EcoNote 2000-3 **July 2000**

Introduction to Attribute Selection in PEM

by Colleen Jones[1] and Del Meidinger[2]

Introduction

In a previous extension note (EcoNote 2000-1), we describe what a Predictive Ecosystem Map is and generally how one is prepared. Essentially, existing digital inventory map coverages are overlain and through GIS processing, resultant polygons with associated databases are produced. Ecological relationship tables, called ***knowledge tables***, indicating the ecological association of a feature to the ecosystem types in an area, are prepared by an ecologist. In the knowledge table, each attribute is coded with a numeric weighting indicating the importance of the attribute in predicting each of the ecosystems. The polygon database and the knowledge table weightings are then processed through the knowledge engine (in our case, EcoNGen) to produce the predicted site series for each polygon.

In EcoNote 2000-2, we discussed many of the issues that need to be addressed when selecting and preparing the map overlays and dealing with complex polygon attributes in the computer environment. In this EcoNote, we'll examine the creation of the knowledge tables—selecting the attribute combinations and filling in the ecological weightings.

Selecting Digital Input Data Sources

In most cases, all digital forest resources inventory data that are available should be considered for use in the PEM. In EcoGen PEM applications, a primary polygon is selected—either forest cover or vegetation resources inventory, or bioterrain. This primary polygon is split into slope and aspect subpolygons with all attributes compiled for the resulting polygons. Although the overlay process may vary somewhat, depending upon the primary polygon, available data, and objectives of the project, a discussion of the main digital inventory sources is discussed below and summarized in Table 1.

The digital inventory data sources that are available throughout the province are:

1. **Biogeoclimatic mapping** (scale of 1:250,000 or larger). Although these data are available in a provincial coverage (http://www.for.gov.bc.ca/research/becmaps/BECMAPS.HTM), its scale is inappropriate for direct use in PEM. It can be useful though to assist in the development of larger-scale biogeoclimatic mapping (http://www.for.gov.bc.ca/research/bigbgc).

EcoNotes are produced by the Ecology and Earth Sciences Section, Research Branch, BC Ministry of Forests. EcoNotes are available online at **http://www.for.gov.bc.ca/research/ecogen/furinfo.htm**. Contact Del Meidiger, Research Ecologist at **Del.Meidiger@gems2.gov.bc.ca** for further information.

---

[1] Shamaya Consulting, Smithers, BC
[2] Research Branch, BC Ministry of Forests, Victoria, BC

2. **TRIM and associated digital elevation models.** TRIM digital data needs to be converted into a digital elevation model using the GIS software for the project or by using the gridded DEM (http://www.elp.gov.bc.ca/gdbc/griddem.htm). TRIM data are available at http://www.elp.gov.bc.ca/gdbc/trim.htm.
3. **Forest Cover (FC1) and FIP files.** Information on acquiring the maps is available at http://www.for.gov.bc.ca/resinv/products/DigData/brochure.htm and the data dictionary is located at http://www.for.gov.bc.ca/resinv/reports/Rdd/Rdd.htm.

Other common inventories available for parts of the province include:

1. **Large-scale biogeoclimatic mapping** (scale of 1:20,000–1:50,000 for selected areas). See http://www.for.gov.bc.ca/research/bigbgc for a progress map, and http://www.for.gov.bc.ca/research/temalt/bigbec.pdf for the methodology.
2. **Vegetation Resources Inventory (VRI).** This mapping is becoming more available and will replace forest cover mapping.
3. **Terrain and terrain stability mapping.** Terrain mapping to support terrain stability mapping is available from http://www.em.gov.bc.ca/Mining/Geolsurv/TerrainMapLibrary.
4. **Bioterrain mapping.** Generally only available where terrestrial ecosystem mapping is already completed. Obtainable from the custodian of terrestrial ecosystem mapping (http://www.elp.gov.bc.ca/rib/wis/tem/index.htm).
5. **TRIM II.** Progress maps are available at: http://www.elp.gov.bc.ca/gdbc/products.htm.
6. **Soils Mapping**. Digital soils maps are available at: http://www.bcsoils.gov.bc.ca.
7. **Bedrock Geology Mapping.** See http://www.em.gov.bc.ca/Mining/Geolsurv/MapPlace/Default.htm.
8. **Satellite imagery**. Images are available in several formats. Analysis of the image is required before use.

