TFL 23

PREDICTIVE ECOSYSTEM MAPPING (PEM)

YEAR 2 FINAL PROJECT REPORT

A Report and Maps

For:

Frances Swan
Pope and Talbot Ltd.
Arrow Lakes Timber Division
P.O. Box 2000
Nakusp, BC V0G 1R0
(250) 265-3741

By:

Maureen V. Ketcheson, M.Sc. R.P. Bio.
Tom Dool B.ES.
JMJ Holdings Inc.
208-507 Baker Street
Nelson, BC V1L 4J2
(250) 354-4913
jmj@netidea.com

Steve Wilson, M.Sc R.P.Bio.
EcoLogic Research
406 Hemlock Avenue
General Delivery
Gabriola Island, BC V0R 1X0
(250) 247-7435

March 31, 2001

POPE & TALBOT
Arrow Lakes Timber Division
T.F.L. 23

FOREST RENEWAL BC

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ACKNOWLEDGEMENTS

This project was made possible through funding by Forest Renewal British Columbia (FRBC) in partnership with Pope and Talbot Ltd., Arrow Lakes Division and cooperation with the Arrow Innovative Forest Practices Agreement (IFPA).

We would like to thank the following people who assisted in the completion of this project: Francis Swan of Pope and Talbot Ltd., Arrow Lakes Division, for coordinating the project; Tom Dool of JMJ Holdings Inc. for GIS analysis; Steve Wilson of Ecological Research for statistical analysis; Del Meidinger of MOF Research Branch and Dave Clark of MOELP Inventory Section for quality assurance; Tom Braumandl, MOF Nelson Region for review; all the JMJ employees who collected data and helped put this project together, including Gareth Kernaghan, Keyes Lessard, Grant Burns, Donna Ross, Rayanne McKay, Bruce Sinclair, Vicky Lipinski, Mary Searchfield, and Catherine Littlewood; plus Neil Munro and Doug Seaton of Nanuq Consulting for assisting with field work

1.0 INTRODUCTION

Predictive Ecosystem Mapping (PEM) was initiated in Tree Farm Licence 23 (TFL 23) in conjunction with the ARROW IFPA Site Index Adjustment Project in November 1999 (Ketcheson et al 2000, 2001). The mapping model developed for the Arrow project was also applicable to the Biogeoclimatic subzones that occur in TFL 23. Pope and Talbot Ltd. Arrow Lakes Division engaged an ecological mapping consultant, ***JMJ Holdings Inc***. to use the PEM mapping model developed for the Arrow IFPA in TFL 23 to generate site series maps for that area. The process took two years; in year one the mapping model was adjusted to the landscapes of TFL 23 and run with a few VRI plots used as an assessment of the accuracy of the model. In year two a field effort combined model development plots with randomly located accuracy assessment plots, as well as plots from Caribou habitat priority areas. This report documents the methodology used to produce those maps and site series summaries. The TFL 23 PEM is one of many developmental projects investigating the utility of computer mapping models as an alternative to Terrestrial Ecosystem Mapping (TEM Alternatives website).

This PEM model relies heavily on landscape shape to allocate site series. Landscape shape is a product of bedrock and surficial geological processes within the study area. Superimposed on those landscape shapes are the biogeolimatic subzones and variants mapped for the area by MOF (Braumandl and Curran, 1992). Within each variant, ecosystems, called site series, are identified based on landscape shape, slope position, exposure and a number of other criteria. Using a raster-based approach the landscape of TFL 23 was allocated to site series.

Wildlife species habitat capability can be related to ecological classification (RIC Wildlife standards 1999). The site series classification is used as the basis to assign a seasonal habitat capability rating to four species in TFL 23, they were Woodland Caribou, Mountain Goat, Grizzly Bear and Northern Goshawk.

PEM is in its early developmental stages and these models could benefit from continuous refinement and improvement. The TFL 23 PEM's accuracy varies with the site series. The model is better in more widespread site series and variants and poorer in the areas with less data. It should be used with a level a caution and in a manner suitable for the interpretations being derived from its results.

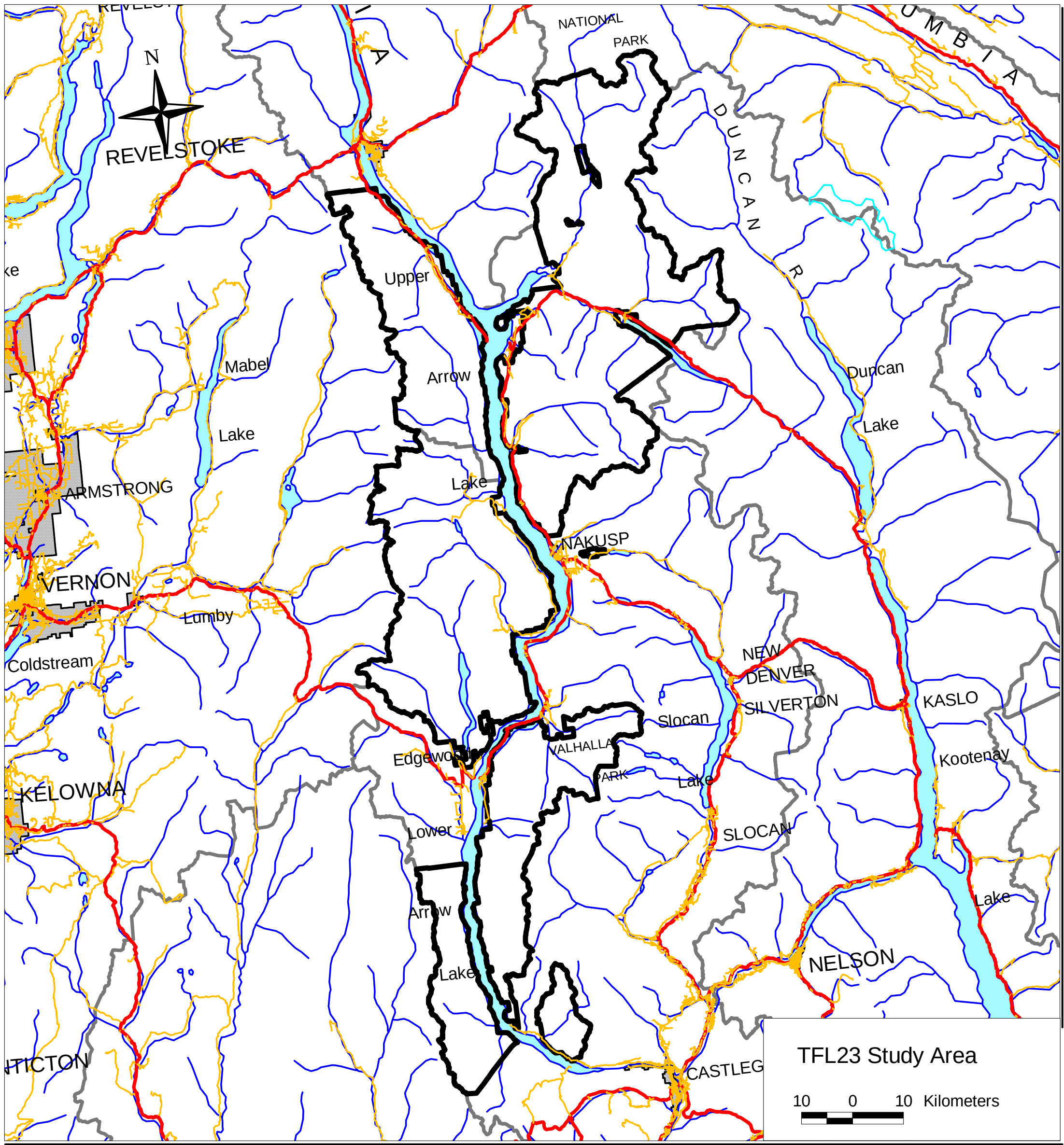

Figure 1. TFL 23 Study Area

1.1 GEOLOGY, SURFICIAL DEPOSITS, and SOILS

Tree Farm Licence 23 is a huge and diverse landscape, which spans much of both sides of Upper and Lower Arrow Lakes. The area of TFL 23 to the west of the Arrow Lakes lies within the Monashee Mountains, and the area to the east of the lake lies within the Selkirk Mountains (Figure 1).

A complex geology is found in TFL 23, with bedrock ranging from plutonic to marine origins. The region surrounding the Upper Arrow Lake is more complex with numerous thrust and extension faults and bedrock types. In this area the east side of the lake is dominated by quartz monzonite (granite) Middle Jurassic in age, and resistant dolomite Cambrian-Devonian in age. The dolomite is part of the Rocky Mountains passive continental margin sediments. Remnants of oceanic marginal basin volcanics and sediments (basalt, tuff, and breccia) occur in a few localized areas, with the majority along the east shore of the lake. On the west side of the Upper Arrow Lake metamorphized granite (gneiss), thought to be early Proterozoic in age, dominates. Palezoic metamorphized sedimentary rocks (phyllite, siltstone, and sandstone) are next in dominance. Inclusions of granite group rock types (granodiorite, quartz monzonite, and quartz diorite) of late Mesozoic to early Tertiary in age are also common (Wheeler, 1991).

The geology surrounding the Lower Arrow Lake is less complex. In this area the east side of the lake is comprised mainly of quartz mozonite and calc-akalic (feldspar rich) syenite of early Tertiary in age. Syenite has less resistant quartz and an abundance of less resistant pink feldspar allowing it to erode more easily than granite. On the west side of the Lower Arrow Lake similar geology can be found. The exception is the portion of TFL 23 south of Deer Park. Here minor amounts of metamorphized marine sediments limestone, sandstone, and shale in origin can be found (Wheeler, 1991).

During the Pleistocene Epoch (2,000,000 to 10,000 years before present (BP)), this area was subjected to multiple episodes of glaciation. Most of the landscape features visible today are the result of the most recent (Fraser) glaciation and the subsequent alpine glaciations. Where glaciers did not cover the highest mountain peaks and ridges, "horns and cirques are common features and ridges have sharp crests formed by the weathering of the jointed granitic blocks" (Little, 1960). Since the end of the Fraser Glaciation, further alteration of the landscape has occurred as a result of the ongoing processes that remove, transport, and re-deposit materials. These include mass movement (slope processes) and fluvial (stream) activity.

Mass movements (gravitational slope processes) such as rockfall, debris flow, debris slide, and avalanching have lead to the accumulation of colluvium on lower slopes and valley bottoms throughout TFL 23. Where slopes are more uniformed morainal material is found (Howes and Kenk 1997). Glaciofluvial terraces and kames are common in major valley bottoms, such as Arrow Park (Mosquito Creek), and Halfway and Whatshan River Valleys. Upper slopes are comprised primarily of colluvium overlying bedrock in varying thickness amounts, due to glacially over-steepened valley walls. There are extensive rocky plateau areas west of Lower Arrow Lake.

The majority of TFL 23 is dominated by coniferous forests, which receive moderate to significant amounts of precipitation, underlain by a medium to coarse-textured parent material. These moist and acidic conditions at higher elevations give rise to podzolic soils, which have light gray eluviated Ae horizons, found overlying enriched B horizons that range from orange-red to dark brown. The diagnostic podzolic B-horizon is enriched with varying amounts of amorphous aluminum and iron as well as organic material leached for the Ae horizon above. Humo-Ferric Podzols are the most dominant soil found in the TFL (Jungen 1980).

