Protocol for Accuracy Assessment of Ecosystem Maps

2003

Ecology and Earth Sciences

BRITISH COLUMBIA

Ministry of Forests Forest Science Program

Protocol for Accuracy Assessment of Ecosystem Maps

Del Meidinger

BRITISH COLUMBIA

Ministry of Forests Forest Science Program

The use of trade, firm, or corporation names in this publication is for the information and convenience of the reader. Such use does not constitute an official endorsement or approval by the Government of British Columbia of any product or service to the exclusion of others that may also be suitable. Contents of this report are presented for discussion purposes only. Funding assistance does not imply endorsement of any statements or information contained herein by the Government of British Columbia.

National Library of Canada Cataloguing in Publication Data
Meidinger, Dellis Vern, 1953-
Protocol for accuracy assessment of ecosystem maps

(Technical report ; 011)

Includes bibliographical references: p.
ISBN 0-7726-5054-3

1. Ecological mapping -- Standards. 2. Environmental mapping -- Standards. 3. Ecological surveys -- Standards. 4. Vegetation mapping -- Standards. 5. Vegetation surveys -- Standards. 6. Biotic communities -- Mapping -- Standards. I. British Columbia. Forest Science Program. II. Title. III. Series: Technical report (British Columbia. Forest Science Program) ; 011.

QH541.15.M26M44 2003 577 C2003-960210-9

Citation
Meidinger, D.V. Protocol for accuracy assessment of ecosystem maps. 2003. Res. Br., B.C. Min. For., Victoria, B.C. Tech. Rep. 011.
<http://www.for.gov.bc.ca/hfd/pubs/Docs/Tr/Tr011.htm>

Prepared by
Del Meidinger
B.C. Ministry of Forests
Forest Science Program
PO Box 9519 Stn Prov Gov
Victoria, BC V8W 9C2

Copies of this report may be obtained, depending on supply, from:
Crown Publications
521 Fort Street
Victoria, BC V8W 1E7
(250) 386-4636
http://www.crownpub.bc.ca

For more information on Forest Science Program publications, visit our web site at
http://www.for.gov.bc.ca/hfd/pubs/index.htm

ABSTRACT

This report outlines a protocol for assessing the thematic accuracy of large-scale ecosystem mapping (i.e., Predictive and Terrestrial Ecosystem Mapping). The protocol presents a statistically unbiased approach to evaluate the acceptability or accuracy of the mapping. The thematic content of randomly selected map polygons (or small areas) is assessed by various means—the methods varying in precision and objectivity. The protocol requires the development of a sampling plan that articulates decisions about protocol level, sample size, assessment method, target error, etc. Although costly, assessing the accuracy of complex thematic maps is critically important to determining appropriate uses for ecosystem mapping. The approach outlined in this protocol provides a means of obtaining some overall, statistically valid scores to rate the accuracy of TEM, PEM, or other ecosystem maps. The results can be used as a component of quality assurance or for presenting statistics on the accuracy of mapping.

ACKNOWLEDGEMENTS

Del Meidinger prepared a draft of this protocol immediately following a workshop session (May 1999) attended by R. Sims, D. Moon, A.Y. Omule, K. Jones, B. Love, D. Clark, T. Lea, and D. Meidinger. In July 2000, the initial protocol document was prepared after the draft protocol had been applied for a year. Other individuals providing valuable input into the versions of this protocol include: Marvin Eng, who was available for technical discussions and provided information on creating a confusion matrix; Amanda Nemec, who with the advice and assistance of A.Y. Omule, prepared the sampling tables; Peter Ott provided statistical advice on the analysis of the data; A.Y. Omule provided reviews of comments and advice on incorporation into the manuscript. This technical report also represents views expressed in contract reports by R. Sims and K. Iverson, the "Sampling Team" of the British Columbia Forest Service's Vegetation Inventory Working Group, and previous work by Moon et al. Everyone's assistance is greatly appreciated.

A.Y. Omule, Dave Clark, Shikun Ran, Colleen Jones, David Moon, Dennis Lloyd, and Corey Erwin provided useful review comments.

Workshop Contributors and Affiliations (May 1999):

Clark, D., Resources Inventory Branch, B.C. Ministry of Environment, Lands and Parks, Victoria, B.C.

Eng, M., Research Branch, B.C. Ministry of Forests, Victoria, B.C.

Jones, K., K. Jones and Associates, Victoria, B.C.

Lea, T., Resources Inventory Branch, B.C. Ministry of Environment, Lands and Parks, Victoria, B.C.

Love, B., B.C. Ministry of Forests, Smithers, B.C.

Meidinger, D., Research Branch, B.C. Ministry of Forests, Victoria, B.C.

Moon, D., CDT Core Decision Technologies, Vancouver, B.C.

Nemec, A., I.S.R. Corp., Brentwood Bay, B.C.

Omule, A.Y., J.S. Thrower and Associates, Victoria, B.C.

Ott, P., Research Branch, B.C. Ministry of Forests, Victoria, B.C.

Sims, R., EBA Environmental Consultants, Vancouver, B.C.

CONTENTS

Abstract......iii

Acknowledgements iii

1 **Introduction**......1

2 **Features of the Protocol** 1

3 **Protocol Design** 2
3.1 Determining Appropriate Areas to Sample 3
3.2 Selecting Mapping Entities for Assessment......3
3.3 Determining Level of Sampling 3
3.4 Determining Sample Size and Selecting Polygons 4
3.5 Preparing a Sampling Plan 7
3.6 Sampling......7
3.6.1 Assessment using air photo interpretation 8
3.6.2 Assessment using air calls 8
3.6.3 Assessment using ground calls 9
3.6.4 Assessment using large-scale mapping 9
3.6.5 Assessment using plot sampling 9
3.6.6 Assessment using line-intercept transects 10
3.7 Summarizing Field Data 11
3.8 Scoring of Data and Reporting Assessment Statistics......11
3.8.1 Scoring ties......12
3.8.2 Graphing and evaluating ecosystem unit proportions for map 12
3.8.3 Determining proportion of dominant entity "correct" for polygons 13
3.8.4 Determining percent overlap for polygons 14
3.8.5 Determining percent acceptable overlap for polygons 15
3.8.6 Creating a confusion matrix 16

4 **Conclusion**......18

APPENDICES

1 Summary of accuracy assessment requirements for use of ecosystem mapping in timber supply modelling 19

2 Determining sample size......20

3 Methods for scoring small areas 21

4 Using Microsoft Excel's CRITBINOM function 22

References 23

TABLES

1 Sampling levels for accuracy assessment 4

2 Approximate sample sizes for assessment of map accuracy 5

3 Attributes collected for plot samples using ground inspection form 9

4 Compilation of field sample data and map polygon site series data for a TEM project 11

5 Evaluation of percent overlap for map area......13

6 Evaluation of TEM data to determine dominant site series 14

7 Calculation of average and weighted average dominant site series correct scores......14

8 Evaluation of TEM data to determine percent overlap 15

9 Calculation of average and weighted average polygon overlap scores 15

10 Evaluation of TEM data to determine percent acceptable overlap 16

FIGURES

1 Design of line-intercept transects to sample a hypothetical polygon 10

2 Designs of line-intercept transects to sample small areas 10

3 Graph of site series proportions......12

1 INTRODUCTION

This report outlines a protocol designed for assessing the accuracy of large-scale ecosystem mapping (i.e., Predictive and Terrestrial Ecosystem Mapping [PEM and TEM]). It also can be applied to any ecosystem mapping such as Broad Ecosystem Inventory (BEI), for example. The protocol should be used in determining the acceptability or accuracy of a map area with respect to the ecosystem units mapped. The resulting assessments present unbiased information to users regarding the accuracy of the mapping. In addition, the results can be used to evaluate acceptance of the mapping for contract administration purposes.

The protocol can be used for accuracy assessment or as a component of quality assurance. The difference between the two evaluations is the amount of on-the-ground field sampling required to evaluate the map. In either case, it is the thematic accuracy[1] of the map that is being evaluated. This is accomplished through an unbiased sample of the map's polygons and then an assessment of the thematic content through air photo interpretation, air calls, and (or) ground assessments.