The first step is to identify which inventories are available and to identify the ones that are useable and useful.

All input data sources should be assessed for data quality, as this can be an issue with any of the inventories, even recent ones. The PEM Inventory Standard (RIC 1999) requires an assessment of input data quality. Although it is fairly straightforward to assess the spatial accuracy (see PEM Inventory Standard, 1999, “Spatial Reconciliation”), it is not as easy to assess thematic accuracy. If at all possible, the accuracy of the main input attributes should be assessed during field sampling for the PEM.

In some areas, limited inventory data availability will restrict the accuracy of the predicted site series map. In other areas, there may be redundant data sets (e.g., FC1 and VRI).

As the ecologist evaluates the inventory data sources and their compilation, he/she may decide that limited inventory sources or poor input accuracy either preclude the production of an accurate predictive ecosystem map or that he/she should only attempt to predict broad ecosystem types, rather than specific site series. General PEM maps could still offer the necessary ecological information for the scale of management planning desired.

Once all the input data are compiled and evaluated, the data are overlain in a GIS environment. As indicated earlier, in EcoGen, a primary polygon is usually selected and then subdivided into secondary polygons using slope and aspect classes. The decision of whether to split the secondary polygons further or to compile the data for these polygons then becomes an issue. Each inventory 'layer' needs to be evaluated to see if it makes sense to split based on the linework in that layer, or to compile the data for the secondary polygons. As a guide, simple attributes, ones that have a single value for their polygon, can easily be used to subdivide the secondary polygons. Complex polygons (i.e., ones with several values of an attribute within a polygon), pose some issues. These are discussed in Meidinger and Jones (2000).

The objective is to end up with the most accurate ecosystem map possible. In EcoGen, a vector analysis method is used in the GIS analysis to overlay the various inventory inputs. A raster analysis could be used, but would not be compatible with the principles used in the EcoPrep module of EcoGen. In EcoPrep, the influence of various topographic features is assessed using cutoffs and buffers (see EcoNote 2000-1, Meidinger *et al.* 2000). This is more amenable to a vector approach to analysis.

Table 1. Forest resource digital data in PEM

Digital inventory data source	**Useful PEM attributes**
1. Biogeoclimatic mapping (1:250,000)	Not appropriate for PEM but can provide guidance for large-scale biogeoclimatic mapping.
2. TRIM digital elevation model	Slope, aspect and elevation classes; and TRIM features programming (e.g., ridges, gullies, stream density, riparian benches).
3. Forest cover	Species or inventory type group; age and height class, basic class or non-forest descriptor.
4. Large-scale biogeoclimatic mapping	Necessary for PEM. If not available, need to create for project.
5. Vegetation resources inventory	In addition to forest cover attributes, could have soil moisture regime and/or soil nutrient regime, land cover class, and meso slope position.
6. Terrain	Surficial material and surface expression, terrain texture, soil drainage.
7. Bioterrain	Surficial material and surface expression, terrain texture, soil drainage.
8. TRIM II	Use if available for improved hydrography and DEM.
9. Soils	Will not usually be of use due to complex map units.
10. Bedrock geology	Surface rock types for mineralogy of parent materials.
11. Satellite imagery	Potential to be useful for non-forested classes such as rock and broad wetland or other vegetation types. May also assist in modeling structural stage.

