	2	28		

Results and Discussion

Table 4.2 presents a summary evaluation of QA scores for the combined data from the three regions and for each region individually across all years. Figure 4.1 presents the frequency distribution of the QA scores for all the regions. Figures 4.2-4.4 present the frequency distributions of QA scores for each region separately. These frequency distributions facilitate evaluating how close each region is to achieving threshold of 65 percent QA scores. While the left three bars contain plots that did not meet the threshold, the relative proportion of plots in these three bars indicates the degree of work needed to bring the QA scores up to the 65 percent threshold. For instance, if all the plots that are below the threshold are in the 55-64.9 percent group, less improvement is needed to meet the 65 percent threshold than if all are in the 0-44.9 percent group. Figures 4.5-4.7 present scatter plots of QA scores by year for each region separately, to facilitate evaluation of differences between years and whether they relate to sample size in a year. Figures 4.8-4.10 present scatter plots of QA scores for all years for each region separately, by number of species for QA specialist. This facilitates evaluation of differences between plots with few species versus many species. Note that Table 4.2 contains the compliance rate for each region, i.e., the percentage of the plots that are to the right of the vertical line in Figures 4.5-4.10.

We are not currently meeting our MQO of 90 percent of QA plots meeting minimum tolerance (Table 4.2) in any region. Frequency diagrams in Figs. 4.1-4.4 indicate that overall we are trending in the right direction, and the IW and PNW regions are close. If we could boost crew achievement to move all QA scores below 65 percent up one category, it appears we would meet the MQO for QA in those two regions. A large proportion of QA scores in those two regions are well above minimum tolerance, which is an excellent outcome.

Table 4.2—The percentage of QA scores which met the minimum tolerance that the number of species found by the production crew is at least 65 percent of the number of species the lichen QA specialist found.

Area	% MQO compliant	No. of QA scores
Combined	65	133
IW	74	47
NE	51	51
PNW	74	35

Scatter plots of QA score by year (Figs. 4.5-4.7) show there were peak performance years in each region, but they show no evidence that performance differs markedly in years with few versus many QA scores. Relatively few QA scores in 1999 and 2000 are a consequence of difficulty with logistics of QA sampling in those early years.

Scatter plots of QA score by number of species found by the QA specialist (Figs. 4.8-4.10) do not suggest any strong differences in proportion that met tolerance for very species-rich versus very sparse plots, suggesting training for how to search for lichens is equivalent for these very different kinds of plots.

We have not evaluated QA scores by individual or by years of crew experience. It appears from qualitative examination of certification scores and QA scores over the years that some individuals do perform better than others. It also appears from qualitative examination of QA scores over the years that many individuals perform better in their second and later years. Retention of well performing crew members for multiple years would likely boost QA scores in general. Some preselection of crew members for collection of lichens based on expressed interest and/or other criteria to be developed might also be considered; this might also boost QA scores in general.

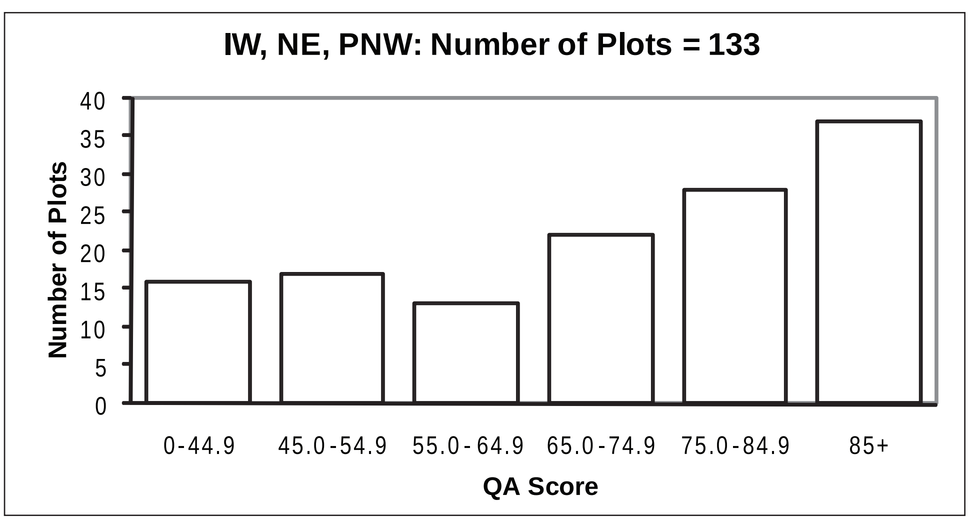


Figure 4.1—Frequency distribution of the QA scores for all regions. QA score = 100 * (Number of Species production crew)/(Number of Species QA specialist).

Figure 4.2—Frequency distribution of the QA scores for Interior West region. See Figure 4.1 caption for definition of QA score.

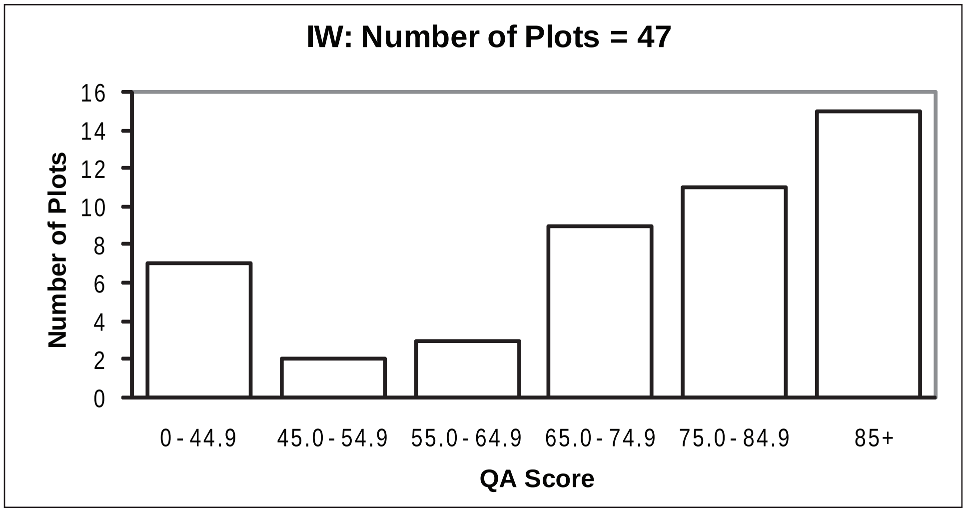


Figure 4.3—Frequency distribution of the QA scores for Northeast region. See Figure 4.1 caption for definition of QA score.

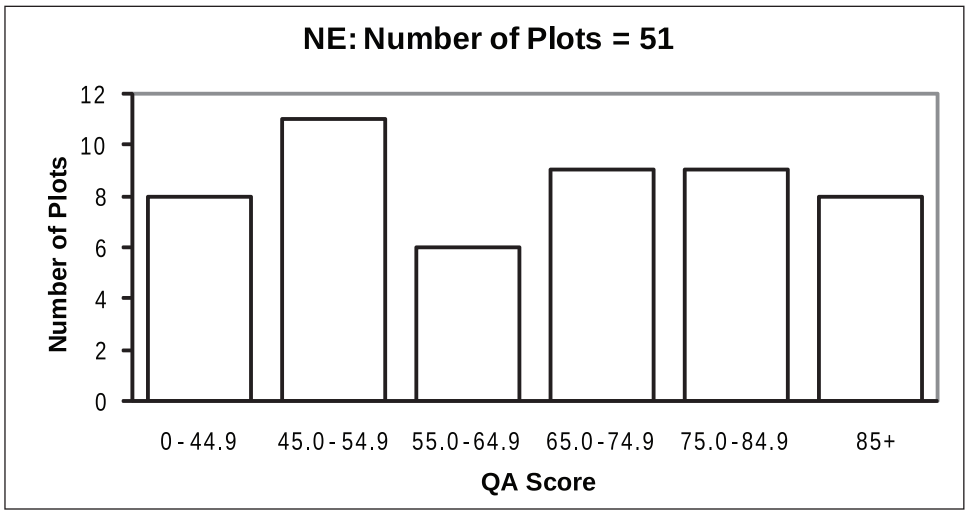


Figure 4.4—Frequency distribution of the QA scores for PNW. See Figure 4.1 caption for definition of QA score.

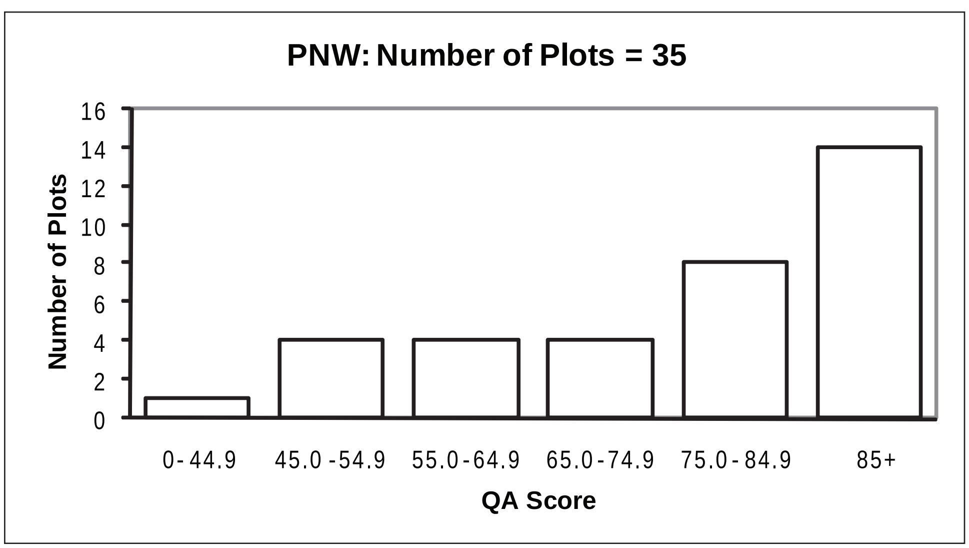


Figure 4.5—Scatter plots, by year, of the QA scores for IW. Points to the right of the vertical line QA Score = 65 achieve minimum tolerance. See Table 4.2 for overall compliance rate.

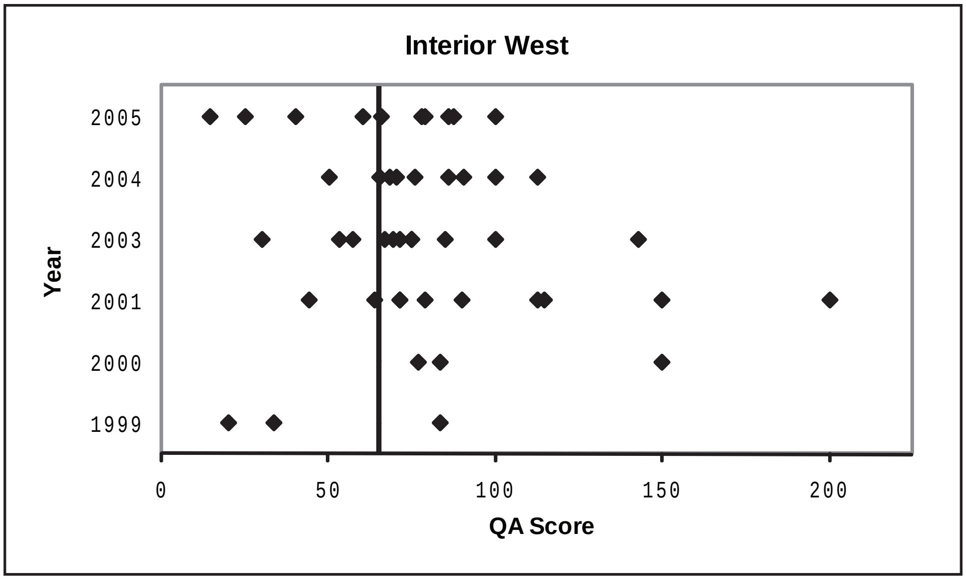


Figure 4.6—Scatter plots, by year, of the QA scores for NE. Points to the right of the vertical line QA Score = 65 achieve minimum tolerance. See Table 4.2 for overall compliance rate.