Regosols have formed where soils have had less time to form and show poor to very poor horizon development. They form in young materials such as river gravels, fresh colluvium and recently deglaciated soils or in disturbed materials subject to flooding or slope processes. Brunisols form on similar young geological sediments, but the soil has undergone moderate development. Brunisols, which occur at lower elevations in TFL 23, can be distinguished from Regosols based on their diagnostic Bm horizon. This horizon exhibits the development of soil structure and removal, by leaching, of soluble salts and carbonates from the A horizon. In the field it is recognizable by its browner to redder colour when compared with the underlying parent material (Lavkulich and Valentine, 1978). Brunisols are commonly found complexed with podzolic soils on steep valley sides, common in TFL 23.

Gleysols and Organic soils have developed where drainage is imperfect to very poor. These soil types can be found along floodplains such as, the Incomappleux River and Beaton Creek Valleys, where periodic to prolonged saturation occurs. Another common site for Gleysols is at toe slopes. Organic and Gleysol soils are also found at upper elevations, particularly in plateau areas, where surficial material is thin over undulating bedrock.

1.2 ECOSECTION AND BIOGEOCLIMATIC CLASSIFICATION OF TFL23

Ecoregions are large regional-sized, ecological land units that have similar macroclimate, physiography, vegetation and wildlife potential. Five levels of Ecoregion Classification are recognized including Ecodomain, Ecodivision, Ecoprovince, Ecoregion and Ecosection. Following the ecological land classification hierarchy set forth by Demarchi (1996), TFL 23 is located within the Humid Temperate Ecodomain, the Humid Continental Highlands Ecodivision, the Southern Interior Mountains Ecoprovince, and in the Northern Columbia Mountains and Selkirk – Bitterroot Foothills Ecoregions.

Ecosections are subregional units within ecoregions that are similar in climate, landforms, bedrock geology, soils, and plant and animal distributions. Demarchi (1996) classifies TFL 23 as being located within the Central Columbia Mountains (CCM), Selkirk Foothills (SFH) and Northern Kootenay Mountains (NKM) Ecosections. (see Figure 2).

The Central Columbia Mountains (CCM) Ecosection is an area of high ridges and mountains, but the valleys and trenches are narrow. Precipitation is high from the valley bottoms to the upper slopes.

The Selkirk Foothills (SFH) Ecosection is an area of high ridges and mountains interspersed with wide valleys and trenches. Precipitation is high on the mountain slopes but rain shadows are common in the southern valleys.

The Northern Kootenay Mountains (NKM) Ecosection is an area of high, rugged mountains, many of which are ice-capped. It has the highest precipitation and coolest temperatures.

Biogeoclimatic Zones, Subzones and Variants occur within each Ecosection and are classified using the Ministry of Forests Biogeoclimatic Ecosystem (BEC) system (Braumandl and Curran 1992). These units represent groups of ecosystems under the influence of the same regional climate. Most of TFL 23 is located in the Moist Climatic Region but it also contains some areas in the north that are in the Wet Climate Region (ICHvk1, ICHwk1, ESSFvc, ESSFvcp). The TFL contains thirteen biogeoclimatic subzones that are briefly described below (see Figure 3).

ICHdw The Dry Warm Interior Cedar - Hemlock Subzone occurs along Lower Arrow Lake at lower elevations below approx.1200m on warm aspects and 1000m on cool aspects. This zone is characterized by very hot, moist summers and very mild winters with light snowfall (Braumandl and Curran 1992). Climax zonal sites have stands of western redcedar and western hemlock. However, due to extensive fires at the beginning of the 1900's, mixed seral stands of Fd, Ep, Lw, and Pw are much more common. Also, fire induced stands of old growth ponderosa pine provide important wildlife protection. The extensive seral forests are important winter range for deer and elk and support a wide diversity of other wildlife species.

ICHmw2 The Columbia – Shuswap Moist Warm Interior Cedar – Hemlock Variant occurs between approx.1200 to 1550m on warm aspects and 1000 to 1450 on cool aspects. It is found above the ICHdw in the southern portion of the TFL and from the valley bottoms north of Needles where ICHdw does not occur. This zone is characterized by hot, moist summers and very mild winters with light snowfall (Braumandl and Curran 1992). Climax zonal sites have stands of western redcedar and western hemlock. Recurrent fires have led to a mosaic of climax and seral stands of Fd, Lw, Sxw and Cw. Old growth stands are very important for some wildlife species while early seral stages provide important forage for many wildlife varieties including grizzly bear, deer, moose and elk.

ICHmw3 The Thompson Moist Warm Interior Cedar – Hemlock Variant occurs from Galena north along the valley floor to approx.1400m.This zone is characterized by hot, moist summers and mild winters with moderate snowfall (Lloyd et al 1990). Zonal sites have dense stands of western redcedar and western hemlock. Wildfires and harvesting have led to a mosaic of climax and seral stands of Fd, Lw, Sxw and Cw. Old growth stands are very important for some wildlife species while early seral stages provide important forage for many wildlife varieties including grizzly bear, deer, moose and elk.

ICHvk1 The Mica Very Wet Cool Interior Cedar – Hemlock Variant occurs from valley floors in the Incomappleux Valley to approx. 1400m on warm aspects and to 1200m on cool aspects. It also occurs in the upper reaches of cool aspect valleys above Trout Lake. This zone is characterized by warm, wet summers and cold winters with heavy snowfall (Braumandl and Curran 1992). Climax zonal sites have western redcedar and western hemlock while hybrid white spruce is the most common seral species. The extensive old growth forests support a wide variety of dependent species especially caribou and grizzly bear.

ICHwk1 The Wells Gray Wet Cool Interior Cedar – Hemlock Variant occurs in the northern part of the TFL. It first appears on cool aspect slopes north of Slocan Lake from approx. 1100 to 1400m and then appears on valley floors farther north in Trout Lake, Beaton and the Incomappleux Valley to 1400m. This zone is characterized by warm, wet summers and cool winters with moderately heavy snowfall (Braumandl and Curran 1992).). Climax zonal sites have western redcedar and western hemlock while hybrid white spruce is the most common seral species. The extensive old growth forests support a wide variety of dependent species, while early seral stages support species such as ungulates and bears.

IDFun The Undifferentiated Interior Douglas-fir (Arrow Lake) Unit is located in a small part of the TFL near Deer Park and occurs from Arrow lake to approx. 800m below the ICHdw. It has not been described in Braumandl and Curran (1992) but site series have been developed by Ketcheson and Marcoux (1995). It is drier and warmer than the ICHdw. Zonal sites are characterized by open Fd and Py stands. This zone provides important winter habitat for ungulates.

ESSFvc The Very Wet Cold Engelmann Spruce – Subalpine Fir Subzone occurs in the very north of the TFL on cool aspects from approx.1600 to 1900m. This zone is characterized by cool, very wet summers and cold winters with very heavy snowfalls (Lloyd et al 1990). Zonal sites have Engelmann spruce, subalpine fir and mountain hemlock. There are extensive old growth stands that are important late winter range for caribou.

ESSFvcp The Very Wet Cold Engelmann Spruce – Subalpine Fir Parkland Subzone occurs in the very north of the TFL on cool aspects from approx.1900 to 2350m and above the ESSFvc. It is a transition between the continuous forest and the alpine tundra. This zone is characterized by cool, very wet summers and cold winters with very heavy snowfalls (Lloyd et al 1990). Zonal sites have Engelmann spruce, subalpine fir and mountain hemlock, often in krummholtz stands. Late lying snow and frost pocketing create a landscape of scattered tree islands and permanent meadows.

ESSFwc1 The Columbia Wet Cold Engelmann Spruce – Subalpine Fir Variant occurs between approx. 1550 to 1700m on warm aspects and between 1450 to 1650m on cool aspects. It is found in a thin band above the ICHmw2 or ICHwk1 and below the ESSFwc4. This zone is characterized by cool, moist summers and cold, wet winters with moderately heavy snowfall (Braumandl and Curran 1992). Climax zonal stands have Engelmann spruce and subalpine fir with Cw and Hw often present as understorey or intermediate trees. Due to a less frequent fire cycle, there are fewer seral stands. The abundant old growth stands support a range of dependent wildlife species.

ESSFwc4 The Selkirk Wet Cold Engelmann Spruce – Subalpine Fir Variant occurs between approx. 1700 to 2100m on warm aspects and between 1650 to 2100m on cool aspects. This zone is characterized by cool, moist summers and cold, wet winters with moderately heavy snowfall (Braumandl and Curran 1992). Climax zonal sites have stands of Engelmann spruce and subalpine fir. Long fire cycles have produced many old growth stands and few seral stands. The abundant old growth stands support a range of dependent wildlife species.

ESSFwcp4 The Selkirk Wet Cold Engelmann Spruce – Subalpine Fir Parkland Variant occurs between approx. 2100 to 2350m. It is a transition between the continuous forest and the alpine tundra. This zone is characterized by short, cool, moist summers and long, cold, wet winters with heavy snowfall (Braumandl and Curran 1992). Mature zonal sites support patchy stands of Engelmann spruce, subalpine fir and subalpine larch, often in krummholtz stands. Late lying snow and frost pocketing create a landscape of scattered tree islands and permanent meadows.

AT The Alpine Tundra Zone occurs at above approx.2350m. It is characterized by short, cool, moist summers and long, cold, wet winters with heavy snowfall. It encompasses the high, treeless peaks of the Selkirks and Monashee Mountains. Much of this subzone is non-vegetated with herb dominated meadows on zonal sites.

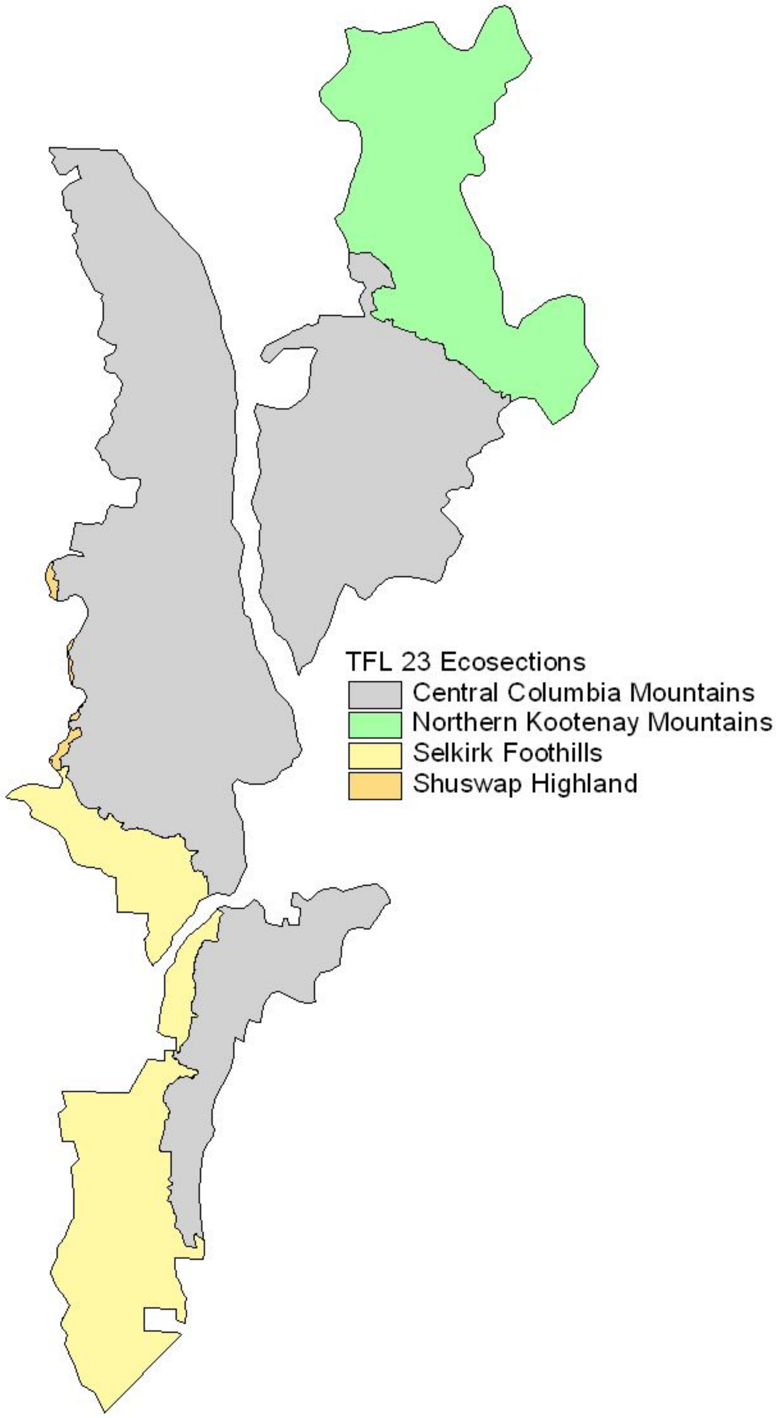

Figure 2. Ecosections of TFL 23, near Nakusp British Columbia.