Identifying Useful Attributes

Once the digital inventory inputs are available, the next step is to identify which data attributes are useable and useful. Using the forest cover inventory as an example, each polygon has a large number of associated attributes (stored in the FIP files). Beyond the label displayed on the map, the FIP files contain data on species composition down to a sixth rank (roughly 2% species composition of the stand), the species percent, environmentally sensitive area descriptors, projected age and height, mean stand diameter, volume adjustment factors, and so on. Each of these pieces of data is an attribute assigned to that polygon and available for use. Once the overlay is done, the attributes from each of the digital inventories are available.

When considering which inventory inputs to use, identify the attributes that have the best correlation with each ecosystem unit to be mapped. Make a list of the attributes you think are useable. Next consider how they relate to each other within the polygon, the goal is to have one value, rather than multiple values, for each attribute type in the polygon. Where an attribute is mapped as a complex of values, it can pose challenges to the use of the attribute in a computer environment (see EcoNote #2000-2, Meidinger and Jones 2000).

Determining the usefulness of all of the attributes is the challenge. At first, consider how the polygon was created. In forest cover mapping, an airphoto interpreter delineates a stand of like composition and characteristics. Some of the attributes describe what that interpreter saw and measured or estimated directly from the airphoto. These attributes include species, percent composition, age, height, crown closure, and ESAs. Other attributes are calculations derived from the seen or measured characteristics, such as site index, projected growth, and volume indices. Any of the attributes can be used, but avoid using calculated attributes if the map is going to be used for interpretations related to the calculated values. For example, do not use site index to predict site series which will then be used to assign site index. This becomes circular logic, using derived information to determine an answer which will be used to derive the same information.

Next, consider the reliability of the attributes on your list. Some attributes are better generated by the computer than by human interpretation—namely slope, aspect and elevation classes—due to an increase in accuracy and consistency. Where possible, use the computer-derived attributes.

Even though many attributes in digital databases were assigned by observation or measurement from airphotos, there are some interpretation errors associated with them. Using, for example, forest cover mapping, many users would suggest that it can only be relied upon in a broad sense—that individual attributes are not often highly accurate, but that, in general, the map attributes are reasonably accurate. Therefore, use attributes that the interpreter was most likely to measure or describe correctly. For example, some tree species are more characteristic in their shape, colour or texture than others. Stands dominated by less common species are likely to be picked out by the interpreter. Trees that when mature are taller or shorter than the norm are more likely to be interpreted correctly. Species composition should be easier to identify in open stands. Non-productive stands or brush patches contrast strongly with surrounding forest and will easily be delineated. The unusual forest cover characteristics stand out on the airphotos and, as such, interpreters are more likely to be correct in these descriptions. Coincidentally, the unusual forest cover characteristics are most often also indicators of less common ecosystem types.

From your list of attributes, identify the range of values of each attribute that you think would be useful. For example, for heights, consider which ecosystem types are associated with tree heights taller or shorter (at maturity) than the average canopy. Then identify the height classes that represent these unique ranges. We generally use the classes for age, height, or crown closure, rather than actual values, since the class ranges are usually adequate for use in a knowledge table. Using crown closure as an example, it may be that only *open* crown closure is worth using as it is a characteristic of certain ecosystems and, therefore, is a reliable feature to interpret.

Your list of attributes should, by now, have been narrowed down. Beginning with a list of all possible attributes, the list was reduced to those that are useful for identifying an ecosystem type, then reduced again to those that are reliable for usage, and finally reduced again by the range of values to each attribute that are indicative of an ecosystem type.

Knowledge Table Formats

Attributes, and their range of values to be rated, are now entered into a set of columns in the knowledge table (see Table 2). Attributes can be entered alone, or in combination with other attributes. When combining attributes, there may be a tendency to combine values into long strings (e.g., “if *this*+*this*+*this*+*this*, then the answer is *that*”). This is referred to as a unique statement approach. It is possible to do this, but such “rule sets” require that every possible combination of attributes, and their values, that could be present in the polygon overlay database be considered. If a combination of attributes is missed, the model will not be able to assign an ecological label to the polygon. The more attributes there are to deal with, the more unwieldy this approach can become.

