Standards for Predictive Ecosystem Mapping

Inventory Standard

Prepared by

Terrestrial Ecosystem Mapping Alternatives Task Force
for the
Resources Inventory Committee

Version 1.0

November 1999

Published by the
Resources Inventory Committee

Canadian Cataloguing in Publication Data
Main entry under title:
Standards for predictive ecosystem mapping [computer file]

Available through the Internet.
Issued also in printed format on demand.
Includes bibliographical references.
ISBN 0-7726-4084-X

1. Ecological mapping - Standards - British Columbia. 2. Biotic communities - British Columbia. 3. Ecological surveys - Standards - British Columbia. I. Resources Inventory Committee (Canada). Terrestrial Ecosystem Mapping Alternatives Task Force.

QH541.15.M26S73 1999 577'.09711 C99-960411-2

Additional copies of this publication can be purchased from:

Government Publications Centre
Phone: (250) 387-3309 or
Toll free: 1-800-663-6105
Fax: (250) 387-0388
www.publications.gov.bc.ca

Digital copies are available on the Internet at:
http://www.for.gov.bc.ca/ric

Preface

The Resources Inventory Committee members are resource specialists from a number of professional disciplines and represent Provincial, Federal, First Nation and private sector agencies and other resource interests. RIC's objectives are to develop a common set of standards and procedures for provincial resource inventories, as recommended by the Forest Resources Commission in its report "The Future of our Forests."

Funding of the Resources Inventory Committee work, including the preparation of this document, is provided by the Corporate Resource Inventory Initiative (CRII) and by Forest Renewal BC (FRBC). Preliminary work of the Resources Inventory Committee was funded by the Canada–British Columbia Partnership Agreement of Forest Resource Development FRDA II.

For further information about the Resources Inventory Committee and its various Task Forces, please visit the RIC website at http://www.for.gov.bc.ca/ric.

Acknowledgments

This document was prepared by Dave Moon, CDT Core Decision Technologies Inc. with contributions from Dave Clark, Resources Inventory Branch, Ministry of Environment, Lands and Parks (MELP); Del Meidinger, Research Branch, Ministry of Forests (MoF); and Keith Jones, R. Keith Jones and Associates. Technical input from Evert Kenk, Geographic Data BC, MELP and technical edit by Sheila Jeck, CDT Core Decision Technologies, were especially valuable.

The report was prepared following three workshops sessions held July 6–7, July 20, 1999, and September 8–9, 1999. The first workshop established the terms of reference and specifications for the PEM standard. The second workshop reviewed mapping and map-entity concepts. Following the second workshop, D. Moon prepared a prototype standard for review and comment. The third workshop reviewed the prototype and recommended revisions to create a 1st approximation standard. This version incorporates the results of the third workshop and subsequent technical review by PEM practitioners.

Contributors and affiliations:

Cadrin, Carmen, Resources Inventory Branch, MELP, Victoria

Clark, Dave, Resources Inventory Branch, MELP, Victoria

Eng, Marvin, Research Branch, MoF, Victoria

Gunning, Terry, Resources Inventory Branch, MELP, Victoria

Jeck, Sheila, CDT Core Decision Technologies Inc., Richmond

Jones, Keith, R. Keith Jones & Associates, Victoria

Kenk, Evert, Geographic Data BC, MELP, Victoria

Kot, Ron, Resources Inventory Branch, MELP, Victoria

Lea, Ted, Resources Inventory Branch, MELP, Victoria

Love, Bobby, MoF, Smithers

McKellar, Darren, Resources Inventory Branch, MELP, Victoria

Meidinger, Del, Research Branch, MoF, Victoria

Moon, Dave, CDT Core Decision Technologies, Vancouver

Schaefer, Claudia, Independent Consultant, Vancouver

Simpson, Keith, Keystone Wildlife Research, White Rock

Tolkamp, Calvin, Resources Inventory Branch, MELP, Victoria

Vassov, Rob, Timberline Forest Inventory Consultants Ltd., Prince George

Summary

Predictive ecosystem mapping (PEM) is a new and evolving program designed to use available spatial data and knowledge of ecological-landscape relationships to automate the computer generation of ecosystem maps. It typically involves the spatial overlay of mapped themes and the processing of the resultant attributes against a formalized knowledge base using automated inference methods. It offers the promise of providing surrogate terrestrial ecosystem maps. The evolving nature of PEM and the high expectations associated with it create the following conditions relevant to creating a standard.

1. Methods of creating predictive ecosystems maps are not stable.
2. The positional accuracy and thematic accuracy of terrestrial ecosystem mapping (TEM) and PEM maps are untested. This makes cost and effectiveness comparisons of the two approaches inconclusive.
3. The spatial and thematic quality of land-based resource maps such as early soil and terrain that are potentially useful to PEM, typically have unknown accuracy. Available information suggests, therefore, that the quality of these inventories can be problematic.
4. Predictive ecosystem mapping requires a broader range of knowledge and experience than traditional TEM. In addition to the knowledge and skills required for traditional TEM, PEM requires knowledge and skills in geographic data analysis; spatial data analysis; data, information, and knowledge management; and knowledge engineering.
5. The range of experience and qualifications of potential PEM practitioners is highly variable.
6. Promises of low PEM costs make the high cost of quality assurance unpopular.

These conditions make the establishment of standards difficult, but necessary. While standards can support consistent data formats and data exchange and can mandate quality control and quality assurance (QC/QA), they can also inhibit progress, especially in a new and rapidly evolving technology like PEM.

The *PEM Inventory Standard* document attempts to strike a balance between the need to ensure the quality and consistency of PEM products and the need to allow for the adaptation and evolution of PEM procedures. It assumes that a qualified PEM practitioner implements the PEM project. Finally, the standard refers to the creation of 1:20 000 PEM maps designed to emulate 1:20 000 TEM maps. While much of the standard would be applicable to smaller scale maps, smaller scale PEM maps are not addressed directly by this standard.

The *PEM Inventory Standard* specifies mapping concepts, documentation standards for assessing the quality of input maps, required documentation for the knowledge base, and the required quality assurance (QA). The early stage of development in PEM procedures precludes the establishment of rigorous, tested, and validated quality control (QC) procedures. The intent of this document is to ensure the documentation of PEM procedures in sufficient detail for a qualified PEM practitioner to evaluate the quality of a PEM product.

As PEM is an emulation of TEM, where the PEM standard has not been explicit, the reader is referred to the TEM standard and associated discipline standards. Because of their codependence, their respective change management processes are being dealt with together.

In addition to the *PEM Inventory Standard*, practitioners will need to acquire and understand the *Method for Large-scale Biogeoclimatic Mapping*, the *Protocol for Quality Assurance and Accuracy Assessment of Ecosystem Maps*, the *PEM Digital Data Standard* (in preparation), and the *Protocol for Structural Stage Modeling in PEM* (in preparation).

The mapping entities for PEM are:

- Ecosection
- Biogeoclimatic unit (zone, subzone and/or variant)
- Site unit (site series or approximate equivalent)
- Slope and aspect modifiers (from TEM).

Options for lumping site series, describing or mapping more general site units, or defining compound site series combinations are described. The reason for providing these options is to minimize "ties" between site series predictions, to allow for the unit being predicted to match the resolution of the input data, and to allow for the prediction of map entities that best describe what is on the ground.

Structural stage will be a separate digital overlay product to be delivered with the PEM product. This will facilitate interpretations requiring structural stage such as old forest site series or wildlife suitability and will allow the structural stage layer to be updated easily and overlaid with the "permanent" ecosystem map.

Documentation standards are not intended to be onerous but are designed to ensure that it is clear how the input data has been used in the predictive model and that the PEM practitioner has carefully evaluated input data quality. A qualitative evaluation of various input data types is provided.

Documentation and evaluation of the knowledge base is critical to the success of any PEM project. Therefore, the knowledge base must be well documented and an independent validation dataset must be run through the knowledge base. The results of the knowledge base validation is reported as project meta-data.

All PEM projects also require large-scale biogeoclimatic mapping to be developed for the project area.

Each PEM project is required to report QA or accuracy assessment statistics following the *Protocol for Quality Assurance and Accuracy Assessment of Ecosystem Maps*.

The inventory standard outlines requirements and recommendations. It attempts to present minimum standards for the present state of development of PEM. The objective is to ensure that the best quality mapping is conducted and that digital products are delivered that can be used for strategic level planning throughout British Columbia. The products of a PEM project are, in summary:

1. Input and input processing documentation
2. Knowledge base documentation
3. Quality assurance (validation) of knowledge base
4. Digital output of predicted ecosystem map (see *Digital PEM Data Standards*)
5. Structural stage/age class map (see *Protocol for Structural Stage Modeling in PEM*)
6. Quality assurance/accuracy assessment of predicted ecosystem map.

Table of Contents

List of Tables

List of Figures

1 Introduction

Predictive ecosystem mapping (PEM) is a new and evolving inventory approach designed to use available spatial data and knowledge of ecological-landscape relationships to automate the computer generation of ecosystem maps. It typically involves the spatial overlay of mapped themes and the processing of the resultant attributes against a formalized knowledge base using automated inference methods. It offers the promise of providing surrogate terrestrial ecosystem maps.

PEM methods can be consistent with and complementary to traditional TEM procedures. However, there are tradeoffs between cost, resolution, and accuracy that must be measured against the intended use of the information. PEM inventory systems by their very nature can provide a consistent treatment of available inventory data and knowledge. They offer the potential for increasing production rates and capacity. Being digital and algorithmic, they offer the prospect of becoming more integral to corporate information and knowledge base systems. Over time, PEM approaches and conventional TEM methods will likely converge.

The evolving nature of PEM and the high expectations associated with it create the following conditions relevant to creating a standard.

1. Methods of creating predictive ecosystems maps are not stable.
2. The positional accuracy and thematic accuracy of TEM and PEM maps are untested. This makes cost and effectiveness comparisons of the two approaches inconclusive.
3. The spatial and thematic quality of land-based resource maps such as early soil and terrain, that are potentially useful to PEM, typically have unknown accuracy. Available information suggests therefore, that the quality of these inventories can be problematic.
4. Predictive ecosystem mapping requires a broader range of knowledge and experience than traditional TEM. In addition to the knowledge and skills required for traditional TEM, PEM requires knowledge and skills in geographic data analysis; spatial data analysis; data, information, and knowledge management; and knowledge engineering.
5. The range of experience and qualifications of potential PEM practitioners is highly variable.
6. Promises of low PEM costs make the high cost of quality assurance unpopular.

These conditions make the establishment of standards difficult, but necessary. While standards can support consistent data formats and data exchange and can mandate quality control and quality assurance (QC/QA), they can also inhibit progress, especially in a new and rapidly evolving technology like PEM.

This document attempts to strike a balance between the need to ensure the quality and consistency of PEM products and the need to allow for the adaptation and evolution of PEM procedures. It assumes that a qualified PEM practitioner implements the PEM project. Finally, the standard refers to the creation of 1:20 000 PEM maps designed to emulate 1:20 000 TEM maps. While much of the standard would be applicable to smaller scales maps, smaller scale PEM maps are not addressed directly by this standard.

1.1 Proviso

This is the first attempt at a PEM inventory standard. It is not operationally tested and in many cases, specification of required formats, codes, values, and criteria is premature. Therefore, much of the required meta-data is not formally structured at this time. Operational testing will guide modifications and specifications for a more formal structure.

1.2 Objectives

The objectives of the PEM standard are:

1. To ensure delivery of PEM information in a standard TEM-like form that allows for merging, integration, comparison or interpretation of multiple data sets.
2. To ensure documentation and meta-data are sufficient to evaluate the accuracy of the PEM product.
3. To effectively manage change to the standard. The standard will be subject to a change management procedure that incorporates an impact analysis of proposed changes.

In keeping with this objective, the following document provides background discussion, guidelines and standards, and rationales for the guidelines or standards presented. *This type-face indicates either a guideline or a required standard.*

1.3 Background

The specification for a PEM standard arises from four preceding works commissioned by the Terrestrial Ecosystem Mapping Alternatives Task Force of the Resources Inventory Committee (RIC). These were:

1. *Towards the Establishment of a Predictive Ecosystem Mapping Standard: A White Paper*, by Keith Jones, R. Keith Jones & Associates; Del Meidinger, BC Ministry of Forests, Research Branch; Dave Clark, BC Ministry of Environment Lands and Parks, Resources Inventory Branch; and Fern Schultz, BC Ministry of Forests, Resources Inventory Branch.
2. *Problem Analysis on Data Quality Assessment Issues* by Dr. David Moon, CDT–Core Decision Technologies Inc.
3. *Situation Analysis for Knowledge-Based Systems* by Dr. David Moon, CDT–Core Decision Technologies Inc.
4. *Problem Analysis on Reliability, Quality Control and Validation of Predictive Ecosystem Mapping (PEM)* by Dr. Richard Sims and Jeff Matheson, R.A. Sims & Associates.