Users may also want to assess a map's reliability for application to their particular requirement. Data collected using this protocol can be used to determine map reliability for selected interpretations of the ecological information.

2 FEATURES OF THE PROTOCOL

The objectives of this protocol are:

- to provide users with an unbiased approach to assessing the thematic accuracy of ecosystem maps; and
- to provide the flexibility necessary to conduct either a quality assurance or accuracy assessment and to accommodate various budgets.

The protocol presents a statistically unbiased approach to evaluate the acceptability or accuracy of the mapping. The thematic content of randomly selected map polygons (or small areas) is assessed by various means—the methods varying in precision and objectivity. These within-polygon assessment approaches are presented as multiple "levels" of the protocol. Users select a level based on their intended use of the data and the project budget.

The protocol requires the development of a sampling plan that articulates decisions about protocol level, sample size, assessment method, target error, etc. Some other features of the protocol include:

- The assessment is conducted by an independent party.
- The assessment is conducted by "experts"[2] and is backed up with field data.

1 Thematic accuracy is the correctness of polygon labelling. A polygon is correctly labelled when the attributes of the polygon fall within the defined attribute ranges of the map unit and its components.

2 An "expert" is someone who: 1) is very skillful; having much training and knowledge in some special field; 2) excels above all others in that special field of knowledge or capability; 3) is able to be relied upon to give the "correct answer" (modified from on-line definition at: http://www.medicalibration.com/our-expert-systems-(tm)/expert.doc/expert.htm).

- The assessment is generally done after the mapping is complete and the polygons are finalized; however, with an appropriate design, the assessment could be conducted concurrent with the mapping.
- The assessment is "blind" (i.e., the sampling crew does not know the map labels).
- The entire project map area is sampled; this may be a partial or full map sheet or an entire project area.
- The project area can be stratified (e.g., alpine and below alpine), but samples must be distributed in all strata.
- The variation within polygons is assessed by multiple plots, mapping at a larger scale, or transects, as most PEM and TEM polygons are complexes of ecosystem mapping units.
- The preparation of a "confusion matrix" of errors of omission[3] and errors of commission[4] is recommended when a map consists of simple units.
- The final scores provide data on both the accuracy in portraying the dominant map unit components and all map unit components (for compound map unit polygons).
- A score for the accuracy of "correct" plus "close" categories can also be reported.
- The assessment can be used for the pass or fail of a contractor's work when the protocol level and pass mark for each assessment score is specified in the contract.
- The results of the assessment are non-spatial—that is, they identify the level of accuracy, but they do not show, within the map or project area, where errors or inconsistencies occur. (Some indication of inaccuracies is shown by the results, but the sample size is often too small to know with certainty where all the errors occur.)

3 PROTOCOL DESIGN

The protocol involves selecting a set of polygons (or small areas) and assessing their thematic content using procedures consistent with the selected sampling "level." The steps are outlined here and described in more detail in sections 3.1–3.8.

1. Determine whether polygons or small areas will be assessed (Section 3.1).
2. Select which mapping entity will be assessed (i.e., site series, site modifiers, structural stage) (Section 3.2).
3. Determine which sampling level will be applied (Section 3.3).
4. Determine the number of polygons (or small areas) to be assessed and select them using a simple random (or stratified) sample, with replacement, proportional to size (Section 3.4).
5. Prepare a plan to sample polygons (or areas) that is consistent with the sampling level selected and other sampling decisions (Section 3.5).
6. Conduct sampling (Section 3.6).

3 Pixel or polygon incorrectly omitted from a class (e.g., site series).

4 Pixel or polygon incorrectly assigned to a particular class when it actually belongs in another class.

7. Summarize data in a table comparing sample results with map predictions (Section 3.7).
8. Evaluate whether site series proportions from the map predictions are similar to those determined by the sample (Section 3.8.2).
9. Score each polygon, and average for the map or project (or each stratum, as well), using protocol scoring to determine: the proportion of map where the dominant mapping entity is correct (Section 3.8.3); and the "overlap" score between the samples and the map entities (Section 3.8.4).
10. Score (if desired) the "close overlap" between the sample and the map entities, which allows acceptable "adjacent" mapping units to score one-half that of the correct unit (Section 3.8.5).
11. Develop a confusion matrix (if map of simple polygons and sufficient data collected) (Section 3.8.6).

3.1 Determining Appropriate Areas to Sample

Large-scale TEM (and most PEM) maps contain polygons that can be sampled using this protocol. That is, taking into account both Terrain Resource Information Management (TRIM) and global positioning system (GPS) errors, the polygon boundaries can still be determined, and the sampling crew can be certain that they are within the polygon being assessed. However, some PEM maps contain very small polygons with boundaries that do not follow observable features. In these cases, the approach is to sample small areas and then compare their thematic content to the polygons they intersect. If used, this method should be designed so that individual polygon scores can be calculated in addition to the scores for the small areas. To ensure that small areas can be adequately sampled, their size should not be too large, but they need to be large enough to sample a reasonable area and minimize the impact of TRIM and GPS errors (e.g., a minimum of 3–5 ha and no larger than an average TEM polygon [10–15 ha]).

3.2 Selecting Mapping Entities for Assessment

Mapping entities can vary somewhat, depending on the map or project objectives, available data, and methods used. In TEM, the mapping entities are site series, site modifiers, and structural stage. In PEM, site series are mapped, generally with slope or aspect site modifiers, and with structural stage as a separate layer. Therefore, all assessments can evaluate site series mapping accuracy; however, selecting other mapping entities for assessment depends on project objectives and funding. Accuracy scores should be calculated and reported separately for each mapping entity (e.g., a set of site series scores, a set of structural stage scores). In most cases, all entities that are mapped should be assessed.

3.3 Determining Level of Sampling

To accommodate the protocol's use for either accuracy assessment or as a component of quality assurance, and to allow for various levels of sampling effort, six sampling levels are designated (Table 1). Levels 1–3 are most appropriate as quality assurance checks of thematic accuracy; levels 4–6 are intended for accuracy assessments, but may also be used for quality assurance. A level is selected by balancing the objectives with the budget available for field sampling and analysis. Appendix 1 contains a summary of accuracy assessment requirements for maps intended for use in timber supply modelling.

TABLE 1 *Sampling levels for accuracy assessment*

Level	Primary application	Characteristics
1	Quality assurance	• 100% of sample polygons or areas evaluated by air photo interpretation, preferably at larger scale using computer-based, softcopy photogrammetric mapping technology • 10–25% of sample polygons or areas field-assessed by air or ground calls
2	Quality assurance	• 100% of sample polygons or areas evaluated by air photo interpretation, preferably at larger scale using computer-based, softcopy photogrammetric mapping technology • 26–50% of sample polygons or areas field-assessed by air and ground calls at a 75:25 ratio
3	Quality assurance	• 100% of sample polygons or areas evaluated by air photo interpretation, preferably at larger scale using computer-based, softcopy photogrammetric mapping technology • 51–75% of sample polygons or areas field-assessed by air and ground calls at a 75:25 ratio
4	Accuracy assessment	• 100% of sample polygons or areas assessed by ground checks[a] • 3–5 sample plots in polygon or small area, randomly or systematically located; map simple map entities at large scale (e.g., 1:5000); or conduct single-line, line-intercept sampling
5	Accuracy assessment	• 100% of sample polygons or areas assessed by ground checks[a] • 6–20 sample plots in polygon (area) randomly or systematically located; or two line-intercept transects located
6	Accuracy assessment	• 100% of sample polygons or areas assessed by ground checks[a] • 21–50 sample plots in polygon (area) randomly or systematically located; or three or more line-intercept transects located

a Accuracy assessment requires that polygon content be assessed with high confidence. Obvious or very general map units may allow for remote determination (see Section 3.6.1).

