

United States Department of Agriculture

Forest Service

Rocky Mountain Research Station

General Technical Report RMRS-GTR-181

October 2006

Forest Inventory and Analysis National Data Quality Assessment Report for 2000 to 2003

James E. Pollard James A. Westfall Paul L. Patterson David L. Gartner Mark Hansen Olaf Kuegler

Pollard, James E.; Westfall, James A.; Patterson,Paul L.; Gartner, David L.; Hansen, Mark; Kuegler, Olaf. 2006. **Forest Inventory and Analysis National Data Quality Assessment Report for 2000 to 2003**. Gen. Tech. Rep. RMRS-GTR-181. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 43 p.

Abstract ___

The Forest Inventory and Analysis program (FIA) is the key USDA Forest Service (USFS) program that provides the information needed to assess the status and trends in the environmental quality of the Nation's forests. The goal of the FIA Quality Assurance (QA) program is to provide a framework to assure the production of complete, accurate and unbiased forest information of known quality. Specific Measurement Quality Objectives (MQO) for precision are designed to provide a window of performance that we are striving to achieve for every field measurement. These data quality goals were developed from knowledge of measurement processes in forestry and forest ecology, as well as the program needs of FIA. This report is a national summary and compilation of MQO analyses by regional personnel and the National QA Advisor.

The efficacy of the MQO, as well as the measurement uncertainty associated with a given field measurement, can be tested by comparing data from blind check plots where, in addition to the field measurements of the standard FIA crew, a second QA measurement of the plot was taken by a crew without knowledge of the first crew's results. These QA data were collected between 2000 and 2003 and analyzed for measurement precision between FIA crews.

The charge of this task team was to use the blind-check data to assess the FIA program's ability to meet data quality goals as stated by the MQO. The results presented indicate that the repeatability was within project goals for a wide range of measurements across a variety of forest and nonforest environments. However, there were some variables that displayed noncompliance with MQO goals. In general, there were two types of noncompliance: the first is where all the regions were below the MQO standard, and the second is where a subset of the regions was below the MQO standards or was substantially different from the other remaining regions. Results for each regional analysis are presented in appendix tables. In the course of the study, the task team discovered that there were difficulties in analyzing seedling species and seedling count variables for MQO compliance, and recommends further study of the issue. Also the task team addresses the issue of trees missed or added and recommends additional study of this issue. Lastly, the team points out that traditional MQO analysis of the disturbance and treatment variables may not be adequate.

Some attributes where regional compliance rates are dissimilar suggest that regional characteristics (environmental variables such as forest type, physiographic class, and forest fragmentation) may have an impact on the ability to obtain consistent measurements. Additionally, differences in data collection protocols may cause differences in compliance rates. For example, a particular variable may be measured with a calibrated instrument in one region, while ocularly estimated in another region.

The Authors

James E. Pollard is FIA Quality Assurance Advisor, University of Nevada, Las Vegas.

James A. Westfall is Research Forester, USDA Forest Service, Northeast Research Station, Newtown Square, PA.

Paul L. Patterson is Mathematical Statistician, USDA Forest Service, Rocky Mountain Research Station, Interior West Research Station, Ogden, UT.

David L. Gartner is Mathematical Statistician, USDA Forest Service, Southern Research Station, Knoxville, TN.

Mark Hansen is Research Forester, USDA Forest Service, North Central Research Station, St. Paul, MN.

Olaf Kuegler is Mathematical Statistician, USDA Forest Service, Pacific Northwest Research Station, PortlandForestry Sciences Lab, Portland, OR.

Forest Inventory and Analysis National Data Quality Assessment Report for 2000 to 2003

James E. Pollard James A. Westfall Paul L. Patterson David L. Gartner Mark Hansen Olaf Kuegler

Introduction___

The Forest Inventory and Analysis program (FIA) is the key USDA Forest Service (USFS) program that provides the information needed to assess the status and trends in environmental quality of the Nation's forests. The FIA program reports on status and trends in forest area and location; the species, size, and health of trees; total tree growth, mortality, and removals by harvest; wood production and utilization rates by various products; and forest land ownership. Recent enhancements to the FIA program have added information relating to tree crown condition, lichen community composition, soils, ozone indicator plants, extent of forest damages, complete vegetation diversity, and down woody material. The major purpose of this program is to provide a scientifically sound census of the Nation's forests that meets the policy and program management needs of the USFS and its partners and constituencies. This is accomplished with the implementation of a nationally consistent sampling design and plot configuration (Bechtold and Patterson 2005). The goal of the FIA quality assurance (QA) program is to provide a framework to assure the production of complete, accurate, and unbiased forest information of known quality.

The heart of the FIA QA program is extensive crew training and nationally consistent protocols and procedures used in the inventory. Quality control (QC) procedures include direct feedback to field staff to provide continual real time assessment and improvement of crew performance. In addition to extensive QC activities, data quality is assessed and documented using performance measurements and post survey assessments. These assessments are used to identify areas of the data collection process that need improvements or refinements to meet the quality objectives of the program. Specific measurement quality objectives (MQO) for precision are designed to provide a window of performance that is allowable for any given field measurement. These data quality goals were developed from knowledge of measurement processes in forestry and forest ecology. The MQO consist of two parts: a compliance standard and a measurement tolerance. A detailed description of the MQO is given in the Methods section. The efficacy of the MQO, as well as the measurement uncertainty associated with a given measurement can be tested using QA data from blind checks. The techniques used to analyze the blind-check data are described in the Methods section.

In some instances, the MQO were established as the "best guess" of what experienced field crews should be able to consistently achieve. However, these measurement quality goals have not been rigorously evaluated for meeting FIA program needs. The results of this report are intended to not only to assess whether the current MQO standards are being met, but more importantly to also provide data collection experts with the information necessary to develop recommendations for changes to the current data collection system. These recommendations must be attribute-specific and derive from intimate knowledge of data collection procedures, extent of training, regional environmental variability, etc. As such, specific recommendations for improvement are beyond the scope of this report.

Analysis of blind-check data has been reported in annual quality assurance reports and FIA State reports (Pollard and Smith 1999, 2001; Frieswyk 2001) and has been summarized in national reports for the Forest Health Monitoring program (Conkling and others 2005). This report is a national summary of Phase 2 (P2) variables and MQO analyses from FIA blind check measurements prepared by regional personnel and the national QA advisor. Details regarding measurement protocols and MQO standards for national and regional attributes can be found in the field guides produced by each FIA region (USDA 2003; USDA 2004b-e). Quality of the Phase 3 (P3) indicator data will be addressed in future reports.

As mentioned above, the task team was not charged with making recommendations for specific variables; but in the process of analyzing the data several items arose that were outside the scope of traditional MQO analysis and the task team would be remiss to not address them and make recommendations. First, analysis using the current MQO for the variables seedling count and seedling species can lead to potentially misleading results; an explanation of the issue is in the Methods section with a further exploration in the Discussion section. The task team recommends that an extensive review of the seedling count and seedling species variables be done. Second, during the development of the process used for MQO analysis the issue of extra and missing trees, conditions, and boundaries arose; an explanation of the issue is in the Methods section with recommendations in the Discussion section. Additional analysis for extra trees is given in appendix C. Third, traditional MQO for disturbance and treatment variables do not provide a complete picture; therefore, additional analysis of the disturbance and treatment variables are given in appendix B.

The primary audience for this report is those intimately familiar with FIA data collection protocols and processes. Although the main audience is data collection experts, users of FIA data may also find the information useful to determine if the repeatability of certain attributes is sufficient to meet their analytical or research needs.

The remainder of the paper is organized into four sections and three appendices. The first section, Methods, starts with a discussion of the process used to assess MQO compliance. This is followed by a discussion of the need for a matching algorithm for the tree-, condition-, boundary-, and seedling-level data and an outline of the matching algorithm used. Lastly, the Methods section contains an explanation of the observations used to analyze the Disturbance and Treatment suites of variables. The Results section contains the observed MQO compliance rates for the P2 core variables and results of the analysis of extra trees, seedlings, boundaries, and conditions. The Discussion section starts with an overview of the observed MQO compliance rates followed by subsections devoted to specific suites of variables. The last section contains the conclusions of the task team. Appendix A contains more detailed results for each of the regions, and appendices B and C are devoted to more detailed analysis of the Disturbance variables and extra trees respectively.