The standard draws upon four additional reports:

1. *Specifications for PEM*, version 2.1
2. *Mapping entities*, draft report
3. *Protocol for Quality Assurance and Accuracy Assessment of Ecosystem Maps*
4. *A Method for Large Scale Biogeoclimatic Mapping in British Columbia*.

The white paper (Jones et al. 1999) integrated elements of the supporting studies to develop the framework presented in Figure 1. The PEM framework comprises five main components, each of which contain important elements, features and other related items.

The five main components are:

1. Client requirement assessment
2. Input data quality assessment
3. Knowledge base assessment
4. Output and reliability assessment
5. QC/QA procedures and standards.

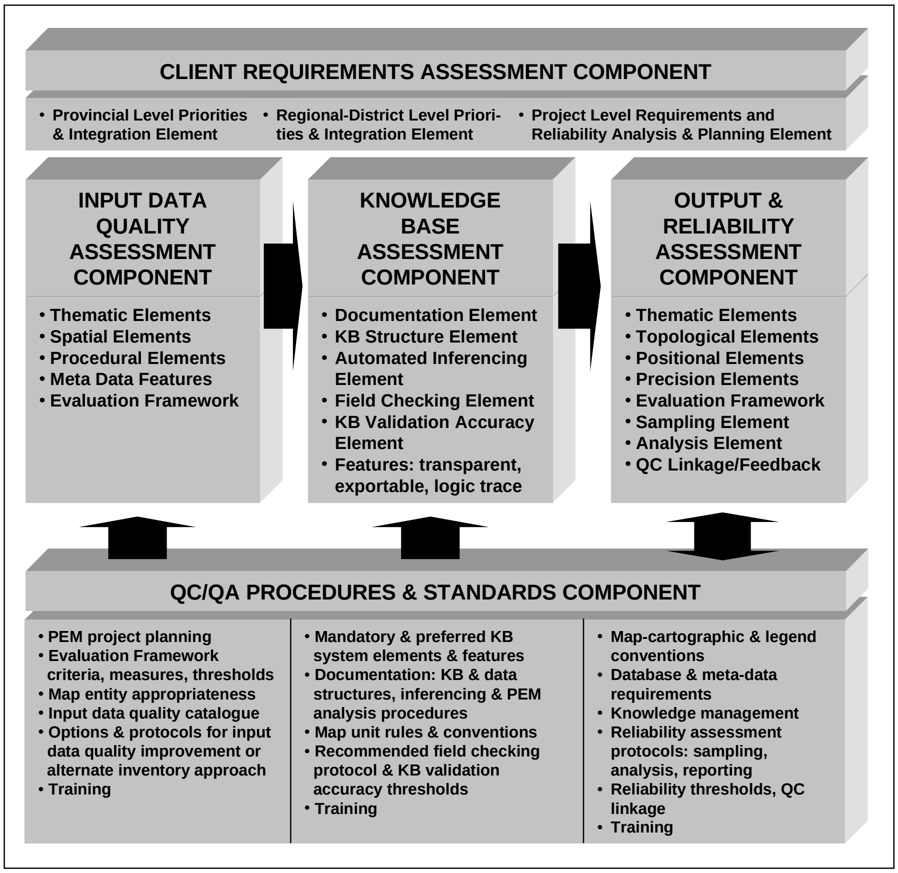

Figure 1. PEM framework

Components 1–4 have a number of defining elements and features, which together characterize their scope and function. The *QC/QA procedures and standards* component contains items that largely correspond vertically, from left to right, to the core *input data quality assessment*, *knowledge base assessment* and *output and reliability assessment* components of the PEM framework. The PEM components, in turn, generally parallel components of the current TEM RIC standards and procedures. With PEM however, the input data quality and knowledge base assessment components are different, but are somewhat analogous to “Mapping and Field Survey Procedures” section of the TEM RIC standard. The *output and reliability assessment* component of the framework is similar for both PEM and conventional TEM.

The PEM standard will address components 2–5 and will provide both standards and guidelines for each element. It will also establish minimum levels of documentation and meta-data standards to support evaluation of the accuracy and utility of the PEM product.

1.4 Principles

The PEM standard will attempt to conform to the following principles.

1.4.1 Conformance to Existing Standards

Wherever possible, the PEM standard will conform to and use existing RIC standards by reference or attachment to existing standards rather than creating new standards requiring correlation or reconciliation.

The PEM standard will follow the general protocol established by RIC for published standards. The Resources Inventory Committee recognizes five classes of standard. These are:

1. **Discipline standard** related to classifications, concepts, and entities in established disciplines such as forest inventory, soil survey, or terrain mapping which should be followed when collecting or using data to support other inventories. For example, the terrain classification standard *Terrain Classification System for British Columbia*, Version 2. (1997) (D. Howes and E. Kenk) should be followed when collecting or using terrain data elements in support of a PEM.
2. **Data collection standard** related to the method of spatial and thematic data capture. These may be field, office, or computer-based methods. For example, *Field Manual for Describing Terrestrial Ecosystems* is the standard for site data collection when used as a component of ecosystem mapping.
3. **Digital data capture standard** specifying entities, attributes, data structures and formats for the submission of digital data to the government repository (e.g., the *Standard for Digital Terrestrial Ecosystem Mapping (TEM) Data Capture in British Columbia*).
4. **Standard for interpretations and interpretive methods** appropriate to the data source to which the standard applies (e.g., the *Standards for Wildlife Habitat Capability and Suitability Ratings for British Columbia*).
5. **Standards for common output products**. Generally, these are specifications for standard analogue map or report products generated from the digital data (e.g., the TEM output standard).

1.4.2 Software Independence

Where possible, the standard will be independent of proprietary software or proprietary software constructs. Where the custodian for PEM requires data in a vendor dependent format or construct, the standard will specify either the form or a vendor independent format that the Ministry can import to their system.

1.4.3 Procedure Independence

Where possible the standard will be procedure independent. However, in some cases the standard will require that a specific procedure, one of a set of specific procedures, or a procedure meeting a minimum set of criteria be followed. For example, the standard requires an estimation and documentation of accuracy. Rather than requiring a specific procedure, the standard requires one of a number of procedures set out in an accompanying document (Meidinger, 1999). In other cases, the standard requires documentation of a procedure. The documentation may be a reference to an existing and publicly available document or an RTF format document appended to the PEM submission. It must be sufficient for a qualified PEM practitioner to replicate and evaluate the procedure.

1.4.4 Professional Accountability

The intended audience for this document is a qualified PEM practitioner. The standard recommends that practitioners or the leader of a team of practitioners be members of a professional society who are bound by a professional code of practices that will be applied to the delivery of PEM projects.

1.5 Conventions Used in This Document

This font designates a requirement of the standard or a guideline.

2 Relationships Between PEM Inventory and Other RIC Standards

All RIC standards in this section can be accessed at the RIC web site: http://www.for.gov.bc.ca/RIC/.

2.1 Other PEM Standards

The *PEM Inventory Standard* has the following supporting standard: Standard for Digital PEM Data Capture. Some concepts illustrated by example in this document will be specified in detail by the Standard for Digital PEM Data Capture. In cases of disagreement, the latter will take precedence.

2.2 TEM Standards

The *Standard for Terrestrial Ecosystem Mapping in British Columbia*, hereafter referred to as the TEM standard, forms the basis for PEM. The associated TEM discipline standards also apply to PEM. Where the PEM standard is not explicit, for example regarding the hardcopy format of map products, the reader is referred to the TEM standard. Because of their codependence, their respective change management processes are being dealt with together.

2.3 Digital TEM Data Capture Standards

The basis for the PEM standard for digital data capture is the Standard for Digital Terrestrial Ecosystem Mapping (TEM) Data Capture in British Columbia; Ecosystem Technical Standards and Database Manual Version 2.0 *(hereafter referred to as the TEM digital standards).*

2.4 Broad Ecosystem Inventory (BEI)

Standards for Broad Terrestrial Ecosystem Classification and Mapping for British Columbia: Classification and Correlation of the Broad Habitat Classes used in 1:250 000 Ecological Mapping, Version 2.0, November 1998, prepared by the Ecosystems Working Group for the Terrestrial Ecosystems Task Force of the Resources Inventory Committee. This standard is hereafter referred to as BEI. *Where the prediction of site series is inappropriate or infeasible, the PEM standard recommends the use of BEI classes as an alternative mapping entity (see Section 4.1).*

2.5 RIC Policy on Change Management

Change management of the PEM standard conforms to the *RIC Change Management Policy for Geographic/Land Related Data Products and Standards* (Version 2.0 – 98/10/13). Projects are expected to follow the standards in place at the time of project initiation.

2.6 Other Relevant Standards

There are a number of other useful standards available or referenced at the RIC standards web site (http://www.for.gov.bc.ca/RIC/standards.htm). These include:

- *British Columbia Standards, Specifications and Guidelines for Resource Surveys Using Global Positioning System (GPS) Technology*, Release 2
- *Policies and Specifications for TRIM (1:20 000) Revision Data Capture,* Version 2.0
- *Standard for the Use of Map Projections in British Columbia for Resource, Cultural and Heritage Inventories*
- *Standard for Developing Digital Data Specification Standards Documents*, Version 1.0
- *Corporate Data Model Framework*, Version 1
- *Corporate Data Modeling Standards and Guidelines Interim* (1996/1997)
- *Field Manual for Describing Terrestrial Ecosystems*
- *Terrain Classification System for British Columbia*, Version 2 (1997)
- *Standard for Digital Terrain Mapping Data Capture in British Columbia — Terrain Technical Standards and Database Manual,* Version 1, June 1998.

3 PEM Process Overview

The following provides an overview of the predictive ecosystem mapping process adapted from the document *Towards the Establishment of a Predictive Ecosystem Mapping Standard: A White Paper*. This material is based on experience in using two of the more developed systems. Figures 2 and 3 illustrate aspects of the process. Individual methods will vary in their specific approaches and protocols.

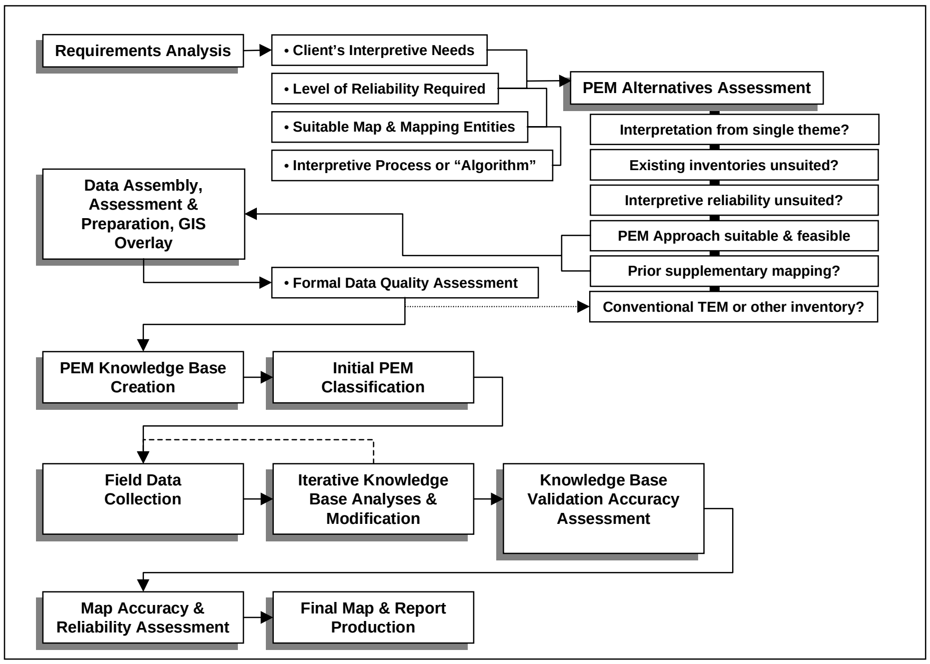

Figure 2. Generic PEM Process: general steps from requirements analysis through to final map and report production.