Statistical rigour and precision in the assessment of sample polygons or small areas generally increases when moving from level 1 to 6. Sample selection at all levels is statistically unbiased, but more objectivity in the within-polygon or area assessment is possible with increasing level. Although the number of polygons or areas sampled is the most important variable in determining the confidence and precision of the estimates, the method of within-polygon or area evaluation is also important.

3.4 Determining Sample Size and Selecting Polygons

Irrespective of protocol level, a random sample of polygons (or small areas) is selected. The larger the sample size, the greater the precision of the estimates. Table 2 shows approximate minimum sample sizes for estimating a proportion of correctly classified polygons. The sample size calculations are based on a selected confidence level, a maximum error of the estimated

TABLE 2 *Approximate sample sizes for assessment of map accuracy*

Confidence level	Maximum error (+/–)	Probability random point correctly classified on map 0.5	0.6	0.7	0.8	0.9
0.80	0.01	4107	2629	1480	658	166
0.80	0.02	1027	658	371	166	42
0.80	0.03	458	293	166	74	20
0.80	0.04	258	166	94	42	12
0.80	0.05	166	106	61	28	8
0.80	0.06	115	74	42	20	6
0.80	0.07	86	55	31	15	5
0.80	0.08	66	42	25	12	4
0.80	0.09	52	34	20	10	4
0.80	0.10	42	28	16	8	4
0.90	0.01	6766	4331	2437	1084	272
0.90	0.02	1693	1084	611	272	70
0.90	0.03	753	483	272	122	32
0.90	0.04	425	272	154	70	19
0.90	0.05	272	175	99	45	13
0.90	0.06	190	122	70	32	10
0.90	0.07	140	90	52	24	8
0.90	0.08	108	70	40	19	6
0.90	0.09	85	55	32	15	6
0.90	0.10	70	45	26	13	5
0.95	0.01	9606	6149	3460	1539	387
0.95	0.02	2403	1539	867	387	99
0.95	0.03	1070	685	387	173	45
0.95	0.04	603	387	219	99	27
0.95	0.05	387	248	141	64	18
0.95	0.06	269	173	99	45	13
0.95	0.07	198	128	73	34	10
0.95	0.08	153	99	57	27	9
0.95	0.09	121	78	45	22	7
0.95	0.10	99	64	37	18	7

proportion, and the probability that a polygon is correctly classified. Because the calculation uses binomial statistics (i.e., correct or not correct), the table values are most appropriate for determining the sample size used for the percent dominant correct score (see Section 3.8.3). For the percent overlap score (see Section 3.8.4), the standard "normal statistics" sampling equation can be used (Appendix 2).

Sample size is designated for the entire map or project area and results in a map or project area estimate. If estimates are required for each map entity or each stratum, then the sample size must be multiplied by the number of classes (if you want to have the same confidence and error in the estimate). This will likely result in a large number, so most assessments will be for the whole project area.

For example, if a maximum error of ±5% is required on the estimate of accuracy, 9 times out of 10, and the probability of any point being correctly mapped is unknown, use the table to look up a confidence value of 0.90, a maximum error of 0.05, and the lowest probability of a correct classification of 0.5. This yields a value of 272; therefore, the sample size is 272 polygons.

To assess the accuracy of TEM or PEM for timber supply modelling (Appendix 1), use a maximum error value of 0.07 (7%), a minimum confidence level of 0.80, and the probability of a point being correct of 0.50. This yields a value of 86. Therefore, in most cases, a sample of at least 86 polygons is required to determine the percent dominant correct score within the required confidence and error. If information is available on the actual probability of a mapped point being "correct," then that probability can be used to determine an appropriate sample size. This information might be for an adjacent, previously assessed area that is ecologically similar and was mapped using the same procedures and the same individuals, or it might be from an initial sample.

To calculate a sample size for the percent overlap score, a set of overlap scores from a sample of polygons is required. By using "normal statistics" to determine the standard deviation of the overlap scores, a target error value of 0.05 (5%), and a confidence level of 0.90, the sample size can then be calculated (Appendix 2). The required sample size will be about the same as that determined for the percent dominant correct using the procedures above. The sample is likely to represent only a small proportion of the total polygons, but, as the cost of ground sampling is quite high, this is acceptable.

For very diverse areas, a sample of 60–100 polygons may not represent a reasonable proportion of the ecological map entities. Therefore, to assess the accuracy of TEM or PEM, and determine their acceptability for timber supply review, the minimum number of sample polygons should be at least twice the number of forested site series (Appendix 1).

Most ecological map users want to know how much of a map is "correct" (i.e., to assess the map's acceptability for timber supply analysis). To do this, select polygons randomly, with replacement (i.e., a polygon could be selected more than once), with the probability of selection proportional to area. Alternatively, select a simple random sample of points or grid intersections (where the polygon or small area that the point falls in is sampled). Some map users may simply want to know how many polygons are "correct." If so, polygons can be selected as a simple random sample, with replacement.

It may be desirable to stratify the map into a few strata and sample some strata more than others; however, all strata of interest require some sampling. Examples of useful strata include:

- alpine and below alpine, where the alpine is a large proportion of the map area and more data are required, proportionately, from the forested areas; and
- high-elevation (inaccessible) areas versus low-elevation areas, where inaccessible areas are suspected of having a lower accuracy; in this case, sample each stratum with a reasonable number of samples.

Note that each stratum may not meet required targets of acceptable error unless the sample size is large enough.

To reduce travel costs between sample polygons, some cluster sampling may be acceptable. Sets of adjacent or nearby polygons are selected at random and all polygons within a sample cluster are checked. On large projects, considerable travel time between sample polygon areas may be involved. Sampling two or more random polygons within a certain distance of each other reduces the costs of field sampling; however, only a few polygons per cluster are recommended to ensure that the sampled polygons cover the geographic and ecological variation within the map area.

Polygons of non-vegetated, sparsely vegetated, and anthropogenic[5] units are usually not ecological units and generally should be excluded from the accuracy assessment. Similarly, broadly classified units, such as "wetland" or "alpine," may be excluded from the assessment because they provide little useful interpretive value. Decisions on types of polygons excluded from the sample should be clarified in the sampling plan.

3.5 Preparing a Sampling Plan

A sampling plan is required for two reasons.

1. It clarifies the assessment objectives, protocol sampling level, and sample size.
2. It provides for the efficient conduct of fieldwork.

The protocol has several options, including accuracy assessment versus quality assurance, sampling level, stratification, sample size, method of assessment, and number of plots or transects per polygon or area. Therefore, the objectives of the assessment should be clearly stated and the sampling decisions outlined. This allows for straightforward planning of the fieldwork. The sampling plan should contain the following information:

- a description of the area to be assessed (e.g., administrative unit, map sheets);
- objectives of the assessment;
- protocol sampling level (from Table 1);
- sample accuracy decisions (confidence level, maximum error, probability, strata);
- sampling decisions (polygon or small area assessment, number of polygons or areas to sample, method of selection, number of polygons for field checking, method of field assessment, description of strata, quality control, map units excluded from assessment);
- for each polygon or area, a sampling plan for the within-polygon or within-area assessment, and the access plan (e.g., air photos; UTM coordinates; closest access, distance, and bearing from access point; difficulty of access); and
- any deviations from this protocol.

The sampling plan should be acceptable to the mapper (as their work is being assessed) and to the client (as they want to use the map). Where the map is for use in timber supply modelling, the sampling plan must also be acceptable to the Ministry of Forests (see Appendix 1).

The following sections elaborate on some issues related to the preparation of sampling plans.

3.6 Sampling

Polygons or small areas can be assessed by air photo interpretation, air calls, ground calls, large-scale mapping, plot sampling, or line-intercept transects (see Table 1). These methods are discussed in sections 3.6.1–3.6.6. The chosen method (or methods) should be consistent with the purpose of the accuracy assessment, stay within the budget available, and ensure that all polygons can be evaluated. The assessment should be conducted "blind"; that is, the sample crew should not know the site series or PEM entity label.