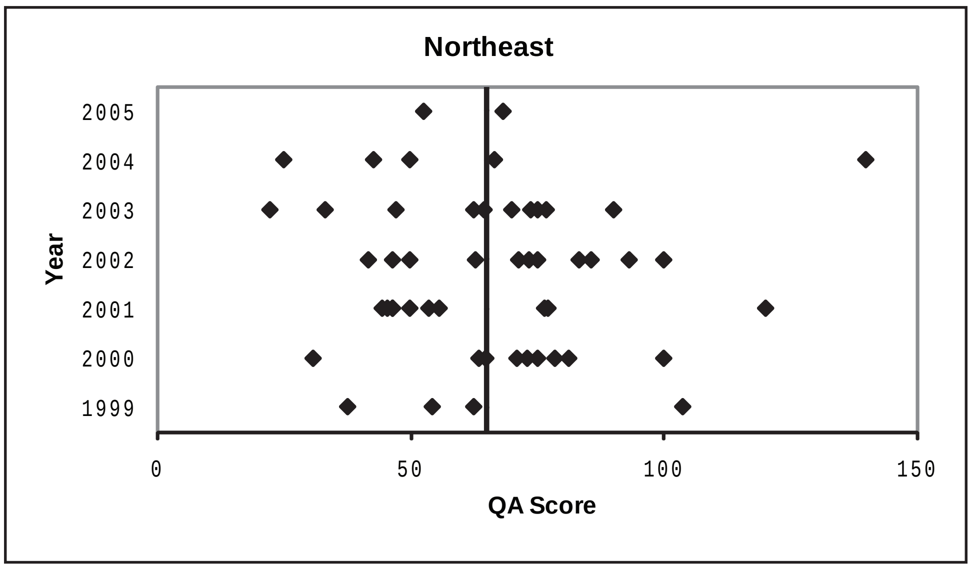


Figure 4.7—Scatter plots, by year, of the QA scores for PNW. Points to the right of the vertical line QA Score = 65 achieve minimum tolerance. See Table 4.2 for overall compliance rate.

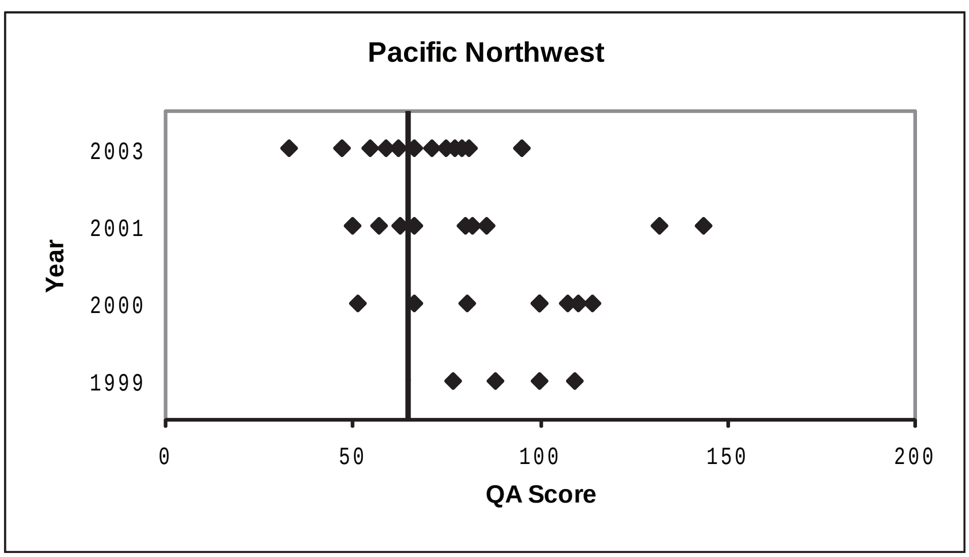


Figure 4.8—Scatter plots, by categories of the total number of species found by the QA specialist, of the QA scores for IW. Points to the right of the vertical line QA Score = 65 achieve minimum tolerance. See Table 4.2 for overall compliance rate.

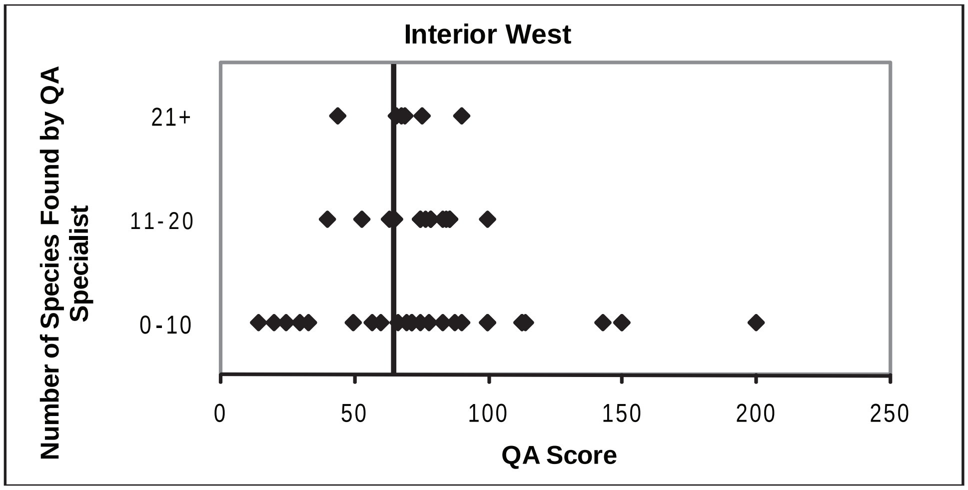


Figure 4.9—Scatter plots, by categories of the total number of species found by the QA specialist, of the QA scores for NE. Points to the right of the vertical line QA Score = 65 achieve minimum tolerance. See Table 4.2 for overall compliance rate.

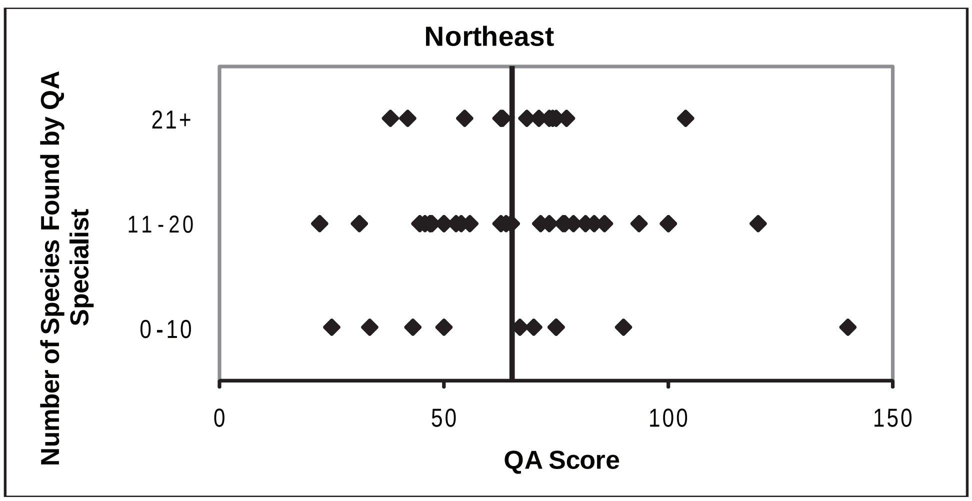
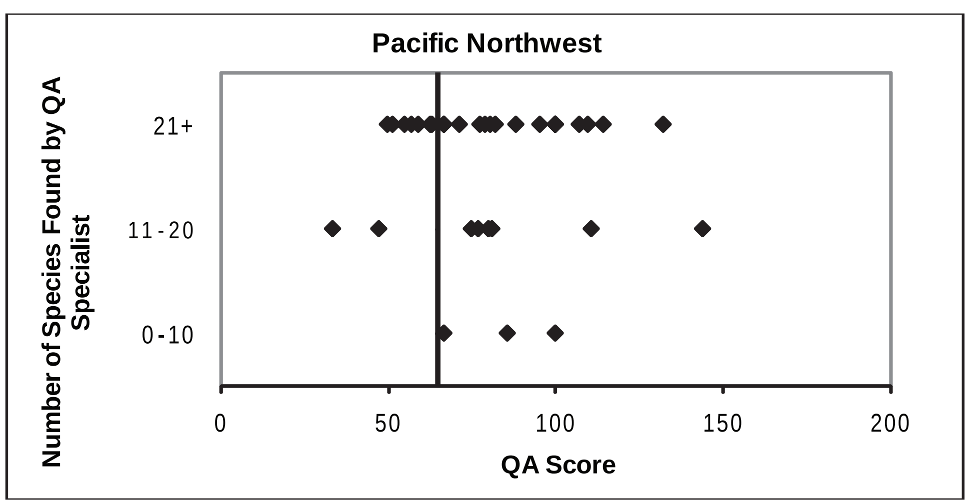


Figure 4.10—Scatter plots, by categories of the total number of species found by the QA specialist, of the QA scores for PNW. Points to the right of the vertical line QA Score = 65 achieve minimum tolerance. See Table 4.2 for overall compliance rate.



Section 5: Ozone Indicator

Olaf Kuegler

Introduction

The collection of ozone data is different from any other FIA data. Ozone biosites are deliberately chosen based on ease of access, optimal (variable) size, ozone-sensitive species, and number of plants (at least 30 individual plants of at least three ozone-sensitive species until 2001; and two ozone-sensitive species from 2002 on) (Smith et al. 2003). Biosites vary in size and are independent of the regular FIA phase 3 (P3) ground plots. A biosite has to be at least 1 acre in size within or alongside a forested area; two nearby open areas within 3 miles of each other can be combined for one biosite. The crew measures up to 30 individual plants of at least three (two beginning in 2002) ozone-sensitive species on each biosite; however, fewer species with at least 10 individual plants are acceptable.

Each plant is systematically selected by the crew so that not every ozone-sensitive plant on a given site is measured. Crews are instructed to select a starting point and move toward the center of the opening in a sweeping pattern, selecting individual plants that grow under similar conditions. Preference is given to plants with high sunlight exposure that are not suppressed. Each plant is rated for the amount and the severity of the suspected ozone damage. Field crews then collect a voucher sample containing three leaves of each injured species evaluated at each location. These vouchers are then mailed to the National Indicator Advisor for validation.

A subset of all ozone sites has been selected to be measured a second time by a QA crew. The QA crew usually consists of more experienced personnel.