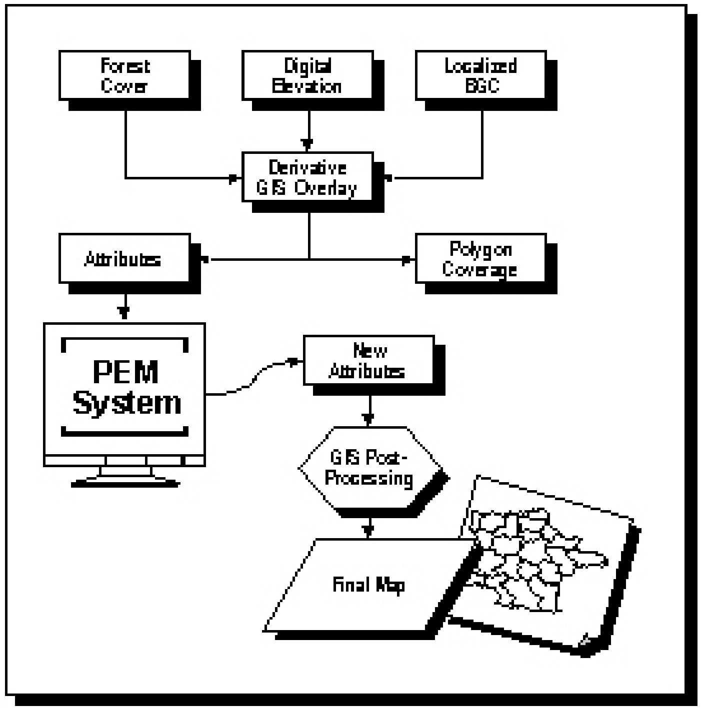

Figure 3. Example of key input themes and processes in the generation of a PEM map product (modified from Downing et al. 1998).

1. **Requirements Analysis.** In consultation with the client, this initial step requires a careful determination of their interpretive needs, including level of reliability required for the ecological mapping. Consideration of how the final PEM outputs will actually be interpreted —the interpretive "algorithm"—is also exceptionally important at this time.

   Depending on the client's inventory requirement, various alternatives may be considered, including those described below.

   - The client's requirements are suitable for a PEM approach using existing digital overages.
   - The client's requirements are suitable for a PEM approach but it is felt that some other precursor mapping is required in order to attain the required level of map reliability. For example, it may be evident that terrain or bioterrain-like mapping of the area would greatly enhance the quality of the predictive map units as the relationships between existing digital data attributes and site series is weak.
   - The quality of existing mapped information is too low for any form of PEM. For example, the relationships between existing digital data attributes and site series is weak, additional data cannot be collected economically, or existing sources are from radically different scales (e.g., a 1:126 000 scale reconnaissance soils map and a 1:20 000 scale forest cover map).

- The client can manage with a straight-forward interpretation of an existing, single theme inventory (e.g., the forest cover map). In some cases, some supplementary point or mapped information may be required in areas of high priority.
- The quality of existing mapped information is too low for any form of PEM and the client's requirements are at a level of detail or decision risk that detailed photo interpretation and intensive ground sampling are the only way to acquire information of sufficient quality (i.e., conventional TEM).

2. **Data Assembly, Assessment and Preparation.** On the basis of the findings from step 1, the next stage involves assembling, assessing, reconciling and preparing all of the various map data sources (e.g., BGC, FC1 or VRI, terrain units, soils, TRIM derivatives like elevation, slope, ridges). The level of effort here depends on the location of the area to be mapped, the state of the available map databases (digital or analog), and ecological relationship knowledge (field guides, map legends, reports, etc.).

   While data quality may have been assessed generally in step 1, at this stage it must be evaluated more carefully for the thematic, spatial and procedural qualities of the legacy inventory. Details on how this is accomplished are discussed in Section 4.4. If the data sources are of acceptable quality then any analog information needs to be digitized or entered in a database. The GIS data are then processed to produce the required layers (e.g., large-scale biogeoclimatic, various TRIM features, forest cover and selected attributes). The layers may be combined in various ways, from a straight GIS overlay of all layers, to a more structured, step-wise process. Either way, a resultant database is produced. This resultant database becomes the input data to be processed by the PEM knowledge base.

3. **Knowledge Base Creation.** The PEM knowledge base model processes the resultant database information from step 2 and classifies or allocates each data record (i.e., the resultant overlay polygon attributes) into the most likely ecological class. The knowledge base structure and inference strategy used varies between PEM approaches, but typically involves some form of automated inferencing to apply the knowledge. The knowledge base is a coding of the relationship between each attribute selected in the data preparation stage to the site series or other ecosystem unit feature. The knowledge base structure is often a data table of site series and attributes and values indicating the relationships. These values may be yes or no values, rankings, or probability or "belief" values. The knowledge base model or "engine" processes the relationships.

4. **Initial Classification.** An initial analysis of the input data through the knowledge base is done using starting knowledge base values for specific attributes, such as the proportion of white spruce that is "allowed" in certain ecosystem types, or the aspect and elevation ranges that are most commonly associated with specific ecological types. These starting values may be obtained from field guides or from previous PEM project models within the same biogeoclimatic unit, and/or from other available plot data. The initial output is used to evaluate the knowledge base relationships and to determine whether other attributes or new data are required.

5. **Field Data Collection.** Often as an integral part of the modeling process, field data provides both the information necessary to modify the knowledge base and to test the accuracy, consistency and sensitivity of predictions. There are three essential aspects to a field program:

   - The field work must be conducted by ecologists who are familiar with the landscapes in the project area. If the field call identifications are incorrect (i.e., if the evidence collected is of inadequate quality or is improperly interpreted), the errors propagate across the entire landscape for all map units with similar attributes.

- Plot locations must be highly accurate. GIS overlays result in smaller polygons in the derivative layer than in any of the contributing themes. If plots are not located correctly, attributes from the wrong polygon could be improperly linked to field results, with the result that errors occur for all polygons with that attribute set.

The initial classification (step 4) can be used to guide the field work (e.g., areas where the system was unable to classify; areas where the classification is suspect and areas where complexes occur). All these situations should be targeted for more intensive surveys. The data collected at each plot must include all attributes required to confirm the ecological classification identification (e.g., site series key criteria). As well, all elements that are used as attributes in the knowledge base tables (e.g., slope, aspect, and cover type) must be recorded so that field values can be compared to the attribute values from the input data sets. Field checking to confirm or evaluate biogeoclimatic boundaries can also be conducted at this time.

6. **Iterative Knowledge Base Analyses and Modification.** Using the field data as further input, the knowledge base is run iteratively. Modifications are made to the knowledge base relationships until an acceptable level of prediction and the lowest possible level of ambiguity (complex predictions of two or more ecosystem types) is achieved. For example, slope and aspect might be more or less influential in the current project area than in a previous project area that may have provided the starting knowledge base for the PEM. Several runs are required usually to determine and correct systematic errors in predictions. The incorrect predictions that remain are almost always a consequence of landscape variations that occur at too large a scale to be captured by the available input themes or inaccuracies in the digital terrain model derivatives (e.g., overestimates or underestimates of slope or incorrect aspects). Most PEM knowledge bases provide a trace of the predictive analysis process to facilitate the identification of problem areas.
7. **Validation Accuracy Assessment of Knowledge Base:** A validation accuracy[1] assessment is carried out where field plots considered to be correct are compared to the modeling system predictions (see Section 4.6.1). A degree of error approach can be used as a measure of distance or how close the predicted class is to the actual class.

   The importance of unambiguous classifications (i.e., one prediction for a given polygon versus two or more possible ecological types) is a preferred outcome of a PEM approach. Much depends on the natural variation in the map area, the nature and quality of the input data sets, and some initial decisions about whether all ecosystem units can be expected to be identified (may require some initial lumping of units). Experience has shown that ambiguities are generally reduced as more data layers are added to the knowledge base. For example, with only biogeoclimatic, all site series in a biogeoclimatic unit would be the only possible outcome. By adding in forest cover, there are fewer ambiguities, adding in digital terrain model derivatives there are fewer again, etc.

   With respect to the representation of complexity and ambiguities, the ability to make predictions that are neither oversimplifications of naturally compound or complex units nor overly complicated representations of simple units is important. A PEM that predicts many complex ecosystem types (two or more site series, for example) when there are in actuality few complex ecosystems is less useful to a resource manager than a system that does not. This is especially true for ecosystem types that have specific interpretive potentials or constraints. If such types are complexed with other ecosystem types that do not have the same potentials or constraints, it is difficult for the resource manager to assess the spatial extent and distribution of, for example, “high” versus “medium” sites for productivity potential.

[1] Validation accuracy is not the same as map accuracy; only subsets of the map polygons are checked.

8. **Map Accuracy or Quality Assurance.** Ideally an independent set of check points using an unbiased sampling approach should be collected in order to assess thematic and spatial accuracy (see Section 4.6.2). These accuracy findings can be used to assess the PEM reliability for the client's intended use. These data can also be incorporated into the PEM to adjust the knowledge base to improve the output quality. However, once these data are used in this way, the results of the accuracy assessment can no longer be cited until another independent sample set is collected and analyzed. If a map accuracy assessment cannot be done, a quality assurance check of the final map is essential.
9. **Final Map and Reports Production.** The predicted ecosystem types are merged back with the original derivative GIS databases, and may be grouped in whatever way is required using GIS database functions for further interpretive analyses or for standard or custom reporting and cartographic presentation. Both standard inventory and interpretive maps and reports are prepared at this stage.

4 PEM Standards

4.1 Mapping Concepts

It is important to clearly define the mapping concepts used in PEM, in the input maps used for PEM, in the PEM maps, and in the PEM standards. Key definitions are hyper-linked to the Glossary for this document.

4.1.1 Map and Mapping Entities

The reference mapping entities[2] (that is, the classes we are trying to predict) for PEM will be the ecosection, the BEC units (BEC zone/subzone/variant/site series), and site series modifier (only those that describe slope and aspect—j, k, w, q and z). In some projects, it may be desirable to predict other site modifiers (such as shallow soils, ridge or active floodplain). These entities will correspond exactly to those defined in the TEM standard. However, PEM presents the practitioner with problems unique to the PEM process. Solutions to these problems require deviation from the TEM standard as outlined below.

In PEM the site modifiers are always specified, where appropriate, in contrast to TEM where the typical site series situation can imply certain modifiers.

Assignment of structural stage to complex map entities is not always straightforward. This PEM inventory standard requires the creation and delivery of a separate structural stage digital product that can be merged with the PEM product for interpretive applications.

In some applications, it may not be appropriate or possible to identify specific site series. In these cases, the standard requires justification for the use of other mapping entities. Sections 4.1.1.1 and 4.1.1.2 identify circumstances where this may occur and indicate the standards to follow.

In other applications, the spatial variability and map scale result in polygons with multiple site series such that the polygons cannot be adequately characterized with 3 or fewer site series. Sometimes these multiple site series occur in a definable pattern that is important to the interpretation of the map unit. Section 4.1.1.3 identifies circumstances in which this may occur and indicates the standards to follow.

4.1.1.1 No Existing Site Series

This situation normally occurs for non-forested sites, including non-vegetated and anthropogenic sites, or for forested sites requiring a new site series.

Standard to be followed:

For areas with no existing site series the following standards apply. They are listed in order of preference.

1. *Mapping entities defined in TEM — non-forested, non-vegetated and anthropogenic map units (in Table 3.1 of the TEM standard or mapcode.xls)*
2. *Mapping entities defined in* Standards for Broad Terrestrial Ecosystem Classification and Mapping in British Columbia: Classification and Correlation of the Broad Habitat Classes*, hereafter referred to as BEI (available at http://www.for.gov.bc.ca/RIC/) if appropriate.*

[2] A reference mapping entity is the standard against which comparable mapping entities are defined. When mapping entities other than the reference are used, they are defined in terms of a reference mapping entity.

3. *Other mapping entities may be used, but must be referenced to a published and available document providing definitions and criteria for the classes or definitions and criteria for the classes must be appended to the submission.*

In all cases, the project meta-data will identify the source for definitions of these classes (Section 4.7.3.1). The custodian for PEM must approve all map codes (for inclusion in mapcode.xls).

4.1.1.2 Inadequate Input Data

The PEM predictive model cannot assign an acceptable confidence of membership for any site series based on available PEM input data. For example, the predictive model may not be able to distinguish reliably between closely related site series.

Standard to be followed:

In these instances, the PEM project will use one of the following standards.

1. *Generalized mapping entities defined in BEI, if appropriate.*
2. *The PEM project may create and document new map entities to describe generalized groupings of site series that cover broader ecological ranges than an individual site series. The definition of these generalized map entities will include the BEC site series or other ecosystem units described in mapcode.xls, or in Table 3.1 in the TEM standard that occur under the expected range of ecological conditions. Definitions for the classes must be referenced in the project meta-data (Section 4.7.3.1).*
3. *The polygon is labeled UN (unclassified).*

None of these options should be implemented unless they add significant value to the interpretations of the PEM. The use of generalized map entities must be approved by the custodian for PEM for correlation and inclusion in map code.xls.