5 Mapcodes for these units are available in the TEM mapcode file at: http://srmwww.gov.bc.ca/ecology/tem/list.html

To ensure that the site series assessments are acceptable to all parties, an expert in the area's ecosystems should lead each sampling crew. In addition, the sampling crew must implement appropriate quality control procedures to ensure that data are collected to a high standard. The location of transects and plots must be accurately determined and the crew well trained and evaluated (i.e., audited) during the field collection. Iverson (2002) looked into quality control standards for an accuracy assessment project that involved line-intercept sampling and recommended at least an 85% "overlap" between the auditor and field crew ecosystem identifications when crews are evaluated.

Polygons that are too dangerous for ground sampling because of extreme slopes may require assessment by air photo interpretation or by an air call. Polygons that are inaccessible because of distance or tenure can be sampled remotely or replaced in the sample with an accessible polygon of the same attributes. The procedure for selecting these "matched substitutes" must attempt to minimize bias (e.g., sample chosen randomly from sorted list of polygons with similar attributes) and must be documented in the accuracy assessment report. To maintain a "blind assessment," the individual selecting the replacement polygons should not be part of the field crew (if at all possible).

All map entities observed or sampled in the polygon or area should be recorded. In some cases (e.g., transitional areas or early successional stands), the person conducting the accuracy assessment may be uncertain of the site series because of the intermediate or inconclusive nature of the assessed features. In such cases, two map entity calls should be made—a "primary" call, as well as an "alternate" or secondary call. The primary call is the field person's determination of the "most likely" ecosystem map entity, whereas the alternate is the next most likely entity. In the accuracy assessment scoring, the alternate or secondary calls can be used to see if they "improve" the assessment (as the mapper should not be penalized if the field person cannot differentiate clearly between two map entities).

When small areas are sampled (see Section 3.1), the assessment procedure should allow for the evaluation of both the small area and its individual polygons. A decision to assess small areas instead of polygons is appropriate in some circumstances (e.g., very small polygons, due to likely TRIM/GPS errors). However, the "scores" for both the small areas and the individual polygons should be determined.

3.6.1 Assessment using air photo interpretation Assessing polygons using air photo interpretation is a viable procedure for quality assurance, but is rarely possible for accuracy assessment unless the map entities can be easily and reliably determined by air photo interpretation. For air photo interpretation, the following materials are necessary:

- for TEM: larger-scale photos, or computer-based, softcopy photogrammetric assessment that allows enlargement of photos.
- for PEM: preferably larger-scale photos, or computer-based, softcopy photogrammetric assessment, but even same-scale photos are acceptable.

When approved for accuracy assessment, the procedure should involve two photo interpreters evaluating the polygons and reconciling any differences in their evaluations.

3.6.2 Assessment using air calls Assessing the composition and proportions of map entities in a polygon using air calls is sometimes possible by hovering over the site in a helicopter. An "air call" is also possible from a viewpoint, if the site is clearly visible and the important features observable.

3.6.3 Assessment using ground calls Ground calls are conducted by traversing enough of the polygon's area to assess the composition and proportions of the map entities.

3.6.4 Assessment using large-scale mapping Mapping a polygon at a larger scale, using procedures similar to those used for silviculture prescriptions, is another way of estimating the composition and proportions of the map entities. Procedures are presented in B.C. Forest Service field guides to site identification and in the *Silviculture Prescription Data Collection Field Handbook* (Curran et al. 2000).

3.6.5 Assessment using plot sampling Sample plots should be randomly or systematically located within each polygon. A suggested procedure is to locate samples on a square grid, starting at a random location within the polygon. Plots are installed at the grid intersections. The spacing between the sample plots depends on the number of plots to be installed in the polygon and can be determined by the following formula:

$$\text{grid spacing in metres} = \sqrt{(10\,000 \times \text{polygon area in ha}) / (\text{no of plots})}.$$

Only the starting point and the grid spacing are required. Another procedure is to plot a transect through the polygon area, and for each required sample, randomly select a distance along the transect, a direction off the transect (right or left), and a distance perpendicular to the transect. If planned in the office, the field stage is straightforward, but it is essential to accurately locate the start of the grid or transect and to correct distances for slope.

At each sample plot, data to identify the mapping entities (site series and possibly site modifiers and structural stage) are collected using the Ground Inspection Form (or equivalent), following the standards in the *Field Manual for Describing Terrestrial Ecosystems* (B.C. Ministry of Environment, Lands, and Parks and B.C. Ministry of Forests 1998). Data are collected in a 10 m radius plot and the mapping entities and their proportions in each plot are recorded.

A list of attributes recommended for collection on each site is presented in Table 3. If the protocol is being implemented to determine the reliability of secondary interpretations, attributes reflecting those interpretations should also be collected.

TABLE 3 *Attributes collected for plot samples using ground inspection form*

1. Air photo number
2. Date
3. Project ID
4. Surveyor(s)
5. Mapsheet
6. Plot no. (polygon + sample)
7. Polygon no.
8. Lat./Long. or UTM co-ordinates
9. Aspect
10. Elevation
11. Slope
12. Soil moisture regime
13. Soil nutrient regime
14. Mesoslope position
15. Drainage: mineral or organic soils
16. Mineral or organic soil texture
17. Surface organic horizon thickness
18. Humus form (to Order level)
19. Depth to and type of restricting layer (if applicable)
20. Coarse fragment content
21. Terrain texture, surficial material, and surface expression
22. Biogeoclimatic unit
23. Ecosection
24. Site series
25. Site modifiers
26. Structural stage
27. Crown closure
28. Using "Ecosystem Polygon Summary" section, record each site series in plot and their proportions (% of plot)
29. Total % cover by stratum
30. Dominant/indicator plant species
31. Percent cover of dominant/indicator species
32. Complete or partial data
33. Notes

3.6.6 Assessment using line-intercept transects Line-intercept sampling along transects within polygons or small areas is a preferred method of assessing map entity composition and proportions. Traverses should be of an appropriate length and should be located so that the range of variation within a polygon or area is sampled. Figure 1 illustrates a design used to assess polygons in the southern Okanagan accuracy assessment (Ryan et al. 2002). Various transect shapes (Figure 2) are available to sample small areas (e.g., triangle, square, radial distribution). Whichever transect shape is chosen for sampling, it should be acceptable to both the mapper and the accuracy project client.

Each site series change along a traverse should be recorded, as well as an accurate determination of the distance along the traverse.

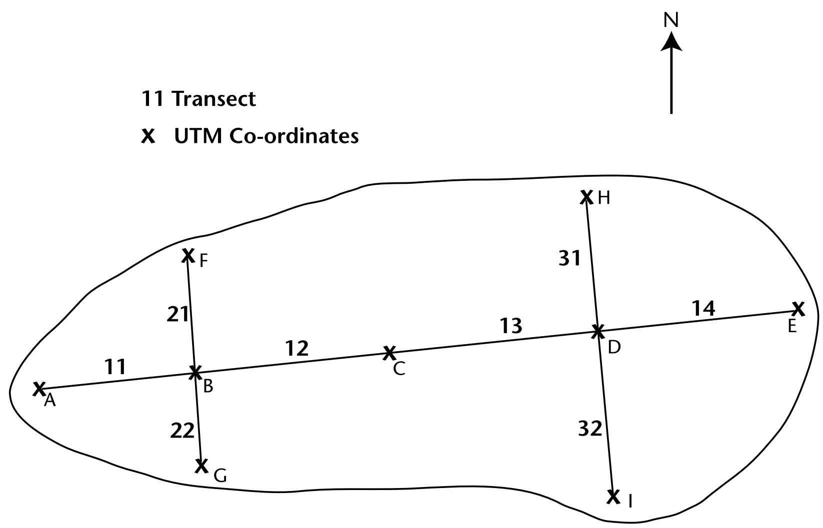

FIGURE 1 *Design of line-intercept transects to sample a hypothetical polygon (numbers identify transect segments within polygon).*

