

## **Reconciliation Report—Late Successional and Old Growth status and trends report**

**Revised March 31 2005, to reflect points made in PNW letter of March 29**

**Original -- March 17, 2005**

Three independent reviewers read the report. They are identified here as Reviewer#1, #2, and #3. All their comments have been numbered consecutively, below.

All reviews were cordial, professional, and very much appreciated by the authors. There were many helpful suggestions for improving the manuscript. Reviewers caught some outright errors that we missed even after several revisions of the paper. Thankfully, there were no comments that indicated any ‘fatal flaw’ in the analyses or reporting.

The main report authors responding to the comments and revising the report were Moeur, Spies, Alegria, and Demeo. Our philosophy was to provide a response to each comment, making every effort to comply with the change or clarification requested by the reviewer. If we decided not to incorporate a comment, we provided a justification for not doing so. Alegria concentrated on the technical comments from #2 regarding map accuracies and statistical analyses. Demeo concentrated on comments about fire ecology and historic range of variation; Moeur and Spies responded to all comments in a general sense.

The comments made in the various reviews were complementary. There was a lot of reviewer overlap in identifying weaknesses in the report. Reviewers #1 and #3 requested additional clarification and explanation, but had almost no substantial comments of a technical nature. Reviewer #2 concentrated more on technical questions, and less on organization or explanation.

The vast majority of comments were requests for additional explanation or clarification, or to make a slight revision in the report’s organization. Only a few comments were of a technical nature. The main areas of concern, other than clarifying details, were:

1. Additional context for the plan--why, what, including expected outcomes by province
2. Low accuracies in the remote sensing and the implications of that for potential uses
3. Differences in methodology between CA and OR/WA, and implications for results/interpretations
4. More ecological detail on significance of results
5. Concern that the connectivity results were in error, and that a general QA check is needed
6. Expansion of section dealing with future monitoring design (improvements/use of other methods)

A summary of our responses follows below. Original comments are attached at the end of this report.

## **Resolution of major comments:**

### **1. Additional context for the plan--why, what, including expected outcomes by province**

We expanded the Introduction to include more Plan background (didn't assume readers knew this already). Referenced the ROD about why reserves were established in the first place. Added a new section to the Introduction on the Plan's Expectations, and moved or reiterated material that was in the Discussion section. Stated explicit assumptions about long-term averages made in the Plan. Clarified that the Plan's goal was to move forests to Outcome 1 or 2 (within HRV). Referred to the synthesis report for additional discussion about whether the Plan's assumptions need to be re-evaluated.

### **2. Low accuracies in the remote sensing and the implications of that for potential uses**

Almost all of Reviewer #2's comments addressed the technical aspects of the report. Major concern was that the basis for most of the results reported was the spatial data—the LSOG maps derived from remote sensing. Reviewer was concerned about implications of making interpretations from the map data results because of low map accuracies for some attributes in some provinces.

We make these points in justifying use of the map data as basis for results (we discussed this strategy earlier with Cindi West and Becky Gravenmeir and got their general agreement).

- 1) Wall-to-wall map-based data covers entire population with equal intensity.
- 2) Plot data does not cover entire population. Leaves out important strata (such as Park Service lands, and to a lesser extent, Ca-BLM). In sampled lands, coverage is sparse, and because it is systematic, it may miss rare conditions (in other words, is not a stratified sample of conditions on the landscape).
- 3) Using plot data for population estimates would require a sound method for expanding to portions of the population missing from the sample. No sound way to do that has been developed, even by FIA (i.e. assumption that non-sample lands are from the same population is not appropriate).
- 4) The derived older forest maps have good accuracy values, except in the Eastern Cascades provinces, and these limitations were fully disclosed in the revision (including a reasonable approximation of the underestimate). Furthermore, consistency with plot estimates for sampled areas helps confirm their accuracy.
- 5) Perhaps most importantly, the maps were necessary for species habitat suitability modeling (Biomapper), because there was no other wall-to-wall driving data available. No simple rule sets existed to describe owl or murrelet habitat (such as acres where QMD ge 20" and multi-storied). It was desirable (necessary) to use the same base data for estimations for all the NWFP monitoring entities.

We dealt with this comment (and related comments by the other two reviewers) by completely re-vamping the section on map accuracy assessment, and taking the strategy of disclosing weaknesses in the remote sensing results in all of the relevant sections of Results and Discussion.

**3. Differences in methodology between CA and OR/WA, and implications for results and interpretations**

Differences between CA and WA/OR were simply the existing situation resulting from datasets derived from different methods, and at different resolutions. We did acknowledge these differences, but feel they did not affect our basic results or conclusions. See comments below related to potential natural vegetation zones, and map resolution (pixels in IVMP vs. polygons in CALVEG). In almost all cases, we avoided combining analyses across provinces. We generally disagree that province comparisons "are not useful", but do agree that it means that you can't expect to use the results to make statements about small to moderate size differences.

**4. More ecological detail on significance of results**

Clarified the scope and objectives of this paper and referred to the synthesis paper for interpretations of results.

**5. Concern that the connectivity results were in error, and that other errors might still be present**

Based on this comment, we performed a 'gut-check' of results from most of the map overlay analyses. The reviewers sharp eyes did identify errors in the connectivity analysis (basically, we misinterpreted the original results), and in the analysis of clearcutting and wildfire effects in the California provinces (only). The latter resulted from an error in re-classifying the GIS data. These corrections were made, and the entire text was checked for consistency with the new results.