4.1.1.3 Complexity

The distribution of site series is too complex or at a scale too large for individual site series to be predicted at the scale of the PEM inventory. The complexity of the input map entities has a strong impact on the complexity of the output polygons and should be considered carefully during the PEM project planning.

Standard to be followed:

1. *Where the majority of the polygons are adequately characterized by up to three site series, the TEM standard applies (only three separate map entities can be identified for each polygon).*
2. *Where the mixture of simple mapping units occurs in a predictable and repeating pattern across the landscape or where many of the polygons are not adequately characterized by three site series, the PEM project may create and document new map entities to describe the polygons. The map-entity description will not consist of a simple listing of individual site series, rather the description will include:*
    - *the BEC site series (or other ecosystem units described in mapcode.xls, or in Table 3.1 in the TEM standard) found in the polygon*
    - *their expected range of proportions*
    - *a description indicating how the component BEC site series are distributed within the map unit. The description should indicate the nature of the spatial pattern or topographic relationships used to predict distribution or indicate that no predictable pattern is evident.*
3. *Descriptions will be referenced in the project meta-data (Section 4.7.3.1).*

None of these options should be implemented unless they add significant value to the interpretations of the PEM and the use of compound or complex map entities must be approved by the custodian for PEM.

4.1.2 Cartographic Conventions

PEM mapping conventions, with the following exceptions, correspond to the TEM output standard. The standard requires:

- *the clear identification of all PEM products and hybrid TEM/PEM products as predictive products*
- *the title of all map products (and reports) will include the qualifier “Predicted” instead of “Terrestrial”*
- *the legend for all PEM maps will contain the qualifier “Predicted” attached to the legend title (e.g., Predicted Ecosystem Units of the Fort Nelson Area, Fort Nelson District Portion of Map Sheet 94I, 94J, 94O and 94P (1:20 000), February 2000).*

4.2 Field Sampling

There is no minimum field sampling requirement for PEM. However, typically some level of sampling is undertaken to confirm relationship information between the input attribute data and the ecosystem classes, to refine Biogeoclimatic boundaries, and to localize and train the knowledge base model to particular local conditions. Field sampling may also be required for the independent validation of the knowledge base (see Section 4.6.1.1).

4.3 Quality Control

The early stage of development in PEM procedures precludes the establishment of rigorous, tested, and validated quality control procedures. The intent of this standard is to ensure the documentation of PEM procedures in sufficient detail for a qualified PEM practitioner to evaluate the quality of a PEM product. The intent is also to compile sufficient data for the creation of more rigorous quality control standards in the future.

Quality control consists of two parts. The first is rigorous, well-documented procedures that, if followed, will produce consistently acceptable results. The second is assurance of implementation of the procedure. An objective of this initial PEM standard is to ensure that sufficient background documentation is assembled to support procedural standards for the following PEM processes.

1. Assessing the quality of existing input maps.
2. Preparing and compiling the input data base.
3. Validating the knowledge base.
4. Implementing the knowledge base against the input data.

To support this objective, the PEM documentation standard (Section 4.7, “Documentation”) ensures that documentation of PEM procedures is sufficient for consistent replication and monitoring of the procedure.

4.4 Input Data Quality

Before the late 1970s, most resource inventories consisted of written reports and hard copy maps and were designed to be used somewhat independently as stand alone products to support planning. With these products, topological accuracy was typically more important than high positional accuracy. These inventories were not intended to support the overlaying process prevalent with most of today's PEM approaches. Through the 1980s and early 1990s geographic information systems became increasingly important tools, but their principal application was for automated cartographic production, and topological accuracy still received more emphasis than positional accuracy. While the lack of positional accuracy may preclude the use of these maps directly in the overlay process in PEM, the associated tacit knowledge often portrayed in legends and reports may be useful in constructing knowledge base relationships. Table 1 provides a summary of the recommended use of these inventory maps and reports for PEM.

Table 1: Summary of recommended use of common inventory maps

Inventory maps	Recommendation on use	Condition/ Consideration/Comment
Digital elevation model from TRIM	Support – consideration	Consider if the DEM resolution is consistent with the dimensions of the landscape features of interest
Biogeoclimatic maps 1:250 000 (legacy)	Do not support	Localization of biogeoclimatic lines required at project scale
Biogeoclimatic maps large scale	Support – conditional	Maps sanctioned by MoF are supported directly; other maps may require demonstration of thematic and possibly positional accuracy
Forest Inventory FC1	Support – conditional	Positional and thematic accuracy should be demonstrated
Forest Inventory VRI	Support – consideration	Caution when using tree species order of dominance and proportions and retro-fitted VRI[a]
Terrain Mapping 1:20 000 pre-1990	Do not support – conditional	Positional and thematic accuracy and reconciliation of complex polygon overlay must be demonstrated
Terrain Stability Mapping 1:20 000	Support – conditional	Thematic accuracy and reconciliation of complex polygon overlay must be demonstrated
Bioterrain Mapping 1:20 000	Support – consideration	Reconciliation of complex polygon overlay should be considered
Soils – 1:20 000	Do not support – conditional	Positional and thematic accuracy and reconciliation of complex polygon overlay must be demonstrated
Soils – 1:100 000–1:126 000	Do not support	Legend and report may provide value in developing knowledge base
Soils – enlarged reconnaissance	Do not support	

[a] "Retrofitted" VRI occurs when the FC1 polygon delineations are re-attributed to VRI standards.

The TRIM program created a well-controlled and positionally accurate base against which thematic inventories could be registered. However, most existing input maps did not use TRIM bases, therefore overlaying these on TRIM can result in significant incongruencies, causing incorrect resultant polygons, followed by incorrect predictions. Depending on the era and base maps used for compilation, the quality of thematic input maps may be suspect. The real challenge is how to take full advantage of useful portions of the information from these data sources. The following sections discuss some prevailing input data sources used in PEM from the perspective of thematic and spatial accuracy.

4.4.1 Thematic Input Data Quality

Formal quality assurance testing of PEM thematic input data is beyond the resources of most PEM projects. However, reasonable inferences about thematic input data quality are possible using the project input meta-data identified in Section 4.7.1 "Input Requirements." Moon (1999) discusses the interpretation of thematic data quality using these kinds of meta-data.

4.4.1.1 TRIM and Associated Digital Elevation Model (DEM)

TRIM maps contain some features (e.g., wetlands, wooded areas, gravel bars, etc.) that may be useful predictive features in a PEM project.

Quality control and quality assurance are well documented for the digital elevation data associated with the TRIM maps. The adequacy of the DEM resolution will be a function of the nature and scale of terrain features in the project area. The adequacy of slope, slope shape, and aspect models will be a function of the DEM resolution, the nature and scale of the terrain features, and the algorithms used.

The standard supports the use of TRIM and associated DEM data, ***but*** *discretion should be applied in the use of slope, slope length, and aspect derivations from these data. The scale and nature of landform variability should be determined and the adequacy of the DEM for modeling ecologically significant landscape features confirmed and documented.*

4.4.1.2 Biogeoclimatic Maps

Provincial digital biogeoclimatic maps, available at 1:250 000, were derived from 1:100 000 to 1:500 000 mapping. This scale is too small for PEM where the biogeoclimatic unit determines the range of possible ecosystem units. Localization of biogeoclimatic units is essential to a quality PEM product.

The standard requires that large-scale biogeoclimatic mapping be developed for the PEM project area following A Method For Large-scale Biogeoclimatic Mapping in British Columbia *at a minimum Level 2 protocol—Reconnaissance Reliability.*

The large scale biogeoclimatic mapping method outlines the steps, field work, and QA required to conduct large-scale biogeoclimatic mapping.

4.4.1.3 Forest Inventory Program Forest Cover Maps

FC1 — Pre-VRI (Vegetation Resources Inventory): These maps are referred to as FC1 maps. The forest inventory program began with the first complete 1 inch to 2 mile provincial inventory in the 1950s. These maps contained basic forest classifications, such as mature, immature, non-productive, alpine and cultivated land. The second complete inventory (1:15 840 and 1:31 680) in the 1960s and 1970s detailed dominant species composition, age, and site class. Subsequent updates continued through to 1993.[3] The basic mandate of the inventory remained constant. The 1960s and 1970s era inventories were designed to provide average strata volumes at a Timber Supply Area (TSA) level with a sampling error of less than ±10 percent at a 95 percent confidence interval. Formal inventory audits begun in 1995, test these objectives but do not test polygon-specific data quality. These audits are available at http://www.for.gov.bc.ca/RESINV/homepage.htm under the heading Audit Reports.

The identification of the three dominant overstorey species in a polygon is generally reasonable but the order of dominance is suspect. Positional accuracy of forest cover polygons should be assumed to be poor. The authors are unaware of any extensive formal evaluations of the thematic and spatial accuracy of individual FC1 map polygons. However, the assumption of poor positional accuracy is based on the following:

1. The Inventory Branch of the BCFS did not design the forest-cover maps as a base for site-specific applications. Discussions about possible site-specific applications and merging forest cover maps with other thematic information such as soil and terrain inventories arose from the program to digitize forest cover maps during the 1980s. Issues of inventory design, base map inconsistencies, and the inability to reconcile spatial discrepancies led to the conclusion that site-specific applications were problematic.
2. In the 1980s, the Land Resources Research Institute of Agriculture Canada attempted to reconcile forest cover polygons for integration with a land resource inventory.[4] At the request of the Vancouver Forest District, they tried to overlay forest cover with soil and vegetation. The thematic content of the forest cover maps was reasonable but positional accuracy was very poor and it was infeasible to reconcile the maps.
3. During the conversion of some digital FC1 maps to the TRIM base, road locations had serious locational and topological discrepancies relative to other ground control features. In some cases, the FC1 road locations were not used.[5] If the forest cover polygons were topologically correct with reference to roads, the shift to the TRIM-based road network will invalidate any topological accuracy between the roads and the polygons. Advances in spatial software, particularly coordinate adjustment to match control points may resolve some of the spatial problems.

The standard conditionally supports use of FC1 forest cover maps in PEM applications. Adequate positional accuracy must be demonstrated using the spatial integrity test outlined in Section 4.7.2.3. This procedure has two applications. The first is an initial check to determine the need for spatial reconciliation procedures and the second is to complete and verify the spatial reconciliation of thematic maps to the TRIM base map.

VRI Forest Cover Maps: VRI forest cover maps are compiled from 1:20 000 photos processed to the same specifications as TRIM base maps. Acceptable levels of both topological and positional accuracy for boundary transfers can be assumed. There are no data on the accuracy of thematic boundary locations and little data on thematic accuracy of polygon contents. A within-polygon variation study[6] suggests overstorey species identification, if not order of dominance and proportion, may be reasonable.

---

[3] Chapter 2 in *The preparation and creation of FRGIS Data Files*, Volume 5. Available at http://www.for.gov.bc.ca/RESINV/STANDARD/volume5/.

[4] Moon, D.E. and J.A. Brierley. 1988. Mill and Woodfibre Creeks Resource Inventory and Planning Base. Agriculture Canada misc. publ. no. 84-01; Land Resource Research Institute, Agriculture Canada; CEF Ottawa.

[5] Jardine, M. and R.A. Sims. 1999. A Compilation Summary of Data Acquisition Standards and Specifications Used in the Forest Inventory Program. Final report submitted to the Resources Inventory Branch, BC Ministry of Forests.

[6] Jahraus, K. and M. Penner. 1997. Within Polygon Variation Study in the Boston Bar Area. Ministry of Forests, Resources Inventory Branch Report.

The standard supports the use of VRI forest cover maps where available, ***but*** *recommends caution when using tree species order of dominance and proportions. The standard conditionally supports retrofitted VRI maps, where the FC1 polygon delineations are re-attributed to VRI standards. These maps should normally be treated in the same manner as FC1 maps (see above).*

4.4.1.4 Terrain and Bio-terrain Maps

Pre-1990 1:20 000 Terrain Maps: The terrain classification system[7] is a scheme designed for the classification of surficial materials, landforms and geomorphological processes. Terrain units reflect inferred mode of deposition and broad classes of depth of deposition and stratigraphy. The mapping entity is the terrain unit. The map entities are complexes of terrain units listed in order of dominance. Broad texture classes are inferred from assumed mode of deposition and landform. Polygons with two to three terrain units are common. The delineation of pre-1990 terrain units was largely air photo interpretation with limited field checking.

Because the location of polygon boundaries between complex map entities can be highly subjective, the assessment of positional accuracy is difficult. Map compilation and boundary transfer procedures were comparable to 1:20 000 soil maps. There are no quality assurance data available for terrain mapping but the broader class ranges and the reliance on visible landform suggest somewhat higher thematic reliability. Spatial reliability is probably comparable to 1:20 000 soil maps.

