Specifications for a Predictive Ecosystem Mapping Standard

1. INTRODUCTION

1.1. OBJECTIVE

The objectives of the PEM Standard are:

1. To ensure that PEM data are delivered in a standard TEM-like form that allows for merging, integration, or comparison of multiple data sets.

2. To ensure sufficient documentation (QC/QA) to evaluate the accuracy of the product.

3. To provide stability in the standard. To achieve stability, changes to the standard will be subject to a change management procedure that incorporates an impact analysis of proposed changes.

1.2. BACKGROUND

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

1. Towards the Establishment of Predictive Ecosystem Mapping Standards: A White Paper, Keith Jones, R. Keith Jones & Associates, Del Meidinger, BC Ministry of Forests, Research Branch, Dave Clark, BC Ministry of Environment Lands & 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., and

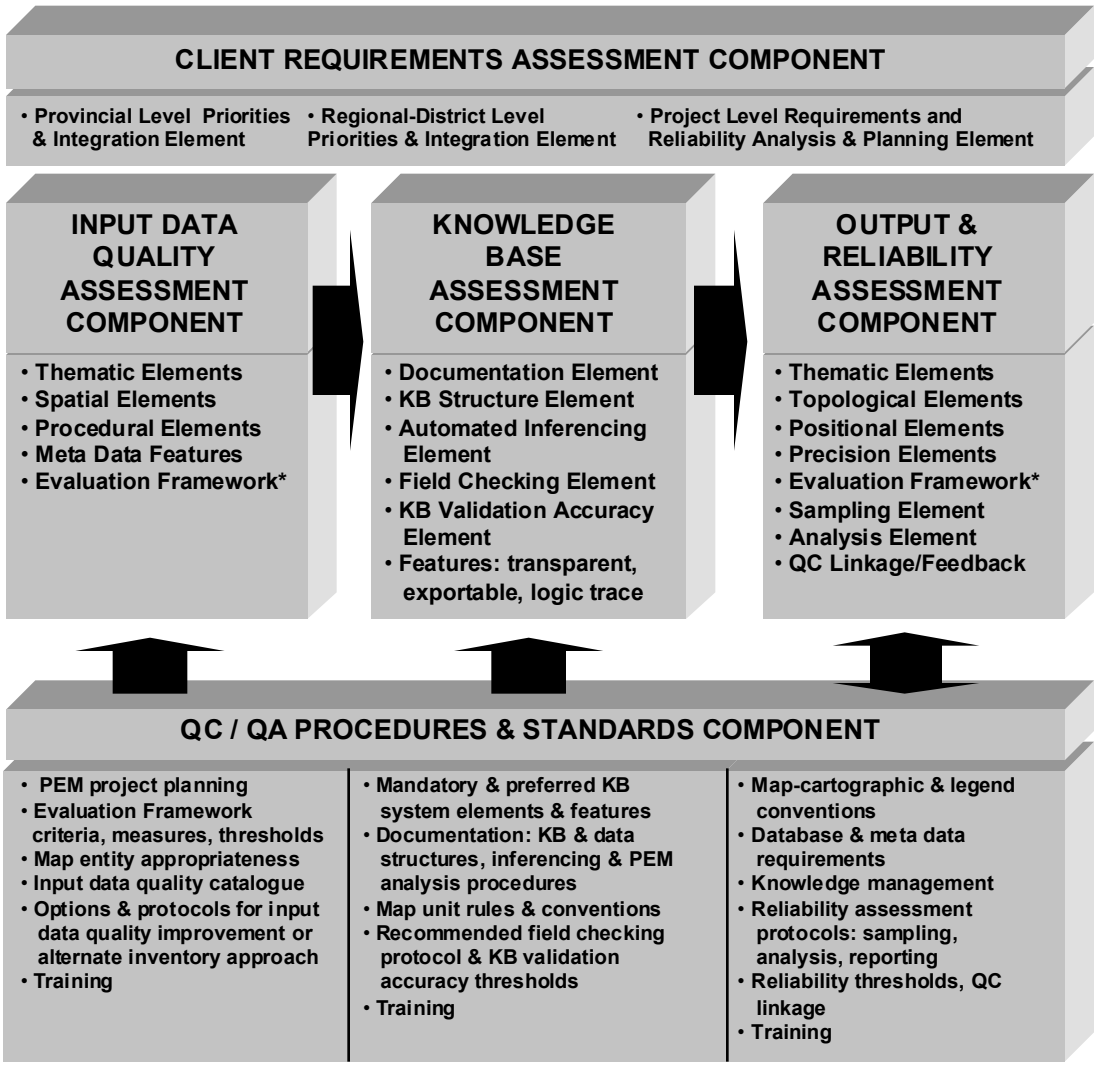
4. Problem Analysis on Reliability, Quality Control and Validation of Predictive Ecosystem Mapping (PEM) by Dr. Richard Sims, and Jeff Matheson, R. A. Sims & Associates.