Each site is mapped, including the starting point for plant selection and approximate location of plant groupings used for evaluation. However, individual plants that were sampled are not mapped and the QA crew often will select different individuals that are close to the plants that were sampled by the production crew. Furthermore, QA crew and production crew may select different species for evaluation. Thus, it is impossible to directly compare data quality at the plant level. Therefore, any site-level difference between production crew and QA crew is a combination of measurement and within-site sampling error. Since these types of errors cannot be separated, the inference that can be drawn from this analysis is limited: If a large difference exists, it is possible that a large (unknown) part is due to sampling different plants.

In this analysis, the data will be summarized at the site level in two different ways: 1) ozone biosite index; and 2) presence and absence of ozone damage. The ozone biosite index was suggested as a standard way of summarizing the FIA ozone data at the site level by Smith et al. (2003). The ozone biosite index accounts for the amount and severity of ozone damage on individual plants within the site.

Methods

Ozone Biosite Index

The ozone biosite index is calculated as

$$BI = 1000\left(m^{-1}\sum_{j=1}^{m} n_j^{-1}\sum_{p=1}^{n_j \geq 10} a_{jp}s_{jp}\right)$$

where

BI = biosite index

m = number of species evaluated

n_j =number of plants of the j^{th} species evaluated (must be at least 10)

a_{jp} =proportion of injured leaves on the p^{th} plant of the j^{th} species

s_{jp} = severity of injury as percentage of leaf area on the p^{th} plant of the j^{th} species

For comparing individual species, a biosite index for individual species can be calculated similarly:

$$BI_j = 1000\left(n_j^{-1}\sum_{p=1}^{n_j \geq 10} a_{jp}s_{jp}\right)$$

where

BI_j = biosite index for species j

For each site, the biosite index differences between production and QA crews are calculated. In addition, for sites where both crews selected the same species, the difference of biosite index for individual species and the general biosite index with only these species is calculated.

Obviously, if neither crew found any ozone damage, both crews will be in complete agreement. States in the IW (Arizona, Colorado, Idaho, Montana, Nevada, Utah) and PNW (Oregon, Washington) regions are good examples of areas with no ozone damage on blind-check plots. Since this is misleading for areas with ozone damage, some of the analyses will exclude sites where neither crew found ozone damage.

Risk Categories

Based on expert opinion, Smith et al. (2003) suggested analyzing ozone damages using four different risk categories.

To evaluate how these categories are affected, the biosite index values for the production and QA crew will be grouped in the four categories (Table 5.1). The least-affected category will be further subdivided into a category "=0" and "0< to <5". A confusion matrix will illustrate the differences between crews (see Table 5.2).

Table 5.1—Classification system for biosite index.

Biosite index	Bioindicator response	Assumption of risk	Possible impact	Relative air quality
0 - <5	Little or no foliar injury	None	Visible injury to highly sensitive species, e.g. black cherry	Good
5 - <15	Light to moderate foliar injury	Low	Visible injury to moderately sensitive species, e.g. tulip poplar	Moderate
15 - <25	Moderate to severe foliar injury	Moderate	Visible and invisible injury. Tree-level response.	Unhealthy for sensitive species
≥25	Severe foliar injury	High	Visible and invisible injury. Ecosystem-level response.	Unhealthy

Presence/Absence

The ozone bioindex can further be reduced to a binary presence or absence of ozone damage. Presence or absence can be either referring to the presence of any ozone damage or to the presence of a (biological) significant amount of ozone damage. A threshold applied to the ozone bioindex could be used to classify the plot into absence of ozone damage if the biosite index is smaller than the threshold or present if larger. Since there is no universally agreed-upon threshold, all possible thresholds between zero and the maximum ozone bioindex were evaluated.

For this binary variable, the percentage agreement between production and QA crews was calculated as well as the difference between production crew and QA crew *presence* (as percentage of total).

Results

Altogether, 222 sites were evaluated. Either the production or QA crew found at least some ozone damage on 107 sites.

Biosite Index

For sites where at least one crew found some damage, the mean biosite index was 14.84 (median: 2.88) and 14.21 (median: 2.61) for the production and the QA crew, respectively. Overall, the mean biosite index difference between production and QA crew was not significantly different from zero (paired two-sided t-test, p-value = 0.8409).

For the absolute difference between both crews' observations, the mean was 13.46, the median equaled 4.05, and the 75 percent quantile was 14.45.

Compared to the mean biosite, this means that about 25 percent of these sites had differences between crews that were larger than the overall mean biosite index.

The correlation between both indices (0.67) was moderately high. However, the moderately high correlation was due to only five extremely large indices (indices larger than 100). Without these five observations the correlation was only 0.38 (Figure 5.1).

The mean of the absolute difference was heavily influenced by a few large observations. Five sites had a difference of more than 50 (51.1, 94.4, 95.9, 154.3 and 227.1). For the four with largest differences, the production crew measured only one species. Three of these plots were collected in Pennsylvania in 1998, one was collected in Maryland in 1999, and one was collected in South Carolina in 2003. The three Pennsylvania plots had been sampled by the QA crew on the same day the production crew measured the site.

Risk Categories

The risk category was the same in 79.3 percent (176 out of 222) of all plots (Table 5.2). After excluding the 115 plots where neither crew found any ozone damage, the percentage fell to 57 percent (61 out of 107 plots).

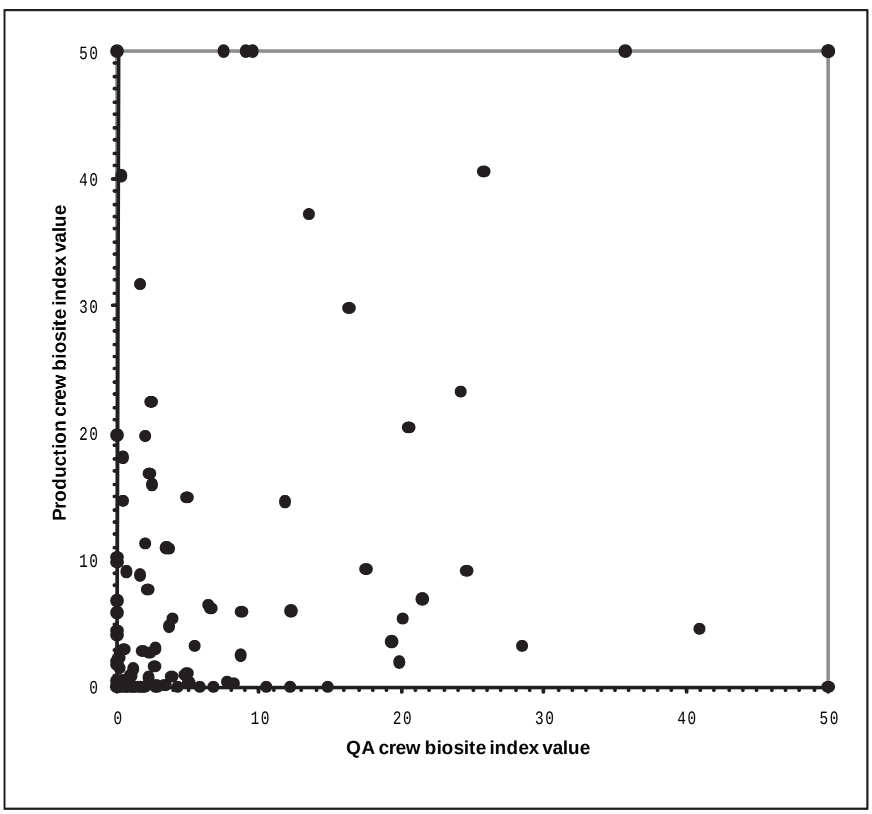


Figure 5.1—Biosite values of the production crew versus QA crew.

Table 5.2—Plot frequency comparing production crew and QA crew observed risk categories. Darker shades indicate number of plots within the same risk category.

	QA crew						
Production crew	**Biosite index**	= 0	>0 - <5	5 - <15	15 - <25	≥25	Total
	= 0	115	21	5	0	1	142
	>0 - <5	6	20	5	2	2	35
	5 - <15	4	9	5	4	0	22
	15 - <25	1	5	0	2	0	8
	≥25	1	2	4	1	7	15
	Total	127	57	19	9	10	222

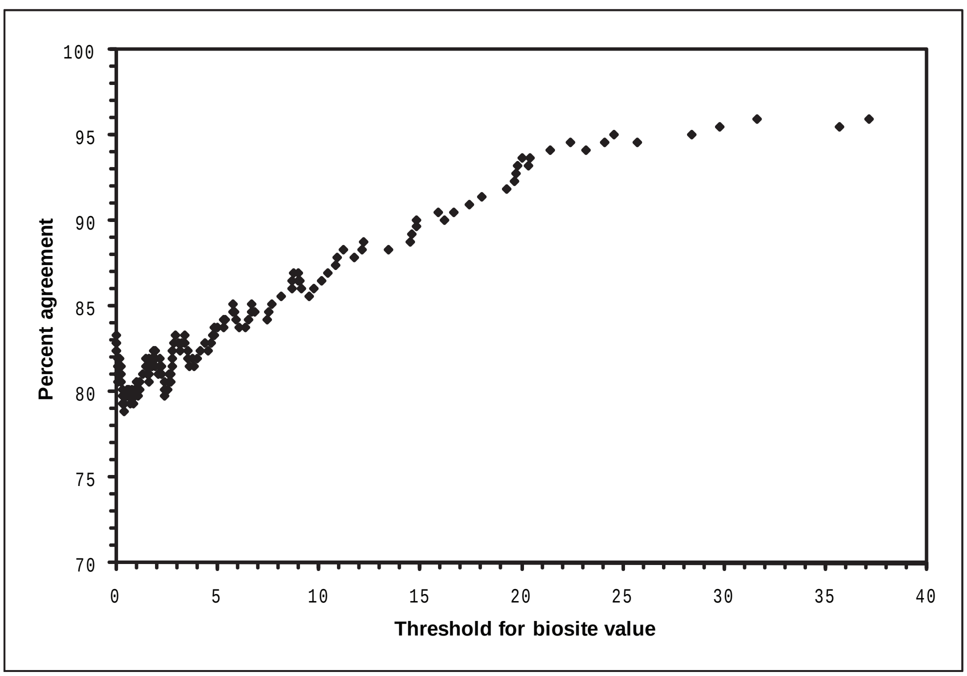
Presence/Absence

When determining the presence or absence of ozone damage, production and QA crews were in agreement on 82 percent of the plots (Figure 5.2: threshold equals 0). The percentage agreement between crews steadily increased with increasing threshold (Figure 5.2). This is as expected, since with an increasing threshold, more sites will be below the threshold until the threshold reaches the maximum biosite index value and all sites are below the threshold.

The Kappa statistics takes into account chance agreement. A Kappa value of 1 means perfect agreement, while a value of zero means no agreement apart from what would be expect by chance alone. The Kappa statistic is calculated as

$$Kappa = \frac{\text{Observed agreement} - \text{Chance agreement}}{1 - \text{Chance agreement}}$$

Figure 5.2—Percentage agreement between both crews at different thresholds.



The Kappa values (Figure 5.3) decrease slightly until the threshold reaches a biosite value of about 12, after which it slightly increases. Comparing the Kappa curve in Figure 5.3 with the percentage agreement curve in Figure 5.2, it appears that increasing agreement is due to chance alone.

Number of Species

Crews are instructed to sample at least three different ozone-sensitive species. Fewer species are accepted if the crews cannot find three species at a site.