Methods__

Field Data Collection Methods

Data that are used to evaluate crew performance are generated by a second measurement of a field plot termed a blind check. This technique involves the re-installation of an inventory plot performed by a qualified crew without production crew data on hand.

In this report, the first measurement of the plot is referred to as that of the field measurement (FM) crew and the second visit is referred to as that of the quality assessment (QA) crew. This type of QA measurement is considered a "blind" measurement because the FM crews do not know when or which of their plots will be remeasured by the QA crew, and cannot therefore alter their performance because of knowledge that the plot will be measured by a QA crew. In addition, the QA crew does not have knowledge of the FM crews' original measurements because this knowledge might bias their measurements. This type of blind measurement provides a direct, unbiased observation of measurement precision from two independent crews. Blind-check plots are randomly selected to be a representative sub-sample of all plots measured.

Blind-check plots can be measured at any time during the field season or panel completion, but are generally planned to be within a 2-week window of the FM crew measurement to avoid the confounding effects of seasonal changes on the plots. All plot measurements are dated and identified as regular or blind-check status so that the results can be interpreted with reference to length of time between FM crew measurement and the QA measurement. Blind-check plot data used for analysis in this report were generated in 2001-2002 for the Northeast, in 2000-2003 for the North Central, in 2001-2003 for the South, in 2001-2003 for the Interior West, and in 2001-2003 for the Pacific Northwest FIA units. The total number of blind-check plots included in the following analyses varied by region. For example, the Northeast, Interior West, and the Pacific Northwest measured a total of 77, 118, and 64 complete plot remeasurements respectively. The North Central and South regions measured 877 and 194 partial plot remeasurements, respectively, using the criteria of completed sub-plots with a minimum of 15 trees in the data sets. This generally resulted in a larger number of plots being visited in the North Central and South, although the number of trees observed was similar in all regions except in the North Central. Sample sizes for the number of observations for a given variable are included in the result tables (table 1, tables A1-A5).

Data Analysis Techniques

Evaluation of regional and national performance is accomplished by calculating the differences between FM crew and QA crew data. Results of these calculations are compared to pre-established measurement quality objectives, which are documented in the FIA National Field Manual (USDA 2004a). Computation and comparison of the MQO compliance rates of various data elements measured by FIA crews is the primary method of analysis used in this report.

Each data element collected by FIA has been assigned a tolerance or acceptable level of measurement error. These tolerances were selected by experts in FIA measurements based on their estimates of the ability for crews to make repeatable measurements or observations within the assigned tolerance. In the analysis of blind-check data, an observation is within tolerance when the difference between the FM crew and QA crew observations do not exceed the assigned tolerance for that data element. For many categorical elements, the tolerance is "no error," thus only observations that are identical are within tolerance. For example, the tolerance for measurement of tree DBH is +/– 0.1 inch for each 20.0 inches of diameter of a live tree with the MQO for DBH set at 95 percent. The quality of the data is evaluated by comparing the desired rate of differences within tolerance (as a percent of observations) to the MQO. In the example above, the objective for DBH would be that 95 percent or more of the DBH observations are within +/– 0.1 inch for each 20.0 inches of diameter for all trees measured by both FM and QA crews. Results can be displayed as a simple percent of difference calculations that fell within the program tolerances. This percent will be referred to as the observed compliance rate.

Table 1 (Continued). **Table 1** (Continued).

 $\overline{5}$

due to weather conditions.

Table 1 (Continued). T**able 1** (Continued).

Data Matching Algorithms

FIA collects a myriad of data at a number of levels of detail. Data collection protocols must be examined at each of these levels and determinations made on appropriate methods of comparison. The most difficult areas to obtain valid comparisons are tree, condition, boundary, and seedling-level data. For these types of data, there is no unique identifier between FM and QA measurements. As such, data matching algorithms are implemented to help ensure that every paired measurement (FM and QA) is an observation of the same tree, condition, boundary or seedlings. Matching of FM crew and QA crew measurements for plot and subplot-level variables are easily accomplished, because there is a unique identifier between FM crew and QA crew measurements (e.g., plot number).

Tree Level Data—For tree level data, the matching algorithm uses weighted distance functions based on azimuth, horizontal distance from plot center, and tree diameter to find the tree from the FM data that most closely matches the QA data on a given subplot. The weights were used to standardize the contribution of each variable in relation to their assigned measurement tolerance. A two-pass process is used to create the list of matched trees. The first pass matches the trees with the smallest weighted distance. Strict decision rules for maximum allowable distance and species differences are implemented to remove any questionable matches; every FM tree is matched to one QA tree. The trees remaining after the first pass are subjected to a second pass, which uses a more complex distance function that helps to account for relatively large differences in any one of the three matching variables (azimuth, horizontal distance, and diameter). Decision rules regarding distance and species are used to prohibit matches that appear to be invalid. After the second pass, each FM tree and QA tree is on one of three lists: a list of matched trees, a list of unmatched FM trees, or a list of unmatched QA trees. The automated part of the matching program is designed to be conservative. As such, there likely exist valid matches in the remaining unmatched data after the algorithm has been run. Generally, a match may still exist when there is a very large difference in one of the matching function variables due to some type of data entry or measurement error. To avoid bias in the MQO analyses, it is imperative that the lists of unmatched items be scrutinized and matched manually where necessary. The program creates output to assist in identifying situations where a match may exist. The manual matching is done by QA personnel. There are three lists when the automated and manual process is completed: matched trees, extra FM trees, and extra QA trees. This MQO analysis was done using the matched tree list only.

Condition Level Data—The condition-level matching program logic is the same as the tree matching program. The execution is based on the condition class delineation variables, which are condition class number, condition status, reserved status, owner group, forest type, stand size class, regeneration status, and tree density. However, for nonforest conditions, condition class number and condition status are the only variables with recorded values; it is impossible to match forest and nonforest conditions in the automated process. After the automated process, the remaining unmatched conditions are inspected by QA personnel for possible matches. When the automated and manual process is completed, there are three lists: matched condition classes, extra FM condition classes, and extra QA condition classes. This MQO analysis was done using the matched condition class list.

Seedling Level Data—Seedling data are different from trees in that individual seedlings are not identified. Instead, seedlings are aggregated by species, with species and count having separate MQO. After discussion with several regional QA supervisors, it became apparent that there was no satisfactory way to match on seedlings as opposed to match on records; therefore, the MQO analysis is based on matched records, as opposed to matched seedlings. It was decided to undertake three objectives in this document with regard to seedlings: (1) recommend that analysts, statisticians, and field representatives conduct an extensive review of the seedling data collection; (2) with the limitation, analyze the current data for MQO compliance; and (3) present the difficulties in analyzing the current seedling data.

The matching of seedling data is done at the subplot level. The only variables available are species and count records. Records with the same species for the FM crew and the QA crew are matched. Records having no species match are retained for manual matching by QA personnel. That person decides to either match a FM seedling record and QA seedling record with different species (indicating that the QA and FM crews disagreed on the species of a particular seedling or group of seedlings), or leave them in the unmatched records (indicating that the QA crew tallied a species that the FM crew didn't while the FM crew tallied a species that the QA crew didn't). Unmatched seedling records are not matched to seedling records that have already been matched to other records.

Boundary Level Data—Due to the fact that boundaries occur on a relatively small number of subplots (or microplots), boundary matching is currently being completed manually.

Disturbance and Treatment Attributes

As noted in the Condition Level Data description within the Data Matching Algorithms section, the MQO analysis of condition level variables is done using the matched condition classes list. As a reminder, a matched condition class is a FM condition class and a QA condition class that, through a combination of automated and manual matching, are deemed to represent the same condition on a plot. This subsection contains a discussion of the matched condition classes that were used to analyze the Disturbance and Treatment variables for MQO compliance. The definitions of the matched condition classes that were used in the analysis of Disturbance and Treatment variables are summarized at the end of this subsection.