**Reference to specific comments (differences between Draft 4.00 and the final--Draft 5.00)**

<u>Refers to reviewer #-comment no.</u>	<u>Explanation of our response</u>
#1-1, #1-5, #1-17	Expanded the Introduction to include more Plan background (didn't assume readers knew this already). Referenced the ROD about why reserves were established in the first place.
#1-2, #1-5, #1-17	In Introduction, stated more explicitly what kind of monitoring is reported (LSOG area, amount, distributions), and what is not (LSOG functionality). Referenced synthesis report and other status and trends report for more discussion of functionality (for example, whether it is providing habitat for owls or murrelets, or contributing to watershed health)
#1-3, #1-55	Added a new section to the Introduction on the Plan's expectations, and moved or reiterated material that was in the Discussion section. Stated explicit assumptions about long-term averages made in the Plan. Clarified that the Plan's goal was to move forests to Outcome 1 or 2 (within HRV). Referred to the synthesis report for additional discussion about whether the Plan's expectations need to be re-evaluated.
#1-22	Expanded results of the change analysis from the re-measured plot data. Expanded table 15 to show Province results Discussed where (i.e. in what provinces) older forest was added, in what proportions.
#1-4, #1-39	Added notes about the specific dates of map data, change detection, and plot remeasurement data to the text, tables, and figures to clarify the range of dates of analyses, whether the baseline or change detection data were being referred to. Identified results that were for the start of the Plan (2004) vs. at the end of the first decade (2003). Changed the tense used to past tense when referring to data and results representing the start of the Plan.
#1-6, #1-33, #1-46, #2-10	Expanded footnotes to explain why FIA plot data was not used (i.e. Ca-BLM), and why no plot data were available for Park Service. Explained that FIA annual inventory data will be available and used in future monitoring analyses.
#1-6	Added more text to inform the reader of differences in methodology and data between Or-Wa and Calif. (plots, veg maps, change maps, potential veg classes, map accuracy methods). Differences between CA and WA/OR were simply the existing situation resulting from datasets derived from different methods, and at different resolutions. Emphasized that differences between CA and WA/OR were simply the existing situation resulting from datasets derived from different methods. This may have led to some minor weaknesses, but nothing that would alter major conclusions. We disagree that comparisons between provinces are not useful, as long as we avoided making statements about small to moderate size differences (which we did).
#1-7, #1-10	Revised the Abstract (shortened it and made it less redundant of the Key Findings section).

Renamed “Summary of Key Findings” as “Executive Summary”

#1-8 Clarified structural elements referred to in “classic” Douglas-fir old-growth (old large trees, complex canopies, snags, CWD), in Key findings and in text. Added additional references from the literature

#1-9 Clarified that stand-replacing harvest means clearcut, but not partial harvest.

#1-11, #2-3, #2-7, #2-9, #3B2 We were much more up-front about disclosing prediction error in the older forest maps. Added discussion about how map errors were assessed in Methods, how they affected results and interpretations, and conclusions. Always attached map error estimates when presenting results (e.g., 7.87 million acres  $\pm$  1.96 million acres). Discussed where we had reason to believe that the results from the map data were underestimates or overestimates, and if possible, a general statement about the magnitude. Discussed the possible effect of map inaccuracies on the conclusions drawn from the landscape pattern analyses.

#3B-4 Included references and made points that accuracies of our remote-sensing results were similar to those found in the literature.

#1-12 Removed reference to ‘3 or 4’ large fires in bullet 5, Key Findings.

#1-13. We were not able to report trends from re-measured plot data for the “large, multi-storied” and “size indexed to vegetation zone” definitions, because we didn’t have all the attributes we needed to classify the plots at the 2<sup>nd</sup> measurement. We explained this in the text.

#1-14 Re-organized Description of Plan area in Introduction according to recommendations. A map of the Plan area is referenced (figure 1)

#1-15 Added figure 1b to identify FS, BLM, NPS management units

#1-16, #1-5 Revised the list of monitoring questions in the sidebar to paraphrase all of the monitoring questions in Hemstrom and others (1998). Added an analysis of stand structure at the plot scale (including new sections in Methods and results, and a new set of figures). Explained that monitoring questions addressing silviculture and salvage effects, and ecological processes and biodiversity were outside the scope of the status and trends report, but that they were discussed at length in the synthesis report.

#-18, #1-52, #3B-1 Expanded the Introductory section “A Continuum of Older Forest Definitions”. Incorporated material that was originally in the Discussion (“An evolving ecological definition of old-growth”). Better explained rationale for choosing these particular definitions. Made the point that these definitions were designed around attributes available in our monitoring data (i.e. they are not ‘ecological’ definitions).

#1-19 Better explained why our Md&Lg definition not exactly the same as in the ROD—Vegetation Strike team changed the standards.

#1-19 Changed the “very large, multi-storied” older forest definition to “large, multi-storied”. (all occurrences in text, tables, and figures).

#1-20 Again, differences between Ca and Wa/Or in the classification of potential natural vegetation zones was a fact we had to live with. However, because we did not cross province boundaries with the “size indexed to Vegetation Zone” definition, we feel it did not have a major affect on our results or conclusions. Diameter thresholds in California were specific to vegetation zones in California, (likewise diameter thresholds in Wa/Or were specific to vegetation zones in Wa/Or), which is evident by comparing figure 3 and table 3.

#1-21 Clarified that Physiographic boundaries are not ecological provinces, because they are cut by state boundaries. Kept provinces as our basic reporting unit—did not summarize value across ecological units (like west Cascades, etc.)

#1-22 Moved material explaining assumption of logging from 1994-1996 negligible, from trend analysis section to IVMP section.

#1-23 Revised/renamed IVMP descriptions (“broadleaf” → “non-conifer” and “non-vegetated” → “non-forested”).

#1-24 Clarified IVMP QMD class intervals in Eastern Cascades Oregon and Eastern Cascades Washington.

#1-25 Removed citation Ducey and Moeur (in prep); in appendix only.

#1-26 Checked statement about CALVEG minimum mapping unit. Corrected value to 2.5 acres. Intend to add metric/english conversion table in final version

#1-27 Discussed implications of Calveg 12” (not 10”) boundary. None for older forests, minor for two smaller size classes.

#1-28, #2-14 Added number of accuracy assessment plots for IVMP and CALVEG, number of plots for various other analyses.