At 9 percent of the sites (20 out of 222 sites), the production crew evaluated only one species, while the QA crew evaluated only one species on 5.4 percent (12 sites) (Table 5.3). Both crews found only one species on 4 percent of the sites (9 sites). This means that in more than half the cases, the production crew could have found at least a second species if they would have looked more carefully, while the QA could have found a second species on 3 out of their 12 sites.

The production and QA crews found less than three species at 66 sites (29.7 percent) and 55 sites (24.8 percent), respectively (Table 5.3). At 31 (14.0 percent) and 15 (6.8 percent) of these sites the other crew found at least one additional species.

The correlation between the indices for sites where both crews found at least 3 species (141 sites) was 0.80.

Figure 5.3—Kappa statistics at different thresholds.

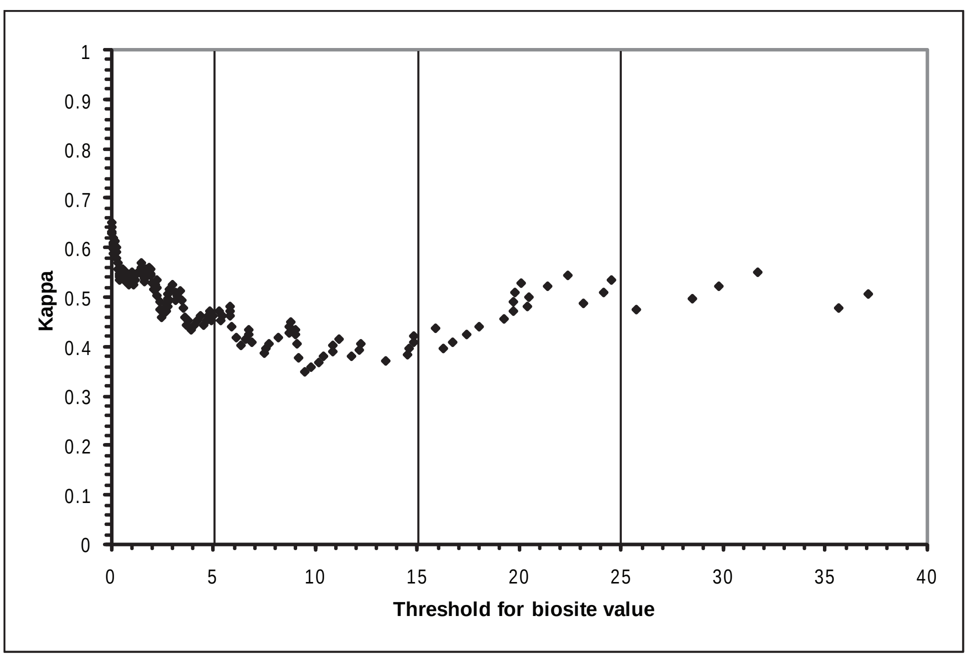


Table 5.3—Number of species found by production crew and QA crew for each blind check site.

		QA crew			
	Number of species found	1	2	≥3	All
Production crew	1	9	5	6	20
	2	0	26	20	46
	≥3	3	12	141	156
	All	12	43	167	222

Discussion and Conclusions

At first glance, the results of the data quality analysis are not very promising. It is, nevertheless, important to keep in mind that Smith et al. (2003) suggested using the average biosite value of several years for an analysis. For the current data quality analysis however, a single site was usually only visited by a QA crew in a specific year. It was therefore not possible to evaluate how much the average biosite index (same plot over several years) would be different between two different crews.

The correlation between the indices for sites where both crews found at least three species was higher than the correlation for all sites (0.80 vs. 0.67). This indicates—not surprisingly—that it is important for field crews to collect at least three species whenever possible. Nearly 50 percent of all plots (31 out of 66; Table 5.3) where the production crew found only one or two ozone-sensitive species, the QA crew found one additional species. To get a more precise ozone biosite index, the production crew should be encouraged to thoroughly search the area for any acceptable ozone indicator species.

Section 6: Vegetation Diversity and Structure Indicator

David Gartner
Bethany Schulz

Introduction

The vascular vegetation composition and structure indicator was one of the first datasets of forested vegetation data collected in a consistent manner over the extent of the United States. As such, there is a great amount of interest in the vegetation indicator data. This section analyzes the repeatability of the vascular vegetation composition and diversity variables. These variables include species identification, species canopy coverage, subplot canopy coverage, ground cover percentages, and species quadrat presence/absence.

Methods

Field Data Collection

The data were collected during the summers of 2001 to 2005. Major changes were initiated in 2004 when the sampling rules for phase 3 (P3) plots were adjusted to match FIA phase 2 (P2) sampling rules, species records on quadrats were limited to presence/absence data, and ground cover variables were collected over subplots rather than quadrats. Another change was an improved accounting of unidentified species. The database used in this exercise reflects these changes (i.e., species cover measures on quadrats have been converted to presence/absence data, and unknown codes have been converted to standard symbols for unknown plants (USDA Nat. Resourc. Conserv. Serv. 2009). Because the understory vegetation/diversity protocols are not implemented in all areas, the amount and spatial distribution of the available data are limited (Table 6.1). Therefore, only a national-level analysis of the data will be considered.

Table 6.1—Number of vegetation production and QA plots by year and region.

Year	FIA Region[a]	Number of production plots	Number of QA plots
2001	N	252	4
2002	N	239	14
2002	IW	45	1
2003	N	182	5
2003	IW	43	2
2004	N	28	1
2004	PNW	65	3
2005	N	32	1
2005	PNW	74	6

[a] N = Northern (combined NE and NC).

The current field data collection methods for the understory vegetation data are documented in Section 13 of the Forest Inventory and Analysis National Core Field Guide version 3.0. (USDA Forest Service 2004b). A subsample of the plots was revisited by QA crews, who remeasured the plots without knowing any of the production crew's data. These blind checks were performed by QA crews on 1.6 percent to 8.1 percent, for an overall average of 3.9 percent (Table 6.1).

From 2001 to 2004, the species data were recorded in subplot level records. During the 2005 field season, the species data were recorded in plot level records, with separate columns for each subplot. For this analysis, the 2005 species data were reformatted into subplot-level records.

Data Matching

Matching the QA data with the production-crew data is easy for quadrat status data and subplot ground cover data because the index variables used, such as state/county/plot/subplot, create a unique identifier. However, matching species-level data is difficult because computer matching requires that the QA crew and the production crew identified the plant as the same species. The species-level data were sorted and matched by combinations of year, state, county, plot, subplot, family name, genus, species, and NRCS code.

The production crew observations and QA observations that had the same genus and species codes were matched and labeled "Match Type 1". Combinations that had the same genus code, with one crew's observation being identified at the species level, while the other crew used a general code for the genus, were labeled "Match Type 2". Combinations that had the same genus code, but different species codes were labeled "Match Type 3". Combinations of observations that had the same family code but different genus codes were labeled "Match Type 4". All remaining unmatched observations with an NRCS code were labeled "Match Type 5". When both crews used the same type of unknown, such as "2gr" for unknown graminoid, the resulting combination was labeled "Match Type 6", but there is almost no basis for claiming that the paired observations actually represent the same species. No observation was used more than once. For Match Types 2, 3, 4, and 6, if more than one of the crews' observations was available for matching with a single observation from the other crew, the observation with the closest agreement in subplot canopy cover and quadrat present/absence was used. Table 6.2 contains a summary of the different match types.

For one plot, the QA crews were not able to measure within a month of the production crew; this plot was dropped because of the dynamic nature of herbaceous vegetation canopy covers. Occasionally QA crews could not remeasure full plots because of weather. One QA blind check plot fell into this category, and only the portions of the plots that were measured by both crews were used in the analysis.

Table 6.2—Match type definitions.

Match Type	Matching basis	Number of pairs of observations
1	Matched at species level	1682
2	Matched at genus level with one crew identified the plant by species while the other crew only identified the plant by genus	399
3	Matched at genus level but identified as different species	221
4	Matched at family level	211
5	Unmatched observations	1550
6	Matched observations where at least one of the crews used an unknown species code	155

Data Analysis

Each measurement has a measurement quality objective (MQO), which consists of a measurement tolerance and a required percent of observations that need to be within the measurement tolerance. For most variables, calculating the percentage of observations within the measurement tolerance is straightforward: differences between the production and the QA data are compared for each observation to determine if the difference is within tolerance, and then the percent of observations within tolerance is calculated. For the canopy cover variables and ground cover variables, the tolerance is one cover class, with the classes being 0, trace, 1-5 percent, 6-10 percent, 11-20 percent, 21-40 percent, 41-60 percent, 61-80 percent, and 81-100 percent. The tolerance for the trampling code is one class, with the classes being 0-10 percent, 10-50 percent, and 50-100 percent. The other variables have no tolerance. For quadrat presence/absence, all species occurring on the subplot were used, but with the results are listed by match type. However, species codes, species canopy covers, and species presence/absences get intermingled. If observations are matched incorrectly due to differences in species codes, then species canopy covers and presence/absences from two different species might be compared against each other. So the problems of comparing species canopy covers and quadrat presence/absences get confounded with the problems of correctly identifying the species.

Results

Quadrat and Ground Cover

The quadrat variables either met the MQO compliance rates, or were missed by only a couple of percentage points (Table 6.3).

With the exception of litter/duff, the subplot ground cover variables either met their MQO compliance rates or were within a couple of percentage points (Table 6.4). The subplot ground cover data variables are relatively new and therefore were measured only on 16 plots, which contained only 48 forested subplots. The low number of subplots causes the observed agreement rates to have low precision.

The subplot total canopy cover variables fell well below their MQO compliance rates (Table 6.5). The highest canopy layer (4) and the lowest canopy level (1) showed the highest observed agreement rates at 81 percent and 73 percent respectively. The observed compliance rates for the middle two canopy layers (2 and 3) were below 70 percent.

Table 6.3—Repeatability analysis of quadrat variables.

Variable	MQO compliance rate	Percent within tolerance	Number of observations
	——— percent ———		
Condition	99	96.5	426
Quadrat Status	99	100	114
Quadrat Status Pre-2004	99	98.7	312
Trampling	90	96.5	405

Table 6.4—Repeatability analysis of subplot ground cover data.

Variable	MQO compliance rate	Percent within tolerance	Number of observations
	——— percent ———		
Cryptobiotic crust	90	100	48
Lichen	90	97.9	48
Litter/duff	90	75.0	48
Mineral soil	90	93.8	48
Moss	90	87.5	48
Road/trail	90	95.8	48
Rock	90	95.8	48
Standing water	90	97.9	48
Stream/lake	90	100	48
Trash/junk/other	90	100	48
Wood	90	95.8	48

Table 6.5—Repeatability analysis of subplot total canopy cover.

Variable	MQO compliance rate	Percent within tolerance	Number of observations
	——— percent ———		
Canopy layer 1	90	73.0	152
Canopy layer 2	90	66.4	152
Canopy layer 3	90	67.8	152
Canopy layer 4	90	80.9	152

Species

Because the ability to meet the MQO compliance rates for species-level data depends partly on the ability to match production crew observations to QA crew observations, the results for the species level are listed by match type.

The MQO compliance rate for species is 99 percent (Table 6.6) and other match types have cumulative maximum compliance rate of 1 percent. Because Match Types 1, 2, 3, 4, and 6 contain two observations (production and QA), those match types are given a weight of 2, while Match Type 5 contains only one observation, and these matches are given a weight of only 1 (Table 6.6). For Tables 6.7 to 6.15, the unmatched observations (Match Type 5) are not used in calculating the overall observed percentage agreement. However the number of unmatched observations has been included in the tables so that the user can calculate an overall observed percentage agreement rate that includes the effect of unmatched observations if they wish.