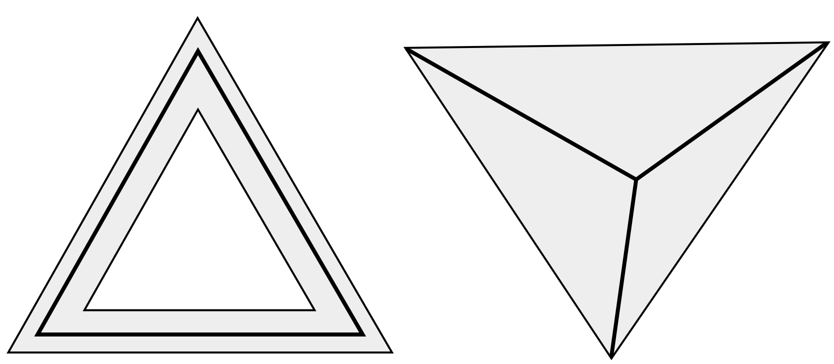

FIGURE 2 *Designs of line-intercept transects to sample small areas (thick lines represent traverses).*

3.7 Summarizing Field Data

After fieldwork, the field data should be compiled (e.g., Table 4) and map entities and their proportions from sampling and from the map polygon data recorded for each polygon or area. Although the map may present the ecosystem map entities only as deciles (10% classes), sample data should be recorded to the nearest percent. Similar tables should be prepared for structural stage or site modifiers (if these are to be assessed). If alternate or secondary calls are made during the field sampling, additional columns with these data will be necessary.

To calculate the weighted averages of some estimates, the areas of sample polygons will be required. To determine map proportions, the total area of each map entity should be determined.

TABLE 4 *Compilation of field sample data and map polygon site series data for a TEM project*

	Map polygon data										Sample polygon summary								
TEM poly #	Zone	Subzone	Vrt	% 1[a]	SS 1[b]	% 2	SS 2	% 3	SS 3	Area (ha)	Zone	Subzone	Vrt	% 1	SS 1	% 2	SS 2	% 3	SS 3
2184	CWH	vm	1	100	AB					8.4	CWH	vm	1	50	AS	50	AB		
2181	CWH	vm	1	70	AB	30	HS			6.6	CWH	vm	1	60	AB	15	AS	5	AF
2205	CWH	vm	1	80	AF	20	SD			4.6	CWH	vm	1	80	AB	20	AS		
2218	CWH	vm	1	100	AB					10.4	CWH	vm	1	85	AB	15	AS		
2183	CWH	vm	1	80	AB	20	AS			6.2	CWH	vm	1	70	AS	30	HD		
2172	CWH	vm	1	100	AB					5.8	CWH	vm	1	75	AB	25	HD		
2163	CWH	vm	1	70	AB	30	HS			3.2	CWH	vm	1	100	AB				
2176	CWH	vm	1	60	AB	40	AF			4.8	CWH	vm	1	80	AS	20	AF		

a %1 refers to percent of first (dominant) ecosystem map entity in polygon; %2 is proportion of second map entity; %3 is percent of third map entity.

b SS 1 contains the TEM map code for first ecosystem map entity (e.g., site series); SS 2 and SS 3 contain the map entity codes for second and third map entities, if present.

3.8 Scoring of Data and Reporting Assessment Statistics

Data from the assessment will be used to provide map users with information on the accuracy of the mapping entities and to compare the accuracy of different mapping methods (e.g., TEM and PEM). To meet these objectives, a set of graphs and several "scores" are calculated and summarized:

- Graph of map entity proportions determined from map predictions compared to proportions determined by sampling[6].
- Proportion of the area where the dominant entity mapped is the same as that determined by sampling.
- Percent overlap in the entities mapped versus those determined by sampling, averaged for the sampled polygons and for the entire map area.
- Percent acceptable overlap, where acceptable mapped entities are compared to those determined by sampling.
- Confusion matrix, where mapped entities are compared to sampled values.

Details about these statistics are presented in sections 3.8.2–3.8.7.

6 The chi-squared test of proportions is replaced by a simple graphing of the proportions and an overall project "overlap."

If some polygon areas required alternate or secondary calls, scores should be calculated for both the primary calls and for the primary plus secondary calls combined (assuming that not all polygons will have alternate calls). If small areas are sampled instead of individual polygons, overlap and dominant correct scores should be calculated for the small areas and for the polygons within the small areas (if possible). This ensures that users have accuracy information for the whole map, small areas within the map, and for individual polygons. Appendix 3 presents examples of methods for scoring small areas (e.g., calculating the small area overlap score and the weighted polygon overlap score).

3.8.1 Scoring ties "Ties" between map entities within a polygon occur in predictive ecosystem mapping where the model determines that two or more entities are equally likely to occur. Ties often happen where the input data are too general to differentiate between the range of ecosystem units being predicted. They are particularly troublesome for scoring and for map interpretation. Ties should be avoided wherever possible either by ensuring that adequate input data exist to predict the ecosystem units or by employing some reconciliation method after the initial mapping. If ties occur in the final map, they must be evaluated consistently in the accuracy assessment.

Many interpretive uses of the map will treat ties as a proportion of the polygon, even though this was not intended. Therefore, for the purposes of accuracy assessment, ties should be treated as equal proportions. For example, a tie of two map entities should be treated as if they are each 50% of the polygon.

3.8.2 Graphing and evaluating ecosystem unit proportions for map A graph should be prepared that compares the proportion of ecological map entities determined from the map predictions (using total map area of ecological entities included in the accuracy assessment) with those determined from the sample data (see Figure 3). This simple visual comparison partially replaces the chi-squared test of ecosystem proportions recommended in the previous version of this protocol (Meidinger 2000). To view the comparisons, a series of graphs may be required (e.g., one for each biogeoclimatic unit).

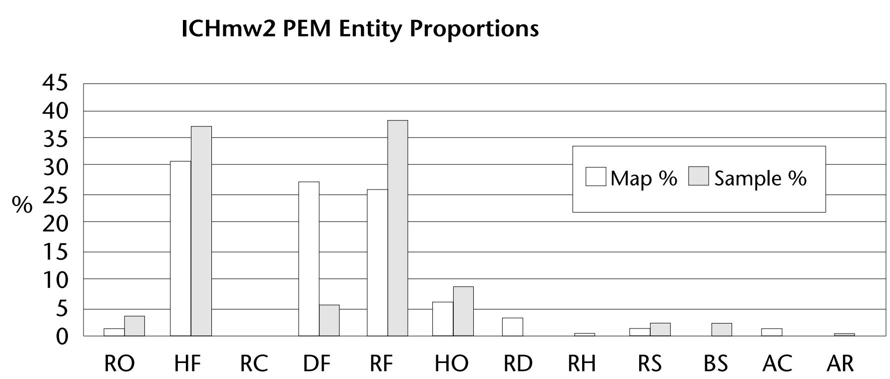

FIGURE 3 *Graph of site series proportions (abbreviations on the x-axis are codes for map entities).*

Areal extent of a site series, for example, is a common interpretive use of an ecosystem map. By examining the graph, it is possible to assess whether or not the proportions on the map are approximately the same as those determined by the random sample. The graph may also demonstrate whether some map entities are over- or underestimated.

In order to further evaluate the ecosystem unit proportions for the map, a percent overlap score can be determined of the proportions of ecological map entities for the entire map compared to the total for all the samples. Table 5 presents an example evaluation of the map overlap score (in which plots are used instead of polygons). If polygons are sampled, the total of all polygon areas for each map entity should be calculated. If line-intercept sampling is used, the map entity proportion is obtained by totalling all line segments of each map entity. This statistic provides an index of similarity between the proportions of the mapped entities and the sampled entities.

