Predictive Ecosystem Mapping Update for TFL 23

Prepared by:

Steven F. Wilson, Ph.D., R.P.Bio.
EcoLogic Research
406 Hemlock Avenue
Gabriola Island, BC
V0R 1X1
sfwilson@shaw.ca

Prepared for:

Pope & Talbot Ltd.
Arrow Lakes Timber
Nakusp, BC

28 February 2005

Executive Summary

Ecosystem mapping is becoming increasingly important in forest management planning. For Pope & Talbot Ltd., it is the primary coverage used in strategic, tactical and operational planning for mountain caribou habitat management, ungulate winter range management, and sustainable forest management related to biodiversity indicators. The current ecosystem map for TFL 23 was part of the first iteration of predictive ecosystem mapping (PEM) completed in 2001. Experience with the coverage suggested that the map was inadequate and should be updated.

I developed an updated PEM for TFL 23 using a base map comprised of "ecological land units" (ELU's) derived from slope, slope position, aspect, and moisture class. ELU's were stratified by BEC and site series were assigned to each ELU-BEC combination according to expert opinion. The resulting expert tables were used to generate a preliminary PEM.

Expert tables were then translated into Bayesian Belief Networks. The associated conditional probability tables were updated according to Bayes Theorem using data from ground plots collected on TFL 23. The "posterior" conditional probability tables were then linked to the ELU-BEC base map to generate a final PEM.

The internal accuracy of the final PEM (i.e., based on the ground plots fitted to the models) varied between 60 and 100%, depending on BEC subzone variant. This represented a considerable increase in thematic accuracy from previous coverages, even after accounting for correct classifications made by chance.

There were a number of distinct advantages to using the Bayesian approach to PEM development:

1. The relative contribution of expert knowledge and of field data is transparent;
2. Use of available field data is maximized;
3. The relative confidence of the ecologist in a site series call can be incorporated explicitly;
4. Models can easily be updated as additional field data are collected; and,
5. The "experience" and "confidence" of the model can be mapped, which can provide a visual representation of PEM reliability and can improve the efficiency of future habitat sampling.

The accuracy of the PEM should increase as additional plot data on the TFL are collected; however, the accuracy will ultimately be limited by the ability of the input data to resolve site series. If accuracy problems arise in the future, additional data layers could be added to distinguish particular site series that are difficult to resolve.

Based on internal accuracy results, the updated PEM should supersede previous PEM coverages for all management and planning purposes. If additional resources become available for field plot sampling, resources should be allocated to sampling those ELU-BEC combinations that have previously not been sampled or where data from <5 plots are available.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I wish to thank Tom Koftinoff (Pope & Talbot Ltd, Nakusp) for all GIS data preparation and analysis, and Evan McKenzie (Evan McKenzie Ecological Research, Nelson) for developing the expert models. Funding for this project was provided by the Forest Investment Account.

Introduction

Ecosystem mapping is becoming increasingly important in forest management planning. For Pope & Talbot Ltd., it is the primary coverage used in strategic, tactical and operational planning for mountain caribou habitat management, ungulate winter range management, and sustainable forest management related to biodiversity indicators (e.g., ecosystem representation in an unmanaged state, habitat supply modelling of important habitat elements). In addition, ecosystem mapping will likely be the future standard for applying site index values to stands (SIBEC) for timber supply modelling.

The current ecosystem map for TFL 23 was part of the first iteration of predictive ecosystem mapping (PEM) completed for the Arrow Forest District in 2001 (Ketcheson et al. 2001). There were several reasons to suggest that the current map was inadequate and should be updated:

- Caribou capability-suitability modelling and site investigations suggested systematic errors in the mapping
- Although the TFL 23 PEM was never assessed, the Arrow TSA PEM required an update before it met accuracy standards and both PEM coverages were developed using identical methods
- There have been considerable improvements in the technology related to predictive ecosystem mapping since the first iteration in 2001
- There have been field data collected since 2001 but the data cannot be used to improve the PEM in its current form

The purpose of this project was to develop a new ecosystem base map that would:

- Improve spatial accuracy for forest management planning
- Be based on existing methods to control costs and provide transparency
- Be simple and inexpensive to update as new data become available
- Provide spatial measures of accuracy and confidence
- Efficiently guide future investments in collection of plot data