The FIA national core field guide says that Disturbance 1 through 3 should be collected for all conditions where the crew records a Condition Class Status = 1 (i.e., Forested). So the natural set of matched condition classes to use for the MQO analysis is the set where both the FM crew and QA crew record Condition Class Status = 1. Using this set of matched condition classes to analyze Disturbance 2 and Disturbance 3 yields artificially high MQO compliance percentages; if both crews record no observable disturbance for Disturbance 1, then clearly both crews will record no observable disturbance for Disturbance 2. The set of matched condition classes that should be used is the set where there is a possibility of there being an observable second disturbance, and this is the set of matched condition classes where at least one of the crews recorded an observable disturbance for Disturbance 1. Continuing with this reasoning, the set of matched condition classes that was used for Disturbance 3 were those where at least one of the crews recorded an observable Disturbance 2.

Similarly a restricted set of matched condition classes was used for analyzing Disturbance Year 1, Disturbance Year 2, and Disturbance Year 3. Both the FM crew and QA crew must record an observable disturbance for there to be any possibility of agreement on the year of the disturbance. So for Disturbance Year 1, the set of matched condition classes used was where both crews recorded an observable disturbance for Disturbance 1. Similar sets of matched condition classes were used for Disturbance Year 2 and Disturbance Year 3.

To analyze the suite of Treatment variables (Treatment 1-3 and Treatment Year 1-3) a similar set of restricted sets of matched condition classes was used. For example, the MQO compliance for Treatment 2 was checked using the set of matched condition classes where at least one of the crews recorded an observable Treatment 1.

These restricted sets of matched condition classes allow one to analyze the conditional agreement between the crews (e.g., how often does either the FM crew or the QA crew record no observable disturbance while the other crew records an observable disturbance). This type of analysis is explored in appendix B.

Description Summary of the Restricted Set of Matched Condition Classes Used in Disturbance and Treatment MQO Analysis*—*The description is limited to Disturbance and Disturbance Year, and the explanation for Treatment and Treatment Year is identical.

Disturbance 1: Matched Condition classes where both the FM crew and the QA crew recorded Condition Status = 1.

Disturbance 2: Matched condition classes where at least one of the crews recorded an observable disturbance for Disturbance 1.

Disturbance 3: Matched condition classes where at least one of the crews recorded an observable disturbance for Disturbance 2.

Disturbance Year 1: Matched condition classes where both the FM crew and the QA crew recorded an observable disturbance for Disturbance 1.

Disturbance Year 2: Matched condition classes where both the FM crew and the QA crew recorded an observable disturbance for Disturbance 2.

Disturbance Year 3: Matched condition classes where both the FM crew and QA crew recorded an observable disturbance for Disturbance 3.

Results___

Core Variable Results

Phase 2 variables collected in all FIA regions using common protocols (Core Variables) were included in this analysis. The use of nationally consistent protocols and standards allows for comparability among FIA regions. The results of MQO achievement for tree, seedling, subplot, plot, boundary, and condition variables are presented in table 1. For each FIA region, results are expressed as the percentage of values observed that fell within the MQO tolerances established for the program.

Regional add-on variables are also collected by each FIA unit, but these regional variables cannot be evaluated from a national perspective. However, the programs used to analyze core variables were also used to compute results for regional variables. These results have been included in appendix A, where each region's analyses are tabled with MQO achievement percentages at one to four times the measurement tolerance levels. These results can be used to evaluate how wider tolerances would affect MQO compliance rates for these regional variables as well as for national core variables measured in each region.

Extra Trees

Following the careful matching of all trees tallied by both crews, any trees tallied by the QA crew and not by the FM crew were identified as extra QA trees, and any trees tallied by the FM crew and not the QA crew were identified as extra FM trees. (The analysis of extra trees did not include data from PNW due to issues with using the macroplot.) The number of extra trees for both crews is given in table 2. Besides the overall total number of extra trees by region, table 2 also contains subtotals of the **Table 2—Tabulation of extra trees observed by FM and QA crews in each region^a.**

a PNW data not available for extra tree analysis.

number of extra saplings and 5-inch+ trees by region. Extra tree information can be used to look for bias in FM crew tallies of trees relative to QA crews using methods and assumptions described in Appendix C. These relative bias values are shown in fig. 1. The values are all small and none are significantly different from zero at the $p = 0.05$ level.

The percentages given in tables 2 through 5 are percents of the observations made by a crew that were not matched. The equation for the percent unmatched observation is:

$$
Percent_Unmatched = \left[\frac{Number_of_extras}{(Number_of_extras + Number_of_matches)} \right] * 100
$$

The percent unmatched for the total number of observations for both crews is a weighted average of the percent unmatched for the two crews. However, the unweighted average will be accurate enough for most uses.

Extra Seedlings

As with the trees, after the FM and QA seedlings records were matched there were extra seedling records tallied by both the FM and QA crews. Table 3 contains the number of extra seedling records by region (except PNW). The rate of extra seedlings is greater than that for trees because of the inherit difficulty in matching the seedling records where the species and count are confounded. A detailed exploration of the analytical problems associated with this type of QA data is included in the Discussion section.

Extra Boundaries and Extra Conditions

Following the careful matching of all boundaries tallied by both crews, any boundaries tallied by the QA crew and not by the FM crew were identified as extra QA boundaries, and any boundary tallied by the FM crew and not the QA crew were identified as extra FM boundaries. The number of extra boundaries for both crews is given in table 4. Similarly, table 5 contains the number of extra conditions by region. The PNW boundary and condition data were not available for this analysis.

Figure 1—Observed bias in the probability that an FM crew will tally an extra tree relative to the probability that a QA crew will tally an extra tree. None of these relative bias values are significantly different from zero at the p = 0.05 level.

Table 3—Tabulation of extra seedling records for FM and QA crews in each region^a.

a PNW data not available for extra seedling analysis.

Table 4—Boundaries: extra records for QA crews and FM Crews^a.

^a PNW data not available for extra boundary analysis.

	NE	NC	SO	IW	Total
Matched records	115	1.191	262	141	1,709
QA extra records	4 (3.4%)	32 (2.6%)	2 (0.8%)	8 (5.4%)	46 (2.6%)
FM extra records	6 (5.0%)	18 (1.5%)	4 (1.5%)	(0.7%)	29 (1.7%)

Table 5-Conditions: extra records for QA crews and FM crews^a.

^a PNW data not available for extra condition analysis.

Discussion

The charge of this task team was to use the blind-check data to assess the FIA program's ability to meet its stated MQO. The results of the analysis of the MQO compliance are in table 1, with the more detailed regional tables in appendix A. The basic results for extra trees, extra seedlings, extra boundaries and extra conditions are in tables 2 through 5. While the discussion in this section will be limited to the above mentioned charge, additional analysis of the data is presented in appendices B and C. The extra analyses will be mentioned at the appropriate places in the discussion below.

Overall MQO Compliance

Two types of noncompliance will be discussed: the first is where all the regions were below the MQO standard, and the second where a subset of the regions was below the MQO compliance of the other remaining regions.

The results indicate repeatability within MQO goals for a wide range of measurements across a variety of forest and nonforest environments. In some cases, this range of environments appears to have little effect on measurement precision (i.e., the percent MQO compliance is similar across all regions). However, there are some attributes where regional compliance rates are quite dissimilar, suggesting that regional characteristics may have an impact on t he ability to obtain consistent measurements. These characteristics may include environmental variables such as forest type, physiographic class, and forest fragmentation. Additionally, differences in data collection protocols may also cause differences in compliance rates. For example, a particular variable may be measured with a calibrated instrument in one region, while ocularly estimated in another region.