These maps have the following characteristics:

- An individual polygon description generally contains from one to three identified terrain units. The precise locations of individual terrain units are impossible to determine in polygons containing multiple terrain units.
- The ecological significance of terrain units is highly variable. For some terrain units such as active colluvial fans, there is a strong correlation between landform and ecological condition. For others such as morainal blankets, the range in ecological conditions is wide.
- The positional accuracy of terrain polygons is probably low to moderate and comparable to that of 1:20 000 soil polygons.

The standard does not support using these maps in PEM applications unless adequate thematic and positional accuracy can be demonstrated and the resultants of complex polygon overlays are appropriately reconciled and documented in the knowledge base. Selected terrain mapping units that have strong correlation with ecosystem condition may be useful in developing the knowledge base.

1:20 000 Bio-terrain: Bio-terrain mapping is a derivative of terrain mapping and is designed to delineate areas of ecological significance. The principal difference between terrain and bio-terrain mapping are the addition of soil drainage[8] as a differentiating criteria for mapping entities and an appreciation for the combined influence of slope and aspect on ecosystem distribution. There are no current published procedures or standards for bio-terrain mapping nor are there any quality assurance data available. However, bio-terrain mapping follows field procedures similar to terrain mapping except that boundary location may be influenced by ecological conditions. They can be expected to show similar characteristics to 1:20 000 terrain maps.

[7] Howes, D.E. and E. Kenk, 1997. Terrain Classification for British Columbia Version 2. Ministry of Environment, Lands and Parks.

[8] The “drainage” does not conform to the soil discipline standard. Instead, it is an estimation of effective moisture, largely during the growing season.

Bio-terrain mapping is a recent innovation and may use ortho-photos consistent with TRIM specifications. Where this is true, bio-terrain maps have the following characteristics:

- An individual polygon description generally contains from one to three identified bio-terrain units. The precise locations of individual bio-terrain units are impossible to determine in polygons containing multiple bio-terrain units.
- The ecological significance of bio-terrain units is less variable than terrain units but this will be dependent on the reliability of the drainage variable. The very low frequency of ground verification, the difficulty in interpreting soil drainage, the inconsistency with which drainage has been determined, and the relatively low accuracy of drainage determination in soils maps suggest that bio-terrain estimates of drainage should be treated with caution.
- The positional accuracy of bio-terrain polygon line transfers from ortho-photos is high but the accuracy of boundary location on the ortho-photo is probably comparable to that of 1:20 000 soil polygons.

The standard supports the use of bio-terrain maps in PEM applications, ***but*** *adequate thematic accuracy must be demonstrated and the resultants of complex polygon overlays appropriately reconciled and documented in the knowledge base meta data.*

1:20 000 Terrain Stability Mapping: Like bio-terrain mapping, terrain stability mapping is a derivative of terrain mapping. Unlike bio-terrain mapping, map delineations are drawn to reflect comparable levels of terrain stability.

The standard conditionally supports using these maps in PEM applications. Adequate thematic accuracy should be demonstrated and the resultants of complex polygon overlays appropriately reconciled and documented in the knowledge base meta data.

4.4.1.5 Pre-1990 Soil Maps

For many surveys, particularly pre-1990 soil surveys, much of the requested meta-data are unavailable. Soil Inventory Methods for British Columbia, 1995 (view or download at http://www.for.gov.bc.ca/ric/Pubs/teEcolo/Soil) presents the concepts and procedures used in soil survey from the 1970s to the 1990s.

1:15 000–1:20 000: The mapping entities used in these maps are soil series. The map entities are single soil series or more commonly complexes of soil series. Data from Agriculture and Agri-Food Canada (Moon and Selby, unpublished[9]) indicate that the SIL 2[10] map areas having at least one ecologically significant property that contrasts[11] to that indicated by the map generally exceeds 60 percent.

Map compilation and boundary transfers from photo to base map ranged from the use of epidiascope to zoom transfer stereoscope. Quality control was limited and base maps ranged from enlarged 1:50 000 topographic maps to forest cover maps and uncontrolled air photo mosaics. Topological accuracy is good but positional accuracy is low.

---

9 Study of soil survey reliability on Vancouver Island.

10 Survey Intensity Level 2 (SIL 2), Mapping Systems Working Group is generally mapped at 1:5000 to 1:40 000 (most commonly 1:20 000), has a minimum delineation size of 0.1 ha to 10 ha (most commonly 2 ha) and is expected to provide information for many purposes down to the level of local planning for groups of farms, stream catchments, small parks or irrigation management. It will not likely be used for specific site selection.

11 The term contrast refers to a difference between mapped and found property values large enough to change most interpretations using the evaluated property. For example an area mapped as imperfectly drained but found to be very poorly or well drained is contrasting.

These maps provide:

- reliable information on the proportion of soil units found in the map area
- moderately reliable information on the general properties of the soil found in a given polygon and low to moderate reliability for the specific soil units found in a given polygon
- low to moderate positional reliability for individual polygons.

The standard does not support using these maps in PEM applications unless adequate thematic and positional accuracy can be demonstrated.

1:100 000–1:126 000: These are reconnaissance scale maps. They provide information on the general types and distribution of soils in the survey area. The mapping entities are soil associations (recurring patterns of soil series generally corresponding to topographically controlled drainage sequences). The map entity is generally a complex of two to three soil associations with each soil association comprising two to four soil series. In addition, mapping standards of the time allowed up to 30 percent of the area to be occupied by unidentified inclusions of soils or non-soil areas. The generalized information content of the map entities has a thematic accuracy of around 60 percent. Because the location of polygon boundaries between compound and complex map polygons can be highly subjective, positional accuracy is very difficult to assess.

Map compilation and boundary transfers from air photo to base map ranged from epidiascope to zoom transfer stereoscope. The compilation and publication base frequently differed and boundaries were transferred to a 1:50 000 base map and then photo-mechanically transferred to a 1:100 000 or 1:126 000 NTS topographic base map for publication. Quality control for mapping and boundary transfer was low but quality control for cartographic production was generally high. Topological accuracy is good but positional accuracy (especially when merged with 1:20 000 base maps) is low.

Reconnaissance soil maps have the following characteristics:

- an individual polygon description can contain from 3 to 12 identified soil series, the precise locations of which are impossible to determine from the map. In addition, up to 30 percent of the polygon may be unidentified soil or non-soil types.
- the landscape/soil relationships used to predict the distribution of soils within an association may provide a useful model for predicting soil properties from topographic and terrain unit information at larger scales.
- positional accuracy and precision for individual soils or soil associations is very low.

The standard does not support using reconnaissance soil maps for PEM applications. However, the landscape models used to define soil associations should be relevant and applicable to the development of a PEM knowledge base.

Enlarged Reconnaissance Soil Maps

A number of 1:100 000, 1:125 000 maps were digitized and enlarged to 1:15 000 or 1:20 000 and may be available in regional Forest Service offices. In addition to the digital enlargement, the dominant soil from each association was used to characterize the polygon.

The standard does not support using enlarged reconnaissance soil maps for PEM applications.

4.4.1.6 Thematic Data Compiled in Support of a PEM Project

Where a PEM project conducts inventories, such as bioterrain, or compiles thematic input data, such as satellite imagery, the relevant discipline and digital data standards apply. In addition, where data is not in full compliance of an appropriate discipline or digital standard, the PEM standard requires the meta-data described in Appendix 1.

4.4.2 Spatial Input Data Quality

Many agencies and organizations follow rigorously defined procedures for the compilation of base and thematic maps. These organizations publish both the procedures and the minimum levels of spatial accuracy (e.g., *US National Map Accuracy Standards* for 1: 24 000 maps are reported as a 90th percentile error of approximately 12 metres). *Where such standards and documentation are available, it is sufficient to cite the accuracy standard in the meta-data and to reference the agency procedures.*

Where appropriate standards and documentation are not available, surrogate measures of spatial accuracy relative to the TRIM base level of documentation are required. Two elements contribute to spatial accuracy. The first relates to the accuracy of the base map used in the compilation of the thematic information. Section 4.7.2.3 presents a procedure for evaluating the spatial quality of the base map. The second is the accuracy of the thematic boundary transfer to the base. Moon (1999) discusses the implications of boundary transfer techniques and ground control. The base map section of Appendix 1 identifies meta-data that will provide an indication of general positional accuracy.

Where the PEM procedure uses thematic maps not supported by the standard for use in PEM application, the standard requires that the use of the maps be supported by the meta-data listed in Appendix 1 and by the testing and reconciliation described in Section 4.7.2.3.

4.5 Base Maps

The standard for all PEM base maps is the B.C. Government TRIM map series. Where available, the TRIM II series is the standard, otherwise the TRIM I series is the standard. Projects wishing to use other base maps must demonstrate conformance to TRIM II standards documented in Policies and Specifications for TRIM II (1:20 000) and (1:10 000) Revision Data Capture Version 2.0, May 1997 *available at* ftp://ftp.env.gov.bc.ca/dist/gdbc/TRIM2/trm2spcs.zip.

The same standard also applies to input data inventories conducted in support of PEM mapping.

4.6 Quality Assurance

Quality assurance tests input, intermediate, or final products to ensure a specified standard of quality. The test is **independent** of the methods used to produce the product and can test different stages of product completion. Section 4.6.2 identifies a protocol for quality assurance and accuracy assessment of ecosystem maps, (Meidinger, 1999). *This protocol can be adapted to assess any map-based thematic data and should be applied to any input, intermediate, or final map product for which quality assurance is required.*

4.6.1 Knowledge Base and Algorithm Validation

The quality of a PEM product will be a function of:

- the input data used to predict ecosystems
- the degree of correlation between input data and the predicted entities

- the knowledge base used to make the predictions (including quality of field samples and experience of the ecologist), and
- the procedures used to apply the knowledge base to the input data.

This section refers to the validation of the knowledge base (i.e., the result of applying the knowledge base to the input data). As such, it does not test the quality of the PEM product because it assumes correct input data. Section 4.6.2 discusses the assessment of the ecosystem map quality.

4.6.1.1 Validation Procedures

The knowledge base validation procedure requires processing a validation data set by the PEM practitioner and reporting the results to the custodian for PEM. The validation data set must not be used in the development of the knowledge base.

The data set can be compiled from:

- pre-existing ecological plots converted to polygon-like data
- project-specific plots collected as a validation data set.

The data set is comprised of data with the suite of attributes used in the knowledge tables and a known outcome (i.e., site series and site modifier). The knowledge base validation results will be used to assess the quality of the knowledge base before its acceptance into the provincial data warehouse.

The validation requires a minimum of 30 samples, although more than 30 will be required for an adequate validation of the knowledge base in projects that cover large areas. Samples should be random or they should represent the range of site conditions found in the project area. In the case of representative sampling, the sampling procedure will be as follows.

1. Criteria defining representative samples must be established prior to data collection.
2. Sampling routes, designed to cover the range of conditions in the project area, are identified by a predetermined starting point, predetermined bearings, and predetermined distances.
3. Field sampling follows a pre-selected route. The areas conforming to the selection criteria (step 1) are sampled.
4. The minimum attribute set is collected. The minimum attribute set consists of those attributes used in the PEM knowledge-based procedure and the results of field-based classification to BEC site series or acceptable mapping entity where BEC site series is inappropriate. Additional attributes may be collected but are not required. Attribute data are collected according to the standards under which the input data in the knowledge base were collected, or as defined in *Field manual for describing terrestrial ecosystems* or other RIC-approved manuals such as the *Vegetation Resources Inventory Ground Sampling Procedures*.

The knowledge base will be subject to quality assurance by the custodian for PEM before acceptance into the provincial data warehouse and the results of that quality assurance process will be captured as project meta-data. The standard requires a minimum of 30 validation samples.

4.6.1.2 Reporting

The test data sets and a table showing the sample plot number, the predicted class, and the observed class must be appended to the report. Report results of the knowledge base validation as indicated in Table 2.

Table 2: Knowledge base validation

Sample plot number	Predicted class	Observed class	Result
1	Site series X	site series Z	0
2	Site series Q	site series T	0
3	Site series M	site series N	0.5
4	Site series Q	site series Q	1

1.0 = an exact match
0.5 = the predicted class is adjacent ecologically to the correct class (e.g., in the case of site series use the edatopic grid in the BEC classification to determine adjacency). For compound mapping entities, determining this intermediate score is more difficult; in these cases discuss the scoring with the custodian for PEM.
0 = the predicted class has one or more intervening class between it and the correct class

4.6.2 Map Data Quality Protocol

Meidinger (1999) presents a set of procedures for assessing the quality of ecosystem maps. The protocol is also applicable to quality assurance of other maps (e.g., input maps, however this is not a requirement).