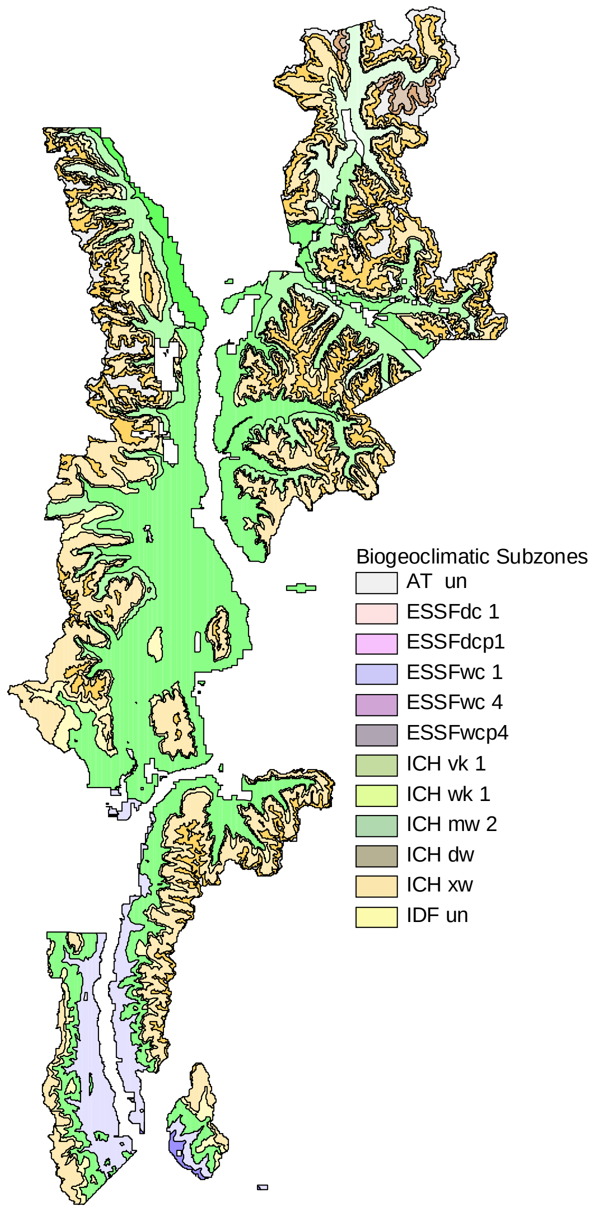

Figure 3. Biogeoclimatic Subzones and Variants of TFL 23 Near Naksup, British Columbia.

1.3 OBJECTIVES

We accomplished the following objectives in this project:

1. To produce a site series map using raster-based PEM methodology over the entire area of TFL 23 using an ARCINFO based GIS model.

2. To utilize, where appropriate, existing ecological data sources for both knowledge table relationships and model verification, as well as collect data for model verification and accuracy assessment.

3. To create habitat capability look-up tables for four species and relate them to the ecological classification of TFL23.

4. To document methodology and results in a report and to provide digital copies of site series mapping in seamless coverage, 1:20,000 and 1:50,000 plot files.

2.0 METHODS

2.1 GIS INPUT DATA ASSEMBLY, ASSESSMENT AND PREPARATION

2.1.1 RASTER DATA FORMAT

A raster data model was selected as the processing format over a vector (polygon) based approach. There are several advantages to using raster data for predictive ecosystem modeling. The raster format provides more efficient processing, especially in multivariate analysis, over vector data since it does not have the topological overhead to maintain. Raster layers are analyzed with numeric calculations on a pixel-by-pixel basis, whereas vector analysis is based more on the geometry of polygons. Raster data maintains a high level of spatial resolution since the landscape at its largest scale is a collection of individual pixels of relatively small size. A 25-meter pixel size was chosen as the standard cell size for the PEM output.

PEM layers portraying landscape character, including, slope, aspect, and shape, raster data are represented by a continuous surface. Digital values will increase and decrease in gradients from pixel to pixel. Neighborhood analysis (moving window) is used to smooth and filter input layers and analyze the gradients between pixel values. Filters are used to reduce minor noise and smoothing with different window sizes is used to adjust layers to an appropriate scale for the landscape model. A majority filter was used for removing noise and a mean filter was used for smoothing. The use of filters is discussed below.

A raster model permits flexibility for assigning and adjusting class breaks since the raw data will remain in a continuous form. Non-linear changes in gradient on a surface can be measured. For example, the rate of change of elevation is measured to extract profile morphology and derive toe slopes and terrain shape.

The software environment used for the raster processing was Arc/Info GRID version 7 and PCI Image Analysis version 6.

2.1.2 SOURCE DATA

The GIS inputs for the predictive ecosystem mapping model were derived from the following five source layers:

- TRIM - 1:20,000
- Forest Cover - 1:20,000 pre-VRI
- Landsat 7 - 30 meter multi-spectral satellite imagery
- Geology - 1:250,000

2.1.2.1 DIGITAL ELEVATION MODEL

A Digital Elevation Model (DEM) was the primary layer used to produce landscape layers. From the TRIM contour, elevation, and break-line layers a TIN (Triangular Irregular Network) was built. There was minimal weeding of TIN nodes in order to preserve the elevation detail of TRIM data. The TIN was sampled on a 50 meter pixel grid to create a raster DEM. Two DEM's were used for deriving landscape layers. The first was the raw output from the TIN to raster DEM conversion. This DEM represents the highest level of terrain complexity. A second DEM was produced from a 3x3 mean filter applied to the raw DEM to provide a low pass smoothing of the elevation model. The DEM smoothed out micro variations in the terrain and produced smoother derivative output.

2.1.3 PEM INPUT LAYERS

There were sixteen input layers created (see Table 1). Layers had a range of one to twelve classes. Each class was assigned a numeric value, which in turn was assigned to the pixel values for the raster layer. The layer's names and the numeric values of each layer relate to the knowledge tables. A zero value was the NULL class. For single class layers, such as wetland, a value of one represented the presence of wetland, and zero represented no wetland was present that pixel location.

Table 1. Input and Derived GIS Attribute Layers Used in the TFL 23 PEM

SOURCE	LAYER	GIS NAME	CLASSES
BEC	BEC	BEC	12 classes
TRIM DEM	SLOPE	SLP	6 classes
TRIM DEM	ASPECT	AS	2 classes
TRIM DEM	SOLAR RADIATION	SRD	2 classes
TRIM DEM	SHAPE	SHP	4 classes
TRIM DEM	TOE SLOPE	TOE	1 class
LANDSAT	SATELLITE CLASSIFICATION	SAT	4 classes
TRIM	WETLAND	TRIM	1 class
FOREST COVER	ALPINE FOREST	B	1 class
FOREST COVER	INVENTORY TYPE GROUP	ITG	3 classes
FOREST COVER	FOREST HEIGHT	HT	2 classes
GEOLOGY	MATERIAL TEXTURE	GEO	3 classes
SOIL	QUATERNARY DEPOSIT	SOIL	1 class
TRIM	STREAM DENSITY	STRMS	2 classes
NEURAL NETWORK	NN CLASS	NEU_NET	12 classes
TRIM	GULLY+AV PATH	AV	2 classes

2.1.4 LANDSCAPE LAYERS

From the 50-meter DEM the following layers were produced:

Slope - percent slope

Slope was classified into the following six classes:

0-5%
6-25%
26-50%
51-70%
71-100%
over 100%

Aspect – warm/cool/neutral

The DEM with a 3x3 mean filter was used to produce the aspect layer. This helped to reduce small amounts of noise and speckle in the output. Aspect was classified into the following three classes:

Warm	135 to 285 degrees azimuth
Cool	285 to 135 degrees azimuth
Neutral	Any aspect with a slope of 25% or less

Solar Radiation

Solar radiation was calculated for Julian days 120, 171, and 273, the start, middle and end of the growing season. An average was taken for the three dates. The model (Kumar, 1997) calculates Kilojoules of energy per square meter per day. The model accounts for the solar azimuth, and elevation variation by the solar calendar. Latitude in decimal degrees is input to adjust sun elevation. The model is useful because it identifies regions with a cool aspect that receive sun due to exposed terrain position, and also identifies regions of warm aspect that are cool because of cast shadows and terrain blockage, such as in deep valley bottoms. These two classes were used as an adjustment layer for aspect.

Shape

Landscape curvature was classified into four categories, concave, straight, convex, and convex-ridge. The DEM with a 3x3 mean filter was used to smooth out micro variations. Pixel values representing curvature range from negative values for concave to positive values for convex. A lookup table with the values was used to classify terrain shape:

concave	-100 to –5
straight	-5 to 5
convex	5 – 15
convex-ridge	15-200

A 3x3 majority filter was run on the classification to reduce noise and speckle and produce more homogenous units.

Toe Slope

The change in slope perpendicular to the direction of the slope was measured from the smoothed DEM. This measure represents regions of increasing and decreasing slope values. Pixel values represent the rate of decreasing slope from steeper to less steep slope values. Through an iterative process and comparison to field plots, a ranges of values were identified as toe slopes areas. A range of values was extracted that represent the flattening out inflection point of the landscape profile. The values used were 80-350.

A 3x3 majority filter was run on the classification to reduce noise and speckle and produce more homogenous units.

2.1.5 LANDSAT LAYER

A Landsat 7 scene from September 9, 1999 was ortho-rectified to TRIM. Thematic bands 3,4,5 representing red, near-infrared and mid-infrared were the source image data for the satellite classification. Digital orthophoto imagery was used as the primary source of ground control training. The overall classification accuracy, based on the training samples, was 81%. There was no field verification of the classification layer. A maximum likelihood classification was trained with the following land cover classes:

Rock
Talus
Avalanche chute
Vegetated rock/soil

The avalanche chute class required post-classification processing since its spectral signature occurred in non-avalanche areas. Logged areas and the area below the operability line were masked out. From forest cover mapping polygons with non-productive code NPBR (non-productive brush) were added. Rock, talus, and operable land were masked out from the NPBR layer.