Selection and training for the production crew has become more sophisticated over the years. Changes include recruitment of crews with better botanical skills, formalizing the process for the tracking identities of unknown species, and more stringent requirements for passing certification. Another improvement includes using data collection software that enables crew trainers to download and quickly score certification tests to determine which crew members may need additional training or are clearly not qualified to collect species data. This change is expected to have resulted in an increase in the percentage of species records that were able to be matched at the species level. Therefore, the 2001-2003 data are analyzed separately from the 2004-2005 data to determine if the expected improvement in match rates were obtained. The percentage of the records that could be matched at the species level increased from 43.1 percent to 61.4 percent (Table 6.6). The normal approximation to the binomial distribution test shows that the improvement in the observed Match Type 1 percentages is statistically significant.

We expected that observations that agree on species identification were more likely to agree on the rest of the species-level variables, therefore, we examined the species presence/absences and species canopy-cover results by match type.

The quadrat presence/absence data fell below the MQO compliance rate for all match types (Table 6.7). The level of species identification agreement contained in match type had very little effect on the observed agreement rates, for matched observations that incorporated at least family level information. Species recorded

Table 6.6—Match type percentages before and after procedural changes.

Match Type	MQO compliance rate	Percent of total observations	
		2001-2003	2004-2005
	percent		
1	99	43.1	61.4
2	<1	13.1	8.3
3	<1	6.4	6.4
4	<1	7.0	4.0
5	<1	25.5	16.0
6	<1	4.9	3.7
Number of Observations		4723	2094

with the '2xx' coding symbols for unknown species (USDA Nat. Resour. Conserv. Serv. 2009) performed particularly poorly.

The species total cover per subplot performed close to the MQO compliance rate for species that were able to be matched at least at the family level (Table 6.8). For species that were matched by family, the level of species identification agreement had relatively little effect on the observed agreement rate. Species with the '2xx' coding performed poorly.

The subplot species layer covers were not recorded until 2004. Therefore there is a large difference in the total number of observations for the subplot species total cover and the subplot species layer cover. Also, the species canopy for canopy layers 1 and 2 were recorded together as one variable and so the results for these two layers can not be separated. The percentage cover for species subplot layers 1 and 2 performed about 10 percent below than the MQO compliance rate for the species matched at the genus level or better (Match Types 1-3) (Table 6.9). The species recorded using the '2xx' format performed poorly.

Table 6.7—Repeatability of quadrat presence/absence by match type.

Match Type	MQO compliance rate	Percent within tolerance	Number of observations
	percent	percent	
All	99	80.8	8469
1	99	89.3	5046
2	99	86.1	1197
3	99	89.6	663
4	99	87.5	633
5			4650
6	99	17.4	930

Table 6.8—Repeatability of subplot species total cover by match type.

Match Type	MQO compliance rate	Percent within tolerance	Number of observations
	percent	percent	
All	90	81.2	5386
1	90	84.5	3364
2	90	88.7	798
3	90	85.1	442
4	90	90.5	422
5			1550
6	90	20.0	310

Table 6.9—Repeatability of subplot species layer 1 and 2 cover by match type.

Match Type	MQO compliance rate	Percent within tolerance	Number of observations
	percent	percent	
All	90	76.9	1816
1	90	79.4	1328
2	90	78.9	180
3	90	85.5	138
4	90	71.1	90
5			347
6	90	22.5	80

The species subplot layer 3 canopy cover exceeded the MQO compliance rate for all match types (Table 6.10).

The species subplot layer 4 canopy cover exceeded the MQO requirements for all match types (Table 6.11).

Subplot Species Layer Covers Using Just the Species with Canopy in that Layer

If a species doesn't have any canopy in a given layer, then both crews should record a canopy cover of 0 percent. They will agree on the percent canopy cover for that layer. Therefore, the agreement rate for species canopy cover by layer will confound the ability to determine the amount of canopy cover of a species in a layer with the ability to determine if a species has canopy cover in that layer. To test the ability of crews to determine the amount of canopy cover for species in a layer, the species subplot layer cover tables were recalculated using just the species that occurred in that canopy layer.

Since most of the matched observations contained canopy in layer 1 and 2, the species cover results for the species occurring in layers 1 and 2 (Table 6.12) are within 0.5 percent of the results using all observations (Table 6.9).

Only about one-quarter of the species had recorded canopies in layer 3. Restricting the analysis to only species occurring in layer 3 dropped the observed agreement rate from 92.6 percent to 60.4 percent (Table 6.13). Restricting the analysis also causes the observed agreement rates to fall below the MQO compliance rate for all match types.

Table 6.10—Repeatability analysis of subplot species layer 3 cover by match type.

Match Type	MQO compliance rate	Percent within tolerance	Number of observations
	percent	percent	
All	90	92.6	1816
1	90	91.4	1328
2	90	96.7	180
3	90	95.7	138
4	90	91.1	90
5			347
6	90	98.8	80

Table 6.11—Repeatability analysis of subplot species layer 4 cover by match type.

Match Type	MQO compliance rate	Percent within tolerance	Number of observations
	percent	percent	
All	90	94.6	1816
1	90	93.6	1328
2	90	95.6	180
3	90	98.6	138
4	90	95.6	90
5			347
6	90	100.0	80

Only about one-seventh of the species had canopy in layer 4. As with layer 3, restricting the analysis only to species that occur in layer 4 drops the overall observed agreement rate from better than the MQO requirement (94.6 percent; Table 6.11) to well below the MQO requirement rate (60.2 percent) (Table 6.14). Also, as with layer 3, this restriction causes the observed agreement rate for each match type to fall below the MQO compliance rate.

Table 6.12—Repeatability analysis of subplot species layer 1 and 2 cover (only species occurring in the layer) by match type.

Match Type	MQO compliance rate	Percent within tolerance	Number of observations
	percent	percent	
All	90	76.4	1780
1	90	78.9	1296
2	90	78.7	178
3	90	85.3	138
4	90	71.1	90
5			339
6	90	22.5	80

Table 6.13—Repeatability analysis of subplot species layer 3 cover (only species occurring in the layer) by match type.

Match Type	MQO compliance rate	Percent within tolerance	Number of observations
	percent	percent	
All	90	60.4	341
1	90	62.0	300
2	90	50.0	12
3	90	62.5	16
4	90	33.3	12
5			16
6	90	0.0	1

Table 6.14—Repeatability analysis of subplot species layer 4 cover (only species occurring in the layer) by match type.

Match Type	MQO compliance rate	Percent within tolerance	Number of observations
	percent	percent	
All	90	60.2	246
1	90	62.2	222
2	90	20.0	10
3	90	80.0	10
4	90	0.0	4
5			11
6	90		0

Number of Species per Subplot

To determine if the species identification/matching problem reflected differences between the ability of the crews to agree on the number of species per subplot, we compared the production crew's number of species on a subplot with the QA crew's number of species on a subplot (Figure 6.1). The QA crew's number of species tended to be higher than the field crew's, both before the changes in 2004 (mean difference = 1.35) and after the changes in 2004 (mean difference = 3.2). If an MQO were created with a 10 percent tolerance, 37.4 percent of the pre-2004 subplots would meet the tolerance, while only 26.7 percent of the 2004-2005 subplots would meet the tolerance.

Ability to Match as Related to Cover Class

One of the hypotheses about why certain species are difficult to match is that plants with smaller amounts of cover are either more likely to be missed in the time constraints related to field work, or are rare so that the crews have a more difficult time recognizing them. Under this hypothesis, the smallest cover classes are expected to have the greatest proportion of unmatched and poorly matched observations. The percent of observations that are matched at the species level is much lower for the smallest canopy cover class (49.4 percent) than for most of the other canopy cover classes (71.8 percent to 85.3 percent) (Table 6.15). This shows that the greatest problem with matching by species occurs in the smallest cover class. However, since none of the cover classes had matched-by-species rates near the MQO compliance rate of 99 percent, separate species identification problems exist in addition to identifying species that have little canopy cover.

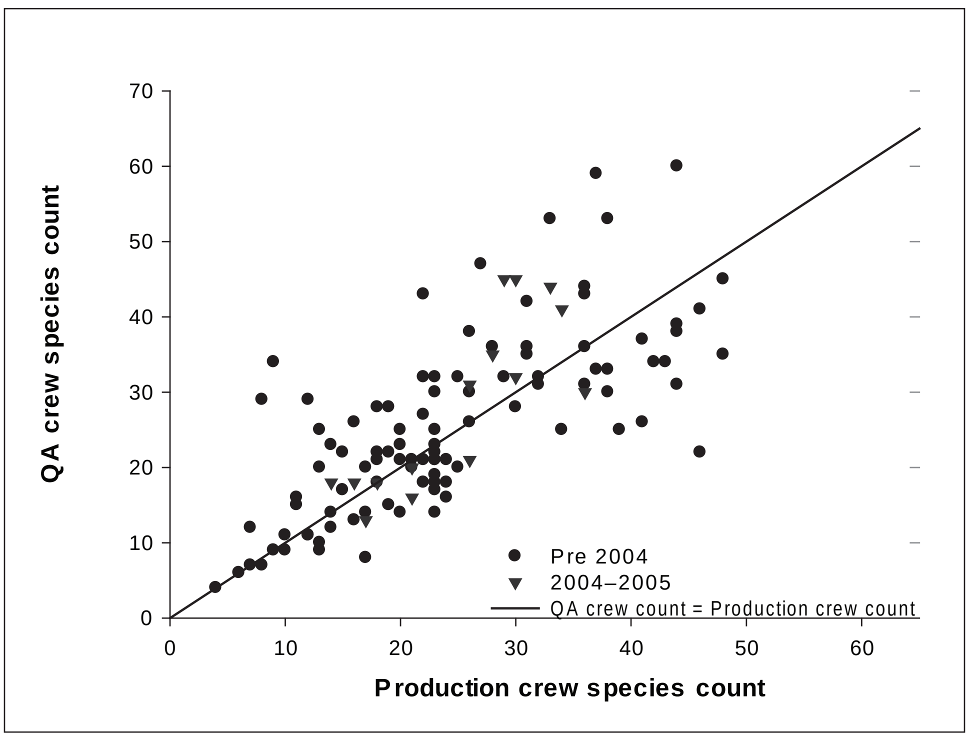


Figure 6.1—Comparison between QA crew and production crew species counts.

Table 6.15—Percentage of match types by production crew cover class.

	Production Crew Cover Class						
	1	2	3	4	5	6	7
Number of Observations	4418	455	321	248	122	78	28
Match Type	percent						
1	49.4	73.0	80.0	84.6	85.3	83.0	71.8
2	15.4	10.6	3.2	7.5	2.8	6.4	15.4
3	7.7	7.9	8.7	5.5	4.2	2.1	10.3
4	8.5	2.6	2.7	0.0	2.8	6.4	0.0
5	16.8	5.7	4.6	2.1	4.9	2.1	2.6
6	2.2	0.2	0.8	0.3	0.0	0.0	0.0

Discussion

The ground cover variables and the other non-vegetation variables, such as quadrat status and trampling, were near or above their MQO compliance rates. The vegetation variables did not perform as well.