The protocol presents a statistically unbiased approach to evaluating the acceptability or accuracy of the mapping. The thematic content of randomly selected map polygons is assessed by various means, with the methods varying in precision and objectivity. These within-polygon assessment approaches are presented as multiple 'levels' of the protocol. Users select the appropriate level based upon their intended use of the data and the project budget.

PEM products must have at least one level of the protocol implemented and the results appended to the product when submitted.

The protocol includes the following features:

- the assessment is conducted after the mapping is complete
- samples are distributed throughout the entire project area
- the project area can be stratified (e.g., alpine and below alpine), but samples must be distributed in all strata
- the assessment is conducted by "experts" and is supported by some field data
- as most PEM and TEM polygons are complexes of ecosystem units, the variation within polygons is assessed by multiple plots or mapping at a larger scale (e.g., 1:5000)
- where a map consists of simple units, the preparation of a contingency table or 'confusion matrix' of errors of omission[12] and errors of commission[13] is recommended
- the final scores provide data on the accuracy in describing the dominant map unit components and all map unit components
- a score for the accuracy of "correct" plus "close" (e.g., classified as an adjacent site series on the edatopic grid) categories can also be reported as an additional statistic.

[12] Polygon incorrectly omitted from a class (e.g., site series).
[13] Polygon that actually belongs in another class incorrectly assigned to a particular class.

The results of the assessment are non-spatial—that is, they identify what the level of accuracy is, but they do not show geographically, within the map or project area, where errors or inconsistencies occur.

*The standard **requires** a minimum of a level 1 protocol (basic quality assurance) be applied to the PEM project area and the results appended to the product submission. The standard **recommends** a minimum of a level 4 protocol (basic accuracy assessment)where feasible. In the sections that follow, data required by the standard are listed as numbered and bracketed items.*

4.6.2.1 Reporting

Report the following assessment statistics[14] for each assessment, whether a QA check or an accuracy assessment:

[1]. Chi-squared test of proportions (e.g., insignificant difference in ecosystem unit proportions).

[2]. Percent dominant correct (e.g., 62 percent), and 95 percent confidence interval (e.g., +/- 6 percent) for each map entity (i.e., site series, site modifier, etc.).

[3]. Percent overlap (e.g., 49 percent) for each map entity.

[4]. Percent acceptable overlap, if assessed (optional).

If all map units are "simple," presentation of a contingency table or confusion matrix is optional.

4.7 Documentation—Meta-data Standards

The PEM standard establishes minimum levels of documentation and meta-data required to evaluate the quality of input data, predictive procedures, and output products of PEM. The meta-data specified below meet three needs.

1. It provides sufficient information about the nature of the input entities, input data, predictive procedures, and output products for a qualified PEM practitioner to understand the limitations of these items for PEM applications.
2. Its compilation by the PEM practitioner ensures that the practitioner has researched the input data and adequately documented the procedures and output products.
3. A longer-term goal of the PEM standard is the eventual integration of PEM/TEM data, information, and knowledge into a single logical data model and repository. The task of integrating TEM with PEM is beyond the scope of this standard but this section will provide the documentation and meta-data necessary to construct such a repository.

4.7.1 Input Requirements

The intent of the meta-data presented below is to document PEM input in detail sufficient for qualified PEM practitioners to evaluate the quality of the input data and PEM products. For a more detailed discussion on the interpretation of meta-data, refer to *A Problem Analysis on Data Quality Assessment Issues* (Moon, 1999).

Most of the meta-data listed below requires reference to a definition or procedure. Where the reference is in published form, the reference is sufficient. Where the reference is not in published form, the standard requires appended documentation of the definition or procedure.

---

[14] A Protocol for Quality Assurance and Accuracy Assessment of Ecosystem Maps (Meidinger, 1999)

For each thematic input data source used in the PEM project, the meta-data identified below are required. For any new, non-RIC inventories (e.g., satellite image analysis) specify the additional meta-data listed in Appendix 1. Meta-data should be submitted following the Standard for Digital PEM Data Capture as a text document.

Project: Input Data Source

[5]. ***Citation*** *– Specify a reference to a formal, published source of the data, if available.*

[6]. ***Consultant/Department*** *– Specify the public or private sector organization(s) responsible for collecting, compiling, and maintaining the data and an appropriate contact within the organization(s).*

[7]. ***Publication scale*** *– Specify the original publication scale of the inventory.*

[8]. ***Period of compilation*** *– Where possible, specify the date range during which the data were compiled. The period of compilation, combined with the responsible agency can be used to identify the compilation and quality control procedures in place during the compilation process.*

Base Map: Input Data Source

[9]. ***Projection –*** *Specify the original projection used.*

Mapping Concepts: Input Data Source

[10]. ***Mapping entities*** *– Describe or reference available definitions of all mapping entities used as input data in the PEM process.*

[11]. ***Map entities*** *– Describe or reference available definitions of the map entities (map symbols) used as input data.*

[12]. ***Entity relationships*** *– Identify how the important input entities relate to each other, particularly those relationships that are pertinent in the knowledge base.*

4.7.2 Input Processing Requirements

The PEM process begins with the preparation, compilation and derivation of the input data used to predict ecosystem units. It ends with the submission of predicted ecosystem maps and associated data (outputs) to the PEM custodian.

For each thematic input data source used in the PEM project, the meta-data identified below are required.

4.7.2.1 Input Map Compilation Quality Control

[13]. ***Edge matching*** *– Specify by reference to published or appended documentation, the procedures for matching polygon boundaries at the edges of multiple map sheets assembled to cover the project area and at the edge of the project area.*

[14]. ***Edge matching error*** *– Specify the minimum, average, and maximum boundary displacement at the joined edges of the maps for the project area.*

[15]. ***Attribute/Label matching*** *– Specify the percentage of polygon labels or attributes that differ on either side of the neat line for the project area.*

[16]. ***Raster size*** *– for raster process, specify raster size used.*

4.7.2.2 Digital Elevation Model Derivation

[17]. Automated digital elevation model methods must be based on the TRIM mapping or the Gridded DEM product. If a digital elevation model based on TRIM or TRIM2 is used, a full description of the method of converting the TRIM format files to a digital elevation model in the processing environment used is required.

4.7.2.3 Spatial Reconciliation

Thematic input maps prepared at different times, scales, and on bases other than the BC Government TRIM series, require evaluation (Section 4.7.2.3.1) and reconciliation (Section 4.7.2.3.2) to the TRIM standard.

The following procedure[15] has two applications. The first is an initial check to determine the need for spatial reconciliation procedures and the second is to complete and verify the spatial reconciliation of thematic maps to the TRIM base map.

4.7.2.3.1 Procedure for evaluating consistency with a TRIM base map

The intent of imposing a positional accuracy standard is to ensure that the various spatial themes (data layers) used are "in the same place." Common problems result from errors in data processing (e.g., the wrong datum or spheroid is used) or inadequate registration of source maps during digitizing or method of transfer of lines from air photos. Detecting locational errors is not an easily automated process. However, relatively easy error detection is possible using check plots.

Use the following procedure to create check plots that demonstrate the positional accuracy of the input data (for all input data except localized biogeoclimatic):

Produce a single map of the entire study area showing the hydrography (streams, rivers, lakes and wetlands) from the TRIM source. Where digital thematic maps do not have an associated hydrographic theme, digitize hydrographic features from the original base map for this procedure.

Select a minimum of eight, temporally stable hydrographic features that are well distributed throughout the study area. Four of the features should be as near the four edges of the area as possible. Record the features in Table 3 and append this table to the documentation.

Table 3: Example recording of control feature shift, initial or adjusted

Feature type	Count and/or average length (m)	Minimum shift (m)	Average shift (m)	Maximum shift (m)
Point/Intersection	3	20	35	70
Linear	2, 200	0	20	55
Polygon	3	30	40	120
Overall	8	25	36	85

Overlay the hydrographic features for each of the other input data sources on the corresponding TRIM features.

Produce 1:20 000 check plot(s) showing the TRIM features and the features from each of the other data sources. Also produce a plot of the entire study area showing the selected features.

Examine the check plots for "shifts" in the data and record the shifts in real world measurements (metres) in Table 3.

While the exact shapes of the features will not be the same, the positional bias or "shift" between the two sources should average less than 50 metres and no shift should exceed 150 metres. If

[15] Procedure adapted from M. Eng and B. Enns, Research Branch, BC Ministry of Forests.

within acceptable limits, append the results as Table 3 to the submitted product. If the estimated average shift exceeds 50 metres or the maximum shift exceeds 150 metres, undertake spatial reconciliation as described in Section 4.7.2.3.2.

If the localized BGC data was based on a digital elevation model using TRIM data then no input data quality testing is required. If the localized BGC data came from another source (e.g., hand drawn and digitized) then its positional accuracy should be checked:

Generate 100 metre contour lines for at least five 1:20 000 map sheets; four at the corners of the study area and one in the middle (or for the entire study area if it is less than five map sheets in size).

Overlay the localized BGC linework on the contour lines and produce check plots showing both.

Examine the check plots for positional bias (as above) and for inconsistencies in elevation-based BGC lines.

4.7.2.3.2 Procedure for reconciling thematic maps to the TRIM base map

Select a minimum of eight, temporally stable control points that are common to the input map base and the TRIM base and that are well distributed throughout the map sheets being reconciled. Four of the features should be as near to the four edges of the area a possible. Use these control points to adjust coordinates for the map such that the control points are congruent. Record and append the following information to the product submission: the software, software version, software function, and function parameters used in the reconciliation.

Following the adjustment process, redo and report the results of the evaluation procedure using hydrography.

4.7.2.3.3 Reporting

[18]. ***Adjusted control feature shift*** *– the standard requires determination and reporting of the control feature shift as in Table 3. Indicate if the numbers are for the initial or adjusted control feature shift. The shift between the two sources should average less than 50 metres and no shift should exceed 150 metres.*

4.7.2.4 Thematic Compilation and Derivation

[19]. ***Mapping entity cross product correlation*** – The overlay of complex or compound map units on other complex or compound map units presents the problem of reconciling which of the possible thematic combinations actually occur in the resultant. For example, the simple overlay of a compound polygon A^7B^3 with X^5Y^5 produces the following possible proportions for each thematic resultant: AX from 21% to 50%; AY from 21% to 50%; BX from 0% to 30%; or, BY from 0% to 30%. There is no a priori method of determining the actual proportional representation unless some combinations can be shown to be impossible. *Specify by reference to a published or appended document, the method used to reconcile the cross product of mapping entities resulting from the overlay, polygonal or raster, of compound or complex map entities on other compound or complex map entities.*

[20]. ***Sliver adjustment –*** *Specify by reference to a published or appended document, the method and criteria used to identify and reconcile sliver or artifact polygons or rasters resulting from the overlay process.*

[21]. ***Attribute extraction –*** *Specify by reference to a published or appended document, the method used to extract or derive and validate attributes (e.g., values, range, qualified list or range, or any combination of these) used as input to the PEM process.* Possible attributes are species presence, species height, stocking, parent material, etc. if used as input to the PEM process. Possible methods are 1) created as a lookup table from the map legend or 2) created as a

lookup table derived from Forest Inventory Planning (FIP) file following the appended algorithm.

[22]. ***Landform feature extraction or derivation –*** *Specify by reference to a published or appended document, the method used to identify, extract, process, and validate special features if used in the PEM process.* Possible features are hill-tops, ridge crests, mid-slopes, toe slopes, gullies, avalanche tracks, stream density, etc. if used as input to the PEM process. Possible methods are 1) gullies identified from terrain symbol for gullies or 2) derived from digital elevation data using the appended algorithm.

[23]. ***Spatial attributes*** *– Specify by reference to a published or appended document, the methods to process and validate spatial attributes used in the PEM knowledge base. Possible examples are adjacency or proximity to features such as lakes, streams, avalanche tracks, etc. and proportion of polygon area occupied by features such as lakes, streams, avalanche tracks, etc. used as input to the PEM process.*

4.7.3 Knowledge Base and Algorithm Requirements

Sections 4.7.3.1 and 4.7.3.2 provide documentation standards for the map entities and the attributes used in a PEM knowledge base. Sections 4.7.3.3 and 4.7.3.4 provide documentation standards for rule-based and/or belief matrix algorithms.

Knowledge base strategies other than rule or belief matrix based systems are acceptable. However, all methods must document the relationship logic between the PEM entities and the input attributes to levels comparable to those shown in Sections 4.7.3.3 and 4.7.3.4.

4.7.3.1 Entities

Define fully the entities predicted (note, not each resultant polygon) by the PEM process. Where the predicted entities are consistent with mapcode.xls, cite the file version number. Where predicted entities differ from mapcode.xls or ecosystem units in Table 3.1 of the TEM standard, the PEM standard requires definition of the entities, using a format similar to that in Tables 4 and 5. New entities require the approval of the custodian for PEM.