2.1.6 TRIM LAYERS

- Wetland

  Wetland boundaries were extracted from the TRIM water layer and wetland polygons were created. The polygons were then rasterized to a 50 meter pixel size.

- Stream Density
  A circular moving window with a 200 meter radius was used to measure the density of streams. Each pixel in the output was assigned a value representing the length in meters of stream within the moving window. The following two classes were created:

  1 – 20 meters of stream/hectare
  > 20 meters of stream/hectare

2.1.7 GEOLOGY and SOIL

- Bedrock geology polygons were reduced to three classes representing the following material texture:

  Fine
  Coarse
  Mix of fine and coarse

- Quaternary deposits were extracted from geology data and put into a separate layer

2.1.8 FOREST COVER

From the Ministry of Forests, Forest Inventory Data, the following layers were extracted:

- Alpine forest – B leading

- Inventory Type Groups
  32 – Yellow Pine Leading
  21 – Spruce Leading
  35 – Cottonwood Leading

- Forest Height Class
  Class 1 – 0.1 to 10.4 meters
  Class 2 – 10.5 to 19.4 meters
- Neural Network Results
  twelve site series classes

2.1.9 OVERLAY

The sixteen input GIS layers were combined into a single raster layer. Each pixel value in the combined grid was assigned a unique number representing the combination of the class values of all the input layers. While there were over ten million possible permutations, in actual number of combinations for the entire TFL 23 was approximately 100,000. Each record in the combined attribute database contained the attribute value for each input layer. This database was the input for applying the knowledge tables for the PEM model.

2.2 FIELD DATA COLLECTION

Field data collection was accomplished in year two of the TFL 23 PEM project between the dates of July 1 and September 15, 2001. Field crews consisted of two, a technician and a vegetation ecologist. In some cases field crews of three were used to gather extra measurements for the Caribou winter range project where there was a cooperative sampling effort between the two projects.

A total of 508 sample plots were recorded using modified Ground Inspection Forms following sampling methodology of BC MOELP (1998). Sample locations were randomly determined within 500 m. of road access. Sampling was biased towards operable areas within TFL 23. The data collected can be found in Appendices I, VI, and IX. A sample area was described from the random start point followed by two more samples 100 m. at a random compass bearing from the start point. The location of each point was determined through a differentially correctable GPS. The Caribou plots did not always include a transect, in that case single points were randomly determined and located in the field.

Data were collected in RIC standard ground inspection forms that are located in appendix VI.

2.3 DATA ANALYSIS AND PEM MODEL DEVELOPMENT

2.3.1 LOGISTIC REGRESSION

In the TFL 23 PEM project we analyzed field data to determine whether the site series assigned subjectively to plots could be derived empirically from various field and GIS attributes and, if so, to determine those variables that were most important in discriminating among different site series. The results are reported in Tables 3 and 4.

We first examined overall data structure with a classification tree analysis (Statistica 1999) stratified by BEC Subzone. Classification trees are used to determine the membership of cases (plots) into different classes (site series) based on a number of ordinal or categorical predictor variables (plot variables). Membership in different classes is based on a splitting method, in this case, the discriminant-based univariate method proposed by Loh and Shih (1997). The result of the analysis is an ordering of plot variables in terms of their importance in classifying plots into site series. The method doesn't provide a measure of statistical significance, but it is useful for examining variables important to the overall classification of plots into site series.

Within site series, we determined important variables with a multiple logistic regression analysis. Logistic regression is best suited to analysis problems involving binary response variables. The method may also be extended to accommodate polytomous dependent variables; however, the interpretation of results is difficult because coefficients must be interpreted in relation to a dependent "reference" category (Menard 1995). Therefore, we used site series as a binary response variable by classifying plots within BEC subzones as either site series *x* or not site series *x*, where *x* was the site series considered in the current analysis. Although this eased interpretation, it resulted in multiple comparisons among plots, and significance values should be interpreted conservatively.

Variables for the analysis were: moisture regime, nutrient regime, slope gradient, aspect (coded to NE and SW as 0-1 variables relative to "flat", i.e., ≤25% slope), mesoslope position, and soil drainage. There were insufficient plot data to analyze several site series. Model fit was considered adequate if the regression model was significantly different from the intercept-only model. Individual coefficients were considered significant if they were significantly different from 0 at an alpha of 0.1 (based on Wald statistics).

We determined the ability of the regression equations to classify plots of different site series by constructing 2 x 2 contingency tables that described the correct classification of plots into site series based on predicted values and cut-offs (to the nearest 0.1) that maximized the sum of correctly classified observations. Complete results of this analysis can be found in Appendix III.

Analysis of plot data based on GIS layers

As an alternative, we took a model building approach to the analysis of plot data based on GIS layers. Rather than interpreting important variables from logistic regression equations based on all input variables (as we had in the above analysis); we derived "reduced" models based on the most parsimonious subset of input variables. Variables considered in the analysis were: slope gradient, aspect (coded to NE and SW as 0-1 variables relative to "flat", i.e., ≤25% slope), solar radiation, "shape" (first derivative of slope), elevation, forest height class, and stream density. The criterion for including variables in the final model was the Akaike Information Criterion (AIC, Akaike 1983). Again, model fit was considered adequate if the regression model was significantly different from the intercept-only model. We relaxed the interpretation of coefficients' significance based on Wald statistics and considered all variables retained in the final model as significant (Menard 1995). Results of this analysis can be found in Appendix IV.

2.3.2 NEURAL NETWORK ANALYSIS

Another method used to assist in the derivation of site series from PEM input layers was a neural network analysis. Neural network analysis is a powerful technique used to model complex functions. A network is composed of an interconnected series of artificial neurons, which function in a way similar to their biological counterparts. A neuron (also known as a "node" or "unit") accepts a series of inputs (with specific input strengths, or "weights") from input data or from other neurons. Each neuron has a threshold value, and if the weighted sum of the inputs exceeds the threshold value, the neuron "fires", and the output value is passed on to the next series of neurons in the network (Bishop 1995). Figure 4 illustrates a very simple network consisting of one decision, or "hidden" node. In this example, the input node of the network accepts a value (0.625) from a single variable, and passes it to a node with a simple threshold activation of 0.5. The threshold is subtracted from the sum of the inputs, and because the result (0.125) is >0, the output node "1" is triggered. As a result, this simple network can accept inputs of numbers between 0 and 1 and classify them into binary categories.

More complex classification problems involving additional input variables and additional output classes can be handled by adding more hidden units. "Training" is the process whereby a network with optimal activations and signal weightings is iterated from a set of "known" cases of input and output data. The result is a very flexible, non-linear classification technique that can model very complex functions.

Typically, a network has an input layer with one node for each input variable, a hidden layer of several nodes, and an output layer with one node for each output class. The nodes of each layer are connected to every node in the preceding and subsequent layers. For this project, the input layer passed the values of GIS variables (*e.g.* slope, aspect classes) to the network for processing, and the output was a set of probabilities that represented the probability that a case belonged to one of the sites series being modeled in the analysis (Figure 5).

There were several steps involved in developing and applying the neural networks. First, network models were "trained" on the GIS data and site series calls associated with ground plots. Data were divided into 3 sets: training, verification, and test sets. The training set was used to train an initial network to associate GIS input values with site series classifications. Weights and thresholds were adjusted iteratively to minimize the sum-of-squares errors between the output activations of the network and the expected activations based on the known site series classifications of the data. The verification data set was then used to test the fit of the model on independent data.

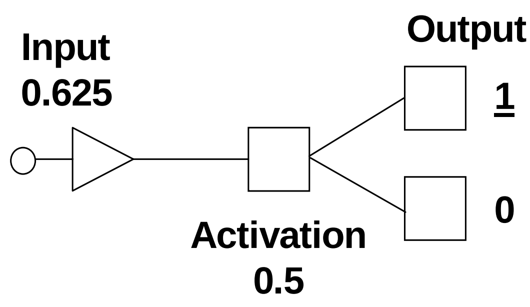

Figure 4. A simple neural network that accepts a single input value between 0 and 1 and classifies it into 1 of 2 categories.

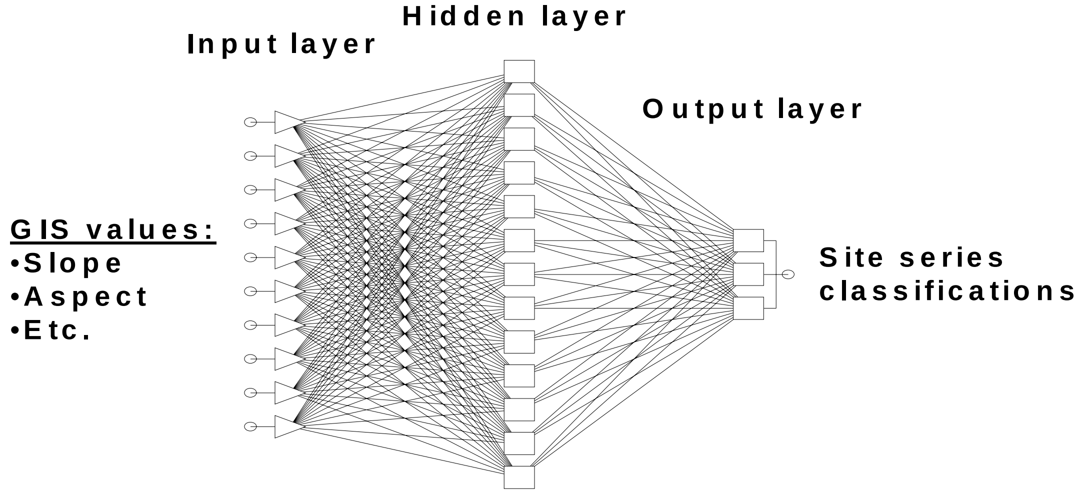

Figure 5. Topology of a typical neural network. The network accepts data from 10 input variables (derived from GIS coverages) and classifies cases into 3 output classes (sites series).

Next, the network topology was changed to include a different set of input variables and a different number of hidden units. Again, weights and thresholds were adjusted to minimize sum-of-square errors in the training set, and then tested independently with the verification set. The process was repeated many times (typically >100) and the network with the lowest error (*i.e.* the closest fit between the predicted and actual site series calls) was selected as the best.

Because the training and verification sets were used repeatedly in developing the model, a test data set was run on the final network to determine whether the network had "over-learned" the data. Over-learning occurs when a model generates a good fit to training and verification data, but generalizes poorly to independent data. Because neural network analysis is a very flexible, non-linear modeling technique, over-learning is a common problem.

Models that fit training and verification sets well, and generalize adequately to an independent test set, can be used to classify novel sets of data. In this project, results from the plot data were generalized to the entire map by running GIS data from each pixel through the models.

2.3.2.1 MODEL BUILDING

Analyses were stratified by BEC subzone, and site series were included in analyses only where there were >10 ground plots. Occasionally, site series with more plots were excluded from model building because they could not be classified correctly with any certainty by the neural network analysis. Plot data were divided between training, verification, and test sets in roughly a 3:1:1 ratio, although we also varied the ratio in attempts to achieve a better fit. Model fit was assessed first by sum-of-squares errors and then by classification frequencies.