Species Identification

Species identification is a problem. Only 61.4 percent of the 2004-2005 observations were able to be matched at the species level (Table 6.4). In 2004, an improved unknown species specimen tracking system was introduced and further refined in 2005. The percent of records that were not able to be matched dropped from 17.4 percent to 15.3 percent, suggesting a result from the improved tracking system. However, this improvement was not sufficient to resolve the species identification problem. The species identification problem occurred in all canopy cover classes and canopy layers. While this problem is most pronounced in the smallest canopy cover classes, the other canopy classes had 15 to 30 percent of their observations that could not be matched at the species level. Observations that were not able to be matched at any level followed the same pattern of having the greatest problem with species with small canopy cover. For readers who are interested in the species identification rates for a specific genus or family, tables with this information have been included in Appendix B.

If a crew misidentifies a species without merging two species into one species name, this misidentification will not change species richness measures. If a crew consistently misidentifies a species without merging two species into one species name then the species compositional change and total number of species observed measures over that crew's plots will not be affected by the wrong species call. However, measures of changes in species presence/absence over time will be affected by wrong species calls. For instance, McNemar's test, which involves the number of plots where a species is found at time 2 but not at time 1 and the number of plots where a species was found at time 1 but not at time 2, could be greatly affected by wrong species calls, especially if the same wrong species calls do not occur consistently over time. Therefore the effects of disagreements in species identification depend on how the data are being used.

Another factor affecting species identification (and thus, species matching) is familiarity with the species encountered. In the production database for 2001-2004, 32 percent of species recorded were observed only once, with an additional 16 percent observed only two times. We speculate that species that are frequently recorded are likely to be identified correctly more often, and we should be less confident in the identification of species observed very rarely unless collected for identification by a herbarium. The vegetation indicator data should not be expected to track changes in the distribution of species that are rarely encountered. The influence of infrequently encountered species should be more thoroughly investigated and factored into the quality objectives for the vegetation indicator.

Some of the MQOs are unable to give an accurate assessment of repeatability. For example, suppose there are two similar species with similar sized patches, A and B, and production crew give the species in patch A species code of RITR (*Ribes triste*) and the species in patch B a species code of RIHU (*Ribes hudsonianum*), and the QA crew reverses the species calls by giving the species in patch A the species code RIHU and the species in patch B as species code RITR. The MQO data process, as currently constructed, will not catch the fact that the two crews did not actually agree. To truly test the species identification MQO, the QA crew would have to be watching the field crew identify individual species. Because this would violate the intent of blind checks in addition to being logistically impractical, the species identification MQO needs to be reconsidered. The species identification MQO carries with it an implicit assumption that if the species identification agreement rate is sufficiently high, that there will be an agreement on the number of species found on each plot. An MQO for the number of species found on each plot should be considered, since none exists at present.

Canopy Cover

The total canopy covers fell below the MQO compliance rate for all four canopy levels. The species canopy 1 and 2 layer fails to meet the MQO compliance rate. How well the species canopy cover layer 3 and 4 measurements performed depends on if the issue is split into ‘Does the species have canopy in the layer?’ and ‘Given the species has canopy in the layer, what is the amount of canopy?’ or not. The crews agree on whether or not the species has canopy in a layer, but have problems agreeing on the amount of canopy in the layer given that the species has canopy in the layer, especially for layers 3 and 4. With the exception of the species with “2xx” codes, there is no evidence that canopy cover agreement rates were affected by the species matching levels.

Conclusions

The initial values for the MQO tolerances and compliance rates are “best guesses” on what can be achieved. The species identification compliance rate of 99 percent is unlikely to be achievable in the field. The compliance rates need to be re-evaluated in view of both, what can be achieved and the effect on the precision of the anticipated analysis. Higher observed agreement rates can be achieved with increased training and crew qualification standards. However, the final level of training will be a trade-off between precision goals and financial constraints.

Because of the scarcity of vegetation data collected in a consistent manner over extensive areas, there will be a great amount of interest in the vegetation indicator data. However, we must be wary of using the vegetation data to draw conclusions for which the vegetation data MQOs are not specifically designed. People are likely to be interested in certain species, such as threatened, endangered species. The problems of assuring accurate (not only repeatable) species identification for vegetation species are much greater than those for tree species. Therefore, some consideration needs to be given into how to assure the accuracy of the species identification.

Experts in the collection of vegetation data should use these results to determine not only how to improve the repeatability of the measurements, and what levels of repeatability are achievable in the field, but also what measures of repeatability will most reflect the needs of their analysis. These results will also give the users of the data a means of determining the effects of the repeatability on their analyses.

Appendix B. Matching Rates by Family and by Genus

The tables in this appendix contain the species identification rates broken out by family and by genus. The objective of the tables in this appendix is to give people interested in specific family or genus information about the reliability of the species identification of the plant group of interest. For example, some users might need to know that we are not very good at getting species identification agreement for the genus *Euonymus* (1 out of 6), but are doing well for the genus *Kalmia* (8 out of 9).

Table B1—Distribution of observations among match types for each family

	Match Type					
Family	**1**	**2**	**3**	**4**	**5**	**Total**
Acanthaceae	0	0	0	0	4	4
Aceraceae	107	0	9	0	22	138
Anacardiaceae	15	0	0	2	8	25
Annonaceae	0	3	0	0	2	5
Apiaceae	18	4	4	3	20	49
Apocynaceae	7	0	1	0	3	11
Aquifoliaceae	7	0	0	0	5	12
Araceae	9	1	0	0	13	23
Araliaceae	15	3	3	0	12	33
Aristolochiaceae	5	0	0	0	3	8
Asclepiadaceae	5	0	0	0	2	7
Aspleniaceae	3	0	0	0	4	7
Asteraceae	82	32	10	47	87	258
Balsaminaceae	5	2	0	0	12	19
Berberidaceae	34	1	6	1	10	52
Betulaceae	40	1	13	7	39	100
Bignoniaceae	0	0	0	0	1	1
Blechnaceae	6	0	0	0	2	8
Boraginaceae	2	0	2	0	6	10
Brassicaceae	4	0	0	0	5	9
Campanulaceae	0	0	0	0	11	11
Caprifoliaceae	45	7	11	4	43	110
Caryophyllaceae	1	0	0	4	9	14
Celastraceae	11	1	0	0	10	22
Chenopodiaceae	0	0	0	0	1	1
Clethraceae	3	0	0	0	1	4
Clusiaceae	0	1	1	4	3	9
Convolvulaceae	0	1	0	0	4	5
Cornaceae	43	6	0	0	22	71
Crassulaceae	0	0	2	0	0	2
Cupressaceae	21	0	0	0	4	25
Cyperaceae	25	40	7	3	88	163
Dennstaedtiaceae	22	5	0	0	10	37

Table B1—Distribution of observations among match types for each family *(continued)*

Family	Match Type 1	2	3	4	5	Total
Dioscoreaceae	0	0	0	0	6	6
Dipsacaceae	3	0	0	0	1	4
Droseraceae	0	0	0	0	2	2
Dryopteridaceae	58	4	3	6	34	105
Elaeagnaceae	0	0	1	0	1	2
Equisetaceae	10	5	0	0	3	18
Ericadeae	72	30	8	6	29	145
Fabaceae	35	15	11	14	36	111
Fagaceae	97	2	8	1	35	143
Fumariaceae	1	0	0	0	0	1
Gentianaceae	3	0	0	0	4	7
Geraniaceae	7	1	0	0	4	12
Grossulariaceae	9	7	2	0	14	32
Hamamelidaceae	13	2	0	0	3	18
Hydrophyllaceae	2	0	0	0	0	2
Iridaceae	0	1	0	0	2	3
Juglandaceae	31	2	5	2	24	64
Juncaceae	2	1	1	0	10	14
Lamiaceae	12	0	0	2	20	34
Lauraceae	28	1	0	1	11	41
Liliaceae	65	5	12	18	52	152
Linaceae	0	0	0	0	1	1
Lycopodiaceae	11	3	1	2	13	30
Magnoliaceae	9	2	0	0	6	17
Melastomataceae	3	0	0	0	2	5
Menispermaceae	1	0	0	0	0	1
Menyanthaceae	4	0	0	0	0	4
Monotropaceae	2	0	0	0	10	12
Moraceae	2	0	0	0	8	10
Myricaceae	2	1	0	0	1	4
Nyssaceae	11	0	0	0	12	23
Oleaceae	16	0	10	0	13	39
Onagraceae	11	6	0	0	21	38
Ophioglossaceae	0	0	1	0	14	15
Orchidaceae	11	0	3	0	24	0
Orobanchaceae	0	0	0	0	1	1
Osmundaceae	4	0	1	0	9	14
Oxalidaceae	5	10	4	0	8	27
Papaveraceae	1	0	0	0	2	3
Passifloraceae	0	0	0	0	1	1
Phytolaccaceae	3	0	0	0	2	5

Table B1—Distribution of observations among match types for each family *(continued)*

Family	Match Type 1	2	3	4	5	Total
Pinaceae	145	0	9	1	22	177
Plantaginaceae	0	4	0	0	1	5
Platanaceae	0	0	0	0	1	1
Poaceae	56	8	8	28	153	253
Polemoniaceae	2	0	3	0	2	7
Polygalaceae	4	0	0	0	0	4
Polygonaceae	2	5	0	0	18	25
Polypodiaceae	5	0	0	0	3	8
Portulacaceae	3	0	2	0	1	6
Primulaceae	22	5	0	0	20	47
Pteridaceae	3	0	0	0	2	5
Pyrolaceae	22	2	1	0	13	38
Ranunculaceae	54	5	0	3	38	100
Rhamnaceae	3	0	0	4	7	14
Rosaceae	145	70	38	31	111	395
Rubiaceae	20	13	5	2	25	65
Rutaceae	0	0	0	0	1	1
Salicaceae	29	10	1	0	13	53
Saxifragaceae	4	0	2	8	18	32
Scrophulariaceae	11	3	0	2	73	89
Selaginellaceae	0	0	0	0	2	2
Smilacaceae	4	10	1	0	17	32
Solanaceae	0	0	0	0	2	2
Staphyleaceae	0	0	0	0	1	1
Stereophyllaceae	0	0	0	0	2	2
Thelypteridaceae	3	0	0	0	12	15
Tiliaceae	4	0	0	0	5	9
Ulmaceae	20	0	9	0	12	41
Urticaceae	5	0	0	1	17	23
Verbenaceae	2	0	0	0	13	15
Violaceae	8	32	2	0	24	66
Vitaceae	27	25	0	4	10	66
Xyridaceae	0	1	0	0	0	1