#2-5, #1-29, #1-30, #2-18, #2-general concerns, #3A-1

**We greatly expanded the section on older forest map accuracy assessment (aa).**

- Repeated the aa for the other two definitions and reported results in table 6 (“large, multi-storied” and “Size indexed to vegetation zone”)
- Added the kappa statistic for testing whether agreement between map and reference values was due to chance alone; discussed results and implications.
- More fully discussed where accuracies were low, and what the potential implications were for conclusions drawn from the results.
- Stated the number of reference plots used for aa by province
- Stated that Calveg Max operator is most similar to IVMP error matrix aa approach
- Explained how canopy layering was used in labeling reference plots

#1-31, Stated unequivocally that map estimates were low in Eastern Cascades Washington and Oregon as a result of modeling weakness in IVMP (Results section), estimated the amount of the underestimate (200,000-300,000 acres), and discussed the implications for interpretations made from the results (both amounts and spatial patterns, i.e. connectivity).

#1-32 Clarified that non-forested plots were included in the analysis. They were classified as ‘potential forest’.

#1-34, #1-36, #1-35 Explained why the ratio of sample plots was so low in WaECW (table 7). Missing panel due to contractor default.

#1-35 Missing portion of Table 8 must have been due to a photo-copying error. (It was present in the electronic version).

#1-37 Expanded the description of older forest map accuracy assessment methods, and added brief explanations of disturbance map accuracy assessments in Ca and Wa/Or

#1-40 Revised the paragraph and removed the statement in error.

#1-38 Moved material from text into FRCC explanation sidebar.

#1-41 Used relative terms, to replace ‘underestimate’ or ‘overestimate’ in talking about PNV-LOSG vs. Md&Lg LSOG

#1-42 Used consistent terminology--changed all occurrences of for PNV “type groups” to “zone groups”

#1-43, #1-44 Revised Fig. 15 to improve point about relative proportions of older forest in different LUAs.

**#2-6, #2-15 Identified and corrected errors in the connectivity analysis.** Thank you to reviewers—in draft 4.00, the results in table 12 were mistakenly identified as the distance between blocks 1,000 acres in size and larger, when they were in fact the results for all blocks. We repeated the connectivity analysis for blocks 1,000 acres and larger, corrected the table and text that refers to it.

Better explained rationale for the particular metrics chosen by making a stronger link to the Plan’s expectations for connectivity (sidebar 1 and appendix IV)

#1-45 Corrected statement in table 12 caption (“edge-to-edge”, not “center”).

**#1-47, #2-6, #1-47, #3B-5 Identified and corrected errors in disturbance map results in tables 13 and 14.** Thank you to reviewers—we identified an error in the map processing in California provinces (only). In draft 4.00, we had errors in these tables, so that the amounts we had reported as harvested or burned were much too high. These have been corrected. Checked entire text for consistency of values in text (harvest and fire losses results)

#1-48, #1-50, #1-51, #2-13, #3C Made corrections as suggested.

#1-49 We believe the statement to be valid and left it as it was originally (about population sampled by grid plots)

#1-53. Made subtle wording changes to suggest that we ‘demonstrated’ rather than ‘established’ an approach to an ecological definition of older forests.

#1-54 (What target should be used to base “older” on?) Question addresses functionality, and is really beyond the scope of this report. Added reference to Synthesis report for additional discussion.

#1-56 Referred to Synthesis report about more discussion of contribution from federal lands to LSOG within regional context

#1-57 Corrected table 3—QMD threshold for hardwood types was 20” (not ‘incapable’ of developing LSOG), as originally stated in the text. In other words, map units in hardwood vegetation zones were mapped as older forest if the value of QMD was 20 inches or larger.

#1-58 Removed confusion about measurement year/growing year distinction by removing statement from the text (not an important detail).

#2-4 Referred to the Synthesis report for additional treatment of the topic of how analysts and policy makers could potentially use the data presented in this section on fire.

#2-1, #2-8, #2-11, #3A-2 Reviewer #2 suggested that that we add a lot of text on the topic of recommendations for changing the future monitoring program. We have an extensive section on monitoring design considerations, but the reviewer has asked us to go to an expanded level of detail that we believe is premature at this point. Basically, we don’t even know what the future holds for the monitoring program

#2-16 In tables and figures, was more explicit about whether results were from map-based or plot-based analyses.

#3A 4 Made the suggested points

## Comments from Reviewer #1

### **Review comments on "Northwest Forest Plan - The First Ten Years (1994-2003): Status and Trends of Late-Successional and Old-Growth Forests", Draft 4.00 (December 20 2004), by Moeur and others**

#1-1 When I began reading this report, I expected to find the following information:

- why did the Plan set aside old-growth forests?
- how much did they intend to set aside?
- how much was set aside in 1994?
- how much is there in 2004?
- is it achieving the goals for which it was set aside?

I found answers to some of these questions, and not to others. General comments are here, and there are also scribbled comments throughout the manuscript.

#1-2 The report needs more information up-front about why these forests were set aside. There's a bit of this in the Discussion at the end, but there seems to be an assumption that readers already know about this. It's also important here to broach the question of what kind of monitoring is reported here, and what type isn't - the scope of the report. Area of forest is reported; functionality isn't (see further comments below). So this manuscript will report on how much forest there is out there, and won't report on whether it's providing habitat or whatever else it's supposed to do. Will other reports? If so, they should be referenced here.

#1-3 As for "how much did they intend ...", I think this is a key question largely overlooked in this report. Clearly it relates to the previous point (Plan intent). Its kind of hinted at in Sidebar 3 ("Plan expectations: Abundance and diversity"). If this is indeed the intent, it should be front-and-centre in the report. This is the goal against which inventories will be evaluated. Is the goal to move forests to Outcome 1? And if so, one of the recommendations of the report should be that these 'Outcome' numbers be improved. Even if these numbers are written in stone from the Plan and can't be revisited without litigation, this also should be discussed. The "Land covered by LSOG" is defined in relation to long-term averages in Sidebar 3 notes. Well, long-term averages are very different throughout the Plan area. You'll have to do a lot of logging to keep the Olympic Peninsula at FEMAT's average of 65%, and a lot of fire suppression on the East Cascades. I suspect East Cascades in Outcome 1 would be a profoundly unhealthy forest, and would certainly not match the Outcome 1 definition. 'Outcome' numbers (targets) should be defined ecologically (e.g., by physiographic province) - they make no sense otherwise. And PNW researchers have done some of the world's best work characterizing natural disturbance regimes. So if the numbers aren't going to be revised in this report, the issue should at least be discussed. This will allow readers to evaluate whether the 600,000 acres of new "older forest" acquired between 1994 and 2004 is good, or bad, happy or sad.