The alternative, and the approach that we recommend, is to combine, at most, a few related attributes. In that way, the knowledge table can easily be modified or new attributes added, and each polygon will be assigned an ecological label by the summation of attribute/value weightings.

The cumulative tally of the weightings assigned to each statement leads to the ranking of the ecological units rated as the first most likely, second most likely, third most likely, and so on. Although this approach allows the ecologist to create simpler knowledge tables, it also requires repeated tests combining multiple statements, as may be found in a variety of polygons, in order to see the cumulative results. There may be cases when a combination of statements leads to an unintended answer, and as such, the ecologist may have to change the weightings in the table to avoid this outcome. Refining the combinations of attributes will help create better knowledge base attribute sets and prevent duplication or double-counting of the weightings which might skew the results.

This approach allows the computer to “weigh” all of the attributes in each polygon and assign an ecological label, regardless of whether there are attributes that should never occur together. Inevitably, the inaccuracies of the inventories compounded by the overlay process will result in attributes that should never be found together in the real world. This is one of the main difficulties of the unique statement approach.

Creating Attributes Statements

The attributes and attribute combinations can be termed 'attribute statements.' In creating the statements, the first and most important ones are those that apply to all polygons and attempt to use the core ecological information in the attribute database. These base attributes must be useful in identifying the general position on the edatopic grid even if no other information exists for a polygon. In the EcoGen program, we use the "water-slope-aspect" statement as the base information. These attributes are determined using the digital elevation model from the TRIM base map, and as such, every polygon has this information. Generally speaking, this base information places the ecological label towards the drier positions if there is no water and/or it is steeper, towards the wetter positions if there is a lot of water and/or it is flatter, and wavers in the middle of the edatopic grid for the remainder of the cases.

The "water-slope-aspect" statements provide some indication of the relative moisture potential of the polygon. The slope and aspect attributes combine to indicate how much water shedding or pooling is likely occurring, and how much sun drying or shading is likely occurring. "Water" is determined by stream density to further indicate how much water is actually in the polygon. This combination of attributes allows us to recognize that though there may be three streams running through the polygon, a hyper-steep south-facing profile indicates it is still likely a drier ecosystem unit. We have recognized, of course, that TRIM maps do not show all streams, so a flat polygon with no streams could still be wet from a high water table—we allow for this in the weightings we assign to the combinations.

The remainder of the knowledge table statements focus on identifying with greater certainty the moisture conditions of polygons and key features of the ecosystems being predicted. Are there any attributes from, for example, forest cover, TRIM, bioterrain, or Landsat imagery which can indicate moisture conditions, especially those towards the outer positions on the edatopic grid? Somewhat like a game of 20 questions, by reviewing the list of attributes that have been selected, ask yourself a set of questions. Do any attributes, on their own, indicate strongly a particular ecological unit, for example, black spruce indicating a bog or swamp ecosystem unit? Can any of the attributes be combined in twos or threes to indicate an ecosystem unit, such as older lodgepole pine that is short and has an open crown closure could indicate a dry ecosystem unit. Can any of the attributes, that would otherwise indicate very little, be combined with other attributes to indicate an ecosystem unit (e.g., NPBr [non-productive brush] on a hyper-steep slope can indicate a slide or an avalanche path). As another example, the presence of a rock polygon on the forest cover map can be used to indicate shallow soils in the adjacent forested polygons. Keep assessing different combinations of attributes until you exhaust all the logical possibilities. Working with other ecologists will likely increase the number of useful combinations.

A knowledge table is typically set up in the format in Table 2. Note that the attributes are identified by code in the first column titled "Category." The range of values for that attribute are listed in the second column titled "Value," and occur in their respective order. No spaces are entered between the items in the category or value columns. The "Notes" column is entered to explain the coding. The ecological weightings assigned to each attribute combination will be entered in the rest of the table. A complete example knowledge base is available at www.for.gov.bc.ca/research/ecogen/furinfo.htm.