Table B2—Distribution of observations among match types for each genus

Genus	Match Type 1	2	3	4	5	Total
Abies	32	0	1	1	3	37
Acer	107	0	9	0	22	138
Achillea	14	1	0	2	5	22
Achlys	7	0	0	0	0	7
Achnatherum	0	0	0	2	2	4
Actaea	8	0	0	0	1	9
Adenocaulon	2	0	0	0	1	3
Adiantum	3	0	0	0	2	5
Agastache	2	0	0	0	2	4
Ageratina	1	0	0	3	7	11
Agrimonia	2	1	0	0	2	5
Agrostis	1	0	0	1	3	5
Alliaria	4	0	0	0	2	6
Allium	1	0	0	0	0	1
Alnus	7	0	0	0	2	9
Ambrosia	2	0	0	1	1	4
Amelanchier	13	7	0	4	21	45
Amphicarpaea	8	0	0	3	7	18
Anaphalis	1	0	0	1	2	4
Andropogon	4	0	0	1	4	9
Anemone	4	0	0	0	10	14
Antennaria	0	0	1	2	4	7
Apocynum	7	0	1	0	3	11
Aquilegia	0	0	0	0	5	5
Aralia	14	3	3	0	11	31
Arctium	0	0	0	0	2	2
Arctostaphylos	5	0	4	0	5	14
Argemone	0	0	0	0	1	1
Arisaema	1	1	0	0	13	15
Arnica	8	0	0	0	0	8
Arnoglossum	0	0	0	0	1	1
Asarum	5	0	0	0	3	8
Asclepias	5	0	0	0	2	7
Asimina	0	3	0	0	2	5
Asplenium	3	0	0	0	4	7
Aster	0	2	0	0	3	5
Astragalus	0	0	0	0	2	2
Athyrium	8	0	0	4	10	22
Balsamorhiza	0	0	0	1	0	1
Berberis	3	0	2	0	2	7
Betula	23	1	0	1	12	37
Blechnum	6	0	0	0	0	6

Table B2—Distribution of observations among match types for each genus *(continued)*

Genus	Match Type 1	2	3	4	5	Total
Boehmeria	2	0	0	0	5	7
Botrychuim	0	0	1	0	14	15
Brachyelytrum	2	0	0	1	3	6
Bromus	1	0	0	4	4	9
Cacaliopsis	4	0	0	0	0	4
Calamagrostis	14	0	0	4	11	19
Calochortus	0	0	0	0	1	1
Caltha	4	0	0	0	0	4
Calypso	0	0	0	0	1	1
Calystegia	0	0	0	0	2	2
Campanula	0	0	0	0	10	10
Capanulastrum	0	0	0	0	1	1
Campsis	0	0	0	0	1	1
Cardamine	0	0	0	0	2	2
Carex	13	39	7	2	85	150
Carpinus	4	0	0	5	7	16
Carya	28	2	5	1	20	56
Castilleja	0	2	0	0	0	2
Caulophyllum	4	0	0	1	1	6
Ceanothus	0	0	0	0	1	1
Celastrus	0	1	0	0	1	2
Celtis	3	0	0	0	4	7
Cerastium	0	0	0	0	3	3
Cercis	6	0	0	0	4	10
Chamaecrista	0	0	0	1	0	1
Chamaecyparis	4	0	0	0	0	4
Chamerion	11	0	0	0	5	16
Chasmathium	1	0	0	0	0	1
Chenopodium	0	0	0	0	1	1
Chimaphila	14	0	0	0	5	19
Chrysosplenium	0	0	0	0	3	3
Chrysothamnus	0	0	0	0	1	1
Cichorium	3	0	0	1	0	4
Cicuta	0	0	0	0	1	1
Cimicifuga	0	0	0	0	4	4
Cinna	0	0	0	0	3	3
Circaea	0	2	0	0	13	15
Cirsium	5	2	0	2	0	9
Claytonia	3	0	2	0	1	6
Clematis	0	0	0	0	1	1
Clethra	3	0	0	0	1	4
Clintonia	12	2	0	0	2	16

Table B2—Distribution of observations among match types for each genus *(continued)*

Genus	Match Type 1	2	3	4	5	Total
Collinsia	7	0	0	0	2	9
Collinsonia	0	0	0	0	1	1
Collomia	2	0	0	0	2	4
Comptonia	2	0	0	0	1	3
Conopholis	0	0	0	0	1	1
Conyza	0	0	0	0	1	1
Coptis	18	0	0	1	2	21
Corallorrhiza	0	0	0	0	1	1
Cornus	43	6	0	0	22	71
Coronilla	4	0	0	0	0	4
Corylus	3	0	13	1	7	24
Crataegus	2	10	0	0	3	15
Crepis	0	0	0	2	0	2
Cryptantha	0	0	2	0	4	6
Cryptotaenia	1	1	0	2	3	7
Cunila	0	0	0	0	1	1
Cynoglossum	2	0	0	0	0	2
Cyperus	0	0	0	0	1	1
Cypripedium	2	0	0	0	2	4
Dactylis	1	0	0	0	3	4
Daucus	4	0	0	0	2	6
Delphinium	0	0	0	1	1	2
Dennstaedtia	5	4	0	0	4	13
Deschampsia	2	0	0	0	1	3
Desmodium	2	4	6	0	5	17
Dicentra	1	0	0	0	0	1
Dichanthelium	1	4	2	2	11	20
Diervilla	0	0	0	1	5	6
Digitalis	1	0	0	0	0	1
Dioscorea	0	0	0	0	6	6
Dipsacus	3	0	0	0	1	4
Disporum	0	0	7	0	0	7
Drosera	0	0	0	0	2	2
Dryopteris	21	4	3	2	10	40
Elaeagnus	0	0	1	0	1	2
Eleaocharis	0	1	0	0	0	1
Elymus	4	0	0	2	7	13
Enemion	2	0	0	0	0	2
Epilobium	0	4	0	0	2	6
Epipactis	0	0	0	0	2	2
Equisetum	10	5	0	0	3	18
Erechtites	0	0	0	0	4	4

Table B2—Distribution of observations among match types for each genus *(continued)*

Genus	Match Type 1	2	3	4	5	Total
Erigeron	4	0	3	1	3	11
Eriogonum	0	0	0	0	2	2
Eriophorum	1	0	0	0	1	2
Erythronium	0	0	0	0	3	3
Euonymus	1	0	0	0	5	6
Eupatorium	0	0	0	1	3	4
Eurybia	11	0	0	9	6	26
Fagus	19	0	0	1	4	24
Festuca	0	0	3	0	1	4
Fragaria	10	7	3	5	9	34
Frangula	3	0	0	4	4	11
Fraxinus	16	0	10	0	13	39
Fritillaria	1	0	0	0	0	1
Galium	12	13	5	0	21	51
Gaultheria	11	0	0	1	4	16
Gaylussacia	0	4	0	0	3	7
Gentiana	3	0	0	0	2	5
Gentianella	0	0	0	0	2	2
Geranium	7	1	0	0	4	12
Geum	0	6	4	2	5	0
Glechoma	3	0	0	0	2	5
Gleditsia	1	0	0	0	0	1
Glyceria	3	0	0	0	0	3
Goodyera	5	0	1	0	8	14
Gymnocarpium	6	0	0	0	2	8
Hamamaelis	5	2	0	0	2	9
Helenium	0	0	0	1	0	1
Helianthella	0	0	0	0	1	1
Helianthus	0	0	0	2	1	3
Hepatica	4	0	0	0	3	7
Hesperis	0	0	0	0	1	1
Hieracium	7	6	4	4	4	25
Huperzia	0	0	0	0	4	4
Hybanthus	1	0	0	0	1	2
Hydrasthis	2	0	0	1	0	3
Hydrocotyle	0	0	0	0	2	2
Hydrophyllum	2	0	0	0	0	2
Hypericum	0	1	1	0	2	4
Ilex	7	0	0	0	5	12
Impatiens	5	2	0	0	12	19
Ipomoea	0	1	0	0	2	3
Iris	0	1	0	0	2	3

Table B2—Distribution of observations among match types for each genus *(continued)*

Genus	Match Type 1	2	3	4	5	Total
Itea	0	0	0	0	1	1
Juglans	3	0	0	1	4	8
Juncus	0	1	1	0	5	7
Juniperus	7	0	0	0	4	11
Kalmia	8	0	0	1	0	9
Koeleria	1	0	0	0	0	1
Krigia	0	0	0	0	2	2
Lactuca	1	1	0	0	2	4
Lamium	0	0	0	0	2	2
Laportea	0	0	0	1	4	5
Larix	0	0	0	0	1	1
Lathyrus	0	0	4	2	4	10
Ledum	2	0	0	0	0	2
Leersia	0	0	0	3	1	4
Lespedeza	0	0	1	0	3	4
Leucanthemum	1	0	0	1	2	4
Leucothoe	0	0	0	0	2	2
Lilium	0	1	0	0	1	2
Linaria	0	0	0	0	1	1
Lindera	7	1	0	0	7	15
Linnaea	8	0	0	1	5	14
Linum	0	0	0	0	1	1
Liquidambar	8	0	0	0	1	9
Liriodendron	4	2	0	0	2	8
Listera	2	0	0	0	2	4
Lolium	0	0	0	4	0	4
Lomatium	1	1	0	0	0	2
Lonicera	3	7	8	0	11	29
Ludwigia	0	0	0	0	1	1
Lupinus	2	5	0	0	1	8
Luzula	2	0	0	0	5	7
Lycopodium	11	3	1	2	9	26
Lycopus	1	0	0	1	4	6
Lyonia	0	0	0	4	0	4
Lysichiton	8	0	0	0	0	8
Lysimachia	2	5	0	0	8	15
Maclura	2	0	0	0	0	2
Magnolia	5	0	0	0	4	9
Mahonia	14	0	4	0	6	24
Maianthemum	31	0	1	2	18	52
Malaxis	0	0	0	0	1	1
Malus	5	1	0	0	0	6

Table B2—Distribution of observations among match types for each genus *(continued)*

Genus	Match Type 1	2	3	4	5	Total
Medicago	0	0	0	2	2	4
Melampyrum	0	0	0	0	2	2
Melilotus	5	0	0	0	1	6
Menispermum	1	0	0	0	0	1
Mentha	0	0	0	0	2	2
Menziesia	10	0	0	0	1	11
Mertensia	0	0	0	0	2	2
Mitchella	8	0	0	2	4	14
Mitella	2	0	2	4	7	15
Moehringia	1	0	0	0	0	1
Monarda	0	0	0	1	2	3
Moneses	1	0	0	0	1	2
Monotropa	2	0	0	0	9	11
Morella	0	1	0	0	0	1
Morus	0	0	0	0	8	8
Muhlenbergia	0	0	0	1	0	1
Mycerlis	3	0	0	0	0	3
Nepeta	0	0	0	0	2	2
Nephrophyllidium	4	0	0	0	0	4
Nothoclenone	0	0	0	0	62	62
Nyssa	11	0	0	0	12	23
Onoclea	6	0	0	0	5	11
Orthilia	7	0	0	0	1	8
Oryzopsis	9	0	0	0	3	12
Osmorhiza	12	0	4	1	2	19
Osmunda	4	0	1	0	9	14
Ostrya	3	0	0	0	11	14
Oxalis	5	10	4	0	8	27
Panax	1	0	0	0	1	2
Panicum	0	0	0	1	6	7
Parthenocissus	25	0	0	0	3	28
Passiflora	0	0	0	0	1	1
Paxistima	10	0	0	0	4	14
Pedicularis	2	0	0	0	0	2
Penstemon	0	1	0	1	2	4
Petasites	2	0	0	0	1	3
Phaseolus	0	0	0	2	0	2
Phegopteris	2	0	0	0	3	5
Phellodendron	0	0	0	0	1	1
Phleum	0	0	0	0	6	6
Phlox	0	0	3	0	0	3
Phryma	2	0	0	0	12	14

Table B2—Distribution of observations among match types for each genus *(continued)*