All analyses were conducted with *Statistica Neural Networks* software (Statsoft Inc., Tulsa, OK). Data were fitted to 3-layer perceptron networks using a second-order conjugate gradient decent training algorithm. The softmax activation and entropy (multiple) error transformations were applied to final models to allow the interpretation of output activations as probabilities (Statistica 2000). Probabilities for each case summed to one.

We used sensitivity analyses to determine the contribution of each variable to the final networks. Models were run in which each variable was excluded in turn, and sum-of-squares errors were calculated for the subset models. Variables were ranked according to the fit of the subset models from which the variable had been excluded.

We summarized classification data with confusion matrices. The goal of the modeling was to minimize classification errors among site series calls in all 3 data subsets (training, verification, and test sets). Similar results among subsets suggested that models generalized well. Significantly higher errors in test sets suggested that over-learning had occurred and models might not generalize well. In practice, over-fitting can be difficult to avoid, particularly with small sample sizes.

2.3.2.2 MODEL APPLICATION

Final models were applied to map data by subzone. We ran GIS data for each pixel on the corresponding subzone model and then mapped the resulting activations. Pixels were assigned to a site series if an activation was >0.75. If no probability was >0.75, the pixel was classified as "unknown."

The neural network pixel classification was added to the input layers and knowledge bases of the PEM. When the neural network gave a probability >0.75 for a site series it received a high score on the knowledge bases, when it was less than >0.75 it received no score and other input layer attributes were used to determine the site series.

2.4 KNOWLEDGE BASE CREATION

Knowledge bases were modeled after those used in the Arrow TSA PEM with columns representing site series and rows representing GIS input data layer attributes (See Section 2.1). The cells were initially filled with zeros and then the zeros replaced with values between 1 and 100, depending on the relationship between that input layer value and site series classification.

Initially each subzone was assigned values based on environmental attributes, summarized by site series, from the collected field data. In units where field data was lacking, environmental descriptions from Braumandl and Curran (1992) or Lloyd et al (1990) were used as the basis for assigning values to the GIS attributes used in the PEM. In some subzones where descriptions lacked the necessary detail needed for the PEM model, we augmented field data and field guide descriptions with environmental data from existing TEM mapping (Ketcheson and Marcoux, 1994, Ketcheson, 1995, Marcoux et al, 1996) where subzones of TFL 23 were represented. These were used to assist in the creation of first draft knowledge bases. These values were then subjectively weighted with the results of statistical in subsequent iterations of the knowledge bases.

Logistic regression was used in the ICHdw, ICHmw2, ESSFwc1 and ESSFwc4, in site series where there was sufficient replication in TFL 23 PEM plot data, to assist in determining the weight to be given to plot environmental and GIS input layer categories by site series. In those same subzones the results of the neural network classification were weighted within the knowledge base and used, with the other input layers, to score for site series. Input data, which were highly correlated to site series, were weighted as high as 25 when directly associated with a single input data layer derived from the neural network. GIS input data which was highly correlated with site series was weighted as high as 15 when logistic regression identified those variables as significant for discriminating between site series within a subzone. Plot environmental data was weighted as high as 30 when that attribute was significantly associated with that site series.

2.4.1 MODEL VERIFICATION

Draft knowledge bases were run with input data layers and 508 plots from TFL 23, to test their success in PEM model prediction of known random plot locations. The results of this process can be found in Appendix II. The percentage of correct allocations was determined and the knowledge bases were adjusted subjectively to improve the performance of the knowledge base, based on the plot site series scores by input data layer.

This process was repeated seven times over the TFL 23. The results vary with the site series, more refinement could be done on wetter sites with low relief and wetter variants with small sample numbers.

The TFL 23 PEM plot data was used for model verification in the PEM mapping exercise. The site series classification of plots was compared to the site series classification given by the model at that same UTM grid coordinate. The results of that process are reported in Section 3.2.

We wanted to examine ecological relationships between site series and assess how "ecological adjacency" affected the results of the model's site series classification compared to field classification. For this exercise we used Huggard's (2000) site series groupings at three levels, fine, intermediate and coarse (see section 3.2.1.2). The results of that assessment of model "goodness of fit" can be found in Appendix III.

The method of assessment of the accuracy of the PEM model was a point assessment, where the computer generated site series classification assigned to a pixel was compared to the site series assigned to a field plot located within that pixel. A percentage of correct assignments was calculated and reported as "goodness of fit" for model building plots and as "accuracy assessment" for independent plots. There are a number of concerns with this type of model assessment. If the plot location is not positionally accurate, then it is possible to get a false wrong answer. If the ecological classification is ambiguous and there are similar site series that can be found on similar combinations of environmental features, then it is possible to be only "partly wrong". However, based on feedback from the year one PEM results (Meidinger pers com, Meidinger 2000), we choose to stick with point assessments of accuracy as our only measure of model success.

2.5 WILDLIFE HABITAT CAPABILITY RATINGS

Wildlife habitat ratings for four species; Woodland Caribou, Mountain Goat, Grizzly Bear and Northern Goshawk were accomplished using the methodology described in RIC 1999. Species accounts for these animals were developed and used in the ratings scheme suggested in the RIC standards for wildlife habitat assessment (RIC1999). They can be found in Appendix VIII.

The ratings completed for each species are found in Table 2.

Field sampling included an assessment of the suitability of each sample plot for winter and growing seasons for living activities of these species. This information was summarized and used to assist in the development of look-up tables that can be found in Appendix VII. These tables were subjectively completed based on the utility of each site series and modified site series to the activity and season of use of the animal being assessed.

The ratings are completed relative to provincial benchmarks for the best habitats for that species. Consequently, they are relative rankings to the potential of all habitats within the province, not relative to the potential of all habitats within TFL 23.

The subjective ratings scale goes from one to six, with one being the best habitat, relative to any others found in the province, and six being, nil, the habitat has no value to the animal being assessed.

Table 2. Species, Activities and Seasons of Use Rated For Site Series and Modifiers in TFL 23.

Species	Activity	Season
Mountain Caribou	Feeding	Early winter
		Late winter
		Spring
		Summer and Fall
	Security thermal cover	Early winter
		Late winter
		Spring
		Summer and Fall
Mountain Goat	Feeding	Winter
		Spring
		Summer
		Fall
	Security thermal cover	Winter
		Spring
		Summer
		Fall
Grizzly Bear	Hibernation	
	Living	Spring
		Summer
		Fall
Northern Goshawk	Living	All
	Reproduction	All

The look-up table provides the information necessary to create interpretive maps for wildlife capability from the PEM BEC, site series and modifiers database.

3.0 RESULTS

3.1 KNOWLEDGE BASE DEVELOPMENT

The refinement of the knowledge bases was based on feedback from PEM runs over the entire Arrow District. What follows is the results of statistical analysis used to assist in the subjective weightings of values in the knowledge bases.

3.1.1 USE OF CLASSIFICATION TREE ANALYSIS AND LOGISTIC REGRESSION IN THE DEVELOPMENT OF KNOWLEDGE TABLES

The plot environmental and GIS data were subjected to classification tree analysis and the results are presented in Table 3, where the numbers represent the relative importance of plot environmental attribute variables for distinguishing between the site series.

The **classification tree** demonstrates that the importance of each plot environmental variable for distinguishing between site series varies with the Subzone. The weightings in the knowledge tables qualitatively reflect these differences.

Logistic regression asks the question: given the values of the X variables (plot environmental data), what is the probability of observing Y, (the site series)? Can environmental attributes measured in plots be used to predict the probability that the site series changes?

The results of the logistic regression of plot environmental variables against site series can be found in Table 4. The results of that analysis were summarized and the statistically significant plot environmental attributes were used to increase the weighting of input data layer characteristics in the knowledge tables. Table 4 outlines the most significant variables for distinguishing between site series and the percentage correct prediction of field plot site series when using the plot environmental variables versus the percent correct prediction of other site series using those same variables.

The same analysis was done on the GIS attributes from the input data layers for the same plots. These results can be found in Table 5. The reported input data layer attributes in Table 5 are all to be considered statistically significant.

These analyses were only useful for site series with sufficient replication to be used in the analysis. However, these were also the most commonly occurring site series in the project area.

The results show that field data attribute soil moisture and drainage to be the most important variables for distinguishing between site series in all subzones. Other field data attributes vary in their importance for differentiating between site series within a subzone. In the ICHdw very few other attributes are useful. In the ICHmw2 slope and aspect are also useful for differentiation between the 01, 03, 04, 05 site series. In the higher elevation subzone, ESSFwc1, site series 01 and 02 are best distinguished by moisture and drainage and to a lesser degree by slope. The ESSFwc4 site series 01, 03, 04 also rely heavily on drainage and moisture, but are also distinguished by elevation and percentage coarse fragments. Table 4 shows the specific relationships and probabilities of some field attributes that are strongly related to site series.

The PEM can only use GIS input variables to distinguish between site series within a subzone variant. Table 5 shows which GIS input variables were statistically significantly related to site series from the year one field data set. The results are less straightforward than the plot field data relationships between field collected attributes and site series. It should be noted that slope and aspect were consistently significant for the differentiation of site series using GIS input data only. To a lesser degree stream density was also significant.

Table 3. Classification Tree Analysis Results For Field Plot Variables Arrow District PEM Projects.

Variable	ICHdw 01a, 01b, 02+	ICHmw2 01, 03, 04, 05+	ESSFwc1 01, 02+	ESSFwc4 01, 03, 04+
Soil moisture	100*	100	100	100
Drainage	82	55	96	69
Soil Nutrient	53	27	12	28
Aspect Class	10	65	14	20
Slope Class	26	86	60	30
Meso Slope Position	25	32	21	57
Surface Shape	3	21	5	36
%Coarse fragments	25	38	42	75
Elevation	10	19	8	94
Overall Error (independent test sample)	12.5%	25%	27.3%	22.2%
N	32	40	18	22

*Number is the relative importance of variables in distinguishing Site Series.
+ Site series with sufficient data for classification tree analysis

Table 4. Logistic Regression Results For Field Plot Environmental Variables.

Site Series	N*	Significant Variables	% Correct prediction of field plot site series	%Correct prediction of other site series
ICHdw-01a (RFa)	68	Soil moisture p=0.0001	91.2%	90.0%
ICHdw-01b (RFb)	73	Soil moisture p=0.004 Meso slope position p=0.018	80.8%	70.5%
ICHdw-03 (RD)	11	Soil moisture p=0.03	90.9%	75.8%
ICHmw2 – 01 (HF)	114	Meso slope position p=0.000	74.5%	59.8%
ICHmw2 – 03 (DF)	89	Soil moisture p=0.000 Warm aspect class p=0.008	75.3%	92.4%
ICHmw2 – 04 (RF)	63	Cool aspect class p=0.000	92.1%	81.1%
ICHmw2 – 05(HO)	40	Drainage p = 0.003 Soil moisture p=0.004 Soil nutrient p=0.038 Cool aspect class p=0.05	82.5%	69.1%
ICHmw2-06 (RD)	16	Soil nutrient p=0.000 Meso slope position p=0.038	81.3%	93.2%
ESSFwc1-01 (FR)	56	Soil nutrient p=0.021 Soil moisture p=0.093	87.5%	63.6%
ESSFwc4-01 (FR)	34	Soil nutrient p=0.048 Meso slope position p=0.063	88.2%	73.1%
ESSFwc4-04 (RF)	15	Slope p=0.028 Soil moisture p=0.045 Soil nutrient p=0.061	86.7	84.4%

*based on year one plot data with sufficient replication within a site series only

Table 5. Logistic Regression Results For Arrow District Plot GIS Environmental Variables From PEM Input Layers.