Methods

Base Mapping

Base-mapping for the updated PEM was derived from a 20-m digital elevation model (DEM) of TFL 23, based on Terrain Resource Information Mapping (TRIM). The DEM was used to derive a map of “ecological land units” (ELU’s), based on methods developed by The Nature Conservancy (TNC 1999). ELU’s are based on models that describe slope and slope position (“landform”), aspect and moisture class. The TNC (1999) methods were applied with the following parameters and modifications:

1. The average ridge to stream distance was estimated to be 600 m (30 pixels).
2. Six slope classes were used instead of 4 to better correspond with BC’s ecosystem classification system (e.g., Braumandl and Curran 2002). Classes used were:
    a. Flat ≤5%
    b. Gentle 5-25%
    c. Moderate 25-45%
    d. Moderately steep 45-60%
    e. Steep 60-90%

f. Cliff ≥90%

3. Aspect classes were adjusted to correspond with BC's ecosystem classification system and were not applied to gentle or flat slope classes:
    a. Warm 136-280°
    b. Cool 281-135°
4. ELU's (combining landform, aspect and moisture classes) were further stratified by BEC (Biogeoclimatic Ecosystem Classification) subzone variant. Lakes and streams from TRIM and rock from forest cover were also added.

The resultant grid was assigned key values based on each pixel's combination of ELU and BEC subzone variant (Appendix).

Ecological Classification

Each ELU (stratified by BEC) was assigned a site series according to the correspondence between the unit and site series descriptions in Braumandl and Curran (2002). Expert opinion was also an important factor in assigning site series because ELU's describe only a subset of the variables that typically distinguish site series from one another. For example, soil and vegetation characteristics (in more detail than forest cover) are important to site series classification but are not included in ELU definitions (because appropriate map layers are not available).

ELU's were assigned more than one site series if the classification was uncertain. Up to 3 site series were assigned along with relative confidences.

Only BEC subzone variants for which site series had been described were assigned to ELU's. As a result, several variants that occur on TFL 23 could not be mapped to site series (ESSFwcp, ESSFvcp, ATunp, ESSFwcw and ESSFdcw).

The resulting expert tables were linked to the ELU-BEC base map via the key value field to generate a preliminary PEM.

Bayesian Network Models

The relationships between landform, aspect, moisture and site series were modelled for each BEC subzone variant (that had site series definitions and associated field plot data) as "Bayesian Belief Networks" using *Netica* 2.17 (Norsys Software Corp., Vancouver, BC; Figure 1). The expert tables from the ecological classification were used to populate the conditional probability tables of the models.

The advantage of modelling the PEM as a series of Bayesian Belief networks was that it allowed the preliminary PEM to be updated explicitly with field data. Data from ground plots collected on TFL 23 during various field projects were compiled by site series according to the same input variables as the expert tables (landform, aspect and moisture). Because ELU mapping was not developed for areas beyond TFL 23's boundaries, plot data collected in the Arrow TSA could not be used. Also, plots that were classified incorrectly by BEC subzone variant (as a result of recent BEC linework changes) were not included in the analysis.

These data were used to update the conditional probability tables of the models according to Bayes theorem:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

Bayes theorem can be summarized as: the "posterior" probability of a site series call (*A*) given the states of the input variables (*B*), is dependent on the likelihood of *A* given a specific value of *B* (P(B|*A*)), and on the "prior" probabilities of both *A* (*P*(*A*)) and *B* (*P*(*B*)).

The result was an updating of the "prior" probabilities (the expert tables that formed the basis of the preliminary PEM) to so-called "posterior" probabilities that incorporated the new information provided by the ground plots.

The posterior conditional probabilities were linked to the ELU-BEC base map via the key values field to create the final, updated PEM.