Rates of repeatability for tree-level variables (table 1) illustrate the aforementioned circumstances. MQO compliance rates are consistently high across all regions for azimuth, species (and genus), tree status, rotten/missing cull, and cause of death. Decay class and DBH had fairly consistent results across all FIA regions, with most regions just below the MQO standard. There are a number of variables for which compliance rates notably vary by region. Horizontal distance compliance is consistent in the eastern U.S. (NE, NC, and SO), but is lower for the western regions (IW and PNW). This is due to the necessity of measuring woodland species in the western regions, where horizontal distance to multi-stemmed trees is based on geographic center. This makes the measurement less repeatable. As an illustration, timberland species in the IW region had a compliance rate for horizontal distance is 95 percent, while woodland species compliance was 68 percent (appendix A). The poor compliance rates for DRC are also attributable to woodland species issues.

Other attributes that have range of repeatability statistics are related to tree height. For tree heights, MQO compliance is poorest in the NE, somewhat better in the SO, IW, and PNW, and best in the NC. For the three tree-crown variables, all regions are below compliance, although there is some variation of the compliance rates among regions for a specific variable and variation of the compliance rates among the three variables for a specific region. The reason for these variations is not apparent; however, differing forest conditions or measurement methods may provide an explanation.

Seedlings

Under the limitation of matching on seedling records, the seedlings results from table 1 show that the Northeast, North Central and Interior West FIA units are making the MQO compliance standard for species but not for counts. On the other hand, the Southern unit is making the MQO compliance standard for counts but not for species. At first glance, these results suggest that the seedling quality issues for the Southern unit are fundamentally different from those of the Northeast, North Central, and Interior West. However, these apparent differences may be an artifact of matching on records.

If there are species with relatively few seedlings per microplot, then differences in species calls between the FM crew and the QA crew will lead to a low observed species compliance rate. Example 1 demonstrates this situation. If there is only one maple seedling and the FM crew calls it a red maple and the QA crew calls it a sugar maple, the two seedling records will be manually matched. This combination would be interpreted as one seedling record with a count agreement and a species call disagreement.

However, if the species have more seedlings per acre, then differences in species call will lead to a low observed count compliance rate. Simply adding two red maple seedlings and two sugar maple seedlings to each crew count illustrates how this happens. The result will be Example 2, where the FM crew and the QA crew have agreed on the species of four of the five seedlings, two red maple seedlings and two sugar maple seedlings, but disagree on the species of the fifth seedling. The records with the same species will be matched. Therefore, this set of records will be interpreted as two records with species call agreements, but with counts disagreements.

Therefore, the low observed count compliance rates for the Northeast, North Central, and Interior West units may have the same root causes as the low observed species compliance rate for the Southern unit.

In addition to the difficulty in interpreting the MQO results due to the fact that species call disagreements can affect observed compliance rates for both species and count, is the difficulty caused by the large percentages of unmatched seedling records (5.8 percent-16.6 percent). Because of these large percentages of unmatched records, additional caution should be used in interpreting the observed compliance rates. The FIA units are meeting one of the seedling MQO compliance standards in table 1 as

percents of matched records. However, if the compliance standards were stated in terms of percent of total records instead of percent of matched records, only the Interior West would be meeting either of the seedling MQO compliance standards—the compliance standard for counts—and they would be just making this compliance standard with an observed compliance rate of 85 percent.

Plot Level Variables

There are few plot-level variables that lend themselves to meaningful MQO analyses. The distance to road attribute has both types of repeatability issues; all regions are below the MQO standard, and there is disparity between the level of repeatability for the eastern regions versus the western regions. Repeatability in the eastern regions was much greater than in the western regions. It is suspected that poorer compliance for western regions is due to the further distances to roads. Statistics for the water on plot variable are close to the compliance standard of 90 percent for all regions except the NE. One possible explanation is that temporary water is more difficult to identify in NE forests.

Subplot Level Variables

Most subplot-level measurements had high levels of repeatability that were near or above the stated compliance standards. Differences between crews for subplot and microplot center conditions are attributable to differences in condition delineation arising from the use of mapped plots. Subplot slope was above the compliance standard for all regions except SO. The most problematic subplot attribute was aspect, with eastern regions having better repeatability than the western regions. It is likely that the more variable topography in Western States presents additional difficulty. Snow and water depth have no MQO standard, but there is clearly a disparity between NC and the other regions, because the majority of the forest land in the NC region is in the northern portions of Michigan, Minnesota, and Wisconsin. FM crews measure plots year-round in these areas and snow is often present on these plots. In other regions snow is typically not present when measurements are taken, which would account for the disparity between NC and the other regions.

Boundary Level Variables

Boundary data are collected when more than one condition class occurs on a sample subplot. This occurs relatively infrequently, so the sample sizes for these analyses are somewhat small. For the NE and IW the numbers of observations were too small to make meaningful statistical comparisons. Variables whose national average exceeds the standard are boundary change and contrasting condition. Left azimuth and right azimuth compliance rates are similar and near the desired compliance level within NC, SO and PNW. Boundary corners are less repeatable than azimuths. Obtaining agreement on the existence and location of a corner point is difficult because the exact path of boundary line is often poorly defined. Note that if one crew observed a corner and the other crew did not, then this was considered an out-of-compliance situation.

Condition-Level Variables

The statistics for the condition-level variables indicate that repeatability can vary widely, depending on the attribute measured. Variables that were consistently close to or above the standard were condition status, reserve status, owner group, regeneration status, regeneration species, tree density, owner class, and owner status. The compliance rates for forest type and forest type group were below the specified standard, but fairly

consistent across regions with the exception of NE, which had poorer repeatability. The NE region extends from Maine to West Virginia, and westward to Ohio, resulting in diverse forests with a large number of forest types, many of which have similar species compositions. The difficulty in identifying particular forest types for the NE is depicted by comparing the increases in percent compliance between forest type and forest type group variables. The increase for the other regions ranges from 2-8 percent, but the increase is 32 percent for the NE, indicating that most of the disparity arises from disagreement among forest types within a forest type group.

Stand size exhibits both noncompliance for all regions and a wide variability across the regions. Of all variables analyzed, stand age exhibited the widest range of compliance statistics. The overall average was 66 percent, with a range of 37 percent for the NE to 94 percent for the IW. These differences are likely the result of differing forest conditions and measurement protocols. There are differences among regions in numbers of trees sampled to obtain stand age. Also, some regions monument the location of trees used to determine age, which allows QA crews to sample the same trees. In other regions, the FM crew and QA crew sample trees independently, which can contribute to poor compliance rates, especially in areas where uneven-aged forests are commonplace. Finally, large trees can be difficult to bore and obtain accurate ring counts, a situation that occurs most frequently in the PNW region. Physiographic class is another example of differences in eastern and western regions. Compliance rates are reasonably similar within the eastern and western regions, but differ notably between the regions. This may be attributed to the more variable landscape of the Western United States.

Disturbances and Treatments

For disturbances and treatments, there are three variables to indicate the type of activity and each has an associated variable to indicate the year of occurrence. The results show that for Disturbance and Treatment most regions are within compliance, but some regions are out of compliance for Disturbance Year and Treatment Year. Much of the compliance for Disturbance and Treatment comes from the crews agreeing that no Disturbance or Treatment has occurred, while for Disturbance Year and Treatment Year the compliance is determined over the smaller set where both crews have recorded a Disturbance or Treatment respectively.

As indicated in the Methods section, subsets of the matched condition classes were used for the analysis of the MQO compliance for Disturbance, Disturbance Year, Treatment, and Treatment Year. Using the these subsets, additional analysis can be done to address questions such as how repeatable is the measurement of an observable disturbance. Since this is not an analysis of MQO compliance, the results are documented in appendix B. The analysis in the appendix shows that the disturbance and treatment variables should be reviewed from perspectives other than MQO compliance.

Extra Trees

One aspect of tree-level measurement that is not specifically addressed by MQO standards is the numbers of trees sampled. Missed or extra trees can bias inventory estimates. Additional analyses found that these extra tally trees were usually those near plot edges or with tree diameters near the 5-inch threshold. Also, differing calls on forest or nonforest status resulted in extra tally trees. Additional analysis of the extra tree data was done to more fully explore the reasons why crews miss or find extra trees (appendix C). Given the assumptions and analysis methods in appendix C, the results of a nonparametric bootstrap test show that overall regional differences in numbers of trees sampled between QA and FM crews are not significant. The analysis performed by reason within each region indicated that there are small (<0.20 percent) but significant ($p = 0.05$) relative biases due to missed trees near plot edges and forest/ nonforest status in some regions. It is suggested that some type of quality assurance standard be implemented for sample tree counts in the future.