Site Series	N*	Significant Variables	% Correct prediction of field plot site series	%Correct prediction of other site series
ICHdw-01a (RFa)	68	Cool aspects p=0.009194 Warm aspects p=0.0000468 Stream density p=0.064404	62.2%	74.1%
ICHdw-01b (RFb)	73	Warm aspects p=0.001155 Elevation p=0.05894	51.9%	80.0%
ICHdw-03 (RD)	11	Slope p=0.011753 Solar radiation index p=0.031965 Height class p=0.042107 Stream density p=0.002654	85.7%	86.3%
ICHmw2 – 01 (HF)	114	Slope p=0.088962 Warm aspect p=0.161523 Stream density p=0.013048	77.3%	42.5%
ICHmw2 – 03 (DF)	89	Slope p=0.0000000127 Cool aspect p=0.0000519 Shading index p=0.000089 Height class p=0.0000641	63.5%	76.2%
ICHmw2 – 04 (RF)	63	Slope p=0.145899 Cool aspect p=0.00000007998 Height class p=0.134596	65.6%	81.6%
ICHmw2 – 05(HO)	40	Slope p=0.0000267 Warm aspect class p=0.03299 Shading index p=0.103864 Shape p=0.010285 Height class p=0.036296	79.5%	64.6%
ICHmw2-06 (RD)	16	Slope p=0.001128 Cool aspect p=0.042359 Shape p=0.009279 Height class p=0.059854	62.5%	87.2%

*based on year one plot data with sufficient replication within a site series only

Table 5. Logistic Regression Results For Arrow District Plot GIS Environmental Variables(continued).

Site Series	N	Significant Variables	% Correct prediction of field plot site series	%Correct prediction of other site series
ESSFwc1-01 (FR)	56	Warm aspect p=0.012438	89.7%	33.3%
ESSFwc1-02 (FF)	16	Warm aspect p=0.084699	31.3%	86.8%
ESSFwc4-01 (FR)	34	Cool aspect p=0.093488 Solar radiation index p=0.046583 Shape p=0.05428 Stream density p=0.02766	50.0%	100%
ESSFwc4-04 (RF)	15	Cool aspect p=0.008923 Shape p=0.0098495 Stream density p=0.048102	53.3%	91.8%

*based on year one plot data with sufficient replication within a site series only

3.1.2 NEURAL NETWORK RESULTS

The complete results of the neural network analysis can be found in Appendix V. A summary of the most important GIS attributes for distinguishing between site series where there was adequate sampling can be found in Table 6. Not all GIS variables were ranked as important by the neural network for all subzone variants. It should be noted that aspect was universally ranked either as the first or second most important GIS attribute for the differentiation between site series in all subzones. Toe slopes were also important in all but the ESSFwc4, as was the drainage call where drainage was noted as a gradient between two classes.

The overall accuracy of the neural network classification, based on 86 independent test case plots is summarized in Table 7. The confusion matrix reporting the results for training, verification and independent test cases is reported in Appendix V.

The neural network, on its own, was the most successful at differentiating between site series in the ESSFwc4. It was weakest in the ICHdw.

Based on the test accuracy scores, it was determined that the neural network classification would be used in the PEM for the classification of pixels allocated to the; ICHdw 01a and 02; the ICHmw2 01 and 02; the ESSFwc1 01 and 02; and the ESSF wc4 01, 03 and 04 site series. This was based on a probability of those site series being >0.75 for the pixel. Pixels with that weighting were allocated to an additional input layer. The weighting for those site series in the knowledge base was increased to as much as 25 points over and above the score that site series achieved for other GIS variables in the input layers.

Table 6. Neural Network Assessment Demonstrating Relative Ranking of the Most Important GIS Variables for Differentiation Between Site Series Within A Subzone Variant –Verification Data Set TFL 23

GIS Variable	**ICHdw** **N=30**	**ICHmw2** **N=38**	**ESSFwc1** **N=18**	**ESSFwc4** **N=15**
ASPECT	1	1	10	8
STREAMS	2	2	6	5
SHAPE	3	6	9	1
HEIGHT	4	9	8	9
QUAT DEPOSITS	5	3		
TOE POSITION	6	7	3	7
SOLAR RAD	7	10	7	4
GEOLOGY	8	8	5	6
SATELLITE	9	4	2	2
SLOPE	10	5	11	10
ALPINE FOREST			4	3
ITG			1	

Table 7. Neural Network Independent Test Plot Accuracy Score For Differentiating Site Series Within A Subzone Variant Where There Was Sufficient Data TFL 23

Site Series	Subzone variant			
	ICHdw **N=30**	**ICHmw2** **N=39**	**ESSFwc1** **N=19**	**ESSFwc4** **N=15**
02	67%	80%	100%	
01a	83%			
01b	53%			
01		99.6%	82%	89%
03		50%	33%	75%
04				100%
Overall	**67.6%**	**76.5%**	**71.7%**	**88%**

3.2 VALIDATION OF KNOWLEDGE BASE

The model verification and model assessment process occurred incrementally, as we proceed through the mapping project. These steps were:

- knowledge base preparation
- input layer preparation
- statistical analysis of the plot environmental and GIS attributes
- year one PEM results

- knowledge base refinement
- addition of neural network classification input layer
- drafts five, six and seven PEM
- final accuracy assessment report
- submission of PEM to clients and QA personnel

Development of knowledge bases was an interactive process, starting with information from field plot data and site series descriptions from Braumandl and Curran (1992) as the basis upon which input data layers were weighted for scoring. The goal was to have the resultant mapping follow the ecological classification set forth in Braumandl and Curran (1992) and Lloyd at al (1990).

Qualitative data from plots and environmental tables from Braumandl and Curran (1992) were augmented with the results of a classification tree analysis and logistic regression analysis of plot environmental data and GIS attributes. Statistically significant attributes were assessed in terms of the input data layers used in the GIS, and where appropriate, received higher weighting in the knowledge bases. A separate input layer using the results of the neural network classification was run and pixels with a probability of 0.75 or greater were allocated to site series where the neural net was considered accurate.

The most common initial qualitative adjustment for all subzones was the weighting of scores for avalanche chutes, talus, rock and vegetated rock from the input layer based on satellite imagery. These attributes were initially weighted to 100 in the knowledge base, but later reduced, as it became evident that the satellite imagery over classified avalanche paths and rock. These types of ecological units are important mapping entities, and often distributed as smaller units within a matrix of more widely distributed site series. The accuracy assessment protocol did not test how well the model mapped these units.

The “point” spread between the correct answer (field classification of site series) and the model’s generated site series was examined for every plot. The scoring of each GIS attribute used in the knowledge base for these plots was noted in relationship to the score of the correct answer. Changes were made to the knowledge base that would better separate individual site series scores. This process was repeated seven times.

3.2.1 RESULTS OF TFL 23 PEM MODEL VERIFICATION AND ACCURACY ASSESSMENT

3.2.1.1 SITE SERIES LEVEL ASSESSMENT

We tested the "goodness of fit" of the PEM model to the TFL 23 plot data and the results are located in Appendix II. They are summarized in Table 8. We also assessed the accuracy of the PEM model using an independent set of randomly located accuracy assessment plots located throughout the Arrow Forest District.

Table 8. TFL 23 Site Series Point "Goodness of Fit" and Arrow TSA PEM Point Accuracy Assessment Scores.

Subzone	Site Series Number	N	Percentage Goodness of Fit	Percentage Accuracy	N
ESSFwc4	01	23	82.1%	45.5%	22
	02			100%	1
	03	1	0%	0%	1
	04	1	0%	0%	4
	05			50%	2
	07			0%	4
	06	1	0%		
ESSFwc1	01	27	88.9%	65%	20
	02	9	0%	50%	8
	03	2	0%	0%	4
ICHmw3	01	12	8.3%		
	02	1	0%		
	04	1	100%		
	06	1	100%		
	07	1	0%		
ICHmw2	01	69	88.4%	58.3%	36
	03	48	50%	37%	8
	04	25	52%	75%	32
	05	7	0%	22%	9
	06	7	0%	0%	1
	99	1	100%		
ICHwk1	01	6	50%		
	05	6	0%		
	06	1	0%		
ICHdw	01a	35	91.4%	67.6%	34
	01b	12	41.7%	47.1%	17
	02	11	63.6%	25%	12
	03	4	0%	0%	3
	04	1	0%		
	99	2	0%		
IDFun	02	2	50%		
	04	8	0%		
	05	6	33%		
	44	3	0%		

The model seems to perform well in the units that are the most widespread within TFL 23. In Table 8 it shows that in the assessment of "goodness of fit" that field plot data collected through a random sampling regime is distributed proportionately to the amount of that site series found within 500 metres of road systems. For example, in the ESSFwc4 92% of the sample plots fell into the 01 site series. The PEM model correctly fit 82.1% of the pixels with field samples in them into the 01 site series. However, over the entire Arrow TSA the model only fit that site series correctly in 45.5% of plots from an independent sample set according to the accuracy assessment protocol of Meidinger (2000) and USGS (1999).

The PEM in TFL 23 was initiated before present standards (PEM Data Committee 2000) for PEM data output existed. We followed the protocol outlined in Moon et al (1999) and the TEM Alternatives Task Force (1999) to produce the output submitted for this project.

The assessments that are reported in Table 5 indicate that the TFL 23 PEM varies in accuracy and model goodness of fit. Generally, it is satisfactory in the most widely distributed units, and needs more refinement in the wetter site series and variants.

The ICHmw3, ICHwk1 and IDFun need more work before the mapping generated by the PEM should be used for any operational interpretations.

3.2.1.2 HUGGARD'S SITE SERIES GROUPINGS

Huggard's (2001) work in the Arrow TSA suggested that the site series level of classification may be too detailed for some interpretations related to stand structure and wildlife habitat assessments. We choose to report the goodness of fit and accuracy figures for Huggard's fine, intermediate, and coarse levels of site series groupings for the results of the TFL 23 PEM. We feel that these results are relevant to the accuracy of the PEM for depiction of wildlife habitat capability. Appendix III reports the goodness of fit of plots classified by Huggard's groupings to the PEM model from TFL 23 and Appendix IV reports the accuracy assessments of independently collected plot data from the Arrow TSA when classified by Huggard's groupings. Table 9reports the results of that investigation.

It appears that the mesic units in all subzones and the drier units in the ESSF are the best classified by the PEM. Wet sites in all subzone variants are not depicted accurately by this PEM model. This pattern is evident in all levels of site series groupings.
Because plot data was collected randomly within 500 metres of road systems, it reflects the relative proportions of site series groupings over the landscape within the operable area of TFL 23. On this basis one can be relatively confident that upland circum-mesic sites are the most common in the ICH, and that they are well mapped by the PEM. The same can be said for mesic and drier ESSF sites. Wet sites are not well depicted by the PEM model, neither are they common on the landscape within 500 metres of road systems.

Table 9. TFL 23 Huggard's Site Series Groupings Point "Goodness of Fit" and Arrow TSA PEM Point Accuracy Assessment Scores.