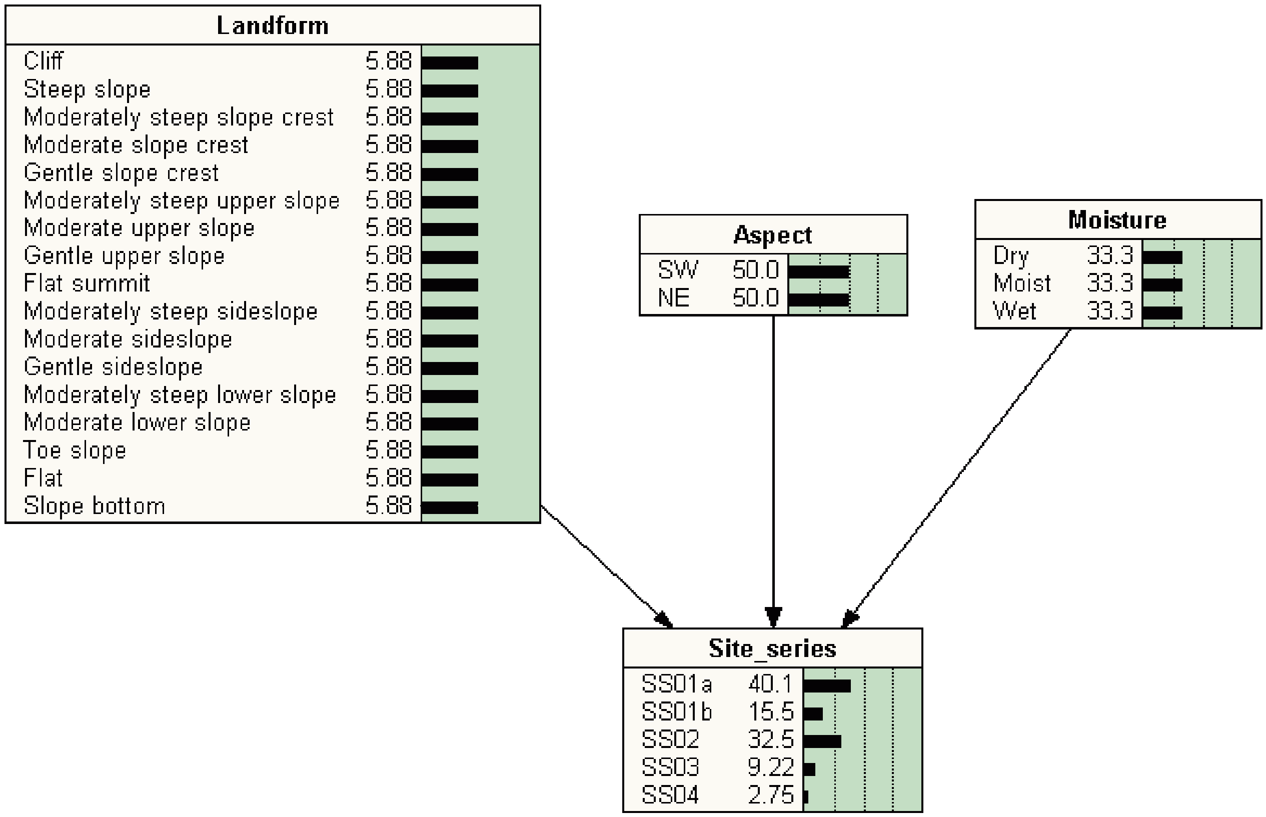

Figure 1. Example of a Bayesian Belief Network developed in *Netica* illustrating relationships between landform classes (from TNC 1999), moisture and aspect in determining preliminary site series for the ICHdw subzone variant.

Another important concept in Bayesian modelling is that of "experience." Experience is a measure of the data used to generate the post probabilities. Experience in the PEM was expressed by the sample size of plots collected at sites with the same combination of landform, aspect, moisture and BEC.

Conditional probabilities that express the likelihood of a particular size being comprised of a certain site series should be interpreted in the context of experience. For example, a site series call might be associated with a 100% conditional probability, suggesting that the call was made with high confidence; however, if the experience was low, there were few data on which to base that confidence.

Experience was also linked to the ELU-BEC base map to provide a visual interpretation of ground sampling intensity throughout the TFL.

Accuracy Assessment

Assessing the accuracy of thematic maps is a relatively common procedure in a number of disciplines and the kappa statistic and its variants is the measure used most often (Rossiter 2004). Kappa is simply a measure of mapping accuracy that controls for correct classifications that occur by chance (Foody 1992).A positive kappa suggests that classification accuracy is better than that expected by chance, while a negative kappa suggests that it is worse.