The white paper integrated elements of the supporting studies to develop the overarching 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, and
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 closely 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 can be the same for both PEM and conventional TEM. Presently this component has not been developed fully for TEM.

The PEM standard will address components 2 – 5 and will provide both standards and recommendations 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.

2. ISSUES IDENTIFIED

A two-day workshop (July 6–7, 1999) identified a number of important PEM related issues. These issues related to required infrastructure (I), contract specifications (C), standards (S) and guidelines (G). While all are important and are identified in the table below, the standard will address only those issues related to standards and guidelines. Table 1 compiles the results of a workshop session that identified issues related to PEM standards and identified functions addressing the identified issues. The issues have been grouped into the following categories: Mapping Entities, Information Management, Procedures, Knowledge Management, and Important issues beyond the scope of the standard. Each issue is flagged with I, C, S, or G to indicate Infrastructure, Contract, Standard, or Guideline.

• Table 1 Issues and Functions Relevant to PEM Standards

Issue	Function	
Mapping Entities		
Product definition	Product definition provides documentation on the nature of the PEM being produced and should include:	S
	- Definition of the PEM or TEM classes used	
	- Definition of the attributes used to define or characterize them	
	- Rationale underlying choice of mapping and map entities	
	- Relation to TEM classes	
	"domain values" (explain)	
Non-forested units	Establishment of standard classes (other classifications) for non-forested unit nomenclature	S
Site-series generalization	Creation of more generalized TEM classes to accommodate areas recognizable as belonging to a group of more detailed classes but not recognizable as to specific class.	S
Information management		
Information delivery	Specification of the form in which the data are delivered to the BC corporate data base (logical data model)	S
	Enhancement of the database to capture user-defined attributes	
Versioning	Identification, labeling, tracking, reconciliation, distribution, documentation	S
Change management	Establishment and implementation of change management procedures	S