TABLE 5 *Evaluation of percent overlap for map area*

Map entity	Area (ha)	Map %	Plots (n)	Plot %	Overlap (%)
RO	2581	1.47	3	3.85	1.47
HF	55319	31.51	29	37.18	31.51
RC	205	0.12			
DF	49847	28.39	5	6.41	6.41
RF	46397	26.42	30	38.46	26.42
HO	10958	6.24	7	8.97	6.24
RD	6730	3.83			
RH	1029	0.59			
RS	2349	1.34	2	2.56	1.34
BS	167	0.10	2	2.56	0.10
	175581	100.00	78	100.00	73.48

3.8.3 Determining proportion of dominant entity "correct" for polygons

The compiled data for each polygon (e.g., Table 4) will show whether the dominant entity from the map corresponds to the dominant entity from the polygon sampling; a yes or no score is recorded (see example in Table 6). In the case of a tie in dominance (e.g., 50:50 or 40:40:20), mapping should be considered correct half of the time and incorrect the other half. This statistic is reported as the "% dominant correct" to the nearest percentage, with the appropriate confidence interval (CI) around the estimate (using the binomial distribution). In the case of a map with only simple polygons (i.e., only one site series for each polygon), this statistic could be labelled "percent correct." If using the map for timber supply analysis, the target for the percent dominant correct score is ±7% with 80% confidence—in these cases, the 80% confidence intervals should be reported.

The statistic should be reported in two ways: as an average of the polygons sampled and weighted by the area of the sampled polygons. The former indicates the proportion of the time that polygons have the "correct" dominance and the latter the proportion of the map area that is considered "correct."

TABLE 6 *Evaluation of TEM data to determine dominant site series*

	Map polygon data									Sample polygon summary									Score
TEM poly #	Zone	Subzone	Vrt	% 1	SS 1	% 2	SS 2	% 3	SS 3	Zone	Subzone	Vrt	% 1	SS 1	% 2	SS 2	% 3	SS 3	Dominant correct
2184	CWH	vm	1	100	AB					CWH	vm	1	50	AS	50	AB			Y
2181	CWH	vm	1	70	AB	30	HS			CWH	vm	1	60	AB	20	AS	20	AF	Y
2205	CWH	vm	1	80	AF	20	SD			CWH	vm	1	80	AB	20	AS			N
2218	CWH	vm	1	100	AB					CWH	vm	1	90	AB	10	AS			Y
2183	CWH	vm	1	80	AB	20	AS			CWH	vm	1	70	AS	30	HD			N
2172	CWH	vm	1	100	AB					CWH	vm	1	80	AB	20	HD			Y
2163	CWH	vm	1	70	AB	30	HS			CWH	vm	1	100	AB					Y
2176	CWH	vm	1	60	AB	40	AF			CWH	vm	1	80	AS	20	AF			N

For example, using the results presented in Table 6, the "average % dominant correct" is 5 out of 8 (x 100) = 62%. The single tie (first row) was scored "correct"; if another tie was found, it would be scored "incorrect." The confidence interval can be calculated using the "CRITBINOM" function in Microsoft Excel (see Appendix 4). In the example above, where the sample size is only 8, the lower 80% confidence interval value is 37.5% and the upper 87.5%. If 100 polygons were sampled, the lower 80% confidence interval value would be 56% and the upper 68% (±6%). The BINOMDIST function can also be used, but is slightly more complicated.

Table 7 shows summary data for the area of each polygon and the calculation for the "weighted % dominant correct" for the polygon data presented in Table 6. The weighted score for percent dominant correct is 69%, which is greater than the simple polygon average score.

3.8.4 Determining percent overlap for polygons By using the compiled data for each polygon (e.g., Table 4), the "% overlap" can be assessed by comparing the map entity proportions to the sample polygon proportions. For this purpose, the sample is assumed correct. The score is presented as both a simple average of the polygons sampled and weighted by the polygon area. Confidence intervals are reported consistent with the confidence selected for the sample size estimate, which in most cases is 90%. For maps used for timber supply analysis, the target is ±5% error, with 90% confidence.

TABLE 7 *Calculation of average and weighted average dominant site series correct scores*

TEM poly #	Area (ha)	Evaluation	Score	Average score contribution	Weighted average contribution
2184	8.4	Y	1	0.125	0.168
2181	6.6	Y	1	0.125	0.132
2205	4.6	N	0	0	0
2218	10.4	Y	1	0.125	0.208
2183	6.2	N	0	0	0
2172	5.8	Y	1	0.125	0.116
2163	3.2	Y	1	0.125	0.064
2176	4.8	N	0	0	0
				0.625	**0.688**

TABLE 8 *Evaluation of TEM data to determine percent overlap*

	Map polygon data									Sample polygon summary									Score
TEM poly #	Zone	Subzone	Vrt	% 1	SS 1	% 2	SS 2	% 3	SS 3	Zone	Subzone	Vrt	% 1	SS 1	% 2	SS 2	% 3	SS 3	% overlap
2184	CWH	vm	1	100	AB					CWH	vm	1	50	AS	50	AB			50
2181	CWH	vm	1	70	AB	30	HS			CWH	vm	1	60	AB	20	AS	20	AF	60
2205	CWH	vm	1	80	AF	20	SD			CWH	vm	1	80	AB	20	AS			0
2218	CWH	vm	1	100	AB					CWH	vm	1	90	AB	10	AS			90
2183	CWH	vm	1	80	AB	20	AS			CWH	vm	1	70	AS	30	HD			20
2172	CWH	vm	1	100	AB					CWH	vm	1	80	AB	20	HD			80
2163	CWH	vm	1	70	AB	30	HS			CWH	vm	1	100	AB					70
2176	CWH	vm	1	60	AB	40	AF			CWH	vm	1	80	AS	20	AF			20

Table 8 presents an example showing the evaluation of the polygon overlap. In the first polygon, the polygon was labelled 100% AB and the sample indicated that the polygon was 50% AB. The overlap would be scored as 50% (i.e., the mapped site series had 50% of the sampled site series for the polygon). The mean percent overlap score for the eight polygons is 390/8 = 49%.

Where all map units are simple (i.e., one map entity) it is still possible to do this assessment as the sample polygon assessments may not be of one map entity. The "simple" polygon proportion is 100% whereas the accuracy assessment proportions are whatever is found within the polygon.

To obtain the proportion of the map area that is considered "correct," the polygon scores are weighted by the area of the sample polygons (Table 9). In this example, the weighted overlap score is 53%, which is greater than the polygon average score.

TABLE 9 *Calculation of average and weighted average polygon overlap scores*

TEM poly #	Area (ha)	Overlap score	Average score contribution	Weighted average contribution
2184	8.4	50	6.25	8.4
2181	6.6	60	7.5	7.92
2205	4.6	0	0	0
2218	10.4	90	11.25	18.72
2183	6.2	20	2.5	2.48
2172	5.8	80	10	9.28
2163	3.2	70	8.75	4.48
2176	4.8	20	2.5	1.92
	50.0		**48.75**	**53.20**

3.8.5 Determining percent acceptable overlap for polygons Using the compiled data for each polygon (e.g., Table 4), the "% acceptable overlap" can be assessed by comparing the map entity proportions to the sample polygon proportions (as in Section 3.8.4), but one-half of the overlap score for "acceptable" adjacent classes is scored relative to the correct one. For site series, this is generally adjacent site series on the edatopic grid; however, not all adjacent site series are ecologically similar so it is best to prepare a matrix of acceptable adjacent site series to assist in scoring. For example, a rich floodplain site series may be adjacent to a poor swamp ecosystem on the grid, but they should not be confused in mapping and have very different interpretations. In this case, a partial score should not be given for mapping one in place of the other. The intent of this score is to allow some flexibility in evaluation of mapping where similar site series are difficult to differentiate accurately. When scoring structural stage, "acceptable" would be one class younger or older.

The score is presented as both a simple average of the acceptable polygon overlap comparisons and weighted by the area of the sample polygons to the nearest percent, with 90% confidence intervals (CI) around the estimate.

Table 10 presents an example evaluation of percent acceptable overlap. In the first polygon, the map label was 100% AB; the sample indicated that the polygon was 50% AB and 50% AS. Because the AS site series is diagonal on the edatopic grid for the CWHvm1 (but is not adjacent), the overlap score would be 50% for the AB portion and zero for the AS portion. For the second polygon, the overlap score is 60% with an additional 10% (one-half of 20%) added for the "acceptable" comparison between the AF and AB site series. The AF site series is adjacent to the AB site series on the edatopic grid for the CWHvm1. The average "% acceptable overlap" score for the eight polygons is 500/8 = 62.5%. The weighted average score can be calculated using the procedure illustrated in Table 9.