Because of the nature of PEM an important modification has to be made to the kappa calculation to accommodate polygons that are assigned probabilities for more than one site series. These polygons bias traditional accuracy assessments because there is more than one correct site series call. For example, a polygon labelled 7 AB 3 AF would be considered correct if a ground plot inside the polygon were called AB *or* AF. Therefore, the kappa calculation treats the polygon as if the polygon were 10 AB *and* 10 AF. This increases the probability of making a correct classification by chance. The bias in accuracy can be compensated for by inflating the proportions of AB and AF in the kappa calculations to 10 from their deciles of 7 and 3, respectively. The omission and commission errors usually calculated with confusion matrices (Meidinger 2003) cannot be calculated using this method, but kappa can be calculated and interpreted in the standard way. I applied this correction where more than one site series call was assigned to an ELU-BEC combination.

Results

Base Mapping

The base ELU-BEC map is available digitally as an Arc grid called ELU_BEC.

Ecological Classification

The results of the expert ecological classification are available digitally in PEM_site_series_lookup.xls.

Bayesian Network Models

Bayesian belief networks for ESSFwc1, ESSFwc4, ICHdw, ICHmw2, ICHvk1, ICHwk1 and IDFun are available digitally in *Netica* (.dne) format. Models reflecting the expert tables (prior probabilities) as well as posterior conditional probabilities (suffixed with "_learned") are provided.

Data collected from 442 plots located on the TFL were used to update the expert models to generate posterior conditional probabilities. Data were available for all BEC subzone variants with site series definitions except ATun.

The posterior conditional probabilities from the Bayesian models are summarized in a lookup table for linking to the base map. The file is available digitally as PEM_site_series_lookup_learned.xls. The final PEM coverage is available digitally as an Arc polygon coverage called TFL_PEM.

Accuracy Assessment

Updating conditional probabilities with data from ground plots significantly improved the accuracy of the PEM over the preliminary PEM based solely on the expert tables (Table 1). Full accuracy statistics are provided digitally as PEM_accuracy.xls. Kappa values ranged between -0.04-0.49 for the preliminary PEM and 0.50-1.00 for the final PEM incorporating information from field data. No independent data were used to test the accuracy of the PEM.

Table 1. Results of the accuracy assessment of the updated PEM based on plot data collected in the field. The proportion of plots expected to be correctly classified by chance, P(chance), and the proportion correctly classified P(correct) are reported for the preliminary PEM (which is comprised of the Bayesian prior probabilities) and for the final PEM updated with field data (which is comprised of the Bayesian posterior probabilities). Also reported is kappa, a statistic that adjusts P(correct) according to the probability of chance classifications (Foody 1992 as modified by Wilson 2004).

BEC Subzone Variant	Ground Plots (n)	Preliminary PEM (Prior Probabilities)			Final PEM Updated with Field Data (Posterior Probabilities)		
		P(chance)	P(correct)	Kappa	P(chance)	P(correct)	Kappa
ESSFwc1	45	0.30	0.27	-0.04	0.42	0.93	0.88
ESSFwc4	50	0.14	0.20	0.07	0.20	0.60	0.50
ICHdw	75	0.42	0.71	0.49	0.38	1.00	1.00
ICHmw2	198	0.23	0.56	0.42	0.25	0.87	0.83

ICHvk1	29	0.24	0.41	0.23	0.26	1.00	1.00
ICHwk1	25	0.19	0.52	0.41	0.21	1.00	1.00
IDFun	20	0.24	0.55	0.41	0.28	1.00	1.00
Total/Overall	442	0.25	0.46	0.29	0.25	0.92	0.89

Discussion

Base Mapping and Ecological Classification

The ELU-BEC map is a static base map that can be easily updated in the future as additional plot data are collected and Bayesian models are updated. Only if the BEC linework changes or if more accurate DEM information becomes available will the map need to be revised. Similarly, the expert tables that define the initial correspondence between ELU's and site series need only be completed once. Additional tables will need to be prepared only if additional site series are classified or if an additional variable is added to the base map (e.g., presence/absence of an indicator tree species from forest cover).

Bayesian Network Models

There are a number of inherent advantages to the Bayesian approach to PEM development:

1. The relative contribution of expert knowledge and of field data is transparent;
2. Use of available field data is maximized;
3. The relative confidence of the ecologist in a site series call can be incorporated explicitly;
4. Models can easily be updated as additional field data are collected; and,
5. The "experience" and "confidence" of the model can be mapped, which can provide a visual representation of PEM reliability and can improve the efficiency of future habitat sampling.