Huggard's Groupings	N	Percentage Goodness of Fit	Percentage Accuracy	N
Fine Groupings				
ICHdw 03	4	0%	0%	3
Subhygric ICH	14	0%	37.5%	8
ICHwk1 06	1	0%	Not sampled	
Mesic-subhygric ICH	6	100%	Not sampled	
ICHmw2 03	44	45.5%	12.5%	8
Mesic ICHdw ICHmw2	138	94.2%	88.6%	114
ICHdw02	11	63.6%	15.4%	13
Xeric ICHmw2 ICHwk1	Not sampled		Not sampled	
Wet ICH ESSF	Not sampled		0%	6
Drier ESSFwc4	1	100%	100%	4
Mesic xeric ESSFwc1	33	100%	87.5%	32
Subhygric ESSF	2	0%	0%	7
ESSFwc4 01	23	100%	50%	20
ESSFwc4 06	1	100%	Not sampled	

Table 9. TFL 23 Huggard's Site Series Groupings Point "Goodness of Fit" and Arrow TSA PEM Point Accuracy Assessment Scores (Continued)

Huggard's Groupings	N	Percentage Goodness of Fit	Percentage Accuracy	N
Intermediate Groupings				
Subhygric ICH	23	0%	27.3%	11
Subhygric hygric ICHmw2 ICHwk1	1	0%	0%	1
Mesic submesic ICHwk1	6	100%	100%	6
Mesic submesic ICHdw ICHmw2	182	93.41%	92.6%	122
Xeric ICH	11	63.6%	15.4%	13
Wet ICH ESSF	Not sampled		0%	8
Drier ESSFwc4	1	100%	100%	4
Mesic xeric ESSFwc1	33	100%	87.5%	32
Mesic ESSF	25	92%	40.7%	27
Subhygric ESSF	1	0%	Not sampled	
Coarse Groupings				
Circum-mesic ICHdw ICHmw2	182	93.41%	92.6%	122
Drier ESSF	34	100%	88.9%	36
Wet ESSF	26	92.3%	40.4%	27
Wet ICH	30	40%	55.6%	18
Wet ESSF/ICH and xeric	11	63.6%	19.1%	21

3.3 SITE SERIES DISTRIBUTION IN TFL23

According to the results of the PEM, TFL 23 supports thirteen subzones and variants and 95 different site series/subzone variant combinations. Table 10. Reports each unit and the number of hectares of each site series unit within the TFL. This ecological classification includes rock outcrops, talus, avalanche chutes and wetlands, as well as, site series as per Braumandl and Curran's (1992) and Lloyd's (1990) classification.

The output from the PEM also modifies each site series by aspects on moderate and steep slope classes. The data bases linked to the site series mapping also include directional exposure modifiers on warm (135 to 285 degrees on slopes >25%), cool (285 to 135 degrees on slopes >25%), very steep warm and very steep cool aspects (greater than 100% slope). These are required by PEM standards. The mapping presented in this report is colour coded by site series. The polygon boundaries are rough, representing the 25 m pixel-based edges, emphasizing that this mapping is generated by a computer model and not yet field verified.

The most widely distributed subzone in TFL 23 is the ICHmw2 (152,744 ha), followed by the ESSFwc4 (143,317ha), ESSFwcp (62,334), the ESSFwc1 (60,556 ha), ICHwk1 (52,183), ICHdw (30,314), AT (21,408 ha), ICHvk1 (19,722 ha), ICHmw3 (7,344 ha), ESSFvc (2,597 ha), ESSFvcp (2360), and IDFun (1,075). TFL 23 is a diverse mountainous landscape dominated by mid slope forested sites.

Within the ICHmw2 the most commonly occurring site series are the 01 (HwCw – Falsebox – feathermoss), the 04 (CwFd – Falsebox) and the 03 (FdCw – Falsebox – Prince's pine). Wet sites are less commonly mapped, that may be consequence of the PEM model underestimating their abundance. Rock outcrops, talus and avalanche chutes account for approximately 5% of the area of the ICHmw2 subzone.

The ESSFwc4 is mapped as being dominated by site series 01 (Bl – Rhododendron – Oakfern) with roughly equal amounts of 02 (Bl – Rhododendron – Falsebox) and 03 (Bl-Rhododendron – Woodrush). Avalanche paths cover approximately 5% of the subzone variant. Rock and talus cover an additional 16% of this variant.

Rock and talus dominate ESSFwcp and AT. There is a significant area of 02 (Bl- Heath) and 04 (Sedge – Pasqueflower) in the parkland. The model needs more field sampling to verify work in the high elevation parkland and alpine subzones. That was not a priority in this project.

The ESSFwc1, according to the model, is dominated by the 01 (Bl – Rhododendron – Oak fern) and the site series 02 (Bl – Falsebox – Grouseberry) together with, in rougly equal amounts the 03 (Bl Devil's club- lady fern) Avalanche chutes occur over approximately 7% of the area of the ESSFwc1.

The ICHwk1 model is weak, to date, it presently maps 04 (HwCw – Falsebox – Feathermoss) as the most common site series, followed by site series 01 (CwHw - Oakfern) and 02 (PlHw – Velvet-leaved blueberry). Avalanche chutes occur over 1% of this subzone. Rock and talus cover approximately 2%.

The ICHdw is dominated by the 01a (CwFd- Falsebox: sx-sm phase). The moister and more productive 01b (CwFd – Falsebox: m-shg phase) has a narrower distribution. The 02 (FdPy – Oregon grape – Parsley Fern) site series is a red-listed ecosystem according to the Conservation Data Centre (1999) of BC and it is mapped with about half the area of the 01b.

The ICHvk1 is dominated by site series 03 (HwCw – Falsebox – feathermoss) and 05 (CwSxw – Devil's Club – Horsetail) followed by site series 04 (CwHw – Oak fern – spiny wood fern) and. Avalanche chutes cover 15% of this subzone. Approximately 5% of the subzone is rock and talus.

The ICHmw3 is dominated by site series 05 (CwHw – Falsebox), 04 (CwFd – Soopolallie – Twinflower) and 05 (CwFd - Falsebox). There is very little rock and talus in this subzone, and a smaller proportion of avalanche chutes.

The remaining subzones ESSFvc, ESSFvcp and IDFun cover a small proportion of TFL 23. The IDFun is an important subzone from a biodiversity perspective, as it is a small subzone with restricted distribution.

Table 10. Area of Subzones and Site Series in TFL 23 Mapped By The Year Two Final PEM Model.

Subzone	Site Series Number	Site Series Name	Hectares	Percent of TFL 23
AT	01	Alpine heath – AH	2959.8	0.53%
AT	44	Talus – TA	3848.8	0.69%
AT	99	Rock outcrop – RO	14,599.8	2.62%
Total AT			**21,408.4**	**3.84%**
ESSFvc	01	BlHm – Rhododendron – Oak fern – MR	338.5	0.06%
ESSFvc	02	BlHm – Rhododendron – Leafy liverwort – ML	872.0	0.16%
ESSFvc	03	BlHm – Rhododendron – Pipecleaner moss – MP	53.5	0.01%
ESSFvc	04	BlHm – Devil's club – Lady fern – MD	10.0	0.0%
ESSFvc	05	BlHm – Horsetail – MH	1.8	0.0%
ESSFvc	44	Talus – TA	274	0.05%
ESSFvc	77	Avalanche chute – AC	531	0.10%
ESSFvc	99	Rock outcrop – RO	516.5	0.09%
Total ESSFvc			**2,597.3**	**0.47%**
ESSFvcp	01		5.0	0.0%
ESSFvcp	02	Bl – Heath – FH	375.3	0.07%
ESSFvcp	03	Juniper – Mountain Hairgrass – JM	9.8	0.0%
ESSFvcp	04	Sedge – Western pasque flower – SW	25.3	0.0%
ESSFvcp	44	Talus – TA	587.8	0.11%
ESSFvcp	77	Avalanche chute – AC	30.5	0.01%
ESSFvcp	99	Rock outcrop – RO	1326.3	0.24%
Total ESSFvcp			**2360.0**	**0.42%**
ESSFwc1	01	Bl – Rhododendron – Oak fern – FR	31,096.0	5.58%
ESSFwc1	02	Bl – Falsebox – Grouseberry – FF	20,286.8	3.64%
ESSFwc1	03	Bl – Devil's club – Lady fern – FD	885.0	0.16%
ESSFwc1	04	Bl – Horsetail – Brachythecium – FH	168.8	0.03%
ESSFwc1	05	Sedge – Sphagnum – SS	10.8	0.0%
ESSFwc1	44	Talus – TA	928.5	0.17%
ESSFwc1	77	Avalanche chute – AC	4,491.3	0.81%
ESSFwc1	99	Rock outcrop – RO	2,688.3	0.48%
Total ESSFwc1			**60555.5**	**10.86%**
ESSFwc4	01	Bl – Rhododendron – Oak fern – FR	54,352.8	9.75%
ESSFwc4	02	Bl – Rhododendron – Falsebox – FF	25,655.3	4.6%
ESSFwc4	03	Bl – Rhododendron – Woodrush – FW	22,945.1	4.12%
ESSFwc4	04	Bl – Rhododendron – Foamflower – RF	10,119.8	1.82%
ESSFwc4	05	Bl – Rhododendron – Lady fern – FL	41.5	0.01%
ESSFwc4	44	Talus – TA	6,227.3	1.12%
ESSFwc4	77	Avalanche chute – AC	7,829.0	1.4%
ESSFwc4	99	Rock outcrop – RO	16,146.5	2.9%
Total ESSFwc4			**143,317.3**	**25.71%**
ESSFwc4p	01	Mountain Heather – MH	111.8	0.02%
ESSFwc4p	02	Bl – Heath – FH	19,481.0	3.49%
ESSFwc4p	03	Juniper – Mountain Hairgrass – JM	1,112.3	0.2%