Genus	Match Type 1	2	3	4	5	Total
Physocarpus	2	0	4	0	1	7
Phytolacca	3	0	0	0	2	5
Picea	24	0	5	0	5	34
Pilea	2	0	0	0	6	8
Pinus	40	0	3	0	11	54
Piperia	2	0	0	0	2	4
Piptatherum	0	0	0	0	4	4
Pityopus	0	0	0	0	1	1
Plantago	0	4	0	0	1	5
Platanthera	0	0	2	0	4	6
Platanus	0	0	0	0	1	1
Poa	8	4	3	2	71	88
Podophyllum	1	1	0	0	1	3
Polygala	4	0	0	0	0	4
Polygonatum	1	0	1	4	7	13
Polygonum	2	5	0	0	14	21
Polypodium	5	0	0	0	3	8
Polystichum	17	0	0	0	7	24
Populus	28	1	0	0	10	39
Potentilla	0	11	0	2	9	22
Prenanthes	0	1	0	3	1	5
Prunella	33	0	0	0	2	5
Prunus	33	0	3	6	22	64
Pseudoroegneria	4	0	0	0	7	11
Pseudostellaria	0	0	0	0	3	3
Psuedotsuga	23	0	0	0	1	24
Pteridium	17	1	0	0	6	24
Pyrola	0	2	1	0	6	9
Quercus	78	2	8	0	31	119
Ranunculus	0	1	0	0	6	7
Rhamnus	0	0	0	0	2	2
Rhexia	3	0	0	0	2	5
Rhus	1	0	0	0	2	3
Rhynchosopra	3	0	0	1	0	4
Ribes	9	7	2	0	13	31
Robina	3	0	0	0	0	3
Rosa	25	8	7	2	9	51
Rubus	35	16	14	8	23	96
Rudbeckia	0	0	0	1	1	2
Ruellia	0	0	0	0	4	4
Rumex	0	0	0	0	2	2
Salix	1	9	1	0	3	14

Table B2—Distribution of observations among match types for each genus *(continued)*

Genus	Match Type 1	2	3	4	5	Total
Sambucus	8	0	0	2	9	19
Sanguinaria	1	0	0	0	1	2
Sanguisorba	0	0	0	0	1	1
Sanicula	0	1	0	0	10	12
Sassafras	21	0	0	1	4	26
Saxifraga	0	0	0	4	0	4
Schizachne	0	0	0	0	2	2
Scirpus	2	0	0	0	1	3
Scrophularia	0	0	0	0	2	2
Sedum	0	0	2	0	0	2
Selaginella	0	0	0	0	2	2
Senecio	3	0	0	0	1	4
Silene	0	0	0	0	2	2
Smilax	4	10	1	0	17	32
Solanum	0	0	0	0	2	2
Solidago	0	16	0	2	10	28
Sonchus	0	0	0	0	1	1
Sorbus	3	0	3	1	2	9
Spiraea	13	3	0	1	1	18
Spiranthes	0	0	0	0	1	1
Stachys	3	0	0	0	0	3
Staphylea	0	0	0	0	1	1
Stellaria	0	0	0	4	1	5
Stereophyllum	0	0	0	0	2	2
Streptopus	4	0	0	6	9	19
Strophostyles	0	0	0	0	1	1
Symphoricarpos	17	0	3	0	5	25
Symphyotrichum	3	3	2	1	9	18
Taraxacum	5	0	0	5	4	14
Thalictrum	9	4	0	0	5	18
Thelypteris	1	0	0	0	9	10
Thuja	10	0	0	0	0	10
Tiarella	2	0	0	0	8	10
Tilia	4	0	0	0	5	9
Tofieldia	1	0	0	0	2	3
Toxicodendron	14	0	0	2	6	22
Tragopogon	2	0	0	0	0	2
Trautvetteria	3	0	0	0	0	3
Triadenum	0	0	0	4	1	5
Trichophorum	2	0	0	0	0	2
Trientalis	20	0	0	0	12	32
Trifolium	4	1	0	1	3	9

Table B2—Distribution of observations among match types for each genus *(continued)*

Genus	Match Type 1	2	3	4	5	Total
Trillium	8	2	2	3	2	17
Tsuga	26	0	0	0	1	27
Ulmus	17	0	9	0	8	34
Urtica	2	0	0	0	3	3
Uvularia	3	0	0	3	5	11
Vaccinium	36	26	4	0	14	80
Vancouveria	5	0	0	0	0	5
Veratrum	2	0	1	0	2	5
Verbascum	1	0	0	1	0	2
Verbena	0	0	0	0	1	1
Vernonia	0	0	0	1	2	3
Veronica	0	0	0	0	2	2
Viburnum	9	0	0	0	8	17
Vicia	0	5	0	3	3	11
Viola	7	32	2	0	23	64
Vitis	2	25	0	4	7	38
Waldsteinia	2	0	0	0	3	5
Woodwardia	0	0	0	0	2	2
Wyethia	0	0	0	1	0	1
Xerophyllum	1	0	0	0	0	1
Xyris	0	1	0	0	0	1

Literature Cited

Amacher, M.C.; O'Neill, K.P.; Dresbach, R.; Palmer, C. 2003. Forest inventory and analysis manual of soil analysis methods. Unpublished information on file with: USDA Forest Service, Rocky Mountain Research Station, Forestry Sciences Lab, 860 North 12th East, Logan, UT 84321. 62 p.

Balakrishnan, N.; Nevzorov, V.B. 2003. A primer on statistical distributions. Hoboken, NJ: John Wiley & Sons. 309 p.

Bechtold, W.A.; Patterson, P.L., eds. 2005. 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, Southern Research Station. 85 p.

Caro, T.M.; Roper, R.; Young, M.; Dank, G.R. 1979. Inter-observer reliability. Behaviour. 39(3-4):303-315.

Geiser, L.H.; Neitlich, P. 2007. Air pollution and climate gradients in western Oregon and Washington indicated by epiphytic macrolichens. Environmental Pollution. 145:203-218.

Jovan, S.; McCune, B. 2005. Air-quality bioindication in the greater Central Valley of California, with epiphytic macrolichen communities. Ecological Applications. 15:1712-1726.

Jovan, S.; McCune, B. 2006. Using epiphytic macrolichen communities for biomonitoring ammonia in forests of the Greater Sierra Nevada, California. Water, Air, and Soil Pollution. 170:69-93.

Kuo, S. 1996. Phosphorus. In: D.L. Sparks, ed. Methods of soil analysis. Part 3. Chemical methods. Madison, WI: Soil Science Society of America: 869-919.

Lischke, H. 2001. Propagation of data uncertainty through models. In: Brassel, P.; Lischke, H., eds. Swiss national forest inventory: methods and models of the second assessment. Birmensdorf, Switzerland: WSL Swiss Federal Research Institute. 336 p.

McCune, B.; Dey, J.P.; Peck, J.E.; Cassell, D.; Heiman, K.; Will-Wolf, S.; Neitlich, P.N. 1997. Repeatability of community data: species richness versus gradient scores in large-scale lichen studies. The Bryologist. 100:40-46.

O'Neill, K.P.; Amacher, M.C.; Perry, C.H. 2005. Soils as an indicator of forest health: a guide to the collection, analysis, and interpretation of soil indicator data in the Forest Inventory and Analysis program. Gen. Tech. Rep. NC-258. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station. 53 p.

Pollard, J.E.; Smith, W.D. 1999. Forest health monitoring 1998 plot component quality assurance report. Research Triangle Park, NC: U.S. Department of Agriculture, Forest Service, Forest Health Monitoring Program.

Pollard, J.E.; Smith, W.D. 2001. Forest health monitoring 1999 plot component quality assurance report. Research Triangle Park, NC: U.S. Department of Agriculture, Forest Service, Forest Health Monitoring Program.

Pollard, J.E.; Westfall, J.A.; Patterson, P.L.; Gartner, D.L.; Hansen, M.H.; Kuegler, O. 2006. Forest inventory and analysis national data quality assessment report for 2000 to 2003. Gen. Tech. Rep. RMRS-181 Ft. Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 43 p.

Reams, G.A.; Smith, W.D.; Hansen, M.H.; Bechtold, W.A.; Roesch, F.A.; Moisen, G.G. 2005. The forest inventory and analysis sampling frame. In: Bechtold, W.A.; Patterson, P.L., 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, Southern Research Station: 11-26.

SAS Institute, Inc. 2003. SAS/STAT User's guide, Version 9.1 [computer program]. Cary, NC: SAS Institute, Inc.

Smith, G.C.; Coulston, J.W.; Jepsen, E.; Prichard, T. 2003. A national ozone biomonitoring program – results from field surveys of ozone sensitive plants in northeastern forests (1994-2000). Environmental Monitoring and Assessment. 87:271-291.

USDA Forest Service. 2004a. Forest Inventory and Analysis national core field guide, volume I: field data collection procedures for phase 2 plots, version 2.0. Washington, DC: U.S. Department of Agriculture, Forest Service. Unpublished report on file with: U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis., Washington, D.C., 20250.

USDA Forest Service. 2004b. Forest Inventory and Analysis national core field guide, volume II: field data collection procedures for phase 3 plots, version 2.0. Washington, DC: U.S. Department of Agriculture, Forest Service. Unpublished report on file with: U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis, 201 14th St., Washington, D.C., 20250.

USDA Forest Service. 2007. Forest Inventory and Analysis national P3 field guide, version 4.0. Washington, DC: U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis. http://fia.fs.fed.us/library/field-guides-methods-proc/ (Accessed 3 June 2009).

USDA Natural Resources Conservation Service. 2009. Plants database. Baton Rouge, LA: U.S. Department of Agriculture, Natural Resources Conservation Service, National Plant Data Center. http://plants.usda.gov (Accessed 17 April 2009).

Westfall, J.A.; Woodall, C.W. 2007. Measurement repeatability of a large-scale inventory of forest fuels. Forest Ecology and Management. 253:171-176.

Will-Wolf, S. In press. Analyzing lichen indicator data in the Forest Inventory and Analysis program. Gen. Tech. Report PNW. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station.

Woodall, C.W.; Westfall, J.A. 2008. Taper and relative size models for coarse woody debris in forests of the United States. Canadian Journal of Forest Research. 38:631-636.

Woodall, C.W.; Monleon, V.J. 2008. Sampling protocol, estimation, and analysis procedures for Down Woody Materials indicator of the FIA Program. Gen. Tech. Rep. NRS-22. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station. 68 p.

Yanai, R.D.; Stehman, S.V.; Arthur, M.A.; Prescott, C.E.; Friedland, A.J.; Siccama, T.G.; Binkley, D. 2003. Detecting change in forest floor carbon. Soil Science Society of America Journal. 67:1583-1593.

Westfall, James A., ed. 2009. **FIA national assessment of data quality for forest health indicators.** Gen. Tech. Rep. NRS-53. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Service. 80 p.

The Forest Inventory and Analysis program of the U.S. Forest Service collects field data to assess the condition and trends of the nation's forest resources. A quality assurance program is implemented to assure that data are collected accurately with consistent protocols. A random subset of field plots is chosen to receive an additional, independent measurement by another field crew. This 'blind check' approach allows for comparison of measurements between the two crews. The measurement differences are evaluated against measurement quality objectives, which specify a level of measurement precision for each attribute. This information should be useful to data collection experts, as variables having poor measurement repeatability can be identified and examined for potential resolution. The results may also be of interest to analysts and researchers wanting to evaluate whether the repeatability of measurements is sufficient for their respective studies.

KEY WORDS: measurement error, classification error, quality assurance, forest inventory

Capitalizing on the strengths of existing science capacity in the Northeast and Midwest to attain a more integrated, cohesive, landscape-scale research program

www.nrs.fs.fed.us