Extra Boundaries and Extra Conditions

Two points are worth noting. First, for both boundaries and conditions the rates of extra observations is quite different for the Southern region than for the other regions. While the NE, NC, and IW have similar overall rates of extra observations, the distribution between QA extra observations and Crew extra observations is different among the three regions. Second, for boundaries there is clearly a high rate of extra observations. A study should be conducted to determine the reasons for and solution to these high rates. The small sample size associated with three of the regions boundary data makes it difficult to interpret. It appears, however, that there are considerably more occurrences of extra boundary calls than there are extra condition calls.

Effects of Having Previous Cycle Information Available

The data analyzed in this report consist of a mixture of observations from new plots (plots installed at locations where there were no previous FIA plots), remeasured plots (plots that are a remeasurement of the national 4-subplot mapped design), and overlaid plots (plots where the national design is being installed at a location where a different plot design was previously measured). As the national annual inventory is fully implemented, most plots will be remeasurements of the national design, and most crews will have some of the data from the first cycle of mapped plots available to them. The actual data available to them will vary from unit to unit, affecting how the MQO data are interpreted.

The North Central unit's current procedures and the Southern unit's plan on how information from the previous cycle will be used for boundaries is presented as an example of how to deal with the above situation. If during the second cycle of mapped plots, the cruiser finds a condition boundary, the cruiser checks to see if the boundary existed during the first cycle of mapped plots. If the boundary existed during the first cycle, the cruiser checks to see if the boundary has changed. If the boundary has not changed, the cruiser accepts the boundary data from the first cycle. Therefore, the QA data become partly a test on the repeatability of the agreement with the data from the previous cycle and not just a test of the repeatability of the measurement and placement of the boundary. While there is a regional variable for boundary change (is it a real change, or a cruiser error, or no change), there will be similar variables (forest type, tree grade, and cubic foot cull) for which there is no regional change variable. The MQO analysis of the data coming from the second cycle of mapped plots will need to try to separate the repeatability of the measurements from the repeatability of agreeing with the previous cycle's data.

Conclusions___

The results presented above indicate that the P2 variables analyzed were generally repeatable within Measurement Quality Objectives for a wide range of measurements across a variety of forest and nonforest environments. However, there were some variables that displayed noncompliance with MQO goals. This was particularly evident with seedling variables. Because of the ambiguity inherent in having the two different levels of observation in the same record for seedlings, the decision on how to evaluate data quality for seedlings was not straightforward. There is a definite need to reestablish less ambiguous MQO standards for seedlings.

In addition, it is recommended that data collection experts, as well as data users, review the results of this study to determine if the current level of data quality is adequate for the FIA program needs. These results need to be evaluated to determine if actions are needed to improve MQO compliance for some inventory variables, or if new measurement quality goals need to be formulated. If corrective action is needed to improve MQO compliance, this may include enhanced training methods, revamping data collection protocols, and respecification of MQO criteria to reflect more realistic repeatability standards.

In addition to some of the more obvious factors that affect MQO compliance rates, as noted above, there are also other more difficult to assess factors that may affect repeatability of crew performance. Examples might include calibration of measurement equipment, measurement protocols, crew experience, and seasonal variability of environmental and forest conditions. The results presented in this report form the basis for evaluation of the adequacy of the P2 FIA data.

References__

- 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.
- Conkling, B.L.; Coulston, J.W.; Ambrose, M.J. 2005. Forest health monitoring 2001 national technical report. Gen. Tech. Rep. SRS-81. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 204 p.
- Frieswyk, T.S. 2001. Forest statistics for Maryland: 1986 and 1999. Resour. Bull. NE-154. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station. 164 p.

Pollard, James E.; Smith, W.D. 1999. Forest health monitoring 1998 plot component quality assurance report, volume I. Research Triangle Park, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station.

- Pollard, James 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, Southern Research Station.
- U.S. Department of Agriculture. 2003. Forest inventory and analysis Southern Research Station field guide, volume 1: field data collection procedures for phase 2 plots, version 2.0. Ashville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. Available: www.srs.fs.usda.gov/fia/manual, [2006, March 23].
- U.S. Department of Agriculture. 2004a. Forest inventory and analysis national core field guide, volume I: field data collection procedures for phase 2 plots, version 2.0. Internal report. On file at: U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis, Washington, DC.
- U.S. Department of Agriculture. 2004b. Forest inventory and analysis national core field guide, volume 1: field data collection procedures for phase 2 plots, version 2.0. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station. Available: http://www.fs.fed.us/ne/fia/ datacollection/manualver2_0/NEp2%202.0_0-8%20Ssections.pdf [2006 March 28].
- U.S. Department of Agriculture. 2004c. Interior West forest inventory and analysis forest survey field procedures, version 2.0. Ogden, UT: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Forestry Sciences Laboratory. Available: http://www.fs.fed.us/rm/ogden/manual/ pdf/2004_manual.pdf [2006, March 23].
- U.S. Department of Agriculture. 2004d. Forest inventory and analysis national core field guide: Volume I: Field data collection procedures for phase 2 plots, version 2.0. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station. 281 p. Available: http://www.ncrs.fs.fed. us/4801/field-guides/Annual/NC-2004-field-guide.pdf [2006, March 28].
- U.S. Department of Agriculture. 2004e. Field instructions for the annual inventory of Washington, Oregon, and California, version 2.0. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Forest Inventory and Analysis. Available: http://www.fs.fed.us/pnw/fia/local-resources/pdf/field_manuals/2004_annual_manual_final.pdf [2006, March 28].

Appendix A: Results of Regional Analyses at 1 to 4 times the MQO tolerance levels__

Table A-1 (Continued)

Table A-1 (Continued)

Right azimuth ±10 degrees 68.8 75.0 81.3 81.3 5 4 3 3 16

÷,

 ∞

 ∞

4

LO

 $\ddot{\cdot}$ J, $\frac{5}{7}$ 115 é Ľ Í ś $\tilde{\mathbf{z}}$ $\frac{1}{1}$ Table A-1 (Continued)
Q**A** results for copr

Table A-1 (Continued)

QA results for seedling-level variables from 1,404 microplots in the North Central QA results for seedling-level variables from 1,404 microplots in the North Central
FIA

Percentage of data within tolerance Number of times data exceeded tolerance

Percentage of data within tolerance

Number of times data exceeded tolerance

QA results for boundary-level variables from 170 boundaries (both subplot and microplot boundaries in the NC Region) **QA results for boundary-level variables from 170 boundaries (both subplot and microplot boundaries in the NC Region)**

Table A-2 (Continued) **Table A-2** (Continued)

QA results for subplot-level variables from 3,497 subplots in the North Central

QA results for condition-level variables from 1,191 conditions in the North Central FIA **QA results for condition-level variables from 1,191 conditions in the North Central FIA**

 $\frac{4}{3}$ Table A-3-South data analysis for core and regional variables. **Table A-3**—South data analysis for core and regional variables. å $\ddot{\cdot}$ \cdot 1 201 tr J. ciabla J, Ń $\frac{3}{4}$ $\ddot{\cdot}$

Seedling count ±20 % 69.6 71.8 72.4 73.3 139 129 126 122 457 Genus No tolerance 99.6 2 457

 $\frac{1}{72.4}$

 $\frac{8}{71.8}$

Seedling count

Genus

 $\sqrt{73.3}$

 122

 $\frac{126}{2}$

 $\sqrt{29}$

 \sim

Corner distance ±1 ft 57.1 57.1 57.1 57.1 3 3 3 3 7 Right azimuth ±10 degrees 88.2 94.1 94.1 94.1 6 3 3 3 51

 $\frac{1}{571}$

 $±10$ degrees ± 1 ft

 $\overline{51}$ $\overline{ }$

 \sim \sim

QA results for condition-level variables from 265 conditions in the South FIA **QA results for condition-level variables from 265 conditions in the South FIA**

Table A-3 (Continued)

Table A-3 (Continued)

Table A-4-Interior West data analysis for core and regional variables. **Table A-4**—Interior West data analysis for core and regional variables.