ESSFwc4p	04	Sedge – Western pasque flower – SW	1,679.5	0.3%
ESSFwc4p	44	Talus – TA	12,536.5	2.25%
ESSFwc4p	77	Avalanche chute – AC	2,582.3	0.46%
ESSFwc4p	99	Rock outcrop – RO	24,830.8	4.45%
Total ESSFwc4p			**62,334.2**	**11.18%**
ICHdw	01a	CwFd – Falsebox:sx-sm phase – RFa	20,954.5	3.76%
ICHdw	01b	CwFd – Falsebox:m-shg phase – RFb	6,145	1.1%
ICHdw	02	FdPy – Oregon-grape – Parsley fern – DO	3,063	0.55%
ICHdw	44	Talus – TA	148.5	0.03%
ICHdw	66	Willow – Sedge – WS	1.8	0.0%
ICHdw	99	Rock outcrop – RO	1.0	0.0%
Total ICHdw			**30,313.8**	**5.44%**
ICHmw2	01	HwCw – Falsebox – Feathermoss – HF	74,405.3	13.35%
ICHmw2	02	Rhacomitrium – Cladonia – RC	4,295.3	0.77%
ICHmw2	03	FdCw – Falsebox – Prince's pine – DF	29,248.6	5.25%
ICHmw2	04	CwFd – Falsebox – RF	33,013.3	5.92%
ICHmw2	05	CwHw – Oak fern Foamflower – HO	1,082.3	0.19%
ICHmw2	07	CwHw – Horsetail – RH	0.5	0.0%
ICHmw2	09	Bluejoint – Sedge – BS	2,484.5	0.45%
ICHmw2	44	Talus – TA	741.5	0.13%
ICHmw2	99	Rock outcrop – RO	7,472.3	1.34%
Total ICHmw2			**152,743.9**	**27.4%**
ICHmw3	01	HwCw – Falsebox – Feathermoss – HF	225.3	0.04%
ICHmw3	02	Fd – Juniper – Cladina – DJ	741.1	0.13%
ICHmw3	03	Fd – Pinegrass – Feathermoss – DP	335.3	0.06%
ICHmw3	04	CwFd – Soopolallie – Twinflower – RS	3,091.3	0.55%
ICHmw3	05	CwFd – Falsebox – RF	1,984.3	0.36%
ICHmw3	06	CwHw – Oak fern	619.8	0.11%
ICHmw3	08	CwSxw – Skunk cabbage – RC	56.0	0.01%
ICHmw3	09	Sedge – Sphagnum – SE	121.5	0.02%
ICHmw3	44	Talus – TA	113.8	0.02%
ICHmw3	99	Rock outcrop – RO	55.1	0.01%
Total ICHmw3			**7,344.0**	**1.32%**
ICHvk1	03	HwCw – Falsebox – Feathermoss – HF	6,221.8	1.12%
ICHvk1	04	CwHw – Oak fern – Spiny wood fern – HO	3,753.3	0.67%
ICHvk1	05	CwSxw – Devil's club – Horsetail – RC	5,737.8	1.03%
ICHvk1	44	Talus – TA	703.3	0.13%
ICHvk1	77	Avalanche chute – AC	2,869.0	0.51%
ICHvk1	99	Rock outcrop – RO	435	0.08%
Total ICHvk1			**19,722.6**	**3.54%**

ICHwk1	01	CwHw – Oak fern HO	11,227.6	2.01%
ICHwk1	02	Rhacomitrium – Cladonia RC	2,711.5	0.49%
ICHwk1	04	HwCw – Falsebox –Feather moss HF	35,263.1	6.33%
ICHwk1	05	CwHw – Devil's Club – Lady fern RD	4.0	0.0%
ICHwk1	06	Cw Swx – Devil's Club – Horsetail RH	1,272.5	0.23%
ICHwk1	07	Act – Dogwood – Twinberry CD	6.3	0.0%
ICHwk1	44	Talus TA	488.5	0.09%
ICHwk1	77	Avalanche Chute AC	704.5	0.13%
ICHwk1	99	Rock Outcrop RO	504.6	0.09%
Total ICHwk1			**52,182.6**	**9.36**
IDFun	02	Bluebunch wheatgrass – Junegrass BJ	356.0	0.06%
IDFun	03	FdPy – Bluebunch wheatgrass – Junegrass FW	417.0	0.07%
IDFun	05	Fd – Common Snowberry PS	71.8	0.01%
IDFun	06	Cw – Hooker's Fairybells RH	124.3	0.02%
IDFun	44	Talus TA	20.8	0.0%
IDFun	99	Rock outcrop RO	85.0	0.02%
Total IDFun			**1,074.9**	**0.19%**
Water			1,229.3	0.22%
TOTAL AREA			**557,418.4**	**100.00%**

3.4 WILDLIFE HABITAT CAPABILITY RATINGS

The ratings tables can be found in Appendix VII. They describe the range of habitat capability found over the 475 site series/modifier combinations that can potentially exist in TFL 23. We were not contracted to produce capability mapping, which is the most effective way to assess the results of capability assessment by species and season.

4.0 DISCUSSION

4.1 RASTER BASED PEM MODELING

The utility of a raster based PEM model is evident. We were allocated a finite amount of time to develop the PEM and to produce and assess a site series map. The raster model permitted flexibility and efficiency for processing the PEM model. The raster approach facilitates the development of input layers, which would not have been possible in a polygon-based approach. It would have been impossible to derive toe slopes, landscape shape, Landsat classifications or stream density from a polygon based PEM model. The downside of the raster-based approach is the "blocky" look of the resultant map. The output can be vectorized, with some loss of resolution, to accommodate polygon-based mapping systems.

4.2 ACCURACY OF PEM MODEL

The model's accuracy is not high, but comparable to the results of other PEM initiatives within BC (Meidinger pers. Com. March 2001). It could be improved in wetter and extremely dry sites. The sampling for accuracy assessment did not include avalanche paths, rock outcrops, talus, or wetlands. The sampling intensity in wet sites was low. The model could be improved with more data from these types of sites and more time to "game" with the results of the seventh iteration in mid slope sites.

4.3 NEURAL NETWORKS AND MULTIPLE LOGISTIC REGRESSION AS A TOOL FOR PEM

This project has provided an interesting opportunity to explore the utility of statistical analyses using both plot and GIS attributes as variables in relationship to the development of a site series classification model for use in PEM.

Early iterations of the model relied heavily on the year one results for multiple logistic regression and classification tree relationships for the weighting of GIS input layer variables in the knowledge bases. Using these results alone reduced the level of model goodness of it in many cases. The knowledge bases still require a combination of qualitative input from an ecologist/mapper familiar with the distribution of site series on the ground and the results of multivariate statistical analysis.

The neural network tended to over fit the PEM input layer data to the site series classified by the neural network. This was reflected in the poor accuracy scores relative to the scores based on the test plots from the neural network. It would be an interesting exercise to let the neural network derive its own classification of pixels, rather than dictate the site series classification as the "correct" answer that the network must achieve. The neural network classification could then be related back to field plots and similarities between plots summarized and developed into a new classification.

5.0 REFERENCES

Akaike, H. 1983. Information measures and model selection. Bulletin of the international statistical institute: proceedings of the 44th session. 1:277-290.

Bishop, C. 1995. Neural networks for pattern recognition. Oxford University Press, Oxford, UK.

Braumandl, T.F. and Curran, M. 1992. A field Guide for Site Identification and Interpretation for the Nelson Forest Region. Land Management Handbook NUMBER 20. BC Ministry of Forests.

British Columbia Ministry of Environment, Lands and Parks, BC Ministry of Forests. 1998 Field Manual for Describing Terrestrial Ecosystems. Land Management Handbook NUMBER 25.

Conservation Data Centre. 2000. Rare Natural Plant Community Tracking List, Arrow Forest District. August. www.elp.gov.bc.ca/rib/wis/cdc/ecology.htm

Demarchi, D.A. 1996. An Introduction to the Ecosections of British Columbia. Pamphlet for BC Ministry of Environment Lands and Parks. Victoria, BC.

Howes, D.E. and E. Kenk (Ed.). 1997. Terrain Classification System for British Columbia. MOE Manual 10 (Version2.) Recreational Fisheries Branch, Ministry of Environment and Surveys and Resource Mapping Branch, Ministry of Crown Lands, Victoria, BC.

Huggard, D. 2000. Grouping Nelson Region Site Series Based on Indicator Species. Unpublished report to the Arrow IFPA.

Huggard, D. 2001. Habitat Attributes in the Arrow IFPA, Non Harvestable Landbase. Unpublished report to the Arrow IFPA.

Jungen, J.R., 1980. RAB Bulletin 20: Soil Resources of the Nelson Map Area (82 F). British Columbia Soil Survey, report No. 28. Queen's Printer, Victoria, British Columbia.

Ketcheson, M.V., G. Smith and S. Wilson. 2000. Arrow TSA Site Index Adjustment Project, Predictive Ecosystem Mapping (PEM) Year One Report. Unpublished report to the Arrow IFPA.

Ketcheson, M.V., T. Dool and S. Wilson. 2001 Arrow TSA Site Index Adjustment

Project, Predictive Ecosystem Mapping (PEM) Year Two Final Project Report. Unpublished report to the Arrow IFPA.

Ketcheson, M.V. and D. Marcoux. 1994. Valhalla Provincial Park Biophysical Mapping Project. Unpublished Maps and Report to BC Parks, West Kootenay District Office. Nelson, BC.

Ketcheson, M.V. and D. Marcoux. 1995. Expanded Biophysical Ecosystem Legend: Deer Creek Watershed. Unpublished report to Arrow Forest District, BC Ministry of Forests, Castlegar, BC.

Kumar, L., Skidmore, A.K.,Knowles, E., (1997) Modeling Topographic variation in Solar Radiation in a GIS Environment. International Journal of Geographical Information Science, 11(5): 475-497

Lavkulich, L.M. and K.W.G. Valentine, 1978. The soil Landscapes of British Columbia: The Canadian System of Soil and Soil Climate Classification. Ministry of Environment.

Little, H.W., 1960. Memoir 308, Nelson Map-Area, West Half, British Columbia (82 F W1/2). Geological Survey of Canada, Department of Energy, Mines and Resources. Crown Publications.

Lloyd, D. K Angrove, G. Hope, C. Thompson. 1990. A Guide to Site Identification and Interpretation for the Kamloops Forest Region. Land Management Handbook NUMBER 23. Parts 1 and 2. BC Ministry of Forests.

Loh, W. –Y., and Y. –S. Shih. 1997. Split selection methods for classification trees. Statistica Sinica 7:815–840.

Marcoux, D., M. Mathers, J. Riddell, M.V. Ketcheson. 1996. Expanded Biophysical Ecosystem Legend: Wilson-Stagleap South Slopes. Unpublished report to Columbia Basin Fish and Wildlife Compensation Program, Nelson, BC.

Meidinger, D. 2001. Personal communication.

Meidinger, D. 2000. Protocol for Quality Assurance and Accuracy Assessment of Ecosystem Maps. Prepared for the TEM Alternatives Task Force. July 2000.

Menard, S. 1995. Applied logistic regression analysis. Series: quantitative applications in the social sciences 106. Sage University Press, Thousand Oaks, CA.

Moon, D., D. Clark, D. Meidinger, and K. Jones. 1999. Standards for Predictive Ecosystem Mapping – Inventory Standard – Review Draft. Version 1.0 November 1999.

Nigh, G., D. Meidinger, and A. Mirza. 1998. SIBEC Sampling and Data Standards. BC Ministry of Forests Miscellaneous Report 075. Version 4.0, Revised November 18, 1998.

PEM Data Committee 2000. Standards for Predictive Ecosystem Mapping (PEM) – Digital Data Capture. Predictive Ecosystem Technical Standards and Data Base Manual. PEM Data Committee for the TEM Alternatives Task Force, Resource Inventory Committee.

Resources Inventory Committee 1999. British Columbia Wildlife Habitat Ratings Standards. Version 2.0. Prepared by the Ministry of Environment, Lands and Parks, Resources Inventory Branch.

Statistica. 1999. Statistica for Windows. Statsoft, Inc. Tulsa, OK.

Statistica. 2000. Statistica neural networks. Statsoft, Inc. Tulsa, OK.

TEM Alternatives Task Force. 1999. Standards for Predictive Ecosystem Mapping. Inventory Standard. Prepared by the TEM Alternatives Task Force for the Resource Inventory Committee. Version 1.0 November 1999.

TEM Alternatives website http://www.for.gov.bc.ca/research/TEMalt)

USGS 1999. NPS Vegetation Mapping Program. Section 5.0 Computational Issues. 5.1 Definitions of Precision and Accuracy. http:/biology/usgs/gov/npsveg/aa/sect5.html

Wheeler, J.O. and McFeely, P, 1991. Tectonic Assemblage Map of the Canadian Cordillera and adjacent parts of the United States of America. Geological Survey of Canada, Map 1712A.