#1-4 If the most recent satellite imagery employed was 2002 in Oregon/Washington, the title of the report should perhaps be "Northwest Forest Plan - The First Nine Years (1994-2002)". (I'm assuming that class changes predicted from inventory plot data are extrapolated to 2003 but depletion due to natural disturbance and logging can only be mapped to 2002 in O/W.)

#1-5 This comes up later in discussion of sidebar 1, but should perhaps be explicitly addressed up-front in the report. The report looks at how much old-growth forest there is and how it has changed over time. But this ignores some broader questions, such as why these reserves were established in the first place. As I recall, it wasn't only for old trees. It was as habitat for the spotted owl, and other creatures. It was a means, not an end. This should be discussed, and other reports (e.g., on owls) mentioned. (There's one sentence about this on p. 66, par. 4: "It is not possible ... not discussed.") One of Hemstrom et al.'s (1998) monitoring recommendations was around stand structure, ecological processes and biodiversity. Where does this monitoring take place?

#1-6 It sounds as if there were two inventories conducted: one for Oregon-Washington, and one for California. This is clear in the report's attempts to combine them into one inventory for the plan area: California BLM ground data are unavailable (for an unexplained reason); potential vegetation classes are different for the two areas; physiographic provinces are different; smaller stem class categories are different; mapping resolution is different (and so Oregon-Washington maps must be 'dissolved' to match the lower-resolution California maps); even methods for assessing map accuracy are different, and incomparable. Change detection methodology is different in Oregon/Washington and California, and the California portion had to remove patches <5 acres to be consistent. The result is that summaries for O-W seem reasonable, and summaries for C seem reasonable, but summaries by physiographic provinces or by potential vegetation classes for the Plan Area are not useful.

#1-7 p. 2. Abstract. This Section repeats much of what's in "Summary of Key Findings" (pp. 8-10), often word-for-word. You don't need both. I suggest beefing up the Abstract a bit to include material from "Key Findings", then deleting "Key Findings".  
"We created baseline maps of older forest conditions at the start of the plan." (p. 2, par. 2) and "We estimated the acres of older forest on federally managed lands at the start of the Northwest Forest Plan." (p. 2, par. 3) say the same thing. You don't need both.  
"Average tree size" (first mentioned p. 2, par. 3) needs to be defined here. (When I first read it, they seemed rather short!)

#1-8 p. 3. Par. 2, ""Very large, multi-storied older forest" retains minimum elements of "classic" Douglas-fir old-growth forests." I'm not sure what this means. (It's +/- explained later in the text, but the Abstract should stand on its own.)

#1-9 par. 3, "stand-replacing harvest", is this all logging? Is all selection logging, e.g., considered "stand-replacing"?

#1-10 p. 8, "Summary of Key Findings". See notes under Abstract.

#1-11 p. 9, bullet 4. Perhaps you should use a statement a tad more cautious: your estimate is 7.87 million acres +/- ~2 million acres (Table 10), so perhaps not an unequivocal confirmation of the Plan.

#1-12 Bullet 5. Are "regeneration harvests" logging? Also, "3 or 4 large fires", is it 3 or 4? (Seems unnecessarily vague).

#1-13 Bullet 6, what are the trends for the other two older forest definitions?

#1-14 p. 10, Introduction. I think the Introduction needs re-working. Tenure information from p. 10, p. 11 ("Lands affected by the Plan ... any non-federal lands."), and p. 12 ("The Plan applies ... the Canadian border.") should be combined at the beginning. Plan outcomes (top of p. 11, bullets on p. 13) should be combined at the end. Also, the introductory material should describe the boundaries of the Plan Area, and include a map as a Figure.

#1-15 Figure 1, I expected a Figure that showed me the 26 Management Units: Figure 1 doesn't. This would be helpful for readers who don't know where the MUs are.

#1-16 p. 14, par. 2 (and sidebar 1). Why have the monitoring questions in Hemstrom et al. (1998) been altered? Why does this report leave out Hemstrom's questions about silviculture and salvage and (most importantly) the question about stand structure, ecological processes and biodiversity? The latter is the step that moves this exercise, this report, from compliance monitoring to effectiveness monitoring.

#1-17 p. 15, end of par. 1. This report should explain somewhere what the Plan goals for old forest are. Is it for habitat? If so, for whom - owls only, or more? Then indicate that this report doesn't cover that topic. (This is closely related to my previous comment.)

#1-18 p. 17. A Continuum of Older Forest Definitions. A lot of what's in "An evolving ecological definition of old-growth" (pp. 63-65) should go here. The latter section explains how you arrived at the definitions used, their strengths and weaknesses. (In fact, the first sentence of the latter section contains the phrase "...a continuum of older forest definitions"!)

#1-19 p. 19, par. 1. "...with the exception ... Record of Decision." Why? Also, par. 2, it's not clear to me why, if a stand has medium and large trees, it's called "medium and large", but if it's got large trees it's called "very large", esp. as "very large, multi-storied older forest" is abbreviated "L-MS" (Figure 15)!

#1-20 p. 20, par. 1. It's not clear why there are 2 separate vegetation classifications. I assume this is a jurisdictional thing? Here and later on, it's not immediately clear how the two compare. For example, is "Douglas-fir" (IVMP) the same as CALVEG's "Douglas-fir - Western Hemlock", "Pacific Douglas-fir", both, neither? This is important

for old growth inventories by type. The unfortunate result (I'm guessing) is that all vegetation types (and associated analyses) stop at the California border. Clearly, that's not true in the real world, and prevents any opportunities to actually conduct old growth inventories by potential natural vegetation.