The PEM entities table (e.g., Table 4) lists predicted entities and their description. Where these do not differ from the site series in *mapcode.xls*, or an ecosystem unit in Table 3.1 of the TEM standard), the table would repeat information from these two sources for the PEM entities used in the project.

1. PEM entity code: is a unique identifier for the predicted entity, used in the project database and legend, and is either in the *mapcode.xls* file or is added to *mapcode.xls* file by the custodian for PEM, or is an ecosystem unit (in Table 3.1 of the TEM standard).
2. Description: is a description of the entity. As this example uses complex and compound entities, Table 5 is required to provide proportions of the component entities. For new compound PEM entities that may have been created, also provide distribution or pattern of the component entities (see also discussion in Section 4.1.1.3.)

Table 4: Example of PEM entities

PEM entity code	Description
OD	Hummocky rock outcrops with intervening shallow to moderately deep soils supporting very xeric to sub-xeric forest communities. [New code OD: Outcrops and Dry forest “compound”map entity]
RF	CwHw – Sword fern
SP	Undifferentiated shallow water [New code SP: Shallow open water and pond complex map entity]

Table 5 further describes the BEC site series or ecosystem units that make up or define the predicted entities. The first three rows indicate that PEM entity OD is made up of site series HwPl – Cladina greater than or equal to 40 percent, ecosystem unit RO less than or equal to 30 percent, and, site series HwCw – Salal less than or equal to 30 percent.

1. PEM entity code: is a unique identifier for the predicted entity, and a key[16] (i.e., index field) pointing "back" (linked) to the PEM entity description in Table 4.
2. Reference entity type: for example, a site series (in *mapcode.xls*) or an ecosystem unit (in Table 3.1 of the TEM standard).
3. Reference entity: is either the code from Table 3.1 or the site series name from *mapcode.xls*.[17] Note that zone/subzone/variant and site series name are mandatory to ensure that each site series is identified uniquely.
4. Proportion: is a description of the acceptable range of relative polygon proportion for the specified reference PEM entity.
5. Cover pattern: for new compound entities only, is a description or code for the typical spatial distribution pattern of the specified reference PEM entity. Codes follow the VRI cover pattern.

Table 5: Example of PEM entity definition

PEM entity code	Reference entity type	Reference entity			Pro-portion (decile)	Cover pattern
		Zone	Subzone/ Variant	Site series/ Ecosystem unit		
OD	Site series	CWH	mm1	HwPl – Cladina	≥4	7
OD	Ecosystem unit	CWH	mm1	RO	≤3	3
OD	Site series	CWH	mm1	HwCw – Salal	≤3	1
RF	Site series	CWH	mm1	CwHw – Sword fern	10	–
SP	Ecosystem unit	CWH	mm1	OW	>5	–
SP	Ecosystem unit	CWH	mm1	PD	<5	–

4.7.3.2 Attributes

Define fully the attributes used in the PEM process to describe, characterize, or infer entities (e.g., site series, ecosystem unit). Attributes include those extracted from thematic input maps and those derived from spatial modeling of digital elevation or climate data (i.e., through thematic compilation and derivation).

All Attributes

*[24]. **Definition and description** – Specify a reference to a published or appended definition of the attribute that includes its data structure (attribute type and format) as illustrated in Table 6.*

*[25]. **Attribute code** – Specify the code used for each attribute. The code may be a standard code from a referenced source or may be project-specific. If applicable, include in Table 6.*

*[26]. **Method** – Specify a reference to a published or appended method of measurement or analysis.*

16 More specifically, a foreign key in terms of database modeling.

17 The site series name is required because it is the only field or combination of fields in *mapcode.xls* that can uniquely identify the site series.

[27]. ***Unit of measure*** *– Specify the units of measure used to determine the attribute value (e.g., centimetres, microns, etc.).*

[28]. ***Entity/Relationship described*** *– Specify for all attributes used in the PEM process, the entity or relation described by the attribute. If the attribute describes a relation, specify the relation described. For example, the map label component for "Slope" may describe the average slope of the polygon, dominant slope of the polygon, or the slope of the dominant terrain unit in a polygon. In the first two cases slope describes the polygon entity. In the second case, it describes the relation of a terrain unit in a polygon.*

Numeric Data (real or continuous variable data)

[29]. ***Precision*** *– Specify the exactness with which a value is measured and recorded (e.g., number of decimal places). The precision should be a function of both the exactness of measurement and the reproducibility of the measurement.*

[30]. ***Statistic*** *– Specify the statistic associated with the value reported (e.g., observation, mean, maximum, minimum, standard deviation) and the number of samples on which the value is based.*

Categorical Data (discontinuous and unranked)

[31]. ***Valid values –*** *list the classes or categories and define or reference definitions for the valid values of the attribute (e.g., surficial material may be morainal, fluvial, lacustrine, etc.).*

Ordered Classes

[32]. ***Rank and limits –*** *identify the rank value for ordered categories (e.g., Rapid = r, Well = w), the rank (e.g., Rapid = 1, Well = 2, …) and class limits (e.g., Min = 1, Max = 9) for ordered class data.*

Table 6: Example of attribute definition

Attribute	Attribute code	Attribute type	Data type and format
Stand age	F_AGE	Real	Number(4)
Soil drainage	T_DRAIN	Ordered class	Char(2)
Surficial material	T_S-MAT	Categorical	Char(2)

4.7.3.3 Rule-based Systems

Rule-based systems identify rules based on one or more attributes whose value ranges conform in total or fail to conform to those required of a specific class (e.g., site series). Many rule-based systems use Boolean comparisons with If, Then, Else program control to identify the class to which an unknown belongs. Others use definite-clause grammars or list processing. The program may compare the attributes of an unknown instance to the rules governing class membership (forward chaining) or alternatively to determine whether the attributes of an unknown instance are consistent with the rules pertaining to a class (backward chaining). Rule-based systems produce binary results (e.g., an instance does or does not belong to a class). Partial or "fuzzy" membership is not recognized.

PEM knowledge bases may or may not use "rules." Rules may be used alone (i.e., a rule-based system), or in combination with another system (e.g., a belief matrix).

The standard requires the following documentation of rule bases to be incorporated into the provincial knowledge base repository:

Rule definitions

The rule definition uses Boolean operators with parentheses to define explicitly the comparison order. Table 7 shows the format for documenting rules. Table 8 and Table 9 list the operators used in building the Boolean condition. The two examples shown are equivalent rules. Note that the rules (e.g., Table 7) should be processing order independent (i.e., the result returned for a class should be the same irrespective of order in which the rules are processed).

Table 7: Example of rule definition format

Rule as a Boolean expression (The two rules shown below use different operators but are equivalent.)	Includes/ Excludes	PEM entity
((V1 >= x AND V1 <= y) AND (V2 >= q AND V2 <= r)) OR (V5 = ‘A’ OR V5 = ‘B’ OR V5 = ‘D’)	includes	AF
		AB
		HS
		RF
(V1 BETWEEN x AND y AND V2 BETWEEN q AND r) OR V5 IN (‘A,’ ‘B,’ ‘D’)	includes	AF
		AB
		HS
		RF

In the examples shown the elements V1, V2, and V5 indicate input data (e.g., FC1 species_1 percent, soil drainage, and terrain surficial material, respectively). The symbols or terms >=, <=, =, IN, BETWEEN are comparison operators. The lowercase letters x, y, q, r represent numeric values or ordered classes (e.g., % cover, or soil drainage classes) against which the input data are compared and ‘A,’ ‘B,’ ‘D’ represent literal or text values (e.g., terrain surficial material values) against which the input data are compared. Within a rule, the parentheses indicate the order of comparison (e.g., comparisons are performed from the innermost parentheses outwards). The second expression reads “IF V1 is BETWEEN the values x and y AND V2 is BETWEEN the values q and r OR IF V5 is one of ‘A,’ ‘B,’ or ‘D’ THEN Classes AF, AB, HS, and RF qualify for membership.”

The comparison operators listed are not an exhaustive listing. However, they represent a set of operators that can, in conjunction with the logical operators, define any relationship. The use of the logical operator NOT (described above) negates the condition to which it is attached. In fact, the comparison operators =, >, < (Table 8) and the logical operators NOT, AND, and OR (Table 9) can describe any Boolean condition.

Table 8: Example of comparison operators

Operator	Test	Condition	Value	Returns
=	Equality	V = 3.5	3.5	True
!=, <>	Inequality	V != 3.5	3.5	False
>	Greater than	V > 3.5	3.5	False
<	Less than	V < 3.5	3.5	False
>=	Greater than or equal to	V >= 3.5	3.5	True
<=	Less than or equal to	V <= 3.5	3.5	True
IN	Equal to any member of list	V IN (A, B, D)	B	True
	This test is equivalent to	V = A OR V = B OR V = D		
NOT IN	Not equal to any member in list	V NOT IN (A,B,C)	D	True
	This test is equivalent to	V != A AND V != B AND V != C		
	A null in the list returns false	V NOT IN (A, null)	D	False
BETWEEN X AND Y	Greater than or equal to X and Less than or equal to Y	V BETWEEN 7 and 11	8	True
	This test is equivalent to	V >= 7 AND V <=11		
IS NULL	This tests for a missing or null value	V IS NULL	null	True
			X	False

Table 9: Example of logical operators

Operator	Function	Example	Value	Returns
NOT	Returns True if the following condition is False. Returns False if the following condition is True. Returns Unknown if the following condition is Unknown	V NOT = 2	2	False
AND	Returns True if both conditions are true. Returns False if either condition is false. Returns Unknown if either condition is Unknown	V1 = 4 AND V2 = 7	V1=2 V2=7	False
OR	Returns True if either condition is True. Returns False if neither condition is True. Returns Unknown if either condition is Unknown	V1 = 4 OR V2 = 7	V1=2 V2=7	True

Logical operators modify the comparison operator used in the Boolean expression.

A knowledge base rule that requires site series X and Z to occur on slopes greater than 40 percent and textures other than clay, or silty clay is shown in Table 10.

Table 10: Example rule

Rule as a Boolean expression	Includes/ Excludes	PEM entity
Slope > 40 AND texture NOT IN (‘clay,’ ‘silty clay’)	includes	site series X
		site series Z

The standard requires output from rule-based knowledge systems compatible with representation in Table 7. The operators must be consistent with those presented in Tables 8 and 9 or they must be defined in terms of the operators identified in Tables 8 and 9.

4.7.3.4 Belief Matrices

A “belief” matrix records the belief in or the probability of an occurrence of a PEM entity (e.g., site series) in relation to a set of attribute values. At this time, there is no single standard format for documentation of a belief matrix. The matrix can be delivered in the format it is generated in, as long as all aspects indicated below are documented. The belief matrix is processed through an inference engine to produce a weighted sum of beliefs for each instance (e.g., polygon) in the GIS dataset.

The standard requires representation of the belief matrix equivalent to the set of tables presented as Tables 11, 12 and 13. In addition, a verbal explanation of the assumptions that guided the assignment of attribute values to map entities is required.

Documentation of individual attributes used in the belief matrix is presented in Section 4.7.3.2. In the following section, documentation of attribute values and combinations of attributes, called “attribute condition sets” are described.

Attribute Values and Condition Sets

Codes for all values or combinations of the various attributes need to be clearly documented. Attribute values and condition sets provide the basis for assigning a belief to a PEM entity. Attribute condition sets are composed of attribute-value combinations.

Table 11 presents a possible documentation format for attribute value codes and attribute condition sets to use in the belief matrix. The example demonstrates concepts rather than a required format.

1. Attribute code: specifies the code used to identify the attribute type
2. Attribute value: specifies the class for the attribute (the combination of attribute code and attribute value must be unique).
3. Attribute value description: provides a name or class limits for an attribute type and value (presented as a convenience as descriptions are recoverable from the cited reference).
4. Reference: provides a reference to a published or appended definition, code, and/or class limits for the attribute.

Table 11: Example of attribute values codes

Attribute code	Attribute value	Attribute value description	Reference to appropriate standard
B	11	NPBr	Ministry of Forests Resources Inventory Branch Relational Data Dictionary 2.0 – Basic Class
B	15	Lake	
B	35	Wetland	
P	1	5 to <20 percent	Appendix X – Percent of feature in polygon.
P	2	20+ percent	
W	00	None	Appendix X – Stream density in polygon.
W	01	1 – 20 m/ha	
W	02	21 – 40 m/ha	
S	01	0 – 9 percent	Section 4.7.3.2 – Slope (average for polygon).
S	02	10 – 25 percent	
AS	00	NA	Appendix X – Aspect classes.
AS	01	135–285 deg	

An attribute condition set may consist of a single attribute value or a combination of attributes and values. The conditions defined by the attribute condition set and the specific entity being predicted comprise the relationship described by the probability or belief value.