	-	Screening of proposed changes	
	-	impact assessment	
	-	notification of changes including what changed	
	-	grandfathering projects and contracts	
	-	timing for upcoming field season, project in progress	
Project definition	Establishment of standards to define the geographic extent, resource discipline supported, strata used (TEM analogue), classes and attributes et cetera for the project.		S
	Develop and maintain a project management system to track ongoing and completed projects		
Attributes	Documentation (rationale, definition, domain)		S
	-	Core attributes	
	-	Interpretation specific attributes	
	-	User defined attributes	
	-	PEMplate of core attributes	
	-	See TEM table 4.1 and 4.2	
Data independence	Programs, validations, interpretations et cetera should be independent of hard coded parameters, parameter files, or other non-data base structures requiring specific non-domain expertise for maintenance.		
Procedures			
Accuracy Assessment	Specification of minimum requirements and characteristics of procedures to assess accuracy (procedure neutral)		S
	-	Thematic	
	-	Spatial	
Biogeoclimatic boundaries	Specification of procedures to refine small scale lines for application at larger scales		S
Cartographic output	Specification of minimum requirements and characteristics		S
	-	Guideline on applicability of TEM Standard	
	-	minimum set (e.g. polygon id, schema for attribution)	
	Creation of standard requirements for edge matching		
	-	When appropriate / mandatory	
	-	Within/Between projects PEM<-->TEM and PEM<-->PEM	
Software independence	Development of specifications for exchange formats or standards		S
Knowledge base	Define minimum content and elements for knowledge base documentation		S
	Define entities and relationships		
Input data	Specification of thematic data requirements		S
	-	Mandatory themes (e.g. TRIM, BEC lines)	
		Mandatory IDQ report as early contract deliverable?	
	-	Appropriate thematic input	
	-	Rationale for non-mandatory thematic selection	
Input data quality	-	Create an input database catalogue	G
	-	Develop procedures to evaluate spatial congruence	
	-	Guidelines for the evaluation of thematic content and format	
Knowledge Management			
Knowledge base applicability	-	Guidelines for determining the applicability of existing KBs to new areas	G
	-	Guidelines for the specification of boundary conditions for a KB when created - project boundary or selected strata within	

Validation	Confirmation of scientific soundness or to corroborate results against a standard		G
	-	Knowledge base	
	-	Corporate validation routines	
	-	Input data	
	-	Define appropriate validation procedures	
		Document the degree of validation effort	
	-	TEM validation rules should show strong relationship with PEM knowledge base	
Important Issues beyond the scope of PEM Standards			
Custodianship	Mandate, responsibilities (as contract specifications), roles (as standards), resources, intellectual equity		I
Responsibilities	Roles and functions (e.g., correlation, validation, certification)		I
Contracting	Templates		C
	-	Schedules A and B	
	-	RFP contents	
	Evaluation criteria for Responses to RFP		
	-	Implications on scheduling and turn around of enforced PEM standards (including limitations on government QC/QA resources, task dependencies)	
	-	Expected efforts/costs associated with normal PEM tasks as a guide to proponents and contract administrators	
	Conflict of interest with QC/QA contracts		
	Full costing of the PEM process to support benefit/cost analysis		
	Documentation of time associated with steps in PEM, including unbilled time for public service employees		
S = standard, G = guideline, I = infrastructure, C = contract			

3. PRINCIPLES

The workshop identified the following principles that the standard will attempt to follow.

3.1. CONFORMANCE TO EXISTING STANDARDS

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

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

1. A 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 TEM.

2. A 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. A digital data capture standard specifying entities, attributes, data structures and formats for the submission of digital data to the government repository.

4. A standard for interpretations and interpretive methods appropriate to the data source to which the standard applies.

5. Standards for common output products. Generally, these are specifications for standard analogue map or report products generated from the digital data.

3.2. SOFTWARE INDEPENDENCE

Wherever possible, the standard will be independent of proprietary software or proprietary software constructs. Where the ministry responsible requires data in a vendor dependent format or construct the standard will endeavor to provide standard conversion procedures or vendor independent formats which can be imported into the ministry system.

3.3. PROCEDURE INDEPENDENCE

Wherever possible the standard will be procedure independent. For example, the standard may require estimated accuracy provided as the attribute(s) being estimated, the estimated value, and statistically defined confidence limits rather than requiring that a specific procedure be followed. In other cases it may provide a minimum acceptable procedure but allow for other more sophisticated procedures that provide additional information.

4. FRAMEWORK FOR THE PEM STANDARD

In keeping with the white paper framework, the PEM standard framework will consist of the following elements.

4.1. RELATIONSHIPS BETWEEN PEM AND TEM STANDARDS

The standard will identify areas of exact correspondence, similarity with, and divergence from existing TEM and, where relevant, other RIC standards. This will include discipline standards, mapping standards, digital capture standards, and reporting standards.