#1-21 p. 22. Physiographic Provinces. Similar comments to those above. These aren't physiographic provinces: they are physiographic provinces cut by state boundaries. For example, I assume that "Washington Western Cascades" and "Oregon Western Cascades" are one physiographic province, as are Eastern Cascades, and Coast Range. However, where having two un-correlated potential vegetation classifications presumably precludes inventories by vegetation type, with physiographic provinces you should be able to combine inventories from, e.g., "Washington Western Cascades" and "Oregon Western Cascades" to report on "Western Cascades". And you should. Incidentally, Nature Trust should be able to give you physiographic province maps that cross the 49th parallel. It would be progressive to at least show the province boundaries extending into Canada.

#1-22 p. 28 (IVMP). Paragraph 1, you mention a 1994 to 1996 logging of "at most, about 0.07 percent of total older forest". I assume you must know +/- where it occurred if you can calculate % to two decimal places. Did you map this (not mentioned)? Also, what about 1992 to 1994 logging? Any idea how much? Did you map it?

#1-23 p. 28, last par. "Broadleaf" seems an odd name for the non-conifer category, esp. as it includes "shrubs, grasses, and forbs". How about "non-conifer"? Also, your "non-vegetated" classes need to be split out a bit, if they were in the classification. Most important is "prairie". "Non-vegetated" prairie??? More importantly, you should be able to track prairie over time to understand issues such as conifer encroachment into grasslands as a result of fire suppression and other factors.

#1-24 p. 29, par. 2, last sentence: explain why "Average tree size is predicted in larger class intervals ...".

#1-25 p. 30, par. 2, "Ducey and Moeur (in prep. Appendix II)". Why are some "in prep" documents here as Appendices, and others not? "In prep" is generally not helpful for a reader. The more of this material that can be appended, the better.

#1-26 p. 31, par. 3. "...an image segmentation process to a minimum mapping unit of 2.5 ha." I assume that this should read 2.5 acres (=1 hectare)? On p. 36 you refer to CALVEG's minimum mapping area of 100m (100m x 100m = 1 hectare). Incidentally, perhaps this report should include a note up front with metric/Imperial conversions, for readers from outside the US?

#1-27 Also, last paragraph, too bad about the IVMP 10"/CALVEG 12" difference. Presumably the CALVEG numbers could be adjusted using vegetation plot data?

#1-28 p. 34, par. 2. "An independent sub-sample of field reference plots ..." How many (25% for IVMP)?

#1-29 It's unfortunate that the accuracy assessments were done so differently for IVMP and CALVEG: it makes it impossible to compare them. I assume CALVEG "Max." is closest to IVMP measurements? Why do you report only CALVEG overall and canopy closure - what about life form and tree size classes? How was "overall" calculated?

#1-30 p. 35, par. 3. "Reference plot values ...": how did "canopy layering" factor into the "older/not older" decision?

#1-31 p. 36, par. 1. You should discuss how, or whether, older forest mapping errors compound CALVEG/IVMP mapping errors. (I assume the CALVEG/IVMP error is incorporated in the "older forest" error?)

#1-32 p. 39, par. 2, last sentence. Did you include non-forested plots in your analysis? If so, how were they classified?

#1-33 p. 40, last 2 lines. "Nor were data from Bureau of Land Management-California available". Why not? Might they be available in the future? Is California holding out, or are the data less-than-useful?

#1-34 p. 41, par. 2 (ref. to Table 7). Why is "Washington Eastern Cascades" so much lower than the rest?

#1-35 p. 31, par. 4, the portion of Table 8 referred to here was not included in my review materials.

#1-36 p. 41, par. 2. Why are ground samples so few in Washington Eastern Cascades? Also, last par., half of Table 8 is missing in my review copy.

#1-37 p. 45, last par. You should briefly describe how you do the accuracy assessment.

#1-38 p. 48, par. 1, "Condition class 1 areas ... class 2 to 1." This material should be moved to the sidebar. It would improve the sidebar and reduce duplication.

#1-39 p. 49, par. 2. You should make it clear, in the text and in the Figures, that all of these numbers are for 1994. This may be even more confusing for readers because of the use of present tense in descriptions. So, "The federally managed land area occupied by older forests ranges from 7.04 million acres ..." might more properly read "The federally managed land area occupied by older forests in 1994 ranged from 7.04 million acres ..."

#1-40 Also, par. 4, I think "ranging from 5 percent in the Western Lowlands of Washington to 47 percent in the California Coast" should read "ranging from 1 percent in the Western Lowlands of Washington to 43 percent in the California Klamath".

#1-41 p. 50, 1st par., "...tended to underestimate older forest amounts ...". Well, perhaps. It underestimates compared to PNV estimates, or perhaps the PNV estimates overestimate - without any actual age estimates it's all relative.

#1-42 pp. 50 (bottom)/51 (top). Can these aggregated vegetation classes be called types (as on p. 50) or groups (as on p. 51) and not zones (as on p. 51 and Figure)? They're not zones, and the terminology gets confusing. Also, why were the vegetation types only examined against the PNV definition? It would be interesting (and less confounding!) to look at them against the other definitions, as you've done for other attributes (e.g., Figs 11a-c, Table 10).

#1-43 p. 51, last par. plus Figure 15. This is a slightly difficult-to-understand way of presenting these data. Could you present them in a Table also so people could look at them different ways? For example, I'd prefer to see a Figure that showed what percent of, e.g., LSR+ lands was 'older' under different definitions, vs. AW/CR and MAT+. The histogram might be similar in format to Figure 15.

#1-44 Also, I don't think this is true: "In general, the proportion of older forest in reserve allocations increased as the definition became more restrictive." It's not true for AW/CR, which I would assume is a "reserve allocation". Also, PNV is more restrictive (for some classes), and it has a lower proportion than M&L for LSR+.

#1-45 p. 54, par. 2. "The average edge-to-edge distance ..." is described in Table 12 caption as the distance between centres of blocks: choose one.

#1-46 p. 56, footnote 2. "...but no Forest Inventory and Analysis plot data were analyzed in the monitoring analysis." Why not?

#1-47 pp. 57-58, check numbers for agreement with Table 13.

#1-48 p. 58. "Almost 90 percent of all stand-replacing fires ..."; should this be "Almost 90 percent of the area burned ..."?