J.

QA results for plot-level variables from 118 plots in the Interior West FIA **QA results for plot-level variables from 118 plots in the Interior West FIA**

Table A-4 (Continued)

Table A-4 (Continued)

Right azimuth \sim 1000 \sim 100

100.0

 $±10$ degrees

Right azimuth

 ∞

Š. J. l. į N -é J, \overline{a} ू
इंग्ल Ń Ι, $\tilde{\zeta}$

Table A-4 (Continued)

Table A-4 (Continued)

Table A-5-West Coast data analysis for core and regional variables. **Table A-5**—West Coast data analysis for core and regional variables.

 L eft azimuth \sim 10 d \sim 10 d \sim 22 $\$ CORNER AZIMUTH = 20 degrees 72.2 degrees 72.2 degrees 72.2 degrees 72.2 degrees 72.4 20.9 86.2 87.2 f 21.2 deg Corner distance ±1 ft 65.5 72.4 72.4 79.3 10 8 8 6 29 B2 degrees 10 degrees 10 degrees 10 degrees 101.0 91.0 91.0 1 degrees 72.2 degrees 72.2 degrees 72.4 7 degrees 1

94.2

 $±10$ degrees ±10 degrees

Left azimuth

 \circ

 $\alpha |\omega |\infty | \sim$

 $|\lambda| \infty$ |이상

 $\frac{36.3}{86.2}$

 $\frac{86.2}{72.4}$ $\frac{1}{93.0}$

 $\frac{3|8|}{8|2|4}$

 $\frac{72.4}{65.5}$ $\frac{74.4}{ }$

35.3

91.9

72.1

±10 degrees \pm 1 ft

Corner distance Corner azimuth

Right azimuth

4

4 4

LO $\overline{\mathcal{A}}$ ∞ $\pmb{\circ}$

USDA Forest Service Gen. Tech. Rep. RMRS-GTR- 1 8 1.2006 3

Table A-5 (Continued).

Table A-5 (Continued).

QA results for plot-level variables from 69 plots in California, Oregon and Washington

QA results for plot-level variables from 69 plots in California, Oregon and Washington

Number of times data exceeded

and Washington **QA results for condition-level variables from 61 conditions in California, Oregon and Washington** conditions in California Oregon Table A-5 (Continued).
OA results for condition-level variables from 61

Table A-5 (Continued).

Appendix B: Additional Analysis of Disturbance and Treatment____________

The sets of matched condition classes used in analyzing the Disturbance, Disturbance Year, Treatment, and Treatment Year facilitate additional analysis of the Disturbance and Treatment variables. This will be illustrated by analyzing the Disturbance 1 and Treatment 1 variables. A short discussion section follows the analysis. The definition of the sets of matched condition classes used in analyzing Disturbance, Disturbance Year, Treatment, and Treatment Year are in the main body of the report (see Disturbance and Treatment Attributes in the Methods section); for ease of reference we repeat the definitions at the end of this appendix.

For easy reference, the pertinent sections of the number of matched condition classes from table 1 have been reproduced in table B-0.

The format of this analysis is to pose a question about the Disturbance 1 variable and then explain how to use the data in table 1 (or table B-0) to answer the question. The explanation of the analysis will be followed by examples of the calculations. Lastly the results of the analysis are presented in a table. The same questions can also be asked about the Treatment 1 variable; the results of the analysis for the Treatment 1 variable will be presented without any explanation.

Table B-0—The number of matched condition classes used in the analysis of MQO compliance. This is replication of entries from the Condition Variables section of Table 1, the columns labeled "# OBS."

	Region					
	NE	NC	SO	IW	PNW	
Disturbance 1	63	921	157	111	48	
Disturbance Year 1	3	40			3	
Disturbance 2	5	73	14	19	12	
Treatment 1	63	921	157	111	48	
Treatment Year 1	8	63	13	2	2	
Treatment 2	10	89	22	4		

Question 1: For what percentage of matched condition classes do both crews record no observable disturbance? This can also be stated as: how much of the MQO compliance is due to both crews agreement that there was no observable disturbance?

Analysis 1: It is the ratio of the number of matched condition classes where both crews did not record observable disturbance divided by the number of matched condition classes where a Condition Status $= 1$ was recorded. The number of matched condition classes where both crews did not record an observable disturbance is the difference between the number of matched condition classes where both crews could have recorded an observable disturbance (i.e., both crews recorded a Condition Status = 1) and the number of matched condition classes where at least one of the crews recorded an observable disturbance (i.e., # Obs Disturbance 2); in equation form this difference is $[(# \t{Obs}$ Disturbance 1) minus (# Obs Disturbance 2)]. Lastly we need to convert the ratio to a percentage. Summarizing,

100* [(# Obs Disturbance 1) minus (# Obs Disturbance 2)]/ (# Obs Disturbance 1) Examples 1: Calculations based on entries from Table B-0.

Interior West: $100*(111-19)/111 = 83\%$

Northeast: $100(63-5)/63 = 92\%$

		Region					
	NE	NC	SO	IW	PNW		
Disturbance 1	92	92	91	83	75		
Treatment 1	84	90	86	96	85		

Table B-1—Of the matched condition classes where both crews recorded a condition status 1, the percentage where both crews recorded no disturbance/treatment.

Question 2: What percentage of the matched condition classes where at least one crew recorded an observable disturbance did both crews record an observable disturbance?

Analysis 2: The matched condition classes where both crews recorded an observable disturbance is the set of matched condition classes used in the MQO analysis of Disturbance Year 1; while the matched condition classes where at least one crew recorded an observable disturbance is the set of matched condition classes used in the MQO analysis of Disturbance 2. So the percentage we want is the ratio of these two numbers converted to a percentage:

100 * (# Obs Disturbance Year 1)/ (# Obs Disturbance 2) Examples 2: Calculations based on entries from table B-0. Interior West: $100*(7/19) = 37\%$ Northeast: $100*(3/5) = 60\%$

Table B-2—Of the matched condition classes where at least one crew recorded an observable disturbance/treatment, the percentage where both crews recorded an observable disturbance/ treatment.

	Region					
	NE	NC	SO	IW	PNW	
Disturbance 1	60	55	50	37	25	
Treatment 1	80		59	50	29	

Question 3: For what percentage of the matched condition classes where both crews record an observable Disturbance 1 do the crews agree what the disturbance is?

Analysis 3: It is easier to calculate the percentage where they disagree and then subtract from 100. The number of matched condition classes where the crews disagreed on the value of Disturbance 1 is the 1X value in the regional tables (tables A-1 through A.5 in appendix A). The crews disagree on the value of Disturbance 1 for two reasons: first, one crew records an observable disturbance and the other crew does not record an observable disturbance; and second, both crews record an observable disturbance but they disagree on what the disturbance is. The number of matched condition classes where the first reason occurs is (# Obs Disturbance $2 - #$ Obs Disturbance Year 1); this number is subtracted from $1X$ to obtain the number of matched condition classes where both the crews recorded an observable disturbance but they disagreed as to what the disturbance was. So the answer of the percentage where they disagree is:

100 * $[(# \text{Obs for } 1X) - \{(# \text{Obs Distance 2}) - (\# \text{Obs Distance Year 1})\}] / (\# \text{Obs Distance Year 1})$ Obs Disturbance Year 1)

The percentage where the crews agree is 100 minus the above number.

Examples 3: Calculations are based on entries from table B-0 and the cell determined by line Disturbance 1 and column 1X from appendix A. Interior West: $100-100*(15-(19-7))/7 = 57%$ South: $100-100*(8-(14-7))/7 = 86%$

Table B-3—Of the matched condition classes where both crews recorded an observable disturbance/treatment, the percentage where both crews recorded the same disturbance/treatment.