4.2. MAPPING CONCEPTS

The standard will identify and specify standards for the identification, documentation, and rationalization of mapping entities and map entities. Mapping entities are the basic ecosystem element being mapped e.g., Biogeoclimatic zone, sub-zone, sub-zone phase, sub-zone variant, or site series, site series structural stage, site modifier, site series seral stage. Map entities are the labeled polygons on the map and may represent different compositions of mapping entities. The standard will also make recommendations or set standards for documenting the relationship between TEM and PEM mapping and map entities.

4.3. MAPPING CONVENTIONS

The standard will identify and make recommendations on the design and implementation of conventions for the display of PEM information in map form. This may consist of symbol, label, and line conventions. It will also make recommendations or set standards for the documentation of the relationship between PEM conventions and TEM conventions.

4.4. QUALITY CONTROL / QUALITY ASSURANCE

4.4.1. Quality Control

Quality control standards will consist of three components. The first is a rigorous, well-documented procedure that if followed will produce consistent results. The second is assurance that the procedure has been followed. Notification is required if significant deviation is contemplated. The third is the determination and monitoring of quality resulting from the applied procedure. The standard will develop documentation requirements for PEM procedures. The documentation will provide sufficient detail for the procedure to be consistently replicated and monitored. Quality control procedures should be developed for application to the determination of:

1. Input data quality. This will be applicable only if new input data are being produced for the PEM project. Where existing thematic maps are being used, quality assurance procedures will be applied to determine the quality of the input maps.

2. Preparation and compilation of the input data base.

3. The validity of the knowledge base.

4. The implementation of the knowledge base against the input data.

4.4.2. Quality Assurance

Quality assurance tests the final product to ensure that a specified standard of quality has been achieved. The test is independent of the methods used to produce the product but it may be phased to test different stages of product completion. Procedures will be developed and documented to evaluate input data quality (thematic and spatial), the knowledge base used to predict PEM units, and the implementation of the knowledge base against the input data to predict PEM units. Thematic input quality on existing thematic maps will of necessity be qualitative since field-testing will be problematic.

4.5. DOCUMENTATION STANDARDS

The PEM standard will establish minimum levels of documentation and meta-data provided to evaluate the quality for the input data, predictive procedures, and output products of PEM. In the initial version of the standards the evaluation may have to be qualitative but the documentation should be sufficient to begin refining procedures and to apply more quantitative ratings to input and output data quality.

4.5.1. Data input

The standard will identify meta-data that should be associated with input data such as forest cover, terrain, soil, and biogeoclimate. The meta-data will characterize content, quality, and format of both the thematic and spatial data.

4.5.2. Data input quality

The standard will specify meta-data relating to the determination of thematic and spatial quality and of the applicability of the input data to PEM. The standard will also reference quality assurance procedures developed for utilizing surrogate information to provide qualitative estimates of data quality in the absence of appropriate meta-data.

4.5.3. PEM process

The standard will specify the materials necessary to document the PEM process in sufficient detail for the process to be evaluated for suitability and applicability. The standard will also specify documentation of the process in sufficient detail for it to be replicated, monitored, and validated.

4.5.4. Analytical procedures

The standard will specify requirements for documentation of and the rationale for the analytical procedures used in the PEM. It will require sufficient detail to evaluate the applicability of the procedure and to replicate and validate the results.

4.5.5. Knowledge base and algorithms

The standard will specify documentation requirements for the knowledge base used in knowledge-based predictive ecosystem mapping. The documentation will include the logic, procedures and assumptions to facilitate meaningful Q/C, as well as the specification of entities, attributes, rule bases, belief coefficients, and algorithms used in the predictive process. . The documentation should also include the pedigree of the KB - the existing KBs that formed the basis of the subject KB, the nature of successive iterations, correlation procedures to which it was subject ...

5. PRIORITY ASSIGNMENTS

Teams have been created to prepare and distribute for discussion the following items by August 15, 1999:

5.1. MAPPING CONCEPTS

A one-day workshop was held to establish standards for the definition, rationale, and documentation of PEM mapping entities and map entities. Provisional mapping and map units will be proposed for discussion. Aspects of section 5.2 will be influenced by this workshop.