#1-49 p. 59, par. 2. "Because they are a systematic sample ... this analysis was conducted." This makes no sense statistically. Just because a sample is systematic doesn't mean it's representative of the population.

#1-50 p. 61, par. 3. "The estimates of older forest amounts developed from the map and plot data were **mostly** consistent ..." (see Fig. 22)

#1-51 "To reiterate the results ... 7.87 million acres, respectively." No, you've got it backwards.

#1-52 p. 62, par. 3. I think almost everything in "An evolving ecological definition of old-growth" should be moved to p. 17 (see notes there).

#1-53 “In this report ... different types of definitions.” If I understand what this paragraph’s getting at, then I disagree with it. I don’t see where in the report the authors “established an approach”. And given that they designed the definitions around the inventory, it’s scarcely surprising that you can use the inventory with the definitions.

#1-54 pp. 64 (bottom) – 65 (top). There’s a certain circularity to developing new “ecologically-based” average tree size criteria without a clearer target of what you’re aiming for. What will these be based on? How will you know when it’s “older”? Actual age (e.g., mean size when a stand’s 250 years old)? Use by spotted owls (e.g., mean size when owls begin to use it as habitat)? This report should provide some direction.

#1-55 p. 66, par. 2 & 3 & embedded bullets. This belongs in the Introduction somewhere: it explains why you're doing this report.

#1-56 p. 70, par. 3. "...the importance of the contribution ... managed at short rotations." True enough - do they?

#1-57 Table 3. You should at least discuss in the body of the report the exclusion of oaks, cottonwoods and willows from the PNV definitions. In this report you've equated 'older' with 'bigger'. However, I'm not sure it's correct that *Quercus garryana* woodlands are "not ... capable of sustaining late-successional and old-growth forest". I've seen some pretty nice old-growth oak forests.

#1-58 Table 8. "Plots with zero years between re-measurement were measured one calendar year apart, but within the same growing season." Not clear: over most of this area, different calendar years are different growing seasons.

## Comments from Reviewer #2

Comments on: Moeur et.al., *Northwest Forest Plan: the first ten years: status and trends of late successional and old growth forests*.

#2-1 This document describes the status and trends of older forests on federal lands in the Northwest Forest Plan between 1994 and 2003. Trends in Older forests were described taking into account both succession and disturbance by both wildfire and logging. Data are presented by fire regime, land use allocations, vegetation zone, physiographic provinces, individual states, and the entire northwest planning region.

#2-2 I find much of this report to be well written and well organized. I think you did a good overall job, but I am concerned about the remote sensing data accuracy and some quality control issues. As you will see from the text below, I encourage you to rethink the conceptual design for future monitoring which could make it more streamline, accurate, and cost-effective.

My detailed comments are centered on the following six review questions:

### **1. The overall conceptual design of the monitoring program. Is it well conceived? Are there any major flaws?**

#2-3 The design of the monitoring program is to use both mapped satellite data and ground based plot data. This design sounds very good at first glance, but it only works if the mapped satellite data is accurate. The authors seem to have downplayed the low map accuracies for some of the provinces, and the impacts of these low accuracies are not fully discussed. For example, how do the low accuracies affect the landscape metrics such as the connectivity analysis or the larger block analysis?

I think much more use could be made of the plot data. Since there are mapped coordinates for each plot, there are many spatial applications for these data. I suggest that whenever possible, and when the plot sample size is high enough, that the authors present more results using plot data. What about summaries by fire regime type, etc.?

In the future, the plots could be more closely tied to the spatial data using imputation methods which would probably increase the mapped accuracies.

### **2. The analytical approaches adopted. Are they well justified and using appropriate methodology?**

#2-4 I applaud the authors on their description of remaining old forests by fire regime. Fire is a huge issue for the existing old forests since fire suppression policies have been in effect for many decades and we are now beginning to see the results of those policies with increasing number catastrophic fires. I would like to see the authors include more on how analysts and policy makers could potentially use the data presented in this section on fire.

**3. The data reported. To the extent that you can tell from the mss, are these well developed, and collected using appropriate techniques?**

#2-5 In order to deal with the map accuracy issues that I bring up, I strongly recommend in addition to producers and users accuracies, that a Kappa statistic be reported in the map error matrices. The kappa is a measure of the difference between the actual agreement between reference data and classified Landsat data, and the chance agreement between the reference data and a random classification. The kappa serves as an indicator of the extent to which the percent correct values of an error matrix are due to a true agreement verses chance agreement. For example the accuracies in Tables III-17 and III-20 are 62% and 61% respectively for binary or two class maps. These results are not much better than what you would expect using a random classifier which would be 50%. Once the kappa values are included, I recommend that the authors fully discuss the low accuracies, the implications of the low accuracies, and options for future monitoring (e.g. using imputation methods, dropping the remote sensing, etc.)

#2-6 The data in Table 12 regarding the connectivity of older forest blocks does not make sense and I am wondering if all or most of the values in the table are in error. The purpose of this table is to show the distance between nearest neighboring older forest large blocks. For example, the mean distance for the Oregon Willamette Valley is 0.45-0.78 miles, but the Willamette Valley does not have any older forest large blocks. I do not understand how you can have a mean without any observations. In addition, the reported mean values for some of the other provinces do not seem correct. This type of problem calls into question the validity of the other data in this large and important report. I understand that this report has been reviewed a number of times, but I wonder about a simple quality control for the data summaries. I recommend that all of the data presented in this report be checked one more time by asking the simple question—“do the numbers make sense and do they look reasonable?”

In addition, I wonder about the author’s selection of the landscape metrics. What was the rationale for including the metrics in the report? What about the pros and cons of the metrics used? The authors should provide more interpretation on the results of the landscape metric analysis. Are these metrics diagnostic of the level of fragmentation?