Discussion of Results

In table 1, the MQO compliance for the Disturbance and Treatment variables is generally good. The above analysis shows that most of the MQO compliance is because of agreement that there is no disturbance/treatment (see table B-1). In four of the regions, if both crews agree that there was a disturbance/treatment, then there is good agreement between the crews as to the type of disturbance (see table B-3). There is poor agreement on whether an observable disturbance/treatment occurred when we restrict to the condition classes where at least one of crews called an observable disturbance/treatment (table B-2).

Summary of the Restricted Set of Matched Condition Classes Used in Disturbance and Treatment MQO Analysis—The description is limited to Disturbance and Disturbance Year, and the explanation for Treatment and Treatment Year is identical.

Disturbance 1: Matched Condition classes where both the FM crew and the QA crew recorded Condition Status = 1.

Disturbance 2: Matched condition classes where at least one of the crews recorded an observable disturbance for Disturbance 1.

Disturbance 3: Matched condition classes where at least one of the crews recorded an observable disturbance for Disturbance 2.

Disturbance Year 1: Matched condition classes where both the FM crew and the QA crew recorded an observable disturbance for Disturbance 1.

Disturbance Year 2: Matched condition classes where both the FM crew and the QA crew recorded an observable disturbance for Disturbance 2.

Disturbance Year 3: Matched condition classes where both the FM crew and QA crew recorded an observable disturbance for Disturbance 3.

Appendix C: Explanation and Examples of Missing/ Extra Tree Analyses

The FIA plot is designed to measure all trees on a fixed area. If trees that are on the plot are not measured (missed trees), the sample underestimates the true population total. If trees that are not on the plot are included in the sample (extra trees), the sample overestimates the true population total. The inclusion of trees on a sample plot will never be perfect; however, missed and extra trees can be kept to a minimum with good field procedures, and estimates will be unbiased if the rate of missed trees is equal to that of extra trees.

The blind checks can be used to estimate the probability that a FM crew will miss a tree or tally an extra tree. Following the careful matching of all trees tallied by both crews, any trees tallied by the QA crew and not by the FM crew were identified as extra QA trees and any trees tallied by the FM crew and not the QA crew were identified as extra FM trees.

Estimates of the probability of a FM crew missing a tree or tallying an extra tree are based on the following assumptions:

- 1. All trees tallied by both crews are truly on the plot and should be tallied (the probability that both crews will tally a tree that is truly not on the plot is equal to zero).
- 2. All trees that are truly on the plot will be tallied by at least one of the two crews (the probability that both crew will miss the same tree is equal to zero).
- 3. Both crews are unbiased in their tally of trees; on average the number of trees they record that are truly outside the plot is equal to the number of trees they miss that are truly inside the plot.

The assumption is not that QA and FM crews miss trees or tally trees outside the plot area at the same rate, but rather that they are both unbiased.

These assumptions are illustrated in the figure C-1. In this hypothetical plot, both the QA crew and the FM crew tallied 33 trees; however, only 30 were matched.

Figure C-1—Hypothetical plot with 30 matched trees, three extra FM trees, and three extra QA trees.

The plot truly contains 33 trees. The QA crew missed one of these trees and tallied one tree that was outside the plot. The FM crew missed two of the trees that were on the plot and tallied two trees that were outside the plot. Both crews are unbiased in their tally; they both tallied the correct number of trees. There are three QA crew extra trees (two are truly on the plot and one was off the plot) and there are three FM crew extra trees (one is truly on the plot and two are off the plot).

Based on these assumptions, the ratio

$$
p_{\text{miss}} = \frac{\text{extra}_\text{QA}_\text{trees}}{\left[\text{matched}_\text{trees} + \left(\frac{\text{extra}_\text{QA}_\text{trees} + \text{extra}_\text{FM}_\text{trees}}{2}\right)\right]}
$$

provides an estimate of the rate at which FM crews are missing trees and the ratio

$$
p_{\text{extra}} = \frac{extra_FM_trees}{\left[\text{matched_trees} + \left(\frac{\text{extra_QA_trees} + \text{extra_FM_trees}}{2}\right)\right]}
$$

provides an estimate of the rate at which FM crews are recording extra trees. The difference between these two ratios (pextra – pmiss) is a statistic that, under the assumptions, has an expected value of zero, with negative values indicating the FM crew is missing trees that are truly on the plot more often than they are recording trees that are truly not on the plot or the QA crew is recording trees that are truly off the plot more often than they are missing trees that are truly on the plot.

In the hypothetical example illustrated in figure C-1, both ratios are equal to 1.5/33 or 4.5 percent and their difference is zero because both crews tallied the same number of trees. In the analysis, the ratios are computed using all of the trees tallied on the blind checks for both the micro plot measurements (trees <5-inches diameter) and subplot measurements (trees >5-inches diameter). Table C-1 provides these estimates based on the blind check tree data but does not include data from the PNW FIA program due to problems related to use of the macro plot. These estimates of the bias in the FM crew tally relative to QA crew can be interpreted as observations of the FM crew bias when the QA crew is assumed to be unbiased. These differences are all quite small. A nonparametric bootstrap test found none of the estimated bias values in table C-1 to be significantly different from zero at the $p = 0.05$ level.

In addition to obtaining estimates of extra QA trees and extra FM trees, each extra tree was assigned a reason for being extra. These reasons were based on the tree's location, size, and the condition land status call of the tree from the crew that tallied the tree. We attempted to identify all of the reasons a crew would have for not recording a tree that was truly on the plot or recording a tree that was truly off the plot. The reasons we identified and the criteria used to assign these reasons to extra trees are:

CODE REASON

⁵⁰ Status—Extra trees due to differences in the condition status. Here one crew determined that the tree existed in a forest condition and should be tallied and the other crew determined that land status of the condition was nonforest and therefore the tree was not tallied.

- 60 Edge trees—Extra trees near the edge of the 6.8-ft microplot (6.5 ft or more) or the 24.0-ft subplot (23.5 ft or more)
- 70 Diameter calls—Extra trees near the diameter limit, (1.2 inch or smaller for microplot trees, 5.2 inches or smaller for subplot trees
- 80 Forked trees—Trees tallied as one tree by one crew and two trees by the other crew that resulted in one matched tree and one extra tree.
- 90 Clumps—Extra trees in a clump of trees (a clump is two or more trees of the same species within 15 degrees and 2 ft for trees $>$ 5 inches and 15 degrees and 1 ft for trees 1 inch-4.9 inches diameter.
- 100 Other—Extra trees that do not meet any of the above criteria.

When multiple reasons applied to a tree, the lowest reason code that is applicable to the tree was assigned. Table C-2 and figure C-2 summarize the number of extra trees by reason for all FIA regions.

Again, as with overall number of number of extra trees, a nonparametric bootstrap test for bias was performed for each region-tree type-reason for a total of 48 tests. In most of these tests, the estimated bias values were not significantly different from zero at the $p = 0.05$ level; however, five resulted in bias values that were significantly different from zero, and in all five cases they showed that the QA crews were recording more trees than the FM crews. The five tests that showed significant bias were:

- 1. Reason = status, region = NC , trees = saplings
- 2. Reason = status, region = NC , trees = >5 inches
- 3. Reason = edge, region = NC , trees = saplings
- 4. Reason = edge, region = NC , trees = >5 inches
- 5. Reason = edge, region = IW, trees = >5 inches

Figure C-2—Number of extra trees tallied by QA and FM crews by identified reason.

Although we were not able to detect an overall bias between QA and FM crews, there may be a small bias in the measurement of edge trees and plots where the land status falls into question. Most of the significant bias was detected in the NC region, which had the most observations giving us the ability to detect very small bias values if they exist.

For both the micro plot and the subplot "Other" is the largest reason for extra trees, with approximately equal number of extra trees reported by both the FM and QA crews. In no case was there significant bias for this reason. This suggests that both crews appear to randomly miss trees at the same rate. We were not able to assign any obvious reason for these extra trees based on the data available. There are four reasons why a tree could be tallied by one crew but not another that we were not able to able to identify:

- 1. Failure of the crews to agree that the tree is alive and needs to be tallied
- 2. Failure of one crew member to communicate to the other crew member which trees are on the plot
- 3. Improper data recording or transfer techniques that result in loosing data for a tree after the data has been entered
- 4. Failure to see a tree

Although it was not possible to assign these reasons to extra tally trees, it appears that the rate at which trees are missed for these reasons is consistent between the QA and FM crews.