Although Bayesian modelling is an effective approach to generating a model-based PEM, it is still inherently limited by the ability of input data to predict site series. The final PEM for this project was developed using slope, slope position, relative moisture and aspect as input variables. All of these inputs were ultimately derived from 1:20,000 contours in TRIM. As a result, the final PEM is limited by the accuracy and resolution of TRIM. In addition, there are many other variables that contribute to the classification of site series that are beyond our current ability to map on a landscape scale; most importantly, soil and detailed vegetation characteristics.

Although the accuracy of the PEM can be improved by adding additional input variables to the models, doing so must be balanced against the cost of development. Adding just one variable with 2 states doubles the size of conditional probability tables and can significantly increase the expert time required to classify site series.

Accuracy Assessment

The accuracy of the preliminary PEM based only on the expert tables was similar to that for other PEM's before fitting to plot data. Updating the conditional probabilities with data from ground plots dramatically increased the accuracy of the PEM. This was not unexpected because the initial model was fitted to the plot data and then the accuracy tested with the same data. An accuracy assessment based on independent plot data would yield more moderate accuracy results.

The reported accuracy was also a reflection of that fact that up to 3 site series could be identified for each polygon and that the site series call was considered correct if any of the 3 site series were classified correctly by a ground plot located within that polygon. This inflated accuracy was controlled by interpreting kappa statistics, which accounted for the probability of making correct classifications by chance. However, for planning and management purposes it is important to consider all site series calls (and their related probabilities) for polygons of interest.

The accuracy of the PEM should increase as additional plot data on the TFL are collected; however, the accuracy will ultimately be limited by the ability of the input data to resolve site series. If accuracy problems arise in the future, additional data layers could be added to distinguish particular site series that are difficult to resolve.

Management Recommendations

1. Based on internal accuracy results, the updated PEM should supersede previous PEM coverages for all management and planning purposes.
2. If additional resources become available for field plot sampling, resources should be allocated to sampling those ELU-BEC combinations that have previously not been sampled or where data from <5 plots are available.
3. If further ground reconnaissance suggests that significant errors in the PEM still exist, additional input variables such as forest cover could be added to the models.

Literature Cited

Braumandl, T. F., and M. P. Curran. 2002. A field guide for site identification and interpretation for the Nelson Forest Region. BC Ministry of Forests Land Management Handbook Number 20.

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

Ketcheson, M. V., T. Dool, and S. F. Wilson. 2001. TFL 23 predictive ecosystem mapping final report and maps. Prepared for: Pope and Talbot Limited, Nakusp, BC.

Meidinger, D. In press. Protocol for accuracy assessment of ecosystem maps. Research Branch, BC Ministry of Forests, Victoria.

Rossiter, D. G. 2004. Technical note: statistical methods for accuracy assessment of classified thematic maps. *Unpublished*.

TNC. 1999. Analysis of ecological land units. Eastern Conservation Science – GIS, Boston, MA. *Unpublished*.

Wilson, S. F. 2004. Suggestions to improve predictive ecosystem mapping and comments on the reliability of the Invermere TSA PEM. Prepared for: BC Ministry of Sustainable Resource Management, Nelson.

Appendix

Table 2. Key values used to classify the grid resulting from the development of ecological land units (ELU's). Key values are summed to give a unique identifier to each ELU-BEC combination. These values are linked to the posterior conditional probability tables to produce the PEM.

Variable	Class	Key Value
Landform	Cliff	10
	Steep slope	11
	Moderately steep slope crest	12
	Moderate slope crest	13
	Gentle slope crest	14
	Moderately steep upper slope	15
	Moderate upper slope	16
	Gentle upper slope	17
	Flat summit	18
	Moderately steep sideslope	19
	Moderate sideslope	20
	Gentle sideslope	21
	Moderately steep lower slope	22
	Moderate lower slope	23
	Toe slope	24
	Flat	25
	Slope bottom	26
Other	Lake	50
	Stream	60
	Rock	70
Moisture	Dry	100
	Moist	200
	Wet	300
Aspect	Warm	1000
	Cool	2000
BEC subzone variant	AT un	10000
	ESSFwcp	20000
	ESSFwc 4	30000
	ESSFwc 1	40000
	ICH vk 1	50000
	ESSFvc	60000
	ESSFvcp	70000
	ICH wk 1	80000
	ICH mw 3	90000
	AT unp	100000
	ICH mw 2	200000
	ESSFwcw	300000
	ICH dw 1	400000
	ESSFdc 1	500000
	IDF un	600000
	ESSFdcw	700000