**4. The conclusions. Are these appropriate given the results reported and the limitations and constraints inherent in the monitoring design and data? Should the conclusions be stronger/less strong?**

#2-7 Although, the pros and cons of plots verses satellite data were described, the authors seem to have minimized major differences between the plots and the Landsat data in some areas that had low classification accuracies. In the conclusion, the authors state, that at the plan scale, estimates between plot and remote sensing data were consistent. In this case, some of the errors among provinces/states cancelled each other out resulting in a 11% difference using the two methods (7.03 verses 7.87 million acres). The mapping overestimated older forest acres in California and underestimated older forest acres in Oregon and Washington. To me, these canceling errors were purely luck and if this did not occur, the overall 11% difference would have probably been much larger. I think the authors have glossed over some of these major problems in this conclusion and I recommend a rewrite that addresses the problems with mapped data accuracy and suggestions for future monitoring work.

**5. Are there changes to the program that should be considered? Is the program adequate to the task of providing effective monitoring for the Plan under the objectives set out?**

#2-8 For the future, I think you should consider using alternative approaches (e.g. imputation) to make more accurate and detailed maps and possibly rethink how you assess fragmentation. Is the fragmentation analysis needed? If so, I suggest more work on developing the appropriate spatial variables and more text on interpreting the fragmentation analysis.

#2-9 Another question that I have is -- Are satellite based maps a must or do the plots provide you with nearly all of the information you need? For plots to work for future updates, the schedule for reading the plots would need to match the schedule needed for updating the older forest database. If you move to a plot only monitoring system, you would not be able to do some of the spatial analysis of the older forests such as characterizing older forest blocks, degree of fragmentation, and connectivity between them. However, I am concerned about the accuracy of these spatial products in the current document (e.g. connectivity, etc.) and I wonder if these data will ever be used much for policy making. The connectivity data in this report (Table 12) are very difficult to interpret, probably have questionable accuracies, and may not be diagnostic of fragmentation levels.

#2-10 I understand that the national parks now have plots in place which would make the plot data set much more complete. (I am not sure if you would have all of the plots that you need for the BLM lands in California.) The big advantage of the plot data is that it provides an unbiased estimate of the population, it is great for analyzing both succession and disturbance, while the remote sensing data is typically only suitable for getting a rough estimate of the population and disturbance, but not succession.

#2-11 Also, in the future, I recommend that the plot data be used for the fire analysis. I think the plot data could provide much more detail for regional fire planning than the remote sensing data. For example, with the plots, data could be displayed by given strata on stocking and changes in stocking, on mortality from insect and disease outbreaks, and fuel loads. Conversely, the remote sensing data may be better for some types of small area analysis, fire risk analysis, and wildlife habitat analysis, but only if it is accurate.

#2-12 If you decide to use remote sensing data in the future, I recommend that you use imputation methods which integrate plots and spatial data. For broad scale assessments, these imputation methods have been shown to get close agreement between map-based and plot-based forest estimates. Another advantage of the imputation methods is that difficult variables can typically be mapped (e.g. dead wood, canopy layering, fuel loads, etc.). Hopefully, you will be able to take advantage of imputation and other newly developed remote sensing and spatial mapping techniques for making more accurate maps by the time you need to do another update.

**6. Other Comments**

#2-13 Class size headings are in error in Tables III-11 and III-23.

#2-14 It would be helpful if the authors would present the number of plots that were used in various summary tables. Also, if more summaries are included for additional strata, please give the number of plots with each presentation.

#2-15 I question the data in Table 12 as being correct and/or useful. This table shows the mean distance between centers of the nearest neighboring older forest blocks (>1000 acres). There is very little range (range = 0.16 to 0.45 miles-M&L) in these values and I find it hard to believe that the overall mean distance between centers of large blocks is only 0.18 miles. Please check the caption and change it to edge-to-edge distance if the caption is in error.

#2-16 Tables 11, 13, 17, and Figures 12,14,15,16,17. You should make it very clear where the data for these tables and figures comes from (is it from plots or remote sensing?). Please state sources in the captions.

#2-17 On page 55, the authors suggest that the reader compare Figure 21 with Figure 12 to compare the estimates of older forest amounts from plot data (Figure 21) and remote sensing data (Figure 12). Since one figure shows percentage and the other acres, I was not able to make this comparison. Why not put both together using the same units so the reader can see how these two compare? Overall, I would like to see a more explicit comparison from the two data sources.

#2-18 I am very troubled by low mapping accuracies in some of the provinces. For example, in the Eastern Washington Cascades province, the number of older forest acres was dramatically underestimated (<200,000 acres mapped verses >400,000 acres plot estimate). Accuracy for the > 20 inch class is 0% and there are fewer than 5 reference plots in the > 20 inch class. The overall accuracy for this two class map is reported as 79.59 % (Table III-11), which seems misleading or at the best confusing. I suggest you put the raw numbers in your error matrices and just like you did in the example error matrix in the beginning of Appendix 3. Also, it is important to add the kappa statistic to these error matrices as discussed above and discuss the resulting Kappa statistics in the text.

### Comments from Reviewer #3

#### **Review of Moeur et al.: “Northwest Forest Plan- the first 10 years: status and trends of late successional and old growth forest.”**

First, I would like to say that despite any criticisms of details, this is a remarkable piece of work for its geographic and ecological scope, its integration of diverse data types, and the overall magnitude of the undertaking. In these terms alone, it is a remarkable achievement and those involved are to be commended for bringing the task to completion. Overall, I find that the authors have made a good effort to achieve their goals and that this is a good first step on the way to a long-term process for monitoring and assessment. The document is generally clear and well written.

#3A. With regard to the general questions posed for the reviewers:

#3A 1. The overall design of the monitoring program is well conceived. It is hard to imagine an alternative to the approach taken that could achieve their goals. I strongly support the two phase strategy of both plot-based and remote sensing-based data sets. As the authors indicate, they are indeed complementary and the project as a whole would be significantly weaker without either one. However, the weakness of the remote sensing based methods for identifying subtle changes in forest status and disturbances less severe than stand replacing is significant, especially for the more fire dependent eastern and southern forest types. I recommend the authors consider increasing the plot based sampling effort, with a focus on these issues, until such time as the remote sensing technology and methods have improved in this regard.