Table 12 presents an example documentation format for attribute condition sets.

1. Attribute condition set: provides a unique identifier for the attribute condition set.
2. Attribute code: identifies the attribute type.
3. Attribute value: as described in Table 11.
4. Attribute value description: as described in Table 11.

Table 12: Example of knowledge base condition sets

Attribute condition set	Attribute code	Attribute value	Attribute value description
BWS-1	B	11	NPBr
	W	00	None
	S	01	0–9 percent
BWS-2	B	11	NPBr
	W	02	
	S	01	0–9 percent
B-15	B	15	Lake
B-35	B	35	Wetland
SAS-1	S	02	10 – 25 percent
	AS	01	135–285 deg

The knowledge base includes a belief or probability value for the occurrence of a site series (or other mapping entity) under the conditions defined by the attribute value or condition set. Table 13 presents an example of the documentation for belief values.

1. Attribute value or condition set: described in Table 11 and Table 12.
2. Map-entity code: identifies the site series or other map entity being evaluated against the condition set; as documented in Section 4.7.3.1.
3. Belief value: represents the belief or probability value assigned to the attribute value or condition set by PEM entity.[18]
4. Authority: identifies the person/organization responsible for providing the value.
5. Date: the date when the value was assigned or last modified.

Table 13: Example belief or probability values

Attribute condition set	Map-entity code	Belief value	Authority	Date
BWS-2	AS	0.6	Jones, K	9/9/99
BWS-2	SS	0.5	Jones, K	9/9/99
B-15	SP	0.8	Jones, K	9/9/99
B-35	RC	0.9	Jones, K	9/9/99
SAS-1	OD	0.2	Downing, D	10/10/99
SAS-1	HS	0.4	Downing, D	10/10/99
SAS-1	LC	0.7	Downing, D	10/10/99
…				

Inference Process

An understanding of the inference process used to assign data sets (e.g., polygons) to site series is important to understand the values assigned in the belief matrix. *To facilitate this understanding, document the algorithm used by the inference system by reference to a published or appended document.*

4.7.4 Output Requirements

The PEM output is required to follow the Standard for Digital PEM Data Capture in British Columbia*.*

In addition to the PEM output, a structural stage age class digital product is required as outlined in the Protocol for Structural Stage Modeling in PEM*, currently under development.*

[18] The scale of the value must be specified (e.g., 1–10 or 1–100).

5 PEM Products Summary

As outlined in these standards, six products are required for a PEM project. These are:

1. Input and input processing documentation (see Sections 4.7.1 and 4.7.2)
2. Knowledge base documentation (see Section 4.7.3)
3. Quality assurance (validation) of knowledge base (see Section 4.6.1)
4. Digital output of predicted ecosystem map (see *Digital PEM Data Standards*)
5. Structural stage/age class map (see *Protocol for Structural Stage Modeling in PEM*)
6. Quality assurance/accuracy assessment of predicted ecosystem map (see Section 4.6.2).

6 References

Downing, D.J., R.K. Jones, and J. Mulder. 1998. Predictive ecosystem mapping: applications to the management of large forest areas. *In* Proc. Ecosystem Management of Forested Landscapes: Directions and Implementation. October 26–28, 1998, Nelson, BC.

Jones, R.K., D.V. Meidinger, D. Clark, F. Schultz, D.E. Moon, and R.A. Sims. 1999. Towards the establishment of a predictive ecosystem mapping standard: A White Paper. BC Min. Environ., Lands and Parks, Resources Inventory Committee, TEM Alternatives Task Force, Victoria, BC.

Meidinger, D. 1999. Protocol for quality assurance and accuracy assessment of ecosystem maps. B.C. Min. For., Research Br., Victoria, BC. Internal Rep.

Moon, D.E. 1999. Problem analysis on data quality assessment issues. BC Min. Environ., Lands and Parks, Resources Inventory Committee, TEM Alternatives Task Force, Victoria, BC. Draft.

7 Glossary

entity: a class or type of thing involved in the mapping process at a given level of the mapping hierarchy. An entity may be a site series, a biogeoclimatic unit, an ecosection, etc.

simple entities: a single entity (e.g., one site series characterizing a polygon).

complex entities: a mix of two or more entities in an unpredictable pattern; the complex entity inherits the properties of its members. Unlike compound entities, the definition of a complex entity does not include a predictable or derivable pattern of the member entities that would allow their specific location at a larger scale.

compound entities: a predictable and recurring association of two or more entities; the compound entity inherits the properties of its member components. Examples include soil associations or concentric wetland ecosystem units.

multiplex entities: are two or more compound or complex entities for which the pattern of all member components cannot be predicted.

mapping entities: any entities used in the mapping process but not necessarily labeled on the map. For example, a soil series forms part of the definition of a mapped soil association or a stratum forms part of a mapped geological formation. At larger scales, the series or strata may be labeled directly on the map. At smaller scales, the series or strata cannot be labeled practically on a map, rather the association or formation that they comprise is labeled. While the series and strata are mapping entities (entities used in the mapping process), they are not map entities (represented as labels on the map).

map entities: represent the code or symbol used to label map features. Map features may be labeled with simple, complex, compound, or multiplex mapping entities.

ground control features: features visible on the base map and aerial photography used in the mapping program and locatable on the ground during mapping. Ground control features are used to reference ground verification and traverses to aerial photographs and to reference thematic boundaries drawn on photographs to the thematic base map.

instance: a specific occurrence or example of an entity (e.g., the Princeton soil association, site series SwAt – Step moss, polygon 192).

map feature: a point, line or polygon representing a site, linear feature, or area on a map.

map legends: information defining and describing the map entities (e.g., symbols and map symbol configuration) found on the map; it may also include other supporting information such as data sources, project objectives, map credits. Legends may be open or closed. In an open legend, the map entities and their modifiers that are found in a map polygon are listed in the polygon label with some indication of proportion or dominance (e.g., rCv/gMbv//Rs). In a closed legend, the polygons have been classified and labeled with reference to a limited number of classes.

map reliability: refers to both thematic accuracy, spatial accuracy, and precision as they effect the specific application of the map information.

positional accuracy: refers to the degree to which map coordinates correspond to the real world coordinates of features shown on the map. Positional accuracy is often stated as the probability of a map feature being represented within a specified distance.

predictive ecosystem mapping: a computer, GIS, and knowledge-based method to assist in the stratification of landscapes into ecologically-oriented map units (typically site series) based on the overlaying of mapped themes and the processing of resultant attributes by inference methods (normally automated software) in association with a formalized knowledge base comprising ecological-landscape relationships.

qualified PEM practitioner: is a registered professional or a team led by a registered professional. The practitioner(s) have a background in and an understanding of ecological principles and practices, terrestrial ecosystem mapping, land resources inventory methods (including fieldwork, mapping, data and map compilation, information processing, and presentation conventions and procedures), spatial and attribute data management, including geographic information systems processing.

spatial accuracy: spatial accuracy includes two components, positional accuracy and topological accuracy.

spatial data: refers to map data stored using a two or three dimensional coordinate system to locate points relative to a common reference system (global projection system, a local projection system, or simple table coordinates).

TEM alternatives: alternative procedures to the RIC Terrestrial Ecosystem Mapping standard, that produce inventory or information products with similarities to TEM products or that provide similar interpretive value. They are typically based on the analysis of existing digital and spatial data.

thematic accuracy: refers to the correctness of polygon labeling and is distinguished from but related to thematic precision. In simple terms, a polygon is correctly labeled if the attributes of the polygon fall within the defined attribute ranges of the map unit and its components. There is generally an inverse relationship between thematic precision and thematic accuracy in mapping projects.

topological accuracy: topology refers to the properties of points, lines, or polygons not affected by changes in size, shape or absolute position. For example, if a point is not located at the correct coordinates but it is located within the correct polygon it is topologically correct with reference to the polygon. Topological attributes are always with reference to two or more features. Of particular concern are the topological attributes of inside, outside, left, right, contiguous, congruent, and connected. Topological accuracy is normally much greater than positional accuracy and because maps are frequently used as a means of locating oneself in the field by reference to topological relationships, positional inaccuracies often go unnoticed.

Appendix 1: Meta-data for new inventories in support of a PEM project

The intent of the meta-data presented below is to document new, "non-RIC" PEM input data, collected for a PEM project, in sufficient detail for other qualified PEM practitioners to evaluate the quality of the data and products. Most of the meta-data listed below requires reference to a definition, procedure, *etc*. Where the reference is in published form, the reference is sufficient. Where the reference is not in published form, the standard requires appended documentation of the definition, procedure, etc.

All the meta-data identified below are required for new inventories in support of a PEM, if RIC standards do not exist for the inventory type. If meta-data is inapplicable to the input data, indicate as such. Additional meta-data may be useful depending upon the type of input data; this may be added at the discretion of the PEM practitioner.

Project

Project meta-data provides project information on the source of the input data and the organization(s) responsible for the project and data.

Citation *– specify a reference to the source of the data.*

Consultant/Department *– specify the public or private sector organization(s) responsible for collecting, compiling, and maintaining the data and an appropriate contact within the organization(s).*

Compilation scale *– specify the scale at which the data were compiled.*

Period of conten*t – specify the date or dates of data collection.*

Period of compilation *– specify the date range during which the data were compiled.*

Base Map

The base map for new inventories will generally be TRIM or TRIM2 maps. If so, reference. If not, *provide the following meta-data.*

Compiling agency *– specify the agency or organization responsible for the compilation*

Year(s) of compilation *– specify the year of compilation.*

Projection *– specify the projection used.*

Ellipsoid *– specify the ellipsoid used.*

Compilation method *– specify the compilation method, (e.g., ortho-photo).*

Datum *– specify the DATUM, (e.g., NAD 83).*

Mapping Concepts

Mapping entities *– describe or reference available definitions of all mapping entities used as input data and all mapping entities used in the PEM process.*

Map entities *– describe or reference available definitions of the map entities (map symbols) used as input data.*

***Entity relationships** – identify how the important input entities relate to each other, particularly those that are pertinent in the knowledge base.*

Inventory Procedures

Data Capture

***Delineation method and criteria** – specify by reference to a published or appended document, the rationale and criteria used to delineate polygon boundaries. The BCFS, VRI Photo Interpretation Procedures Version 2.2 March 1999 (http://www.for.gov.bc.ca/RESINV/standard/fri_man.htm) provides an excellent example.*

***Sampling design** – specify, by reference to published or appended documentation, the sampling design and the location of all sample points and traverses. In the case of selective or modal sampling, document the selection criteria and in the case of stratified sampling document the stratification criteria.*

***Sampling methods** – specify by reference to published or appended documentation, the sampling method. Examples may be, for vegetation, full relevé versus dominant species and, for soils, interval versus horizon sampling.*

***Sampling frequency** – specify, using Table A-1, the type and frequency of sampling.*

Table A-1: Sampling frequency

Method		Frequency
Ground call		
	Full site	
	Short form	
	Class identification	
Air call		
Unknown		

***Attribution** – specify by reference to a published or appended document the method and criteria for assigning attributes to delineated polygons. The BC Forest Service, VRI Photo Interpretation Procedures Version 2.2 March 1999 (http://www.for.gov.bc.ca/RESINV/standard/fri_man.htm) provides an excellent example.*

Quality Assurance

***Validation method** – specify by reference to a published or appended document the method used to validate inferred entities and their boundaries. Possible methods of validation are field traverses, point inspections, aerial inspection, road traverse,* etc.

***Validation criteria** – specify by reference to a published or appended document, the criteria used for verification (e.g., attributes and/or classes verified, allowable error,* etc.*).*

***Validation design –** specify by reference to a published or appended document, the validation sampling design (e.g., selective, modal, stratified random* etc.*) and the criteria for selection or stratification if not random or systematic. In addition, document the locations of sample points and traverses by reference to accessible or appended photographs or maps.*

***Validation results** – report the results of the validation and any changes to the map resulting from the validation results.*

Quality Control

Correlation procedures *–specify by reference to published or appended documentation, the correlation procedures and standards used.*

- *Taxonomy – the taxonomy/classification and version being correlated, the criteria and procedure followed, and the correlation results*
- *Attributes – the attributes correlated, the correlation criteria and procedure followed, and the results*

Map production – *specify by reference to published or appended documentation, the quality control procedures and standards used for the following map-production tasks.*

- *Edge matching*
- *Line edit*
- *Symbol edit*
- *Attribute edit*
- *Legend edit*