5.2. RELATIONSHIP TO TEM STANDARDS

5.2.1. PEM Data Architecture

The existing TEM data architecture and TEM standard for digital data capture will be reviewed for applicability to PEM applications. The specific issues of minimum core data, the ability to address non-standard data, and non-forested ecosystem data will be evaluated and recommendations submitted. (Darren McKellar *et al*., John Johnson, Carmen Cadrin, Dave Clark.)

5.2.2. PEM Mapping Conventions

The existing TEM mapping conventions will be reviewed for applicability to PEM mapping entities, map entities, and map products. If appropriate, a PEM specific standard will be created to be distributed independently or to be appended to the TEM standard. (Bob Maxwell, Ted Lea)

5.3. QUALITY CONTROL / QUALITY ASSURANCE

5.3.1. Input data quality

Required elements for documenting and evaluating the quality of thematic and spatial data inputs will be prepared and distributed for discussion. In addition, a procedure to determine the spatial quality of input data maps will be prepared and distributed for discussion. (Dave Moon, Marvin Eng, Bruce Enns, Ev Kenk)

5.3.2. Biogeoclimatic Unit scale refinement

Procedures for localizing Biogeoclimatic boundaries from smaller scale maps to the scale of the PEM will be developed, documented, and distributed for discussion. (Marvin Eng, Del Meidinger)

5.3.3. Output data quality

A procedure for determining overall PEM output quality for a map will be prepared and distributed for discussion prior to the second workshop. (Del Meidinger)

5.4. DOCUMENTATION / META-DATA REQUIREMENTS

Project definition (Darren McKellar, Dave Moon)

The PEM Process (Keith Jones)

Input Data (Dave Moon)

Knowledge base (Dave Moon)

Predictive Algorithms and Implementation (Dave Moon)

Quality Control / Quality Assurance (Dave Moon, Del Meidinger)

5.5. CHANGE MANAGEMENT

5.5.1. Versioning and version control

A procedure and standard for version management, control, release, and distribution will be prepared. (Sheila Jeck)

5.5.2. Change Management

A requirements analysis and suggested procedure to screen, assess impacts, track versions, time releases, and communicate changes to the PEM community will be prepared and presented for review. This will follow the current RIC policies for change management. (Sheila Jeck)

5.6. CREATION OF A PROTOTYPE ARCHITECTURE TO SUPPORT THE STANDARD

A prototype architecture to organize, populate, manage, and review the standard will be created for use during the second workshop. The prototype will manage the components of the standards, their relationships to each other, and identify the implications of changes to one or more of the components. (Sheila Jeck and Dave Moon)

6. APPROACH AND SCHEDULE

The development of the standards will proceed in five stages.

Stage 1 (July 7)

Stage 1 consisted of a two-day workshop to define the objectives of the standard, the issues to be addressed, and the elements of the standard needed to address the issues identified and meet the defined objectives.

Stage 2 (July 20)

The results of the workshop were synthesized into a specification for a standard that identifies the objectives, issues addressed, structures and functions that the standard must support. This report is published on the web and you are now reading it.

Stage 3 (Aug 15)

The team assignments identified in Section 5 - 0 - are submitted for posting to the web site.

Stage 4 (Sept 8)

Following review of the team assignments and their incorporation into the prototype knowledge framework, a second workshop will be held to review and approve the contents of the framework for preparation as a 1st approximation PEM standard.

Stage 5 (Oct 1)

Decisions taken at the second workshop will be incorporated and a draft standard created for web distribution to the TEM and PEM communities for comment and review. The review period will close October 15.

Stage 6 (Nov 9)

Comments and reviews received from the TEM and PEM communities will be evaluated and, where appropriate, be incorporated into the proposed standard. The standard will then be submitted to the Resources Inventory Committee for acceptance.