TABLE 10 *Evaluation of TEM data to determine percent acceptable overlap*

	Map polygon data									Sample polygon summary									Score
TEM poly #	Zone	Subzone	Vrt	% 1	SS 1	% 2	SS 2	% 3	SS 3	Zone	Subzone	Vrt	% 1	SS 1	% 2	SS 2	% 3	SS 3	% Acceptable Overlap
2184	CWH	vm	1	100	AB					CWH	vm	1	50	AS	50	AB			50
2181	CWH	vm	1	70	AB	30	HS			CWH	vm	1	60	AB	20	AS	20	AF	70
2205	CWH	vm	1	80	AF	20	HD			CWH	vm	1	80	AB	20	AS			50
2218	CWH	vm	1	100	AB					CWH	vm	1	90	AB	10	AS			90
2183	CWH	vm	1	80	AB	20	AS			CWH	vm	1	70	AS	30	HD			35
2172	CWH	vm	1	100	AB					CWH	vm	1	80	AB	20	HD			90
2163	CWH	vm	1	70	AB	30	HS			CWH	vm	1	100	AB					85
2176	CWH	vm	1	60	AB	40	AF			CWH	vm	1	80	AS	20	AF			30

3.8.6 Creating a confusion matrix A confusion matrix, also known as an error or contingency matrix, is commonly used to assess the accuracy of resource classifications. This assessment process is well developed for remote sensing and satellite image analysis. For assessments of ecosystem mapping accuracy, a confusion matrix can be prepared where map entities are simple (i.e., one per polygon). Where map entities are complexes of mapping entities, preparing a confusion matrix can be difficult or impossible because it may not be clear which ecosystems are being "confused" with others.

A confusion matrix can be used to estimate:

- the overall accuracy of a classification,
- the average accuracy for all entities in the classification, or
- the specific accuracy for a given item.

The last two estimates require knowledge of errors of omission and commission (see below).

To create a confusion matrix:

- list categories of "things" in the classification system in both row and column headings.
- label either the row or the column as "Known to be" (usually from ground-based plots in natural resource mapping).
- label the other as "Classified as" (usually air photo interpretation or digital image classification in natural resource mapping).
- enter each "sample" (usually from ground-based plots in natural resource mapping) in the appropriate cell (i.e., what it was known to be and what it was classified as).
- sum the rows and columns.

Commission errors are represented by non-diagonal elements in the matrix where a known "thing" (a ground plot) is classified into a category to which it does not belong. In other words, *the mapper* ***committed*** *the act of getting it wrong. Omission errors* are represented by non-diagonal elements in the matrix where a known "thing" is not classified into the category to which it belongs. In other words, *the mapper* ***omitted*** *the act of getting it right*. Each commission error in a given category is also an omission error for a different category. Therefore, the overall error is the sum of the off-diagonal elements, divided by the total number of elements. For example:

	Known to be:				
Classified as:	Bog	Fen	Marsh	Total	Omission
Bog	10	1		11	9%
Fen	5	10		15	33%
Marsh	1	1	5	7	28%
Total	16	12	5	33	
Commission	37%	17%	0%		

Overall Error (8/33) = 24%
Average Omission Error = 23%
Average Commission Error = 27%

- *Commission Errors:* Six bogs were misclassified (five as fens and one as a marsh). Two fens were misclassified (one as a bog and one as a marsh). No "known" marshes were misclassified.
- *Omission Errors:* Of those classified as bogs, one was a fen; of those classified as fens, five were bogs; and of those classified as marshes, one was a bog and one was a fen.
- *Each Error Is Both:* For the "known" bog "classified" as a marsh (lower left cell of the table) we committed the error of not classifying it as a bog. Our omission error was to classify it as a marsh.

In summary, this example shows that we always know a marsh when we see one, but on rare occasions we think a bog or a fen may be a marsh.

From a confusion matrix, the overall "accuracy" can be determined by assessing the diagonal. In the example above, 25/33 or 76% of the ecosystems were classified correctly. In mapping, a certain number of correct classifications can occur by chance, even in the most uncertain situations. Therefore, a value called the "Kappa index" (Foody 1992) is often calculated to express the degree of agreement by taking into account the chance "correct" classifications. A procedure for calculating the Kappa index is found at http://biology.usgs.gov/npsveg/aa/sect5.html. A statistical test on the degree of agreement (*K*-statistic) can also be calculated.

4 CONCLUSION

Although costly, assessing the accuracy of complex thematic maps is critically important to determining appropriate uses for ecosystem mapping. The approach outlined in this protocol provides a means of obtaining some overall, statistically valid scores to rate the accuracy of TEM, PEM, or other ecosystem maps. The results can be used as a component of quality assurance or for presenting statistics on the accuracy of mapping.

The accuracy assessment focuses on the thematic content of polygons. Polygon boundaries are not assessed. The results are essentially "non-spatial," in that they do not show explicitly where the errors can be found over the whole map. However, information is provided about overall accuracy of selected ecosystem mapping entities. This is useful because it provides potential users of the map or project with a level of confidence in the results.

In summary, the following statistics will be presented for each assessment:

1. Graph of ecosystem map entity proportions.
2. Percent mapped area overlap, for each ecosystem mapping entity—site series, structural stage, etc.
3. Percent dominant correct and appropriate confidence interval (usually 80%); weighted and not weighted by polygon area, for each mapping entity.
4. Percent polygon overlap and selected confidence interval (usually 90%); weighted and not weighted by the polygon area, for each mapping entity.
5. Percent acceptable polygon overlap, if assessed (optional), and confidence interval (usually 90%); weighted and not weighted by the polygon area, for each mapping entity.
6. Confusion matrix, if all "simple" map units (optional).

APPENDIX 1 Summary of accuracy assessment requirements for use of ecosystem mapping in timber supply modelling[7]

For use in "base-case" timber supply analysis:

1. An independent accuracy assessment must be conducted following the steps outlined in this protocol.
2. For PEM, other ecosystem maps, and TEM Level 5 and R survey intensity,[8] the accuracy assessment must be conducted at a minimum of sampling level 4 (see Table 1 in Section 3.3); for TEM Level 4 survey intensity, the accuracy assessment must be conducted at a minimum of sampling level 2.
3. The B.C. Forest Service Regional Ecologist or Provincial Ecologist must approve the sampling plan.
4. The percent dominant correct and percent overlap scores for polygons must be greater than or equal to 65% (focus is on the Forest Area).
5. Sample size (i.e., number of polygons or small areas sampled) must be large enough so that:
   - the estimate of percent dominant correct has a maximum error of 7% with a confidence of 0.8,
   - the estimate of percent overlap has a maximum error of 5% with a confidence of 0.9, and
   - it is at least two times the number of forested site series.
6. The minimum score for "percent overlap" for the entire map is 65%.
7. The ecosystem map must be a "site series PEM" or TEM.

7 From Meidinger (2003). See document for further information.
8 See Table 6.3 in *Standard for Terrestrial Ecosystem Mapping in British Columbia* (Resources Inventory Committee 1998)

Dominant Correct Score

To determine the percent dominant correct score, binomial statistics are used to evaluate the required sample size. Table 2 in this report provides estimates of sample size using the following equation:

Sample size = [(*t*-value)² x (1 – probability of random point being correct)²] / (acceptable error)²

[As the *t*-value is a function of *n* (degrees of freedom = *n* – 1), the *t*-values used in the calculations in Table 2 were determined using the "Solver" function in Excel].

To use an ecosystem map in timber supply modelling, the sample size for an accuracy assessment must be large enough to provide a dominant correct score with an error of no more than 7%, with 80% confidence. Using a probability of a random point being correct of 0.50 and inserting the values into the above equation, the result would be:

Sample size = (1.292)² x (1 – 0.50)² / (0.07)² = 85.2

Therefore, a sample size of 86 should provide an estimate within the acceptable error limits.