#3A 2. The analytical approaches are reasonable and appropriate, given the status of our technology. However, it is clear that in the medium- to long-term, the remote sensing based methods will need to become better at identifying minor changes in forest structure associated with low-severity disturbances (fire, insects, partial harvesting) and developmental changes. This is clearly a limitation of the technology at this point, not the authors or the methods chosen. However, if future monitoring is to fulfill the expectations of spatial coverage that is the potential strength of the remote sensing methods, then these issues will need to be addressed.

#3A 3. The data are largely appropriate to the goals and the techniques are the best available. I would have liked to see more detailed analysis and discussion of the ecological implications of some of the data, but I presume this is being saved for formal publication in journal articles, etc. It would be useful to have a rationale for what is discussed and what is “reserved for later”. There is an enormous amount of data presented here and it is somewhat “under discussed”.

#3A 4. The conclusions are appropriate to the results and limitations of the monitoring design and data. They are stated with appropriate strength. One minor change I would like to see is that in the abstract and summary discussions it should be recognized more explicitly that the increase of about 600,000 acres of “medium and large” older forest does not represent a large ecological change in those acres, but the incremental growth over a (short) timeframe of 10 years moving those areas across an arbitrary analytical threshold. This is discussed appropriately in the body of the text.

#3A 5. I think that the program is appropriate to its objectives, given the scope of the task. Again, it is a rather remarkable effort to deal with an enormous task.

#3 B. Other more general or conceptual points:

#3 B 1. Given the functional ambiguity and arbitrariness of simple numerical definitions of old forest, I appreciated very much the use of the three definitions in this report. They are indeed complementary. If anything, I would like more discussion, both in the introductory and methodological sections and in the discussion of the issues around choosing definitions. There is a literature on this that has not really been referenced (e.g. papers by Spies and co-workers on the use of “indices of old growthness”, Wells et al 1998 in *Natural Areas Journal*). These are important issues and, in principle, I would expect this monitoring effort to move towards more explicitly ecologically based definitions over time. This document should lay the conceptual groundwork for that move. The continuum of possible definitions is referred to on several occasions, but the continuum is not referred to explicitly, other than the definitions used. What are other possible examples of definitions? What are the end points of the continuum? What direction do the authors hope to move in the future? This relates somewhat to what the authors refer to as an inability to address the “functionality of these results” – page 62. We would like to see issues of function addressed and the authors don’t say why they don’t think this can be done. I presume it is because it is difficult to make functional interpretations from definitions that are numerically arbitrary and a rather coarse filter for the ecological problems associated with late seral forests. Again, I don’t think it is reasonable to require different definitions or functional interpretations at this point, but the authors should lay the ground work for doing so in the future.

#3 B 2. The authors provide a lot of information on the accuracy/error inherent in the classification and use of the remote sensing data bases. While some data types are classified with a great deal of accuracy, others are relatively poor (e.g. 40-60%). The authors should provide some discussion of the potential problems arising from  $>1/3$  misclassified polygons in some cases. I realize that these numbers are within the bounds of what is normal for some remote sensing applications, but potential errors in interpretation of the results should be addressed. As well, readers unfamiliar with these types of data should be given some guidance in interpreting these magnitudes of error.

#3 B 3. Similarly, in the Results (p 53), we are given values for the amount of forest in various categories, e.g. 29% of federal forest-capable acres in “medium and large older forest”. Yet it is not really possible to assess the significance of these numbers without reference to some external measure, such as pre-industrial reference conditions. The idea of reference conditions comes up briefly in the discussion (pg 68), but we really need to have the idea introduced earlier. I recommend introducing the concept of reference conditions in the Introduction so the reader will anticipate that is how these numbers will be evaluated when they get to them. In general, we need more conceptual foundations for the application of the reference conditions idea. The idea of reference conditions and the values used here are in fact absolutely critical components of the assessment of whether Outcomes have been achieved discussed on Pg 68, but there has been no context established for their application. How were the reference conditions being used developed? Are they sufficient? Should this be an area for future work? Merely reporting the

numbers and citing a couple of references is not sufficient for such an important component of the project.

#3 B 4. The issue of needing more context raised in the previous item is in fact a somewhat more general issue. I imagine that there is not a strong literature review element to this report because it is expected that will be covered in journal publications arising from it. The vast majority of citations are to methodological items or the political/management context. I missed, however, a sense of conceptual review and context. The reference conditions item is one such, the issue of other efforts at “old growth” definitions in item

#1 is another. Similarly, efforts at using remote sensing to assess the extent and character of late seral forest elsewhere (e.g. in British Columbia) could be usefully reviewed.

#3 B 5. On Pg 57 it is stated that 10% of the medium and large category was regeneration harvested in the California Coast Range Province, and further, that this included 50% of the > 30” larger, multi-storied older forest. These are remarkable numbers, well in excess of the losses in any other area of the study, and very different from the overall conclusions for the Plan area. The authors should discuss the significance of this at least a little.

#3 B 6. Why are 90% of the stand replacing fires in older forest in reserved allocations (Pg 58)?

#3 B 7. Figure 23 needs a more explanatory caption – and perhaps an outline of the perimeter of the burned area.

#3 B 8. On page 68 there are various thresholds referred to (e.g. “rules thresholds for process and functions”). The concept of thresholds is intuitive here, but has not been mentioned before and the text seems to imply some specific known thresholds without providing references or context. There should be more supporting text given to this concept before it is used in this way.

#3C. Other minor or editorial points.

There are a number of cases where inappropriately colloquial, jargony, or otherwise poor language is used. Mostly, I think these reflect terms which are common in informal technical discussions but should be avoided in formal documents. There are a lot of cases where there are long strings of adjectival nouns used. While not formally “incorrect”, this tends to make the language less accessible to non-technical readers. These are minor points. Examples are:

#3C 1. Pg 15. “wall to wall”.

#3C 2. Pg. 20 “cross-walked”

#3C 3. Pg 41 “first occasion sampled plots”

#3C 4. Pg 41, ¶3 there is something missing in “direction of as part”.

#3C 5. Pgs 46, 69, and 70. “Transition” should not be used as a verb (as in “transitioning to ...”).

#3C 6. Pg. 61 “We analyzed ...” -- there is an “a” missing.