Table 2. Example of knowledge table format

Category			Site series				
(attribute)	Value	Notes	01	02	04	06	10
W+S+As	2+3+1	Mod. water, mod. slope, south facing aspect					
Sp+A+H+CC	Pl+5+2+2	Pine, age-mature, short, open					
NP+S	11+5	NPBr, hyper-steep					
NP+Adj3	11+1	NPBr, adjacent to rock					

In preparing the attribute statements, using the same attribute value in different statements is ill-advised, as this could “double-count” the weight of an attribute value in a polygon. There may be times, however, when this cannot be avoided. In a game of 20 questions, this would be the equivalent of asking “is it a Colluvial deposit,” then “is it a Colluvial veneer?” The first question narrows the range of possible ecosystem units to those that are circum-mesic, the second question refines this down to a specific ecosystem unit(s)—the drier position. The second question alone would not assign any weighting for the non-veneer Colluvial deposits, yet this is useful information on its own. Intentional double-counting was created in this case. The risk in double-counting is the possibility of contradicting the result accidentally.

Grouping Attributes with Database Queries

At times, it is easier to manipulate the attribute database to group the range of values for certain attributes than to create individual attribute statements for each value. This is most useful when there are many attribute values that, as a group, have importance, but individually have no difference in the ecological weightings assigned to each member of the group. This allows fewer statements in the knowledge table, overall making the table size more manageable and understandable. For example, if mature forest was an attribute of interest, it can be created by grouping forest cover age classes 5 to 9 rather than treating each age class separately.

Such database manipulations are done as queries. In the query, a set or range of values are assigned a new value label. For example, the value “m” can be the alias for age classes 5 through 9. Table 3 illustrates the difference in the number of statements required in a knowledge table when using the grouped values for age, height and crown classes instead of the original values.

Table 3. Example of value statements equivalent to one statement using grouped values

Statements required if using original values:		
Attributes	**Values**	**Notes**
Sp+A+H+CC	PL+5+1+1	Pl, age 5, height 1, crown 1
Sp+A+H+CC	PL+5+1+2	Pl, age 5, height 1, crown 2
Sp+A+H+CC	PL+5+1+3	Pl, age 5, height 1, crown 3
Sp+A+H+CC	PL+5+2+1	Pl, age 5, height, 2, crown 1
Sp+A+H+CC	PL+5+2+2	Pl, age 5, height 2, crown 2
Sp+A+H+CC	PL+5+2+3	Pl, age 5, height 2, crown 3
Sp+A+H+CC	PL+6+1+1	Pl, age 6, height 1, crown 1
Sp+A+H+CC	PL+6+1+2	Pl, age 6, height 1, crown 2
Sp+A+H+CC	PL+6+1+3	Pl, age 6, height 1, crown 3
Etc. for age ranges 5 to 9, height class ranges 1 & 2, and crown classes 1 to 3	Continue for all possible values (30 lines)	
Equivalent statement using grouped values for age, height and crown classes:		
Sp+A+H+CC	PL+m+s+o	Pl, mature, short, open crown m = alias for age classes 5 to 9, s = alias for height classes 1&2, o = alias for crown classes 1 to 3

Any text abbreviation can be used as the alias for a group of values. The grouped values are usually suggested by the ecologist completing the knowledge table, but need to be created by a database specialist prior to running the EcoNGen module. The new values also need to be defined in the knowledge base legend.

The defined groups should not overlap each other. The ranges of the values should be kept mutually exclusive, otherwise a polygon could belong to two or more groups for a single attribute value. This is likely to cause inaccurate predictions.

It is possible, however, to use the same attribute value or value group in different attribute statements, such as using slope in the base case information statement and again in the forest characteristic statement. However, care needs to be taken to ensure that this duplication is complementary, not redundant or contradictory. For example, in some biogeoclimatic subzones, short, open pine can be a dry ecosystem or a swamp/bog ecosystem; therefore, slope is added to distinguish slope position: Pl+f (f = flat, or slope classes 0 to 1). The use of the slope class here should complement the flat slope classes used in the base case information which would likely have identified the flat slopes as most likely moist to wet, though a slight possibility for dry (see the next section on assigning ecological weightings).