In the tally of saplings, disagreements related to diameter measurement and the location of trees near the edge of the plot were both major reasons for extra trees. This might suggest that an increased emphasis on the careful measurement of the diameter of trees near the 5 inches threshold and the location of trees near the 6.8 inches limiting distance could possibly improve the precision of the tally of saplings.

Edge was found to be a significant reason for bias in three cases (NC-saplings, NC-trees >5 inches and IW-trees >5 inches). There may be a small bias. FM crews may have a slight tendency to record these edge trees less often than the QA crews; however, the bias, if it exists, is so small that it was not significantly different from zero in all cases given the number of measurements we have. In the NC region where the majority of the observations were made, the estimated relative bias is –0.20 percent for saplings and –0.04 percent for trees >5 inches.

Status was found to be a significant reason for bias in both saplings and trees >5 inches in the NC region. Only the QA crews recorded any extra trees where status was identified as the reason for an extra tree. These are trees where the QA crew determined that the tree existed in a forest condition and should be tallied and the FM crew determined that land status of the condition was nonforest and therefore the tree was not tallied. In no cases did the FM crew tally a tree as existing in a forest condition and the QA crew determine that the condition was not forest and the tree should not be tallied. There were subplots and portions of subplots where the FM crew classified the area as forest and the QA crew classified the area as nonforest; however, in none of these areas did this result in an extra tree being tallied by the FM crew.

It is important to note that for most reasons, both the FM and QA crews had relatively equal numbers of trees. QA crews tend to have more experience and training than FM crews. If we are to focus our training to reduce bias caused by missed trees, it appears that we should focus more training on land status determination and the measurement of edge trees. Also, we need to improve the consistency of training among the regions.

Table C-1—Tabulation of extra trees observed by FM and QA crews in each region^a, the estimated rates at which FM crews are missing trees and recording exra trees, and the estimated bias in the FM crew tree tally relative to the QA crew.

	Region							
	NE	NC	SO ₂	IW	Total			
Total matched trees (trees tallied by both crews)								
Saplings	450	3,161	266	374	4251			
5 -inches +	1,424	10,802	1,053	1,962	15,241			
trees								
Total	1,874	13,963	1,319	2,336	19,492			
	Extra QA trees (trees tallied by the QA crew but not by the FM crew)							
Saplings	20	72	5	21	118			
5 -inches +	15	152	$\overline{5}$	66	238			
trees								
Total	35	$\overline{224}$	10	87	356			
Extra FM trees (trees tallied by the FM crew but not by the QA crew)								
Saplings	$\overline{30}$	40	4	$\overline{21}$	95			
5 -inches $+$	13	46	9	72	140			
trees								
Total	43	86	13	93	235			
	Estimate of the probability that an FM crew missed a tree ^b							
Saplings	0.021	0.011	0.009	0.027	0.014			
5 -inches +	0.005	0.007	0.002	0.016	0.008			
trees								
Total	0.009	0.008	0.004	0.018	0.009			
	Estimate of the probability that an FM crews recorded an extra tree ^c							
Saplings	0.032	0.006	0.007	0.027	0.011			
5 -inches +	0.005	0.002	0.004	0.018	0.005			
trees								
Total	0.011	0.003	0.005	0.019	0.006			
Estimate of the bias in the FM crew tally relative to QA crew ^d								
Saplings	0.011	-0.005	-0.002	0.000	-0.003			
5 -inches +	-0.001	-0.005	0.002	0.001	-0.003			
trees								
Total	0.002	-0.005	0.001	0.001	-0.003			

^a Data from the PNW were not available for extra tree analysis due to problems related to use of the macro plot.

		Reason					All	
Region-Tree	Crew	Status	Edge	Diameter	Forked	Clumps	Other	reasons
type	Type		tree	call	tree			
Sapling NE -	QA	θ	2	8	$\boldsymbol{0}$	$\overline{4}$	6	20
NE - Sapling	Crew	θ	5	23	$\boldsymbol{0}$	$\overline{2}$	θ	30
- Tree NE	QA	3	4	$\overline{2}$		θ	5	15
NE - Tree	Crew	$\boldsymbol{0}$	$\overline{2}$	5	$\boldsymbol{0}$	$\mathbf{0}$	6	13
NC - Sapling	QA	6	18	11	$\boldsymbol{0}$	5	32	72
NC - Sapling	Crew	θ	5	9	$\boldsymbol{0}$	$\overline{2}$	24	40
NC - Tree	QA	50	20	9	3	5	65	1,84
NC - Tree	Crew	θ	11	10	$\overline{0}$		24	71
SO - Sapling	QA	$\boldsymbol{0}$	1	3	$\boldsymbol{0}$	$\boldsymbol{0}$		5
SO - Sapling	Crew	θ	$\boldsymbol{0}$	θ	$\boldsymbol{0}$	θ	5	5
SO - Tree	QA	θ	$\boldsymbol{0}$	$\overline{2}$	$\boldsymbol{0}$	1	4	7
SO - Tree	Crew	$\boldsymbol{0}$	3	$\overline{2}$	$\mathbf{1}$	θ	6	12
IW - Sapling	QA	$\boldsymbol{0}$	\overline{c}	$\overline{7}$	$\boldsymbol{0}$	θ	12	21
IW - Sapling	Crew	θ	5	4	$\boldsymbol{0}$	θ	12	21
IW - Tree	QA	θ	11	$\overline{7}$	$\boldsymbol{0}$	θ	48	66
IW - Tree	Crew	$\mathbf{0}$	3	$\overline{2}$	$\boldsymbol{0}$	$\mathbf{0}$	67	72
Total - Sapling	QA	6	23	29	$\mathbf{0}$	9	51	1,18
Total - Sapling	Crew	θ	15	36	$\boldsymbol{0}$	4	41	96
Total - Tree	QA	53	35	20	$\overline{4}$	6	122	2,72
Total - Tree	Crew	θ	19	19			103	1,68

Table C-2—Number of extra trees enumerated by either FM crews (Crew) or QA crews and reason.

Notes

The Rocky Mountain Research Station develops scientific information and technology to improve management, protection, and use of the forests and rangelands. Research is designed to meet the needs of National Forest managers, Federal and State agencies, public and private organizations, academic institutions, industry, and individuals.

Studies accelerate solutions to problems involving ecosystems, range, forests, water, recreation, fire, resource inventory, land reclamation, community sustainability, forest engineering technology, multiple use economics, wildlife and fish habitat, and forest insects and diseases. Studies are conducted cooperatively, and applications may be found worldwide.

Research Locations

- Flagstaff, Arizona

Fort Collins, Colorado* The Reno, Nevada

Albuquerque, N Boise, Idaho **Rapid City, South Dakota** Moscow, Idaho **Logan, Utah** Bozeman, Montana **Canada Boxeman, Montana** Ogden, Utah Missoula, Montana **Provo**, Utah
- Albuquerque, New Mexico

*Station Headquarters, Natural Resources Research Center, 2150 Centre Avenue, Building A, Fort Collins, CO 80526

The U.S. Department of Agriculture (USDA) prohibits discrimination in all its programs and activities on the basis of race, color, national origin, age, disability, and where applicable, sex, marital status, familial status, parental status, religion, sexual orientation, genetic information, political beliefs, reprisal, or because all or part of an individual's income is derived from any public assistance program. (Not all prohibited bases apply to all programs.) Persons with disabilities who require alternative means for communication of program information (Braille, large print, audiotape, etc.) should contact USDA's TARGET Center at (202) 720-2600 (voice and TDD).

To file a complaint of discrimination, write to USDA, Director, Office of Civil Rights, 1400 Independence Avenue, S.W., Washington, DC 20250-9410, or call (800) 795-3272 (voice) or (202) 720-6382 (TDD). USDA is an equal opportunity provider and employer.