Percent Overlap

To determine the percent overlap, normal statistics are applicable and the sample size equation is as follows:

Sample size = [(*t*-value)² x (standard deviation)²] / (acceptable error)²

In the following example, 20 polygons are sampled and the standard deviation of the polygon overlap scores is 25.77%. The table below demonstrates how to calculate a sample size that meets the error and confidence targets.

Mean%	58.92%	average overlap score for 20 polygons
SD%	25.77	standard deviation of scores
***t*-value (90% prob.)**	1.666	*t*-value for 90% confidence using Excel "Solver"
Sample size (to achieve 5% max. error with 90% confidence)	73.8	(*t*-value² x SD²) / 5²
Number of polygons sampled	20	
Additional polygons needed	54	sample size – no. polygons sampled

Note: the TINV function in Excel will give a reasonable approximation of the *t*-value. However, since the *t*-value is a function of the degrees of freedom (df), a predicted sample size should be estimated to determine df. For example, if you assume that the sample size will be similar to that determined for the percent dominant correct score, you could start with a *t*-value for alpha = 0.1 and df = 85 (TINV(0.1, 85) = 1.66). Using the *t*-value of 1.66, the initial predicted sample size in the example above is 73. If you then substitute df = 72 into the TINV function, the *t*-value is recalculated as 1.67, the same value as using Microsoft Excel "Solver." Therefore, the predicted sample size will be 74, as in the above table.

The data from one line transect are summarized as follows:

Transect segment	Length (m)	Bearing	BGC	Site series	Map code
LPH1_1	170	153	SBPSmk	01	LP
LPH1_2	10	153	SBPSmk	06	ST
LPH1_3	95	153	SBPSmk	01	LP
LPH1_4	105	153	SBPSmk	06	ST
LPH1_5	30	153	SBPSmk	00	OB
LPH1_6	20	153	SBPSmk	01	LP
LPH1_7	10	153	SBPSmk	06	ST
LPH1_8	70	153	SBPSmk	07	SH
LPH1_9	30	153	SBPSmk	06	ST
LPH1_10	35	153	SBPSmk	07	SH
Total:	575				

The next table demonstrates the calculation of the small area overlap score for seven areas sampled using line-intercept transects. If all areas were of equal size, then the average small area score would be an average of the individual overlap scores (in this case, 71%).

	Transect data						Map data						
Small	**SBPSmk**						**SBPSmk**						
area	**01**	**04**	**06**	**07**	**08**	**X[a]**	**01**	**04**	**06**	**07**	**08**	**X**	**Overlap**
1	35%	5%	15%	19%	10%	16%	44%	0%	37%	18%	0%	0%	68%
2	70%	3%	14%	3%	2%	8%	78%	0%	15%	8%	0%	0%	87%
3	70%	0%	7%	6%	0%	17%	71%	0%	22%	7%	0%	0%	83%
4	85%	0%	8%	0%	0%	8%	40%	0%	53%	7%	0%	0%	48%
5	83%	0%	15%	1%	0%	1%	79%	2%	17%	2%	0%	0%	94%
6	27%	44%	13%	3%	2%	11%	31%	1%	51%	17%	0%	0%	43%
7	66%	0%	27%	2%	0%	5%	46%	0%	33%	20%	0%	0%	75%

[a]X refers to non-forested ecosystems

The calculation of the individual polygon overlap score for one transect is demonstrated as follows:

				Transect data			Map data			Overlap scores	
Transect segment	**Poly #**	**Polygon area (ha)**	**Length (m)**	**BGC**	**Site series**	**Code**	**BGC**	**Site series**	**Code**	**Average**	**Weighted average**
LPH1_1	123	6	100	SBPSmk	01	LP	SBPSmk	01	LP	100%	7.06%
LPH1_1	345	6	70	SBPSmk	01	LP	SBPSmk	06	ST	0%	0.00%
LPH1_2	154	4	10	SBPSmk	06	ST	SBPSmk	06	ST	100%	4.71%
LPH1_3	256	9	55	SBPSmk	01	LP	SBPSmk	01	LP	100%	10.59%
LPH1_3	257	9	40	SBPSmk	01	LP	SBPSmk	06	ST	0%	0.00%
LPH1_4	465	8	65	SBPSmk	06	ST	SBPSmk	06	ST	100%	9.41%
LPH1_4	467	6	40	SBPSmk	06	ST	SBPSmk	01	LP	0%	0.00%
LPH1_5	576	7	30	SBPSmk	00	OB	SBPSmk	00	OB	100%	8.24%
LPH1_6	564	6	20	SBPSmk	01	LP	SBPSmk	01	LP	100%	7.06%
LPH1_7	558	4	10	SBPSmk	06	ST	SBPSmk	06	ST	100%	4.71%
LPH1_8	597	4	30	SBPSmk	07	SH	SBPSmk	07	SH	100%	4.71%
LPH1_8	589	8	40	SBPSmk	07	SH	SBPSmk	06	ST	0%	0.00%
LPH1_9	458	3	30	SBPSmk	06	ST	SBPSmk	06	ST	100%	3.53%
LPH1_10	655	5	35	SBPSmk	07	SH	SBPSmk	07	SH	100%	5.88%
	Total:	**85**	**575**							**71%**	**66%**

Enter the values in the CRITBINOM function as follows:
if, for example, n=8,
% dom correct=0.62 (62%)

For 80% confidence intervals (CI):
lower CI: 0.1
upper CI: 0.9
=(CRITBINOM(8,0.62,0.1))

The answer is 3; as a proportion of 8, 3/8 is 0.375, therefore the lower 80% CI is 37.5%.
=(CRITBINOM(8,0.62,0.9))

The answer is 7; as a proportion of 8, 7/8 is 0.875, therefore the upper 80% CI is 87.5%.

For 95% confidence intervals:
lower CI: 0.025
upper CI: 0.975
=(CRITBINOM(8,0.62,0.025))

The answer in this case is 2; as a proportion of 8, 2/8 is 0.25, therefore the lower 95% CI is 25%.
=(CRITBINOM(8,0.62,0.975))

The answer is 7; as a proportion of 8, 7/8 is .875, therefore the upper 95% CI is 87.5%.

With small sample sizes, the CI's may not be exactly symmetrical, as the function is discrete.

REFERENCES

B.C. Ministry of Environment, Lands and Parks and B.C. Ministry of Forests. 1998. Field manual for describing terrestrial ecosystems. Victoria, B.C. Land Manage. Handb. No. 25.

Curran, M., I. Davis, and B. Mitchell. 2000. Silviculture prescription data collection field handbook: interpretive guide for data collection, site stratification, and sensitivity evaluation for silviculture prescriptions. B.C. Min. For., Victoria, B.C. Land Manage. Handb. No. 47.

Foody, G.M. 1992. On the compensation for chance agreement in image classification accuracy assessment. Photogrammetric Engineering and Remote Sensing 58(10):1459–1460.

Iverson, K. 2002. Terrestrial ecosystem mapping reliability assessment quality assurance procedures for ecological components. Prepared for Lignum Ltd. Unpubl. rep.

Meidinger, D. 2000. Protocol for quality assurance and accuracy assessment of ecosystem maps. Res. Br., B.C. Min. For., Victoria, B.C. < http://www.for.gov.bc.ca/hre/temalt/public.htm >

———. 2003. Accuracy assessment requirements for use of ecosystem mapping in timber supply modelling. Res. Br., B.C. Min. For., Victoria, B.C. < http://www.for.gov.bc.ca/hre/becweb/standards.htm >

Resources Inventory Committee. 1998. Standard for terrestrial ecosystem mapping in British Columbia. Prepared by Ecosystems Working Group, Terrestrial Ecosystems Task Force, Resources Inventory Committee, Victoria, B.C.

Ryan, M.W., D.A. Lloyd, and D. Meidinger. 2002. Predictive ecosystem mapping: accuracy assessment procedures for the Okanagan Study Area. B.C. Min. For., Kamloops, B.C.