Assigning Ecological Weightings to the Attribute Statements

Once the attributes and their values are determined and the basic format of the knowledge table has been created, the ecological weightings are entered. The basic coding is 0 to 3, whereby 0 equals no chance of occurrence, 1 = slight chance of occurrence, 2 = average chance of occurrence, and 3 = high chance of occurrence. For each statement in the knowledge table, the ecologist must evaluate the likelihood of each ecological unit for that biogeoclimatic unit occurring in a location with that attribute value or values. Given a statement that is designed to identify a wet site, what is the likelihood for each of the possible ecological units of occurring on the type of site? In ecology, it is much easier and more appropriate to use the "relative" concepts of slight versus high chance of occurrence, rather than specific percentages. Tables 4a and 4b illustrate how the ecological weightings are filled in.

The result for a polygon is determined by a cumulative tally method (i.e., adding up the weightings for the set of attribute statements found in the polygon data). In some cases, the basic weighting of 0 to 3 is insufficient to achieve the intended ecological outcome. In order to ensure that obvious, unintended outcomes are impossible, an extreme weighting (e.g., "-100") is used. In most knowledge tables this is used to ensure that forested polygons are not determined to be non-forested ecosystems, and *vice versa*. In Tables 4a and 4b, the non-forested ecological units are coded with -100 where the "basic class" (B) has a value of "0" for "forest."

There are cases when an attribute value strongly indicates a particular ecological unit but the weighting of 3 is not enough to end up with a certain result in the cumulative tally. For example, in the SBSmc2, black spruce is a strong indicator of a fen edge or swamp forest ecosystem unit. However, as it is not always present and because these units are uncommon in the subzone, the units may have received a low weighting for other attribute value statements to prevent them from coming out too often as the answer. When Sb is present, its strong indicator value should be indicated with a higher weighting, such as 5 or 10. The choice of 5 or 10 depends on how the ecologist views the certainty. If there is no other ecosystem unit possible, then the use of 10 would be preferable to ensure a high score. However, if there is a second possible choice, though less likely, the use of 5 is preferable to avoid too large a separation between the two. Some other attribute may switch the cumulative score from one to the other.

When testing the knowledge table against known data, other possible unintended outcomes could occur. Adjusting the knowledge table weightings can usually "correct" these problems. Small negative weightings can be useful at this time for separating out ecological units. In one knowledge table, we used a " -1" weighting in an elevation attribute value to separate two grassland units, one of which is commonly at higher elevations than the other. If all attributes were equal, we wanted one of the units to be selected as the most likely rather than a tie, and we wanted the other grassland unit to be a close second. We have also used "-5" under one of the attribute statements in a case where an unlikely ecological unit came out as the secondary answer. Though this ecological unit was not out of the question, it overtopped the "best" secondary label. We used the "-5" under one of the attribute statement–ecosystem unit cells to ensure that this ecosystem did not score in the top two.

Table 4a. Example of ecological weightings for a wetland bench site

Category	Value	Notes	Site series				Oakfern/		Devil's	Horsetail		Scrub grass		Alder/	Riparian	Slide/			
			Mesic	Sub-mesic	Dry	Sb	thimbleberry	scrub	club	flats	bog	steppe	meadow	willow	shrub	avalanche	rock	bogs/fen/marsh	
			1	1c	2	3	5/6	7	9	10	12	81	82	AW	RS	SA	Ro	31/32	Devm't
W+S+As	0+1+1	no water, flat, south	3	2	1	1	2	1	1	1	1	2	3	2	0	0	1	1	0
LB+LB_P	1+2	wetland bench	1	0	0	0	1	2	1	3	2	0	0	2	1	0	0	3	0
Sp+A+H+CC+S	PL+m+s+o+f	Pine, short, open, flat	0	0	2	2	0	2	0	0	3	0	0	0	0	0	0	0	
B	0	treed	0	0	0	0	0	0	0	0	0	-100	-100	-100	-100	-100	-100	-100	-100
Total:			4	2	3	3	3	5	2	4	6	-98	-97	-96	-99	-100	-99	-96	-100

Table 4b. Example of ecological weightings for a hill top site

Category	Value	Notes	Site series				Oakfern/		Devil's	Horsetail		Scrub grass		Alder/	Riparian	Slide/			
			Mesic	Sub-mesic	Dry	Sb	thimbleberry	scrub	club	flats	bog	steppe	meadow	willow	shrub	avalanche	rock	bogs/fen/marsh	
			1	1c	2	3	5/6	7	9	10	12	81	82	AW	RS	SA	Ro	31/32	Devm't
W+S+As	0+1+1	no water, flat, south	3	2	1	1	2	1	1	1	1	2	3	2	0	0	1	1	0
HT+HT_P	1+2	hill top	1	3	2	0	1	0	0	0	0	2	2	1	0	2	3	0	0
Sp+A+H+CC+S	PL+m+s+o+f	Pine, short, open, flat	0	0	2	2	0	2	0	0	3	0	0	0	0	0	0	0	0
B	0	treed	0	0	0	0	0	0	0	0	0	100-	100-	100-	100-	100-	100-	100-	100-
Total:			4	5	5	3	3	3	1	1	4	-96	-95	-97	-100	-98	-96	-99	-100

Note, that the only difference between these tables is the evaluation of a wetland bench versus hill top. Notice how the resultant answers change to predict unit 12 (a bog) in Table 4a to a tie between unit 1c (submesic ecosystems) and unit 2 (dry ecosystem) in Table 4b.

Preparing the Knowledge Tables for the EcoNGen

Once the knowledge tables are complete, the polygon database and the knowledge table weightings are processed through the knowledge engine, EcoNGen, to produce the predicted site series for each polygon. The EcoNGen documentation (*EcoNGen Version 1 User's Guide*), included with the EcoNGen program module, describes how to load the knowledge tables in the program. However, before running through EcoNGen, the knowledge tables need to be converted to a database format. This format requires that the spreadsheet version be modified, as follows, to ensure that:

- All attribute and value codes correspond exactly between knowledge table and polygon input data (the input database programmers have reviewed the legend and have grouped attribute values correctly and have used the same new attribute value names as in the knowledge table).
- No spaces are present between attributes or values in complex attribute statements.
- There are no blank fields in the ecological weighting fields. All fields with no weightings should be zero entries.
- All blank lines used to separate blocks of attribute statements are removed.
- All notes and their columns are deleted from the knowledge tables (save the working version as an Excel workbook, separate from the database version).
- All formatting, colour coding and cell alignment is removed.
- The BGC subzone/variant is recorded on the same line as the site series labels—this line will be kept while the rest of the descriptors will be deleted.

It is common to duplicate the working spreadsheet into a new worksheet within the project workbook. Before inputting into EcoNGen, each worksheet, with the modifications, has to be imported into an ACCESS database table. Each knowledge table, for each biogeoclimatic unit, is a separate database table.

Acknowledgements

The authors would like to thank Dave Clark for reviewing this manuscript.

Literature Cited

Meidinger, D. and C. Jones. 2000. Dealing with Complex Polygon Attributes as PEM Input. Research Branch, BC Min. For., Victoria, BC. EcoNote 2000-2.

Meidinger, D., B. Enns, A. Banner, C. Jones and S. Reed. 2000. EcoGen – A Model for Predictive Ecosystem Mapping, Research Branch, BC Min. For., Victoria, BC. EcoNote 2000-1.

TEM Alts Task Force. 1999. Standards for Predictive Ecosystem Mapping – Inventory Standard. BC Min. For., Resour. Inv. Comm., Victoria, BC.